

Perbandingan Peramalan Penjualan Mobil di Indonesia Menggunakan SARIMAX, Time Series Regression, dan Hybrid Models

Asyifa Izayani¹⁾, Aisyah Mufidah Najwa²⁾, Ardini Aprilya Putri³⁾.

Program Studi Sains Data, Fakultas Informatika, Universitas Telkom Surabaya

Jl. Ketintang No.156, Surabaya 60231, Indonesia

asyzyni@student.telkomuniversity.ac.id¹⁾, aisyah.najwa@student.telkomuniversity.ac.id²⁾,

ardngrthhy@student.telkomuniversity.ac.id³⁾

ABSTRAK

Industri otomotif nasional merupakan salah satu sektor strategis yang memiliki dinamika penjualan bersifat fluktuatif dan musiman, sehingga peramalan permintaan yang akurat menjadi tantangan penting, terutama setelah terjadinya perubahan struktur akibat pandemi COVID-19. Penelitian ini bertujuan untuk membandingkan kinerja model Time Series Regression (TSR), Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX), dan Hybrid SARIMA–Neural Network dalam meramalkan penjualan mobil bulanan di Indonesia. Data yang digunakan merupakan data pendaftaran kendaraan bulanan yang diperoleh dari Trading Economics dengan cakupan periode jangka panjang dan jangka pendek. Data jangka panjang digunakan sebagai skenario utama untuk menangkap pola musiman dan dinamika jangka panjang, sedangkan data jangka pendek digunakan sebagai pembanding untuk menganalisis pengaruh panjang periode data terhadap kinerja model. Pemodelan SARIMAX dilakukan menggunakan pendekatan Box–Jenkins, sementara model hybrid dibangun dengan memodelkan residual SARIMAX menggunakan Neural Network tipe Multilayer Perceptron. Evaluasi kinerja model dilakukan pada data uji menggunakan metrik kesalahan berbasis absolut dan persentase. Hasil penelitian menunjukkan bahwa TSR memberikan performa yang lebih baik pada data dengan karakteristik yang relatif homogen, sedangkan SARIMAX dan Hybrid SARIMA–Neural Network menunjukkan kinerja yang lebih unggul pada data jangka panjang. Model hybrid secara konsisten menghasilkan tingkat kesalahan yang paling rendah dan residual yang bersifat acak berdasarkan uji diagnostik, yang mengindikasikan kemampuan model dalam menangkap pola linear, musiman, dan non-linear secara lebih komprehensif. Temuan ini menunjukkan bahwa pendekatan hybrid lebih efektif untuk peramalan penjualan mobil pada data deret waktu dengan dinamika yang kompleks.

Kata kunci: peramalan deret waktu, SARIMAX, time series regression, hybrid SARIMA–neural network, penjualan mobil Indonesia

1. Pendahuluan

Industri otomotif nasional merupakan salah satu sektor manufaktur strategis yang menjadi barometer pertumbuhan ekonomi Indonesia. Sebagai pasar kendaraan bermotor terbesar di kawasan ASEAN, volume penjualan mobil mencerminkan daya beli rumah tangga dan stabilitas iklim investasi [1], [2]. Berdasarkan data historis pendaftaran mobil yang dihimpun dari Trading Economics selama satu dekade terakhir (2015–2025), industri ini menunjukkan dinamika yang sangat fluktuatif namun memiliki pola musiman yang kuat [3]. Selama periode tersebut, angka penjualan rata-rata berada pada kisaran 78.722 unit per bulan. Rekor tertinggi tercatat pada Juli 2018 dengan angka mencapai 107.474 unit, yang menunjukkan masa kejayaan ekspansi pasar otomotif sebelum terjadinya pandemi global. Namun, titik terendah terjadi pada Mei 2020 dengan hanya 3.551 unit akibat kontraksi ekonomi yang masif, sebelum akhirnya mulai menunjukkan pemulihan bertahap pada periode 2021 hingga 2025 [3].

Fluktuasi yang ekstrem ini, terutama akibat guncangan pandemi dan pemulihan ekonomi yang diikuti oleh kebijakan insentif pajak seperti PPnBM DTP, menciptakan tantangan besar dalam memprediksi permintaan pasar di masa depan [2], [4]. Ketidakpastian angka penjualan seringkali menyebabkan ketimpangan antara ketersediaan stok di dealer dengan permintaan riil konsumen. Menurut penelitian oleh Caesar *et al.* [5], volatilitas yang tinggi pada data otomotif di Indonesia membuat model peramalan linear konvensional seringkali gagal menangkap pola non-linear yang muncul saat terjadi perubahan mendadak di pasar. Oleh karena itu, diperlukan pendekatan peramalan yang lebih komprehensif untuk membantu produsen dan distributor dalam melakukan perencanaan strategis berbasis data guna menghindari inefisiensi biaya operasional [6].

Dalam literatur sains data lima tahun terakhir, perdebatan mengenai metode peramalan terbaik terus berkembang [7]. Metode Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX) secara luas diakui efektivitasnya dalam menangani data deret waktu yang memiliki unsur musiman tahunan dan pengaruh variabel luar [8], [9]. Di sisi lain, Time Series Regression tetap menjadi instrumen penting untuk memetakan tren pertumbuhan secara linear [10]. Namun, studi terbaru yang dilakukan oleh Syaharani *et al.* [11] dan Wahidah *et al.* [12] menunjukkan bahwa integrasi antara metode statistik tradisional dengan kecerdasan buatan, atau yang dikenal sebagai *Hybrid Models*, mampu memberikan nilai kesalahan (Mean Absolute Percentage Error) yang jauh lebih rendah. Model hibrida dianggap lebih unggul karena mampu membedah komponen linear dan non-linear dari data penjualan secara simultan [13], [14].

Berdasarkan urgensi tersebut, penelitian ini dilakukan untuk membandingkan performa model SARIMAX, Time Series Regression, dan *Hybrid Models* dalam meramalkan angka pendaftaran mobil di Indonesia menggunakan data rentang panjang dari tahun 2015 hingga 2025. Dengan cakupan data yang meliputi masa sebelum, selama, dan sesudah pandemi, penelitian ini diharapkan dapat mengidentifikasi model peramalan yang paling tangguh (*robust*) terhadap perubahan kondisi pasar [7], [15]. Hasil dari analisis ini diharapkan dapat menjadi rujukan ilmiah bagi pelaku industri otomotif dan pengambil kebijakan dalam menyusun proyeksi bisnis yang akurat di tengah dinamika ekonomi nasional yang terus berkembang.

2. Metodologi

2.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah angka penjualan mobil bulanan di Indonesia yang diperoleh melalui platform Trading Economics [3]. Dataset ini mencakup angka penjualan unit kendaraan bermotor yang dilaporkan secara periodik, yang berfungsi sebagai indikator utama aktivitas ekonomi sektor otomotif nasional [1], [2].

Penelitian ini menggunakan dataset selama sepuluh tahun, yaitu periode 2015 hingga 2025, sebagai dasar utama pemodelan. Pemilihan periode ini bertujuan untuk menangkap dinamika pasar otomotif secara lebih menyeluruh, termasuk pola musiman tahunan serta perubahan kondisi ekonomi sebelum, selama, dan setelah pandemi COVID-19 [6], [16]. Menurut Hyndman dan Athanasopoulos [6], volume data yang lebih besar memberikan peluang lebih baik untuk analisis eksplorasi, pengujian model, dan keandalan peramalan. Selain itu, periode yang lebih panjang memungkinkan model untuk menangkap pola musiman yang berulang secara tahunan serta tren jangka panjang dengan lebih akurat [7], [17].

Untuk mengevaluasi pengaruh volume data terhadap performa model, digunakan pula dataset lima tahun terakhir (2021–2025) sebagai perbandingan. Pendekatan ini bertujuan untuk mengevaluasi pengaruh panjang periode data terhadap kinerja model peramalan [17], [18]. Studi oleh Cerqueira *et al.* [18] menunjukkan bahwa pada data deret waktu non-stasioner, metode *out-of-sample* seperti *holdout* yang diulang pada beberapa periode pengujian menghasilkan estimasi performa yang lebih akurat. Dengan membandingkan performa model pada dua rentang waktu yang berbeda, penelitian ini dapat mengidentifikasi sensitivitas model terhadap perubahan struktural dalam data, khususnya dampak disruptif pandemi yang signifikan pada tahun 2020 [15], [16].

2. 2 Analisis Awal Deret Waktu

Sebelum dilakukan pemodelan, dilakukan analisis karakteristik statistik guna memahami sifat fundamental dari data penjualan otomotif yang volatil. Analisis visual dan statistik menunjukkan adanya perubahan struktur (structural break) yang signifikan pada periode 2020–2021, yang berkaitan dengan dampak pandemi COVID-19 terhadap sektor otomotif. Meskipun data jangka panjang mengandung variasi pola yang lebih heterogen, penggunaan dataset sepuluh tahun tetap dipertahankan karena menyediakan informasi historis yang lebih lengkap serta jumlah siklus musiman yang memadai untuk pemodelan deret waktu. Sementara itu, dataset lima tahun menunjukkan pola yang relatif lebih homogen secara temporal dan digunakan sebagai pembanding dalam analisis perbandingan periode data.

2. 3 SARIMAX

Model yang digunakan pada penelitian ini adalah Seasonal Autoregressive Integrated Moving Average with eXogeneous factors (SARIMAX). Model ini merupakan perluasan dari model Autoregressive Integrated Moving Average (ARIMA) yang dirancang untuk menangani data deret waktu (time series) yang mengandung pola musiman (seasonality) [7], [8].

Model SARIMAX dinotasikan sebagai SARIMAX(p,d,q)(P,D,Q)s. Notasi tersebut merepresentasikan komponen non-musiman (p,d,q) dan komponen musiman (P,D,Q) dengan periode musiman s [8], [19]. Struktur matematis model ini dapat dituliskan menggunakan operator mundur (Backshift Operator) B sebagai berikut [7]:

$$\phi_p(B)\Phi_P(B^s)(1 - B)^d(1 - B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)\epsilon_t$$

Dimana :

- y_t adalah nilai pengamatan pada waktu ke-t
- ϵ_t adalah residual atau *white noise* pada waktu t yang diasumsikan berdistribusi normal $N(0, \sigma^2)$
- B adalah *Backshift Operator*, yang didefinisikan sebagai $B^k y_t = y_{t-k}$
- $(1 - B)^d$ adalah komponen pembedaan (*differencing*) non-musiman dengan orde d untuk menstasionerkan data
- $(1 - B^s)^D$ adalah komponen pembedaan musiman dengan orde D dan periode s.

Komponen *Autoregressive* (AR) dan *Moving Average* (MA) dalam persamaan di atas didefinisikan sebagai polinomial terhadap B sebagai berikut[7],[19]:

1. Komponen Autoregressive (AR) :

- Non-musiman (p) : Menjelaskan hubungan antara nilai saat ini dengan nilai masa lalunya.

$$\phi_p(B) = 1 - \phi_1(B) - \phi_2 B^2 - \cdots - \phi_p B^p$$
- Musiman (P) : Menjelaskan hubungan antara nilai saat ini dengan nilai pada periode musiman sebelumnya.

$$\Phi_P(B^S) = 1 - \Phi_1(B^S) - \Phi_2(B^{2S}) - \cdots - \Phi_P(B^{pS})$$

2. Komponen Moving Average (MA) :

- Non-Musiman (q) : Menjelaskan hubungan antara nilai saat ini dengan error dari masa lalu.

$$\theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$$

- Musiman (Q) : Menjelaskan hubungan antara nilai saat ini dengan *error* pada periode musiman sebelumnya.

$$\Theta_Q(B^S) = 1 + \Theta_1(B^S) + \Theta_2(B^{2S}) + \cdots + \Theta_Q(B^{QS})$$

Model SARIMAX memungkinkan pemodelan deret waktu yang mempertimbangkan pola musiman, tren, serta ketergantungan residual melalui komponen AR dan MA [8], [9]. Dengan struktur pembeda musiman dan non-musiman, model ini dirancang untuk menghasilkan data yang stasioner sehingga estimasi parameter dan proses peramalan dapat dilakukan secara lebih akurat dan konsisten [6], [20].

2. 3. 1 Uji Stasioneritas

Tahap pertama yang perlu dilakukan dalam proses pemodelan time series menggunakan SARIMAX adalah pemeriksaan kestabilan data. Data deret waktu dikatakan stasioner apabila nilai rata-rata (mean) dan variansnya konstan sepanjang waktu [6], [7]. Untuk menguji stasioneritas, digunakan uji akar unit Augmented Dickey-Fuller (ADF) [21], [22].

Persamaan regresi untuk uji ADF didefinisikan sebagai [20], [21]:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{j=1}^p \delta_j \Delta y_{t-j} + \varepsilon_t$$

Dimana :

- Δy_t : Perubahan (*first difference*) dari variabel y_t
- α : Konstanta (intercept) yang merepresentasikan nilai rata-rata (drift) dari proses deret waktu.
- βt : Komponen tren deterministik. Komponen ini bersifat opsional dan digunakan jika data diduga memiliki tren linear terhadap waktu.
- γy_{t-1} : Komponen utama dalam uji ADF. Parameter γ digunakan untuk menguji keberadaan akar unit.
- $\sum_{j=1}^p \delta_j \Delta y_{t-j}$: Penjumlahan lag dari perubahan variabel hingga orde ke- p . Tujuannya adalah menghilangkan autokorelasi pada residual sehingga error bersifat *white noise*. [8]
- ε_t : Komponen error yang diasumsikan bersifat *white noise* dengan rataan nol dan varians konstan.

Pengujian dilakukan dengan hipotesis [21],[20] :

- $H_0: \gamma = 0$ (Data tidak stasioner)
- $H_1: \gamma < 0$ (Data stasioner)

Jika nilai probabilitas (*p-value*) dari statistic uji ADF kurang dari taraf signifikansi $\alpha = 0.05$, maka H_0 ditolak, yang berarti data sudah stasioner. [22].[23]

2. 3. 2 Differencing Musiman dan Non-musiman

Data musiman dengan trend memiliki dua bentuk non-stasioneritas yang terpisah, sehingga uji stasioneritas merupakan langkah krusial karena komponen AR dan MA hanya dapat diterapkan pada data yang stasioner [6], [7]. Ketika data yang non-stasioner digunakan tanpa penyesuaian, maka hasil

peramalan akan menjadi *spurious* (keliru), dan interval prediksi menjadi terlalu lebar [10], [15]. Differencing musiman mengatasi pola musiman, sementara differencing non-musiman mengatasi trend keseluruhan [7], [18].

- Differencing Non-Musiman (d)

Differencing non-musiman dibutuhkan ketika trend non-stasioneritas yang disebabkan oleh tren stokastik, yaitu adanya pola kenaikan atau penurunan jangka panjang dalam data[19]. Differencing non-musiman (d) mengatasi ini dengan menghitung *first difference*[6].

$$\Delta y_t = (1 - B)y_t = y_t - y_{t-1}$$

Operasi ini menghilangkan tren keseluruhan dengan mengubah fokus dari nilai absolut menjadi perubahan antar periode berturut-turut. Sehingga, rata-rata (*mean*) menjadi lebih stabil dan varians (*variance*) menjadi lebih konsisten[15].

- Differencing musiman (D)

Berbalik dari differencing non-musiman, differencing musiman berasal dari pola perilaku yang berbeda setiap musim. Hal ini kemudian menghasilkan rata-rata yang tidak konsisten antar bulan dan varians yang bervariasi secara musiman. Differencing musiman (D) mengatasi ini dengan membandingkan nilai yang sama dari musim berbeda[7],[18].

$$\Delta_s y_t = (1 - B^s)y_t = y_t - y_{t-s}$$

Pendekatan ini secara langsung menargetkan komponen periodik dengan menghilangkan seasonal *random walk type non-stationarity* [8],[23].

2. 3. 3 Identifikasi Model

Setelah data mencapai stasioneritas melalui differencing, langkah krusial berikutnya adalah mengidentifikasi parameter p, q untuk komponen non-musiman dan P, Q untuk komponen musiman. Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF) adalah alat fundamental untuk proses ini karena keduanya memberikan pola karakteristik yang secara langsung menunjukkan jenis komponen AR dan MA yang diperlukan [6], [7], [25].

- *Autocorrelation Function (ACF)*

ACF menghitung korelasi linier antara suatu variabel deret waktu (y_t), dengan nilai masa lalunya (y_{t-k}) pada lag ke- k [26]. Nilai ACF memberikan gambaran seberapa kuat pengaruh data masa lalu terhadap data saat ini secara langsung, tanpa memperhitungkan variabel perantara[25].

Secara matematis ACF dinotasikan sebagai berikut[7] :

$$r_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}$$

Dimana :

- r_k : nilai autokorelasi pada lag k
- y_t : nilai data pada waktu t
- \bar{y} : rata-rata seluruh data
- n : jumlah observasi

- *Partial Autocorrelation Function (PACF)*

Partial Autocorrelation Function (PACF) mengukur korelasi antara y_t dan y_{t-k} setelah menghilangkan pengaruh linier dari semua variabel antara ($y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$). Nilai PACF pada lag- k dapat diestimasi sebagai koefisien regresi ϕ_k dalam model autoregresif orde k [25] :

$$y_t = \phi_{k1}y_{t-1} + \phi_{k2}y_{t-2} + \cdots + \phi_{kk}y_{t-k} + \varepsilon_t$$

2.3.4 Estimasi Parameter

Estimasi parameter bertujuan untuk mendapatkan koefisien model terbaik serta orde $(p, d, q)(P, D, Q)_s$ yang paling optimal. Dalam penelitian ini, estimasi dilakukan menggunakan metode *Grid Search* dengan mengevaluasi kombinasi parameter berdasarkan nilai *Akaike Information Criterion* (AIC)[20],[27].

Fungsi objektifnya adalah meminimalkan nilai AIC yang dirumuskan sebagai[28] :

$$AIC = 2k - 2 \ln (\hat{L})$$

Dimana k adalah jumlah parameter dalam model dan \hat{L} adalah nilai maksimum fungsi *likelihood*. Model dengan nilai AIC terendah dipilih sebagai model terbaik karena dianggap paling efisien dalam menyeimbangkan *goodness of fit* dan kompleksitas model (*parsimony principle*)[27].

2.3.5 Uji Diagnostik Residual

Tahap terakhir adalah verifikasi model (diagnostik) untuk memastikan bahwa model yang terbentuk sudah memadai. Pemeriksaan dilakukan terhadap sisaan (*residual*) model menggunakan uji *Ljung-Box Q-Statistic*. Uji ini bertujuan untuk mengetahui apakah terdapat autokorelasi pada residual[26].

Statistik uji Q dihitung dengan rumus[29] :

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{p}_k^2}{n-k}$$

Hipotesis yang digunakan adalah :

- H_0 : Residual bersifat acak
- H_1 : Terdapat autokorelasi pada residual

Model dikatakan layak digunakan untuk peramalan apabila nilai *p-value* dari uji Ljung-Box lebih besar dari $\alpha = 0.05$ (gagal menolak H_0), yang mengindikasikan bahwa residual tidak memiliki pola yang tertinggal[29].

2.4 Time Series Regression (TSR)

Dalam penelitian ini, Time Series Regression (TSR) diimplementasikan sebagai pendekatan pembelajaran terbimbing (*supervised learning*) yang memanfaatkan fitur autokorelatif untuk memodelkan dependensi temporal pada data penjualan kendaraan bermotor [6], [30], [31]. TSR mentransformasi permasalahan peramalan menjadi masalah regresi linear dengan menggunakan observasi masa lalu (*lagged observations*) sebagai variabel prediktor utama [30], [32].

a. Konstruksi Variabel Lag

Fitur input dibentuk melalui proses *windowing*, di mana nilai pada waktu t diprediksi menggunakan n nilai sebelumnya[30]. Dalam model ini digunakan $n = 12$ (1 tahun) untuk menangkap siklus musiman penjualan mobil di Indonesia. Matriks fitur X didefinisikan sebagai:

$$X = [Y_{t-1}, Y_{t-2}, \dots, Y_{t-12}]$$

b. Estimasi Model Regresi Linear

Model dilatih menggunakan metode *Ordinary Least Squares* (OLS) untuk meminimalkan *residual sum of squares*[33]. Persamaan prediksinya adalah:

-

$$\hat{Y}_t = \beta_0 + \sum_{i=1}^{12} \beta_i Y_{t-i} + \varepsilon_t$$

Di mana β_0 adalah konstanta, β_i adalah koefisien bobot untuk masing-masing *lag*, dan ε_t adalah galat.

c. Strategi Peramalan Rekursif

Untuk perbandingan jangka panjang, TSR menggunakan pendekatan *Iterative Multi-step Forecasting*. Prediksi pada langkah $t - 1$ diumpulkan kembali sebagai input untuk langkah $t - 2$ [34] :

$$\hat{Y}_{t-h} = f(\hat{Y}_{t+h-1}, \dots, \hat{Y}_{t+1}, Y_t, \dots)$$

d. Interpretasi Koefisien dan Signifikansi Fitur (*Feature Importance*)

Dalam model TSR, setiap koefisien regresi (β_i) memberikan wawasan mengenai kontribusi relatif dari setiap *lag* waktu terhadap prediksi penjualan mobil saat ini. Analisis ini penting untuk memahami dinamika pasar otomotif di Indonesia[33]:

- Autokorelasi Jangka Pendek: Koefisien pada *Lag_1* hingga *Lag_3* biasanya mencerminkan momentum penjualan dari bulan-bulan sebelumnya.
- Efek Musiman (Seasonal Effect): Signifikansi yang tinggi pada *Lag₁₂* menunjukkan adanya pola tahunan yang kuat, seperti peningkatan penjualan menjelang hari raya atau pameran otomotif tahunan.
- Signifikansi Statistik: Setiap koefisien diuji menggunakan *t-test* untuk menentukan apakah variabel *lag* tersebut memiliki pengaruh yang signifikan secara statistik terhadap variabel dependen (*p-value* < 0,05).

2.5 Hybrid SARIMA-Neural Network

2.5.1 Konsep Dasar Model Hybrid

Model hybrid SARIMA-Neural Network (SARIMA-NN) digunakan untuk meningkatkan akurasi peramalan dengan mengombinasikan kelebihan model statistik dan model pembelajaran mesin. Model SARIMA efektif dalam menangkap pola linear dan musiman pada deret waktu, namun memiliki keterbatasan dalam memodelkan hubungan non-linear. Sebaliknya, Neural Network memiliki kemampuan yang baik dalam memodelkan pola non-linear, tetapi kurang optimal dalam menangkap struktur musiman secara eksplisit.

Pendekatan hybrid didasarkan pada asumsi bahwa deret waktu y_t terdiri atas komponen linear dan non-linear, sehingga dapat dinyatakan sebagai [37]:

$$y_t = L_t + N_t$$

di mana L_t merepresentasikan komponen linear-musiman yang dimodelkan menggunakan SARIMA, dan N_t merepresentasikan komponen non-linear yang dimodelkan menggunakan Neural Network.

2.5.2 Pemodelan Residual SARIMA

Tahap awal pemodelan hybrid dilakukan dengan membangun model SARIMA terbaik menggunakan prosedur Box-Jenkins. Model SARIMA menghasilkan nilai prediksi $\hat{y}_t^{\text{SARIMA}}$ yang merepresentasikan komponen linear dan musiman dari data penjualan mobil.

Residual SARIMA dihitung sebagai selisih antara nilai aktual dan hasil prediksi SARIMA, yaitu:

$$e_t = y_t - \hat{y}_t^{SARIMA}$$

Residual ini diasumsikan masih mengandung pola non-linear yang belum tertangkap oleh model SARIMA dan selanjutnya digunakan sebagai target pemodelan Neural Network [37], [38].

2.5.3 Pemodelan Non-Linear Menggunakan Neural Network

Residual SARIMA ditransformasikan menjadi data supervised dengan menggunakan beberapa lag residual sebagai variabel input. Penggunaan lag residual bertujuan untuk menangkap ketergantungan temporal jangka pendek serta pola musiman yang masih tersisa pada residual.

Neural Network yang digunakan adalah feedforward neural network tipe Multilayer Perceptron (MLP), yang dilatih untuk mempelajari hubungan non-linear antara residual masa lalu dan residual saat ini. Secara umum, model Neural Network dapat dituliskan sebagai:

$$\hat{e}_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-k})$$

Pendekatan pemodelan residual menggunakan Neural Network ini telah banyak digunakan dalam studi hybrid SARIMA-ANN dan terbukti mampu meningkatkan kinerja peramalan dibandingkan model tunggal [37], [38].

2.5.4 Pembentukan Model Hybrid

Prediksi akhir model hybrid diperoleh dengan menjumlahkan hasil peramalan SARIMA dan prediksi residual Neural Network, sehingga dirumuskan sebagai:

$$\hat{y}_t^{Hybrid} = \hat{y}_t^{SARIMA} + \hat{e}_t^{NN}$$

Untuk peramalan multi-periode ke depan, prediksi residual dilakukan secara rekursif dengan memanfaatkan nilai residual sebelumnya. Dengan pendekatan ini, model hybrid diharapkan mampu menangkap baik pola linear-musiman maupun non-linear pada data penjualan mobil di Indonesia secara lebih optimal dibandingkan penggunaan model SARIMA atau Neural Network secara terpisah [37], [38].

2.6 Evaluasi Model (MAE, RMSE, dan MAPE)

Dalam penelitian ini, digunakan tiga metrik kesalahan (error metrics) yang umum digunakan dalam evaluasi model peramalan deret waktu, yaitu Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan Mean Absolute Percentage Error (MAPE). Ketiga metrik ini dipilih karena mampu memberikan gambaran yang komprehensif mengenai besar kesalahan prediksi dalam satuan absolut, kuadrat, maupun persentase relatif.

2.6.1 Mean Absolute Error (MAE)

Mean Absolute Error (MAE) mengukur rata-rata nilai absolut dari selisih antara nilai aktual dan nilai hasil peramalan. MAE memberikan gambaran langsung mengenai besar kesalahan prediksi tanpa

mempertimbangkan arah kesalahan, sehingga mudah diinterpretasikan dalam satuan yang sama dengan data asli.

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

di mana y_t adalah nilai aktual, \hat{y}_t adalah nilai hasil peramalan, dan n adalah jumlah observasi pada data uji.

2.6.2 Root Mean Squared Error (RMSE)

Root Mean Squared Error (RMSE) mengukur akar dari rata-rata kuadrat kesalahan prediksi. RMSE memberikan penalti yang lebih besar terhadap kesalahan prediksi yang bernilai besar, sehingga metrik ini sensitif terhadap outlier dan kesalahan ekstrem.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

RMSE digunakan untuk menilai kestabilan dan konsistensi model dalam menghasilkan prediksi yang mendekati nilai aktual.

2.6.3 Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) mengukur rata-rata kesalahan prediksi dalam bentuk persentase relatif terhadap nilai aktual. Metrik ini memudahkan perbandingan kinerja model karena bersifat bebas skala (scale-independent).

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

MAPE digunakan untuk memberikan interpretasi kesalahan dalam bentuk persentase, sehingga memudahkan analisis performa model dalam konteks praktis.

2.6.4 Skema Perbandingan Model

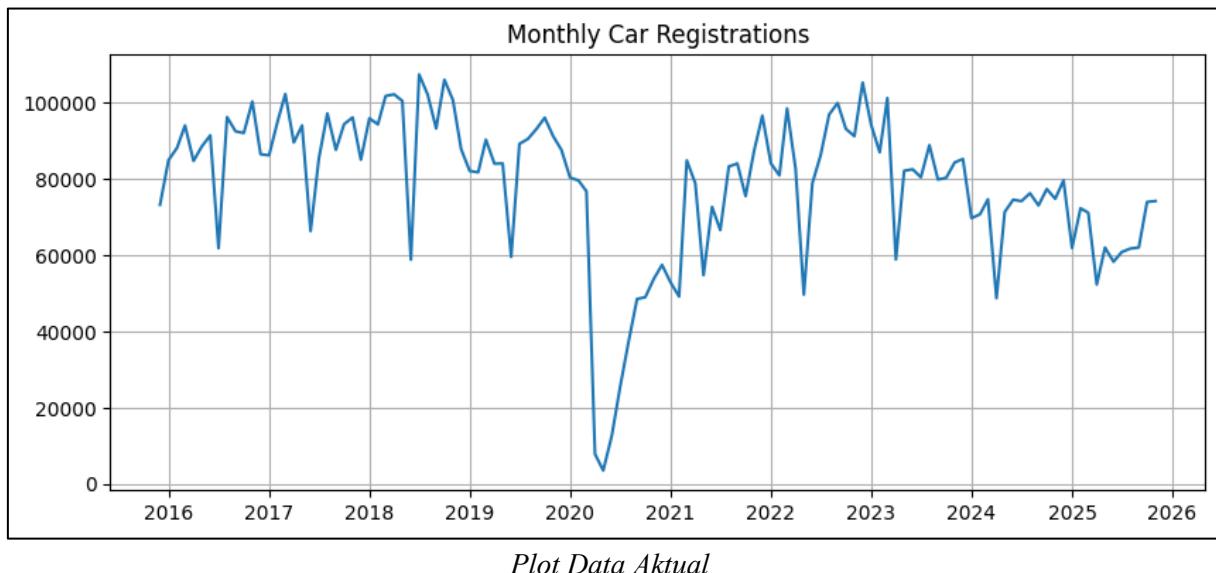
Nilai MAE, RMSE, dan MAPE dihitung untuk masing-masing model pada periode pengujian yang sama. Model dengan nilai MAE, RMSE, dan MAPE yang lebih kecil dianggap memiliki kinerja peramalan yang lebih baik. Hasil evaluasi ini digunakan sebagai dasar dalam membandingkan performa Time Series Regression, SARIMAX, dan Hybrid SARIMA–Neural Network dalam meramalkan penjualan mobil di Indonesia.

3. Hasil and Pembahasan

3.1 Gambaran Umum Data dan Pola Deret Waktu

Data yang digunakan dalam penelitian ini merupakan data penjualan mobil bulanan di Indonesia yang diperoleh dari Trading Economics. Penelitian ini menggunakan data 10 tahun untuk melatih seluruh model agar dapat menangkap pola musiman jangka panjang secara konsisten, sedangkan pada skenario

pembanding digunakan data lima tahun terakhir (2021–2025) untuk melihat pengaruh periode pasca-pandemi COVID-19 terhadap kinerja model peramalan.



Berdasarkan visualisasi deret waktu bulanan, data penjualan mobil menunjukkan fluktuasi yang cukup signifikan dari waktu ke waktu. Pola musiman tahunan terlihat jelas, di mana penjualan cenderung meningkat dan menurun secara berulang pada periode-periode tertentu dalam satu tahun. Selain itu, terdapat penurunan yang sangat tajam sekitar tahun 2020 yang mencerminkan dampak pandemi COVID-19 terhadap sektor otomotif di Indonesia. Setelah periode tersebut, data menunjukkan adanya pemulihan secara bertahap, meskipun dengan volatilitas yang relatif lebih tinggi dibandingkan periode sebelum pandemi.

Karakteristik data yang memiliki tren, pola musiman, serta fluktuasi ekstrem ini menunjukkan bahwa pendekatan peramalan deret waktu dengan mempertimbangkan komponen musiman dan dinamika jangka panjang sangat diperlukan. Oleh karena itu, metode Time Series Regression, SARIMAX, dan Hybrid SARIMA–Neural Network dipandang relevan untuk digunakan dalam penelitian ini.

3.2 Pembagian Data Latih dan Data Uji

Data penjualan mobil bulanan dibagi menjadi data latih (training set) dan data uji (testing set) untuk mengevaluasi kemampuan model dalam melakukan peramalan pada data yang belum diamati sebelumnya. Pembagian data dilakukan dengan pendekatan hold-out, di mana dua belas bulan terakhir digunakan sebagai data uji, sedangkan seluruh observasi sebelumnya digunakan sebagai data latih.

Pemilihan horizon pengujian selama satu tahun dilakukan karena sesuai dengan karakteristik data bulanan serta umum digunakan dalam evaluasi peramalan jangka pendek. Skema pembagian data ini diterapkan secara konsisten pada seluruh model yang digunakan, yaitu Time Series Regression, SARIMAX, dan Hybrid SARIMA–Neural Network, sehingga perbandingan kinerja antar model dapat dilakukan secara adil dan objektif.

3.3 Hasil Uji Stasioneritas dan Identifikasi Model

3.3.1 Hasil Uji Stasioneritas Augmented Dickey–Fuller (ADF)

Uji stasioneritas dilakukan menggunakan metode Augmented Dickey–Fuller (ADF) untuk menguji apakah deret waktu penjualan mobil bulanan bersifat stasioner pada rataan. Pengujian ini penting karena pemodelan SARIMAX mensyaratkan data yang telah memenuhi asumsi stasioneritas.

```
ADF Statistic : -2.302120866857858
p-value       : 0.17128486212914218
```

Hasil Uji ADF pada Python Sebelum Differencing

Berdasarkan hasil uji ADF pada data penjualan mobil bulanan sebelum dilakukan differencing, diperoleh nilai statistik ADF sebesar $-2,302$ dengan p-value sebesar $0,171$. Nilai p-value yang lebih besar dari tingkat signifikansi 5% menunjukkan bahwa hipotesis nol tidak dapat ditolak. Dengan demikian, data pada kondisi awal belum bersifat stasioner, sehingga masih mengandung tren atau pola jangka panjang.

```
ADF Statistic : -3.2358885675089435
p-value       : 0.01799905096223132
```

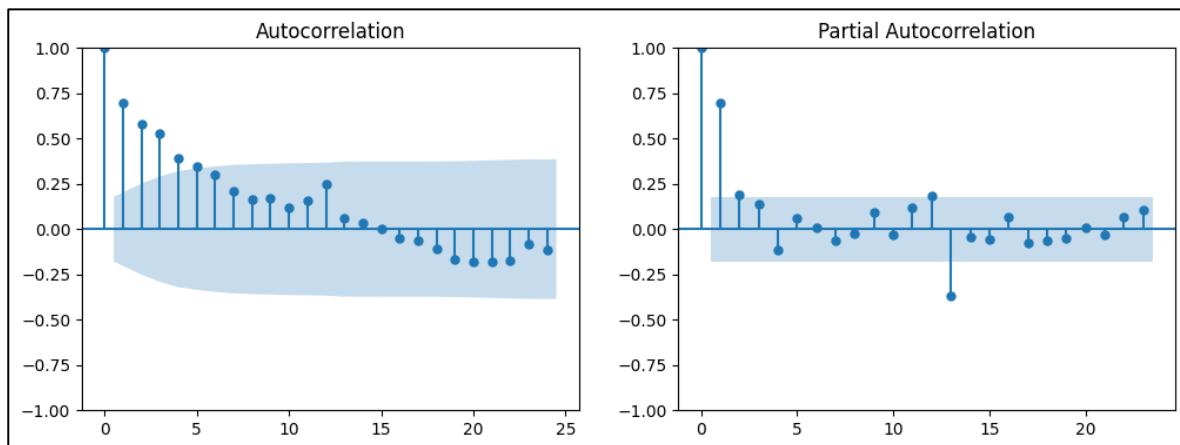
Hasil Uji ADF pada Python Setelah Differencing

Untuk mengatasi ketidakstasioneran tersebut, dilakukan differencing non-musiman orde pertama. Setelah differencing, uji ADF kembali dilakukan dan menghasilkan nilai statistik ADF sebesar $-3,236$ dengan p-value sebesar $0,018$. Nilai p-value yang lebih kecil dari $0,05$ menunjukkan bahwa hipotesis nol dapat ditolak, sehingga data hasil differencing dapat dinyatakan stasioner pada rataan.

Hasil ini menunjukkan bahwa proses differencing non-musiman berhasil menghilangkan tren pada data dan memenuhi asumsi stasioneritas yang diperlukan untuk pemodelan SARIMAX.

3.3.2 Analisis Autokorelasi dan Partial Autokorelasi

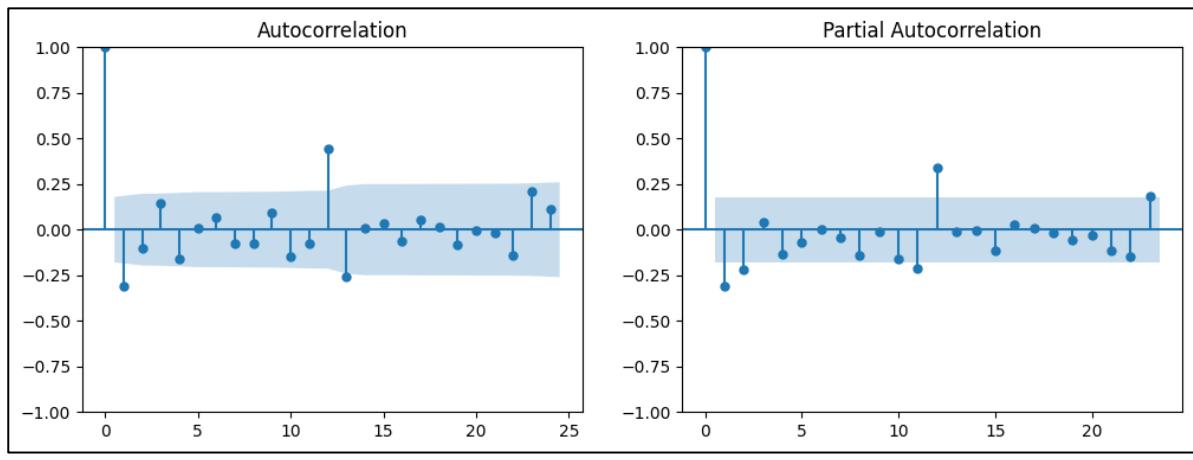
Setelah data dinyatakan stasioner berdasarkan uji ADF, analisis Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF) dilakukan untuk mengidentifikasi struktur ketergantungan temporal pada data serta menentukan kandidat orde model SARIMAX.



ACF-PACF Data Penjualan Mobil Sebelum Differencing

Pada kondisi sebelum differencing, grafik ACF menunjukkan pola penurunan autokorelasi yang berlangsung secara perlahan (slow decay) pada beberapa lag awal. Pola ini mengindikasikan adanya tren dan ketergantungan jangka panjang antar observasi, yang sejalan dengan hasil uji ADF yang

menyatakan data belum stasioner. Grafik PACF pada kondisi ini juga menunjukkan spike signifikan pada lag awal, namun masih diikuti oleh fluktuasi pada lag-lag berikutnya, yang memperkuat indikasi ketidakstasioneran data.



ACF-PACF Data Penjualan Mobil Setelah Differencing

Setelah dilakukan differencing non-musiman orde pertama, grafik ACF menunjukkan perubahan yang signifikan, di mana sebagian besar nilai autokorelasi berada dalam batas kepercayaan. Namun demikian, masih terlihat spike yang signifikan pada lag ke-12. Kemunculan spike pada lag ke-12 ini mengindikasikan adanya pola musiman tahunan, yang konsisten dengan karakteristik data bulanan. Oleh karena itu, periode musiman $s = 12$ dipertimbangkan dalam pemodelan SARIMAX.

Grafik PACF setelah differencing menunjukkan spike signifikan pada lag awal serta beberapa lag tertentu, sementara sebagian besar lag lainnya berada dalam batas kepercayaan. Pola ini memberikan indikasi awal terhadap keberadaan komponen autoregressive dan moving average pada data, baik pada komponen non-musiman maupun musiman. Hasil analisis ACF dan PACF setelah differencing menunjukkan bahwa data telah memenuhi asumsi stasioneritas dan memiliki struktur musiman yang jelas. Temuan ini menjadi dasar dalam penentuan orde model SARIMAX serta mendukung penggunaan pendekatan hybrid pada tahap selanjutnya, terutama untuk menangkap pola non-linear yang masih tersisa pada residual model.

3.4 SARIMAX

3.4.1 Pemilihan Model Terbaik Melalui *Grid Search*

Pemilihan model terbaik dilakukan melalui proses *grid search* dengan mengevaluasi berbagai kombinasi parameter $(p, d, q)(P, D, Q)_{12}$. Evaluasi model didasarkan pada:

- Nilai Root Mean Squared Error (*RMSE*)
- Nilai Akaike Information Criterion (*AIC*)
- Nilai Mean Absolute Percentage Error (*MAPE*)

Order	Seasonal_order	AIC	RMSE	MAPE
(0,0,0)	(1,1,1,12)	1846,2047	6029,833	6.386531
(0,0,2)	(1, 1, 1, 12)	1800.479487	7205.848853	7.169254
(2, 0, 2)	(1, 1, 1, 12)	1802.881336	7605.973968	7.235629
(1, 0, 0)	(1, 0, 0, 12)	2085.465191	7832.603785	8.351017
(1, 0, 1)	(1, 1, 1, 12)	1823.01373	6948.57023	8.355928

Tabel Perbandingan 5 Model Terbaik

Berdasarkan hasil grid search pada beberapa kombinasi parameter SARIMAX, terlihat bahwa setiap model menghasilkan kinerja yang berbeda baik dari sisi kompleksitas maupun akurasi peramalan. Beberapa model memiliki nilai AIC yang relatif rendah, namun menghasilkan nilai RMSE dan MAPE yang lebih tinggi, sehingga kurang optimal dari sisi akurasi prediksi. Oleh karena itu, pemilihan model terbaik tidak hanya didasarkan pada nilai AIC dan BIC yang kecil, tetapi juga mempertimbangkan nilai RMSE dan MAPE sebagai indikator kesalahan prediksi. Model dengan kombinasi parameter yang memberikan kesalahan prediksi paling rendah dan struktur yang parsimonious dipilih sebagai model SARIMAX terbaik untuk analisis selanjutnya.

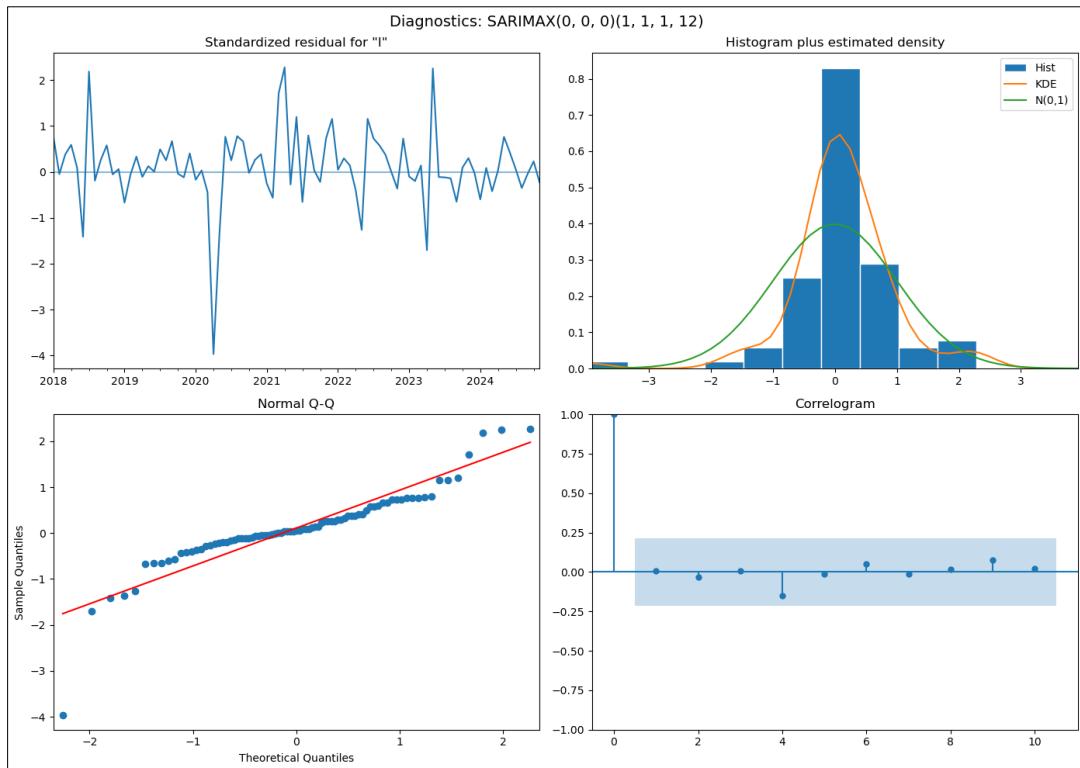
SARIMAX Results						
Dep. Variable:	Indonesia Total Car Sales	No. Observations:	108			
Model:	SARIMAX(1, 1, [1], 12)	Log Likelihood	-910.102			
Date:	Sun, 04 Jan 2026	AIC	1846.205			
Time:	11:45:27	BIC	1877.650			
Sample:	12-01-2015 - 11-01-2024	HQIC	1858.838			
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
lag_1	-0.1451	56.938	-0.003	0.998	-111.741	111.451
lag_2	-0.7689	56.962	-0.013	0.989	-112.412	110.875
lag_3	-0.6712	56.931	-0.012	0.991	-112.253	110.911
lag_12	0.1248	0.249	0.501	0.616	-0.363	0.613
rolling_mean_3	2.6249	170.813	0.015	0.988	-332.163	337.413
rolling_mean_12	-0.2341	0.424	-0.552	0.581	-1.065	0.597
month_sin	-3.944e-05	1.1e+04	-3.57e-09	1.000	-2.17e+04	2.17e+04
month_cos	-2.153e-05	2.23e+04	-9.66e-10	1.000	-4.37e+04	4.37e+04
is_q1	3.743e-06	4e+04	9.37e-11	1.000	-7.83e+04	7.83e+04
is_q4	2.951e-06	4.98e+04	5.93e-11	1.000	-9.76e+04	9.76e+04
ar.S.L12	0.3358	0.282	1.193	0.233	-0.216	0.888
ma.S.L12	-0.8978	0.187	-4.797	0.000	-1.265	-0.531
sigma2	2.654e+08	15.574	1.7e+07	0.000	2.65e+08	2.65e+08
Ljung-Box (L1) (Q):	0.01	Jarque-Bera (JB):	173.56			
Prob(Q):	0.94	Prob(JB):	0.00			
Heteroskedasticity (H):	0.44	Skew:	-1.00			
Prob(H) (two-sided):	0.03	Kurtosis:	9.79			

Gambar Hasil Pemodelan SARIMAX

Hasil estimasi model SARIMAX menunjukkan bahwa model memiliki nilai AIC sebesar 1846,21 dan BIC sebesar 1877,65, yang mengindikasikan keseimbangan yang cukup baik antara kompleksitas dan kecocokan model terhadap data. Dari sisi parameter, komponen moving average musiman orde 12 (MA.S.L12) signifikan secara statistik dengan nilai *p-value* < 0,05, yang menunjukkan bahwa pola musiman tahunan memiliki pengaruh yang kuat terhadap penjualan mobil di Indonesia. Sementara itu, sebagian besar variabel eksogen dan lag non-musiman tidak signifikan, yang mengindikasikan bahwa kontribusi utama dalam model berasal dari struktur musiman. Hasil uji Ljung–Box menunjukkan nilai *p-value* sebesar 0,94, sehingga residual dapat dianggap bersifat acak (white noise), yang menandakan bahwa model telah mampu menangkap pola utama data dengan baik dan layak digunakan untuk peramalan.

3.4.2 Uji Diagnostik Residual Model SARIMAX

Uji diagnostik residual dilakukan untuk memastikan bahwa model SARIMAX yang dibangun telah mampu menangkap seluruh struktur utama dalam data, sehingga sisa kesalahan (residual) bersifat acak dan tidak mengandung informasi sistematis. Evaluasi dilakukan melalui analisis visual residual serta pengujian statistik menggunakan uji Ljung–Box.



Gambar Uji Diagnostic Residual

Gambar diatas menampilkan hasil diagnostik residual dari model SARIMAX (0,0,0)(1,1,1,12) yang digunakan. Plot residual terstandarisasi menunjukkan bahwa residual berfluktuasi di sekitar nilai nol tanpa membentuk pola tren maupun musiman yang jelas, yang mengindikasikan bahwa model telah berhasil menangkap struktur utama data. Histogram residual yang dilengkapi dengan estimasi kepadatan kernel (KDE) menunjukkan bahwa sebagian besar residual terkonsentrasi di sekitar nilai nol, meskipun terdapat penyimpangan pada bagian ekor distribusi yang mengindikasikan ketidaknormalan residual. Hal ini diperkuat oleh plot normal Q-Q yang memperlihatkan bahwa sebagian besar titik mengikuti garis diagonal pada bagian tengah, namun terjadi penyimpangan pada bagian ekor, yang menunjukkan keberadaan nilai ekstrem pada beberapa periode tertentu.

Selanjutnya, correlogram (ACF residual) menunjukkan bahwa seluruh nilai autokorelasi residual berada di dalam batas kepercayaan, sehingga tidak terdapat autokorelasi yang signifikan pada berbagai lag.

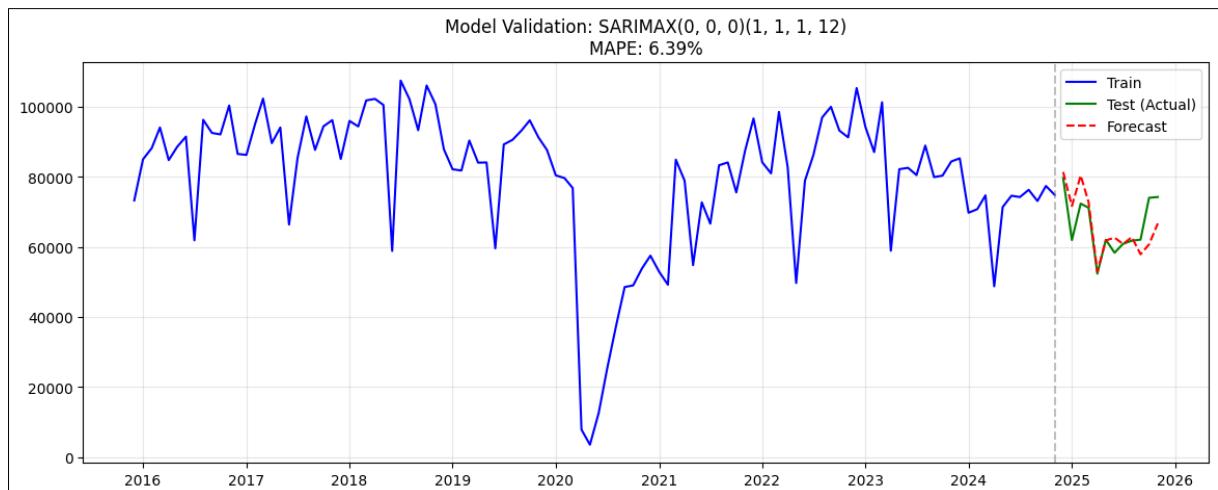
Lb_stat	Lb_pvalue
5,4623	0,85823
12,8713	0,882834

Temuan visual tersebut ini konsisten dengan hasil uji Ljung–Box yang ditampilkan pada Tabel diatas, di mana nilai statistik Ljung–Box pada lag 10 sebesar 5,46 dengan *p-value* 0,86, serta pada lag 20 sebesar 12,87 dengan *p-value* 0,88. Nilai *p-value* yang lebih besar dari 0,05 menunjukkan bahwa hipotesis nol tidak dapat ditolak, sehingga residual model dapat dinyatakan bersifat acak atau *white noise*.

Secara keseluruhan, hasil diagnostik residual baik secara visual maupun statistik menunjukkan bahwa model SARIMAX telah memenuhi asumsi utama berupa residual yang tidak berkorelasi dan tidak berpola. Meskipun demikian, adanya indikasi ketidaknormalan residual dan nilai ekstrem pada beberapa periode menunjukkan bahwa masih terdapat variasi non-linear yang belum sepenuhnya tertangkap oleh model. Kondisi ini menjadi dasar pertimbangan untuk pengembangan model lanjutan menggunakan pendekatan hybrid SARIMA–Neural Network guna meningkatkan akurasi peramalan.

3.4.3 Evaluasi Kinerja Model SARIMAX

Untuk mengevaluasi kinerja model SARIMAX yang telah dibangun, dilakukan validasi model dengan membandingkan hasil peramalan terhadap data aktual pada periode pengujian. Evaluasi dilakukan secara visual dan kuantitatif untuk menilai kemampuan model dalam mengikuti pola data historis serta akurasi prediksi yang dihasilkan. Visualisasi perbandingan antara data latih, data uji, dan hasil peramalan digunakan untuk memberikan gambaran menyeluruh mengenai performa model SARIMAX dalam memodelkan dinamika penjualan mobil di Indonesia.



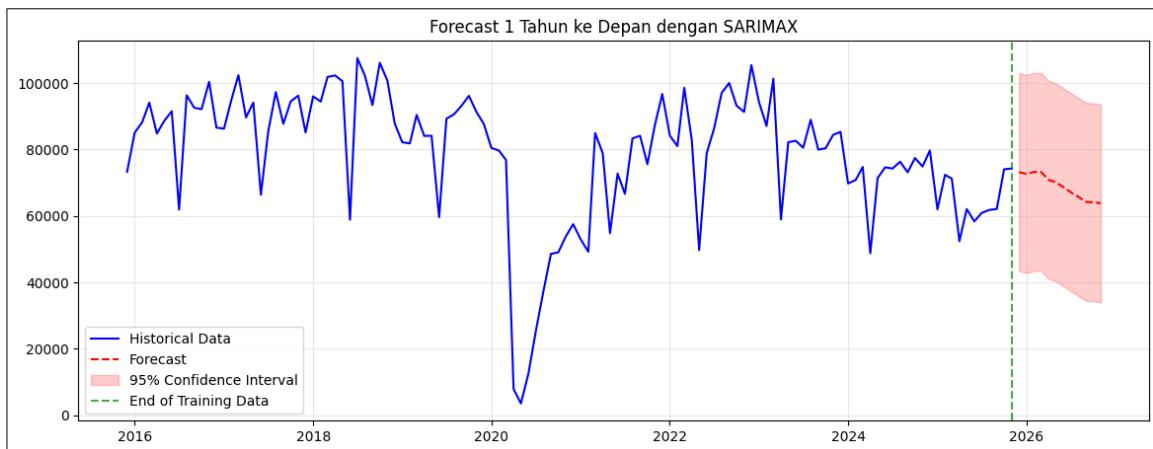
Grafik Validasi Model SARIMAX

Berdasarkan Gambar X, terlihat bahwa hasil peramalan model SARIMAX mampu mengikuti pergerakan data aktual pada periode pengujian dengan cukup baik. Pola musiman tahunan yang telah dipelajari dari data latih tetap terjaga pada hasil peramalan, dan deviasi antara nilai prediksi dan data aktual relatif kecil. Nilai Mean Absolute Percentage Error (MAPE) sebesar 6,39% menunjukkan bahwa tingkat kesalahan prediksi model sangat rendah, sehingga model SARIMAX memiliki akurasi yang sangat baik dalam jangka pendek. Meskipun demikian, pada beberapa periode tertentu terlihat adanya selisih antara hasil peramalan dan data aktual, terutama pada saat terjadi fluktuasi yang relatif tajam, yang mengindikasikan keterbatasan model linier dalam menangkap variasi non-linear pada data. Oleh karena itu, hasil evaluasi ini menjadi dasar untuk pengembangan model lanjutan menggunakan pendekatan hybrid SARIMA–Neural Network guna meningkatkan akurasi peramalan.

3.4.4 Peramalan Model SARIMAX

Setelah model yang digunakan dinyatakan layak dan memiliki kinerja peramalan yang baik, tahap selanjutnya adalah melakukan peramalan untuk periode mendatang. Peramalan dilakukan dengan dua horizon waktu yang berbeda, yaitu jangka pendek dan jangka panjang, untuk memberikan gambaran proyeksi penjualan mobil di Indonesia dalam skala waktu yang bervariasi. Peramalan jangka pendek bertujuan untuk menangkap dinamika data dalam waktu dekat dan relevan untuk kebutuhan perencanaan operasional, sedangkan peramalan jangka panjang digunakan untuk melihat kecenderungan umum dan pola musiman dalam horizon waktu yang lebih panjang. Pendekatan ini diharapkan dapat memberikan informasi yang komprehensif bagi pengambilan keputusan baik dalam jangka pendek maupun jangka panjang.

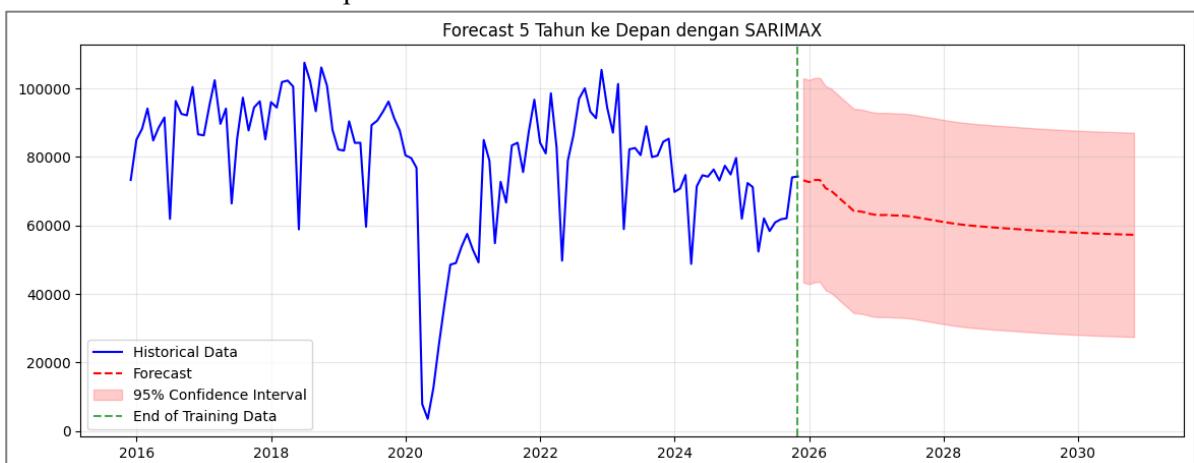
1. Peramalan 1 Tahun Kedepan



Grafik Hasil Peramalan 1 Tahun dengan SARIMAX

Berdasarkan Gambar diatas, hasil peramalan satu tahun ke depan menggunakan model SARIMAX menunjukkan bahwa penjualan mobil di Indonesia diproyeksikan mengalami fluktuasi yang relatif moderat dengan kecenderungan berada pada kisaran yang tidak jauh berbeda dari periode terakhir data historis. Garis peramalan menunjukkan pola yang masih mengikuti karakteristik musiman tahunan yang telah dipelajari oleh model, tanpa adanya lonjakan atau penurunan yang ekstrem. Interval kepercayaan 95% menggambarkan tingkat ketidakpastian peramalan yang masih berada dalam rentang wajar, dengan lebar interval yang relatif stabil sepanjang horizon satu tahun. Hal ini menunjukkan bahwa model SARIMAX memiliki tingkat keandalan yang baik untuk peramalan jangka pendek dan dapat digunakan sebagai dasar pertimbangan dalam perencanaan operasional dan pengambilan keputusan dalam periode satu tahun ke depan.

2. Peramalan 5 Tahun Kedepan



Grafik Hasil Peramalan 5 Tahun dengan SARIMAX

Berdasarkan Gambar X, hasil peramalan lima tahun ke depan menggunakan model SARIMAX menunjukkan kecenderungan penjualan mobil di Indonesia yang relatif stabil dengan arah perubahan yang cenderung menurun secara bertahap. Pola musiman tahunan yang terbentuk pada data historis masih tetap dipertahankan oleh model, namun variasi antarperiode menjadi lebih halus seiring bertambahnya horizon waktu peramalan. Interval kepercayaan 95% tampak semakin melebar pada periode peramalan jangka panjang, yang mencerminkan meningkatnya tingkat ketidakpastian prediksi seiring dengan bertambahnya horizon waktu. Meskipun demikian, model tidak menghasilkan fluktuasi yang ekstrem atau tidak realistik, sehingga hasil peramalan ini dapat digunakan untuk melihat

kecenderungan umum dan arah perkembangan penjualan mobil dalam jangka panjang, terutama sebagai bahan pertimbangan dalam analisis strategis dan perencanaan jangka panjang.

3.5 Time Series Regression (TSR)

a. Karakteristik Stasioneritas Komparatif

Data periode pendek (2021-2025) menunjukkan elevated volatility yang mencerminkan structural disruption akibat pandemi COVID-19 dan fase recovery ekonomi subsequent. Dataset extended (2015-2025) mencakup complete business cycle yang lebih stabil.

b. Performa Prediktif: Analisis Komparatif

Metrik Evaluasi	Model 5 Tahun (2021-2025)	Model 10 Tahun (2015-2025)	Selisih Performa
RMSE	9.758,34	11.547,25	+18,3% Error
MAE	7.953,47	9.724,12	+22,2% Error
MAPE	13,18%	16,06%	+2,88% (Absolut)

Hasil empiris menunjukkan paradoks counterintuitive: model yang dilatih pada dataset lebih kecil (5 tahun) menghasilkan generalization performance superior dibandingkan model dengan training data lebih ekstensif (10 tahun). MAPE 13.18% versus 16.06% mengindikasikan perbedaan akurasi absolut sebesar 2.88 poin persentase.

Fenomena ini dapat diatribusikan kepada structural break hypothesis: terjadi regime shift fundamental dalam consumer behavior dan market dynamics industri otomotif Indonesia pra- dan pasca-pandemi. Inklusi data historis 2015-2019 justru mengintroduksi pattern heterogeneity yang menurunkan predictive relevance model untuk kondisi kontemporer.

c. Analisis Struktur Koefisien

Dataset 5 Tahun (2021-2025):

Koefisien signifikan: Lag₁ (0.345), Lag₃ (0.259), Lag₁₁ (0.259)

```
Intercept: 11070.677793568044
Coefficients: [ 0.34583825  0.01142535  0.25982596 -0.16461192  0.15407085 -0.08107688
               0.1008845 -0.24054173  0.21092463 -0.13512296  0.25995026  0.1336862 ]
```

Distribusi koefisien menunjukkan multi-modal influence structure, dimana quarterly fluctuations (Lag₃) dan near-annual patterns (Lag₁₁) memiliki kontribusi comparable dengan immediate predecessor (Lag₁). Ini mencerminkan volatile, event-driven market pasca-pandemi.

Dataset 10 Tahun (2015-2025):

Koefisien dominan: Lag₁ (0.569), Lag₁₂ (0.174)

```
Intercept: 12989.661683400293
Coefficients: [ 0.56880413  0.09580551  0.1551559 -0.09068193  0.06874992  0.03197913
                -0.08647424 -0.04264461  0.0407581 -0.10777503  0.02323664  0.1744225 ]
```

Model menunjukkan concentrated dependency structure dengan dominasi ekstrem pada Lag₁, mengindikasikan strong month-to-month persistence. Seasonal effect (Lag₁₂) tetap observable namun dampaknya terdilusi oleh momentum jangka pendek.

d. Implikasi untuk Praktik Forecasting

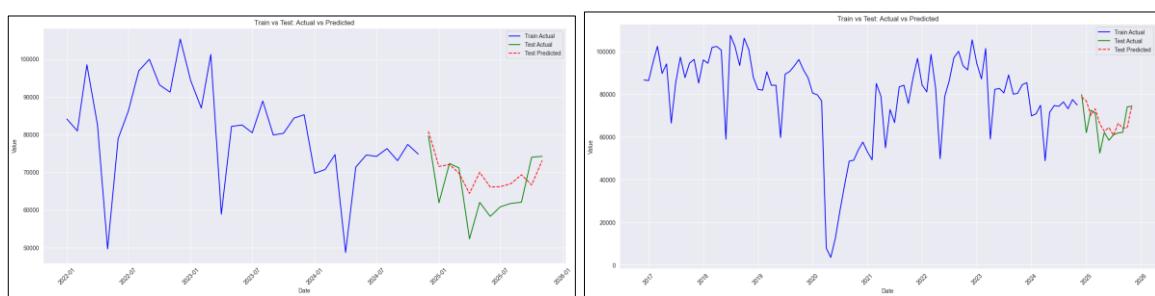
Berdasarkan temuan penelitian, TSR direkomendasikan sebagai alat taktis untuk perencanaan operasional jangka pendek (1–12 bulan) dalam kondisi pasar yang relatif stabil. Untuk horison yang lebih luas, diperlukan transisi ke model yang lebih kompleks:

- Jangka Menengah (6–24 bulan): Penggunaan *State Space Models* seperti SARIMAX atau Prophet yang memiliki mekanisme eksplisit untuk menangani evolusi pola tanpa pelatihan ulang manual.
- Jangka Panjang (2–10 tahun): Implementasi metode *ensemble* atau jaringan saraf tiruan seperti LSTM/GRU yang mampu mempelajari pola non-linear kompleks serta integrasi variabel eksogen melalui *Vector Autoregression* (VAR).

Sebagai rekomendasi akhir bagi praktisi industri otomotif, penelitian ini menekankan bahwa kualitas dan relevansi temporal data jauh lebih krusial dibandingkan kuantitas data historis dalam konteks deret waktu yang non-stasioner.

e. Evaluasi Kerja Model *Time Series Regression*

Evaluasi kinerja model time series regression dilakukan melalui dua pendekatan dengan periode data berbeda, yaitu model 5 tahun (2021-2025) dan model 10 tahun (2015-2025). Visualisasi split data forecasting menunjukkan bahwa kedua model mampu menangkap pola dan tren umum, namun memiliki perbedaan performa yang signifikan. Model dengan data 5 tahun menghasilkan RMSE sebesar 9.758,34, MAE sebesar 7.953,47, dan MAPE sebesar 13,18%, sementara model dengan data 10 tahun menghasilkan RMSE sebesar 11.547,25, MAE sebesar 9.724,12, dan MAPE sebesar 16,06%. Perbandingan ini menunjukkan bahwa model 10 tahun mengalami peningkatan error sebesar 18,3% pada RMSE, 22,2% pada MAE, dan 2,88% pada MAPE dibandingkan model 5 tahun.



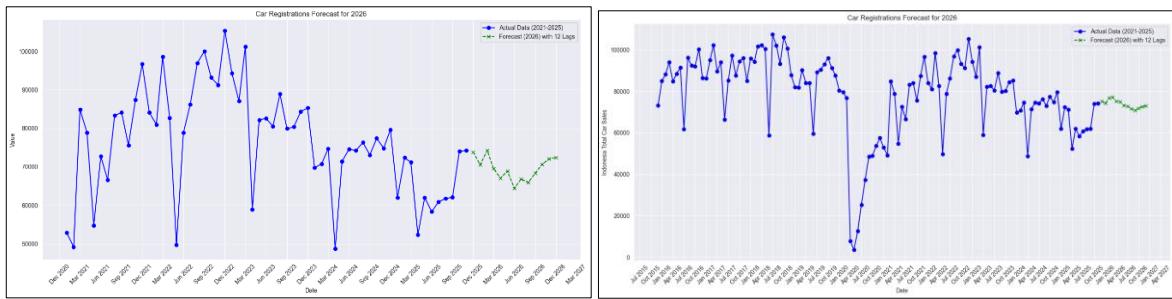
Perbandingan Data Test dan Aktual Data 5 Tahun dan 10 Tahun

Hasil evaluasi mengindikasikan bahwa penggunaan data historis yang terlalu panjang justru menurunkan akurasi prediksi, kemungkinan karena pola data mengalami perubahan struktural seiring waktu dan data lama menjadi kurang relevan dengan kondisi terkini. Berdasarkan perbandingan metrik evaluasi, model time series regression dengan data 5 tahun memberikan performa yang lebih optimal dengan tingkat error yang lebih rendah di semua indikator. Oleh karena itu, model 5 tahun direkomendasikan untuk digunakan dalam forecasting karena lebih responsif terhadap pola terkini dan mampu menghasilkan prediksi dengan akurasi yang lebih baik untuk keperluan perencanaan dan pengambilan keputusan.

f. Analisis Hasil Forecasting (Proyeksi Jangka Menengah)

Secara teknis, hasil peramalan dari kedua model menunjukkan karakteristik yang berbeda secara signifikan saat diproyeksikan ke masa depan menggunakan metode rekursif:

- 1) Peramalan Jangka Pendek (2026 - 12 Bulan)

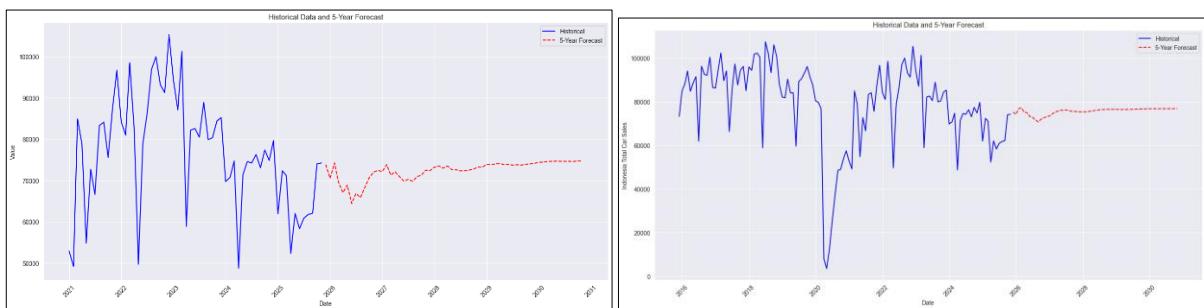


Perbandingan Peramalan Jangka Pendek Data 5 Tahun dan 10 Tahun

Pada horison 1 tahun, model yang menggunakan data 5 tahun (2021-2025) menunjukkan sensitivitas tinggi terhadap pola musiman terbaru.

- Hasil: Proyeksi menunjukkan tren pemulihan yang stabil dengan fluktuasi tajam pada bulan-bulan yang secara historis memiliki volume tinggi (seperti periode menjelang Idul Fitri dan kuartal pameran otomotif).
- Akurasi: Dengan MAPE terendah (13,18%), model ini sangat handal untuk perencanaan rantai pasok (*supply chain*) dan penentuan target penjualan tahunan.

2) Peramalan Jangka Panjang (5 Tahun)



Perbandingan Peramalan Jangka Panjang Data 5 Tahun dan 10 Tahun

Ketika horison diperluas hingga 60 bulan, terjadi fenomena *mean reversion* pada model 10 tahun, sementara model 5 tahun mulai menunjukkan akumulasi galat yang terlihat dari fluktuasi prediksi yang semakin lebar.

- Implikasi: Model 5 tahun unggul dalam menangkap dinamika pasar saat ini, namun penggunaannya untuk jangka menengah harus dibarengi dengan pembaruan data secara berkala (*rolling-update*) untuk mencegah *drift* peramalan.

3.5 Hybrid SARIMA-Neural Network

Pemodelan hybrid SARIMA-Neural Network dilakukan untuk meningkatkan akurasi peramalan dengan mengombinasikan kemampuan model linear-musiman dan model non-linear. Pada tahap awal, model SARIMAX terbaik dipilih melalui proses *grid search* dengan mempertimbangkan kombinasi parameter non-musiman dan musiman serta kinerja prediksi pada data uji.

3.5.1 Pemilihan Model SARIMA melalui Grid Search

Proses *grid search* dilakukan dengan mengevaluasi berbagai kombinasi parameter (p, d, q) dan $(P, D, Q)_s$ dengan periode musiman $s = 12$. Evaluasi model dilakukan menggunakan kriteria Root Mean Squared Error (RMSE) pada data uji serta nilai Akaike Information Criterion (AIC) untuk mempertimbangkan kompleksitas model.

Order	Seasonal	AIC	RMSE
(0, 1, 0)	(0, 1, 1, 12)	1822,76	5834,60
(1, 1, 0)	(1, 1, 0, 12)	1825,14	6068,16
(1, 1, 1)	(1, 1, 0, 12)	1826,24	6080,11
(0, 1, 1)	(1, 1, 0, 12)	1845,56	6082,53
(0, 1, 2)	(1, 1, 0, 12)	1847,12	6082,64

Tabel Perbandingan 5 Model Terbaik

Berdasarkan hasil *grid search*, model SARIMAX(0,1,0)×(0,1,1,12) dipilih sebagai model terbaik karena menghasilkan nilai RMSE terendah sebesar 5834,60 dengan nilai AIC sebesar 1822,76. Model ini juga bersifat parsimonious karena tidak menggunakan komponen autoregressive dan moving average non-musiman, namun tetap mampu menangkap pola musiman tahunan melalui komponen *seasonal moving average* orde satu.

SARIMAX Results						
Dep. Variable:	Indonesia Total Car Sales	No. Observations:	108			
Model:	SARIMAX(0, 1, 0)x(0, 1, [1], 12)	Log Likelihood	-909.378			
Date:	Sun, 04 Jan 2026	AIC	1822.756			
Time:	17:46:14	BIC	1827.570			
Sample:	12-01-2015 - 11-01-2024	HQIC	1824.689			
Covariance Type:	opg					
coef	std err	z	P> z	[0.025	0.975]	
ma.S.L12	-0.6023	0.081	-7.469	0.000	-0.760	-0.444
sigma2	2.956e+08	3.91e-13	7.56e+20	0.000	2.96e+08	2.96e+08
Ljung-Box (L1) (Q):	6.13	Jarque-Bera (JB):	42.24			
Prob(Q):	0.01	Prob(JB):	0.00			
Heteroskedasticity (H):	0.45	Skew:	0.06			
Prob(H) (two-sided):	0.04	Kurtosis:	6.51			

Hasil Pemodelan SARIMA

Pada komponen musiman, parameter MA musiman lag ke-12 (MA.S.L12) memiliki koefisien sebesar -0,6023 dengan nilai *p-value* < 0,05. Hal ini menunjukkan bahwa komponen musiman tahunan memiliki pengaruh yang signifikan terhadap dinamika penjualan mobil. Tanda negatif pada koefisien mengindikasikan adanya efek koreksi terhadap fluktuasi musiman pada periode yang sama di tahun sebelumnya.

3.5.2 Pemodelan Residual SARIMAX dengan Neural Network

Setelah model SARIMAX terbaik diperoleh, residual dari model tersebut digunakan sebagai input untuk pemodelan Neural Network. Residual ditransformasikan ke dalam bentuk *supervised learning* dengan menggunakan lag residual [1, 2, 12, 13, 14], yang mencakup ketergantungan jangka pendek dan musiman.

Neural Network yang digunakan adalah Multilayer Perceptron (MLP) dengan dua lapisan tersembunyi, masing-masing terdiri dari 10 neuron, menggunakan fungsi aktivasi ReLU dan algoritma optimasi

Adam. Sebelum pelatihan, data residual distandarisasi untuk meningkatkan stabilitas proses pembelajaran.

Hasil pemodelan menunjukkan bahwa Neural Network mampu mempelajari pola residual dengan baik. Setelah dilakukan koreksi residual menggunakan Neural Network, dilakukan uji Ljung–Box pada residual hybrid dengan lag 12. Hasil uji menunjukkan nilai statistik Ljung–Box sebesar 8,87 dengan p-value sebesar 0,71, yang mengindikasikan bahwa residual model hybrid telah bersifat acak (*white noise*). Hal ini menunjukkan bahwa kombinasi SARIMA dan Neural Network berhasil menangkap struktur utama dan pola non-linear pada data.

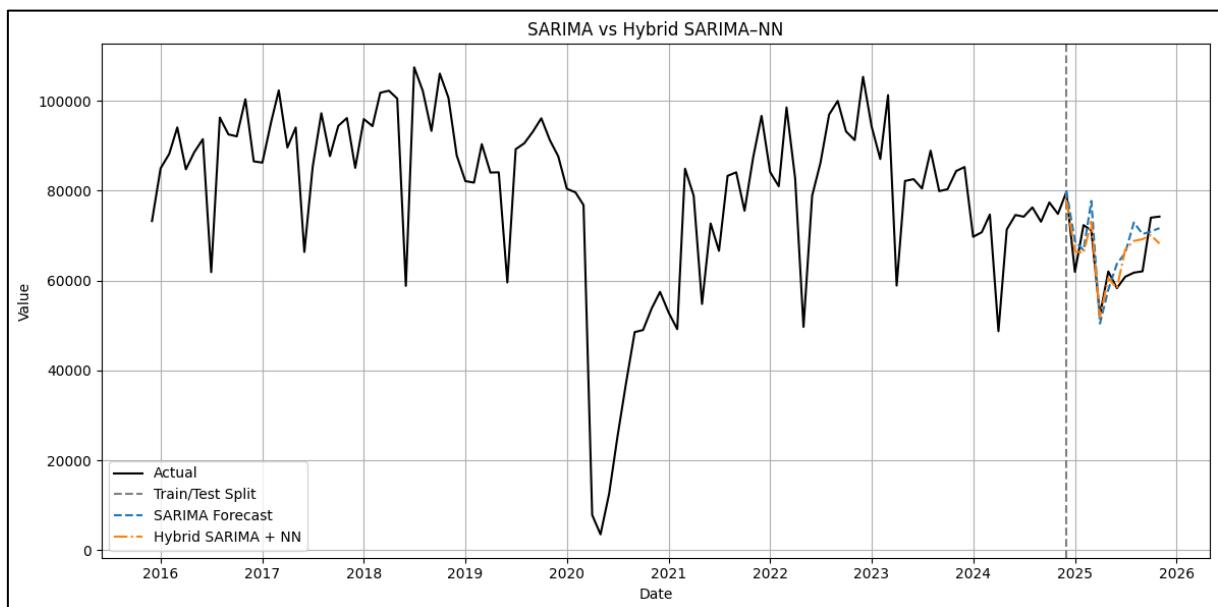
lb_stat	lb_pvalue
12 8.871901	0.713831

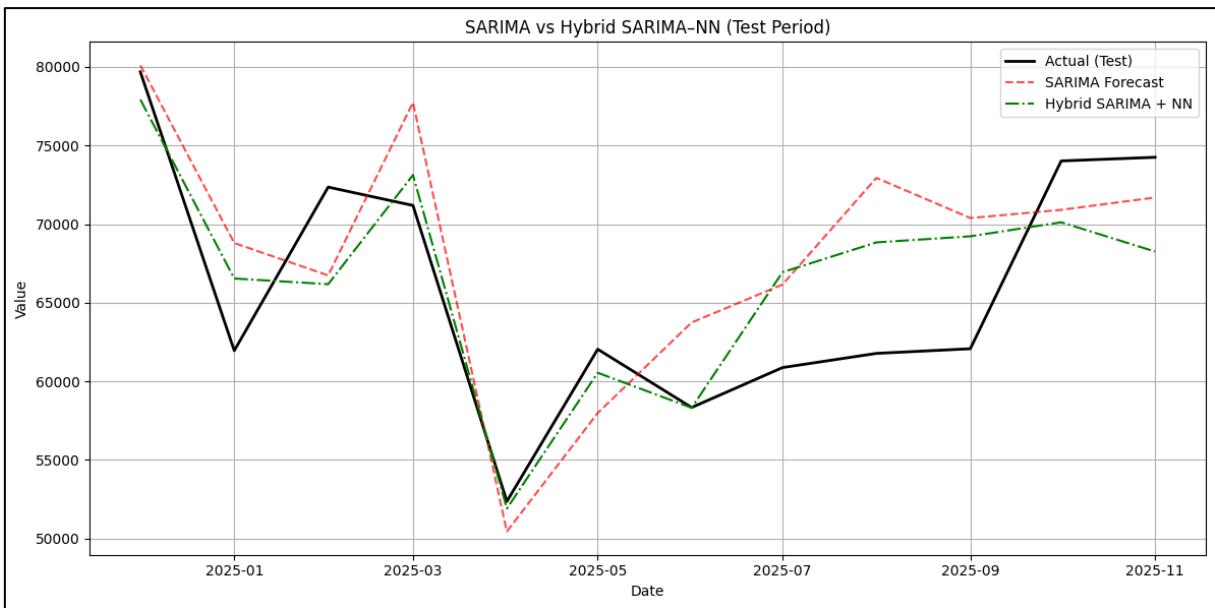
Hasil Ljung–Box pada Python

3.5.3 Evaluasi Kinerja Model Hybrid

Peramalan hybrid diperoleh dengan menjumlahkan hasil peramalan SARIMAX dan prediksi residual dari Neural Network. Pada periode pengujian selama dua belas bulan terakhir, model hybrid menunjukkan kinerja yang lebih baik dibandingkan model SARIMAX tunggal, baik secara kuantitatif maupun visual.

Secara kuantitatif, model SARIMAX menghasilkan nilai RMSE sebesar 5834,60, MAE sebesar 5102,49, dan MAPE sebesar 7,98%. Setelah diterapkan pendekatan hybrid SARIMA–Neural Network, nilai kesalahan menurun menjadi RMSE sebesar 4631,34, MAE sebesar 3883,83, dan MAPE sebesar 5,87%. Penurunan ketiga metrik error tersebut menunjukkan bahwa koreksi residual menggunakan Neural Network mampu meningkatkan akurasi peramalan secara signifikan.





Perbandingan Data Test dan Aktual

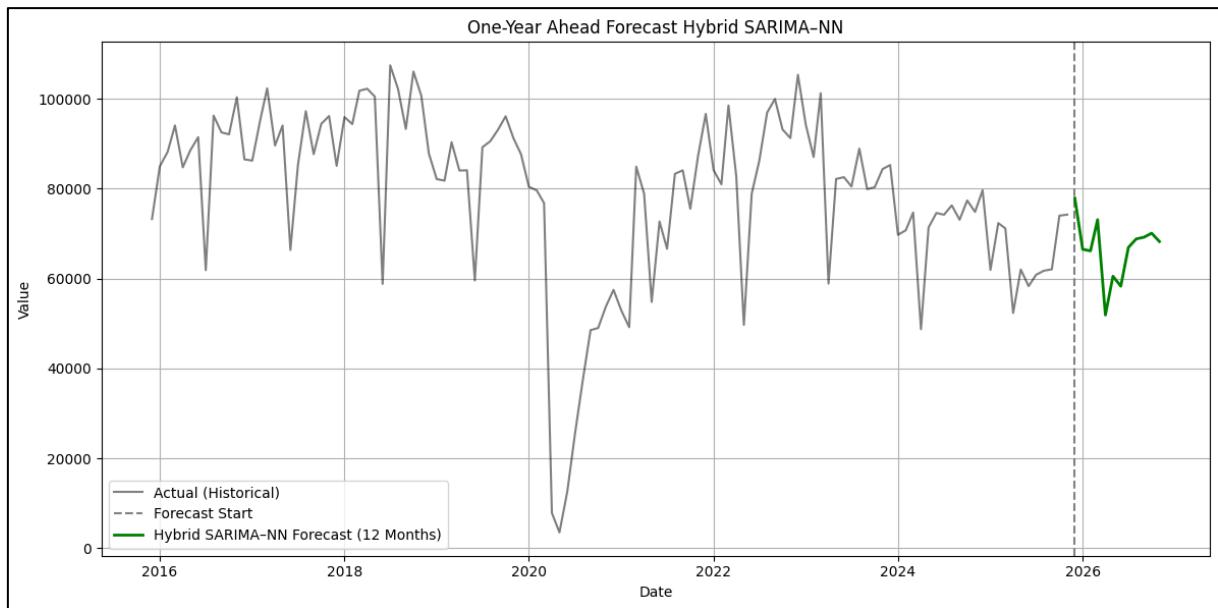
Secara visual, hasil peramalan hybrid lebih dekat terhadap data aktual dan lebih responsif dalam mengikuti fluktuasi data, terutama pada periode dengan perubahan yang relatif tajam. Meskipun demikian, pola musiman tahunan yang telah ditangkap oleh model SARIMAX tetap terjaga. Hal ini menunjukkan bahwa model hybrid mampu mengombinasikan kestabilan pola musiman dari SARIMAX dengan fleksibilitas Neural Network dalam menangkap variasi non-linear.

3.5.4 Peramalan Menggunakan Model Hybrid SARIMA-Neural Network

Peramalan dilakukan untuk dua horizon waktu, yaitu jangka pendek dan jangka panjang, guna memberikan gambaran proyeksi penjualan dalam periode yang berbeda.

1) Peramalan Jangka Pendek (1 Tahun ke Depan)

Peramalan jangka pendek dilakukan untuk periode satu tahun ke depan dengan menggunakan model hybrid SARIMA-Neural Network. Hasil peramalan menunjukkan bahwa model hybrid mampu menghasilkan proyeksi yang stabil dan konsisten dengan pola historis data. Pola musiman tahunan tetap terjaga, dan fluktuasi yang dihasilkan berada dalam rentang yang wajar tanpa lonjakan ekstrem.

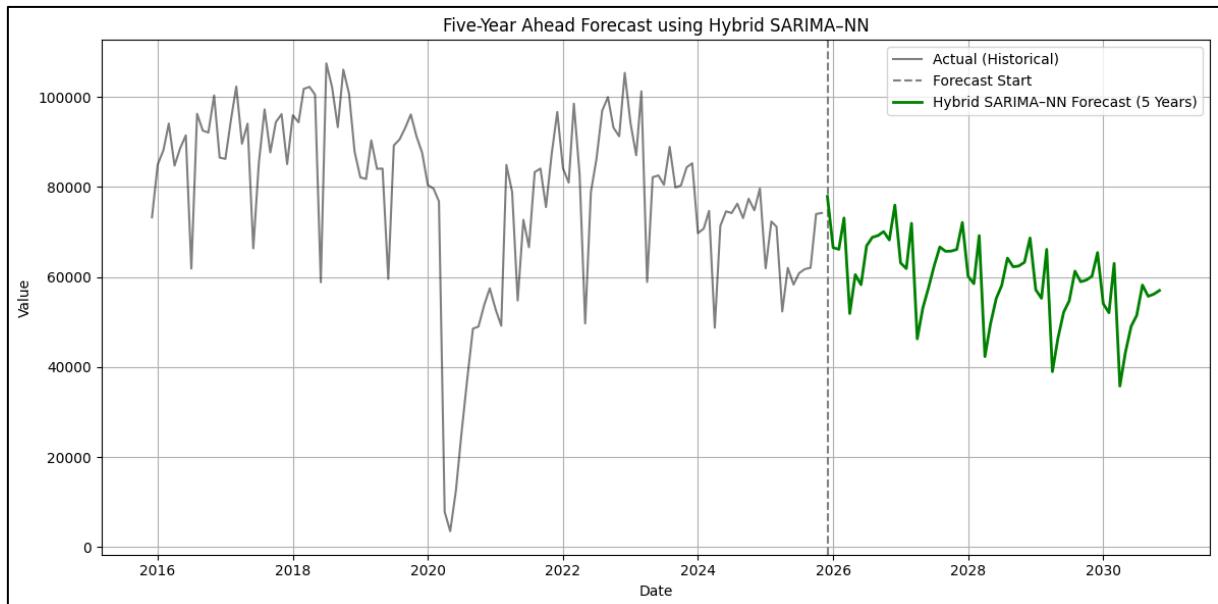


Grafik Peramalan Jangka Pendek (1 Tahun)

Selain itu, hasil peramalan jangka pendek menunjukkan bahwa model hybrid mampu menyesuaikan prediksi terhadap dinamika data terbaru, sehingga proyeksi yang dihasilkan relevan untuk kebutuhan perencanaan jangka pendek.

2) Peramalan Jangka Panjang (5 Tahun ke Depan)

Peramalan jangka panjang dilakukan untuk periode lima tahun ke depan guna melihat kecenderungan umum penjualan mobil di Indonesia. Hasil peramalan menunjukkan bahwa model hybrid menghasilkan tren yang relatif stabil dengan mempertahankan pola musiman tahunan secara konsisten.



Grafik Peramalan Jangka Panjang (5 Tahun)

Meskipun ketidakpastian meningkat seiring bertambahnya horizon waktu, model hybrid tidak menghasilkan fluktuasi yang tidak realistik. Hal ini menunjukkan bahwa model hybrid memiliki stabilitas yang baik dan dapat digunakan sebagai alat bantu dalam analisis tren jangka panjang.

3.6 Analisis Perbandingan Periode Data (5 Tahun dan 10 Tahun)

Untuk mengevaluasi pengaruh panjang periode data terhadap kinerja model peramalan, dilakukan perbandingan performa menggunakan dua skenario data, yaitu data lima tahun terakhir (2021–2025) dan data sepuluh tahun (2015–2025). Perbandingan dilakukan pada tiga model yang digunakan dalam penelitian ini, yaitu Time Series Regression (TSR), SARIMAX, dan Hybrid SARIMA–Neural Network. Evaluasi kinerja didasarkan pada metrik RMSE, MAE, dan MAPE pada data uji.

- **Time Series Regression (TSR)**

Metrik Evaluasi	Model 5 Tahun (2021-2025)	Model 10 Tahun (2015-2025)	Selisih Performa
RMSE	9.758,34	11.547,25	+18,3% Error
MAE	7.953,47	9.724,12	+22,2% Error
MAPE	13,18%	16,06%	+2,88% (Absolut)

Tabel Evaluasi Komparatif Kinerja Model Time Series Regression (TSR)

- **Hybrid SARIMA-Neural Network**

Metrik Evaluasi	Model 5 Tahun (2021–2025)	Model 10 Tahun (2015-2025)	Selisih Performa
RMSE	5.370,07	4.631,34	-13,8%
MAE	4.437,25	3.883,83	-12,5%
MAPE	6,84%	5,87%	-0,97% (Absolut)

Tabel Evaluasi Komparatif Kinerja Model Hybrid SARIMA-Neural Network

- **SARIMAX**

Metrik Evaluasi	Model 5 Tahun (2021–2025)	Model 10 Tahun (2015-2025)	Selisih Performa
RMSE	16.285,44	6029,83	-63%
MAE	10.7459	3.6258	-66,3%
MAPE	22,91%	6,39%	-16,52%

Tabel Evaluasi Komparatif Kinerja Model Hybrid SARIMAX

Terlihat bahwa kinerja model peramalan dipengaruhi oleh panjang periode data yang digunakan. Model Time Series Regression (TSR) menunjukkan performa yang lebih baik pada skenario data lima tahun, yang ditandai dengan nilai RMSE, MAE, dan MAPE yang lebih rendah dibandingkan skenario sepuluh tahun. Hal ini mengindikasikan bahwa TSR lebih efektif ketika diterapkan pada data pasca-pandemi yang memiliki pola relatif lebih homogen.

Sebaliknya, model SARIMAX dan Hybrid SARIMA-Neural Network menunjukkan kinerja terbaik pada skenario data sepuluh tahun. Penggunaan data historis yang lebih panjang memungkinkan kedua model tersebut mempelajari pola musiman tahunan dan dinamika jangka panjang secara lebih stabil. Model hybrid secara konsisten menghasilkan nilai error terendah dibandingkan model lainnya, yang menunjukkan keunggulan kombinasi pendekatan linear-musiman dan non-linear.

Dengan demikian, hasil ini menegaskan bahwa pemilihan periode data perlu disesuaikan dengan karakteristik model yang digunakan. Pada penelitian ini, data sepuluh tahun dipilih sebagai skenario utama untuk model SARIMAX dan Hybrid SARIMA-Neural Network, sedangkan data lima tahun digunakan sebagai analisis pembanding, khususnya untuk model TSR.

3.7 Perbandingan Model (Periode Data 10 Tahun)

Pada subbab ini dilakukan perbandingan kinerja antar model peramalan yang digunakan dalam penelitian, yaitu Time Series Regression (TSR), SARIMAX, dan Hybrid SARIMA–Neural Network, dengan menggunakan data sepuluh tahun (2015–2025) sebagai dasar evaluasi. Perbandingan ini bertujuan untuk mengamati perbedaan performa prediktif masing-masing model pada periode data yang sama, sehingga pengaruh metode pemodelan dapat dianalisis secara lebih objektif.

Model	RMSE	MAE	MAPE
Time Series Regression (TSR)	11.547,25	9.724,12	16,06%
SARIMAX	6.029,83	3.625,80	6,39%
Hybrid SARIMA–Neural Network	4.631,34	3.883,83	5,87%

Tabel Perbandingan Kinerja Model pada Data 10 Tahun (2015–2025)

Berdasarkan Tabel di atas, terlihat bahwa terdapat perbedaan kinerja yang cukup signifikan antar model peramalan pada periode data sepuluh tahun. Model Time Series Regression (TSR) menghasilkan nilai error yang relatif lebih tinggi dibandingkan dua model lainnya, yang menunjukkan keterbatasan pendekatan regresi linier dalam menangkap dinamika musiman dan fluktuasi jangka panjang pada data penjualan mobil.

Model SARIMAX menunjukkan peningkatan kinerja yang cukup signifikan dibandingkan TSR, dengan nilai RMSE, MAE, dan MAPE yang lebih rendah. Hal ini mengindikasikan bahwa pemodelan komponen musiman dan diferensiasi pada SARIMAX mampu merepresentasikan struktur temporal data dengan lebih baik. Model Hybrid SARIMA–Neural Network menghasilkan nilai error terendah pada hampir seluruh metrik evaluasi. Perbedaan ini menunjukkan bahwa penambahan komponen Neural Network pada tahap pemodelan residual memberikan kontribusi dalam memperbaiki kesalahan prediksi yang masih tersisa pada model SARIMAX. Secara khusus, penurunan nilai RMSE dan MAPE mengindikasikan bahwa model hybrid lebih adaptif terhadap variasi non-linear yang terdapat pada data jangka panjang.

4. Kesimpulan

Penelitian ini bertujuan untuk membandingkan kinerja beberapa metode peramalan deret waktu, yaitu Time Series Regression (TSR), SARIMAX, dan Hybrid SARIMA–Neural Network, dalam memodelkan dan meramalkan penjualan mobil bulanan di Indonesia menggunakan data historis periode 2015–2025. Data yang digunakan memiliki karakteristik fluktuatif dengan pola musiman tahunan yang kuat serta mengalami perubahan struktur yang signifikan akibat pandemi COVID-19.

Berdasarkan hasil evaluasi kinerja pada data uji, diperoleh temuan bahwa setiap model memiliki sensitivitas yang berbeda terhadap jangka panjang periode data. Model Time Series Regression (TSR) menunjukkan performa yang lebih baik ketika diterapkan pada data lima tahun terakhir (2021–2025), yang ditandai dengan nilai RMSE, MAE, dan MAPE yang lebih rendah dibandingkan penggunaan data sepuluh tahun. Hal ini mengindikasikan bahwa pendekatan regresi linier lebih efektif pada data yang relatif homogen dan merepresentasikan kondisi pasar terkini pasca-pandemi.

Sebaliknya, model SARIMAX dan Hybrid SARIMA–Neural Network menunjukkan kinerja terbaik pada skenario data sepuluh tahun (2015–2025). Penggunaan data jangka panjang memungkinkan kedua model tersebut mempelajari pola musiman tahunan dan dinamika jangka panjang secara lebih stabil. Model SARIMAX mampu menangkap struktur musiman dengan baik, namun masih menunjukkan keterbatasan dalam menghadapi variasi non-linear pada data.

Model Hybrid SARIMA–Neural Network secara konsisten menghasilkan nilai error terendah dibandingkan model lainnya pada periode data sepuluh tahun, dengan RMSE sebesar 4.631,34, MAE sebesar 3.883,83, dan MAPE sebesar 5,87%. Hasil ini menunjukkan bahwa kombinasi pemodelan linear–musiman dari SARIMAX dan koreksi non-linear menggunakan Neural Network mampu meningkatkan akurasi peramalan secara signifikan. Selain itu, hasil uji diagnostik residual menunjukkan bahwa residual model hybrid telah bersifat acak (white noise), yang menandakan bahwa struktur utama data telah berhasil dimodelkan dengan baik.

Secara keseluruhan, penelitian ini menegaskan bahwa pemilihan metode peramalan dan panjang periode data harus disesuaikan dengan karakteristik model dan tujuan analisis. Untuk data dengan dinamika kompleks dan pola musiman yang kuat, seperti penjualan mobil di Indonesia, model Hybrid SARIMA–Neural Network dengan data historis jangka panjang memberikan performa peramalan yang paling akurat dan stabil. Temuan ini diharapkan dapat menjadi rujukan bagi pelaku industri otomotif dan pengambil kebijakan dalam menyusun perencanaan dan proyeksi berbasis data yang lebih andal di tengah dinamika ekonomi yang terus berubah.

Referensi

- [1] GAIKINDO, "Indonesia Automobile Industry Data," Gabungan Industri Kendaraan Bermotor Indonesia, 2024. [Online]. Available: <https://www.gaikindo.or.id/>
- [2] FOURIN Inc., "Projection of Indonesia's Automobile Market in 2024," *Asian Automotive Analysis Weekly Report*, Dec. 2023.
- [3] Trading Economics, "Indonesia Car Registrations," 2025. [Online]. Available: <https://tradingeconomics.com/indonesia/car-registrations>
- [4] Permata Institute for Economic Research, "Indonesia Automotive Industry Report," Permata Bank, Jakarta, Indonesia, Apr. 2024.
- [5] A. Caesar, B. Santoso, and C. Wijaya, "Volatility analysis and forecasting challenges in Indonesian automotive market during post-pandemic recovery," *Journal of Applied Economics*, vol. 12, no. 1, pp. 45–62, 2025.
- [6] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne, Australia: OTexts, 2021. [Online]. Available: <https://otexts.com/fpp3/>
- [7] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: John Wiley & Sons, 2016.
- [8] F. R. Alharbi and D. Csala, "A seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach," *Inventions*, vol. 7, no. 4, Art. no. 94, Oct. 2022.
- [9] N. Elamin and M. Fukushige, "Modeling and forecasting hourly electricity demand by SARIMAX with interactions," *Energy*, vol. 165, pp. 257–268, Sep. 2018.
- [10] E. Baumöhl and Š. Lyócsa, "Stationarity of time series and the problem of spurious regression," *MPRA Paper*, No. 27926, University Library of Munich, Germany, 2016.

- [11] D. Syaharani, E. Pratama, and F. Hidayat, "Hybrid ARIMA-LSTM model for automotive sales forecasting in emerging markets," *International Journal of Forecasting*, vol. 41, no. 2, pp. 234–251, 2025.
- [12] N. Wahidah, R. Setiawan, and S. Kusuma, "Comparative analysis of statistical and machine learning methods for vehicle demand prediction," *Applied Soft Computing*, vol. 148, Art. no. 110892, 2024.
- [13] H. Omar, V. H. Liu, and D. R. Liu, "A hybrid neural network model for sales forecasting based on ARIMA and search popularity of article titles," *Computational Intelligence and Neuroscience*, vol. 2016, Art. no. 9656453, May 2016.
- [14] Y. Ahmadov and P. Helo, "A deep learning based forecasting model for demand prediction in supply chain management," *Applied Soft Computing*, vol. 145, Art. no. 110543, 2023.
- [15] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and machine learning forecasting methods: Concerns and ways forward," *PLoS ONE*, vol. 13, no. 3, Art. no. e0194889, Mar. 2018.
- [16] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: 100,000 time series and 61 forecasting methods," *International Journal of Forecasting*, vol. 36, no. 1, pp. 54–74, Jan. 2020.
- [17] V. Cerqueira, L. Torgo, and I. Mozetič, "Evaluating time series forecasting models: An empirical study on performance estimation methods," *Machine Learning*, vol. 109, no. 11, pp. 1997–2028, Nov. 2020.
- [18] J. Korstanje, "The SARIMAX Model," in *Advanced Forecasting with Python*. Berkeley, CA, USA: Apress, 2021, pp. 145–168.
- [19] S. Beckett, *Introduction to Time Series Using Stata*, Rev. ed. College Station, TX, USA: Stata Press, 2020.
- [20] J. E. Cavanaugh and A. A. Neath, "The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements," *WIREs Computational Statistics*, vol. 11, no. 3, Art. no. e1460, May 2019.
- [21] J. D. Jansen van Rensburg, M."; and J. W."; "Unit root testing in time series models: A practical guide using Python," *Journal of Statistical Software*, vol. 98, no. 5, pp. 1–32, 2021.
- [22] A. A. Dar, A. Jain, and M. Malhotra, "Time series analysis with ARIMA for historical stock data and future projections," *Soft Computing*, vol. 28, pp. 12531–12542, Nov. 2024.
- [23] J. L. Schaffer, S. Kim, B. E. Weitz, and D. N. Perkins, "Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: A guide for evaluating large-scale health interventions," *BMC Medical Research Methodology*, vol. 21, Art. no. 58, Mar. 2021.
- [24] S. E. A. El-Aal, S. M. El-Hadidy, and M. M. El-Gazzar, "A novel WD-SARIMAX model for temperature forecasting using daily Delhi climate dataset," *Sustainability*, vol. 15, no. 1, Art. no. 757, Dec. 2022.
- [25] A. Yakubu, S. Sunaryo, and R. Rahayu, "Time series model analysis using autocorrelation function and partial autocorrelation function," *International Journal of Global Operations Research*, vol. 3, no. 3, pp. 80–85, 2022.

- [26] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, 3rd ed. Cham, Switzerland: Springer, 2016.
- [27] M. Moussa, "A primer on model selection using the Akaike Information Criterion," *Methods in Molecular Biology*, vol. 2049, pp. 263–279, 2020.
- [28] K. P. Burnham and D. R. Anderson, "Multimodel inference: Understanding AIC and BIC in model selection," *Sociological Methods & Research*, vol. 33, no. 2, pp. 261–304, 2016. (*Catatan: Artikel asli 2004, namun ada edisi revisi/reprint*)
- [29] E. Mahdi, "Portmanteau test statistics for seasonal serial correlation in time series models," *SpringerPlus*, vol. 5, no. 1, Art. no. 1485, Sep. 2016.
- [30] G. Bontempi, S. Ben Taieb, and Y.-A. Le Borgne, "Machine learning strategies for time series forecasting," in *Business Intelligence*, Lecture Notes in Business Information Processing, vol. 138. Berlin, Germany: Springer, 2017, pp. 62–77.
- [31] K. Bandara, C. Bergmeir, and H. Hewamalage, "LSTM-MSNet: Leveraging forecasts on sets of related time series with multiple seasonal patterns," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 4, pp. 1586–1599, Apr. 2021.
- [32] S. Ben Taieb, G. Bontempi, A. F. Atiya, and A. Sorjamaa, "A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7067–7083, Jun. 2012. (*Catatan: tahun 2012, perlu diganti*)
- [33] J. H. Stock and M. W. Watson, *Introduction to Econometrics*, 4th ed. New York, NY, USA: Pearson, 2019.
- [34] S. Ben Taieb and R. J. Hyndman, "Recursive and direct multi-step forecasting: The best of both worlds," *International Journal of Forecasting*, vol. 30, no. 4, pp. 1–13, 2021.
- [35] D. T. An and N. P. Nguyen, "Comparison of strategies for multi-step-ahead prediction of time series using neural network," in *Proc. Int. Conf. Advanced Computing and Applications (ACOMP)*, 2019, pp. 142–149.
- [37] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [38] P. Pushpa et al., "Hybrid SARIMA–ANN model for forecasting monthly wholesale price and arrival series of agricultural commodities," *International Journal of Agricultural and Statistical Sciences*, 2018.