

Analisis Prediksi Stok Obat Berdasarkan Data Historis Penjualan Menggunakan Model ARIMA (*Autoregressive Integrated Moving Average*)

Sri Rahayu Hutapea¹⁾, Nova Noor Kamala Sari²⁾, Widiatry³⁾

¹⁾²⁾³⁾Jurusan Teknik Informatika, Fakultas Teknik, Universitas Palangka Raya
Jl. H. Timang Kampus UPR Tanjung Nyaho Palangka Raya, Kalimantan Tengah

¹⁾srihayuhutapeaa@gmail.com

²⁾novanoorks@it.upr.ac.id

³⁾widiatry@it.upr.ac.id

Abstrak

Permasalahan penumpukan obat kadaluwarsa dan kekurangan stok akibat ketidakakuratan estimasi kebutuhan menjadi kendala utama dalam pengelolaan persediaan di Toko Obat Basaria. Kondisi ini berdampak negatif pada efisiensi operasional dan kepuasan pelanggan. Penelitian ini bertujuan mengevaluasi kinerja model Autoregressive Integrated Moving Average (ARIMA) dan Seasonal Autoregressive Integrated Moving Average (SARIMA) dalam memprediksi penjualan obat. Pemilihan model didasarkan pada hasil dekomposisi komponen dataset: ARIMA digunakan jika dataset tidak menunjukkan pola musiman, sedangkan SARIMA dipertimbangkan jika pola musiman teridentifikasi. Selanjutnya, penelitian ini menganalisis akurasi kedua model menggunakan metrik Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan Mean Absolute Percentage Error (MAPE) untuk menentukan model prediksi terbaik yang akan digunakan memproyeksikan penjualan di masa mendatang. Data yang digunakan adalah deret waktu penjualan obat berdasarkan jenisnya, mencakup dataset penjualan Bodrex, Vitamin, dan Paramex. Hasil penelitian menunjukkan bahwa performa model optimal bervariasi untuk setiap jenis obat. Untuk dataset penjualan Bodrex, model ARIMA (1,0,0) memberikan akurasi terbaik dengan MAE (18.14), RMSE (22.16), dan MAPE (27.07%). Sementara itu, dataset penjualan Vitamin paling baik diprediksi oleh model SARIMA (1,0,2),(3,1,2)₁₂ dengan MAE (18.92), RMSE (25.37), dan MAPE (47,68%). Untuk penjualan Paramex, model SARIMA (2,0,2),(1,1,2)₁₂ menjadi yang terbaik dengan akurasi MAE (15.57), RMSE (26.13), dan MAPE (12.61%). Prediksi penjualan ini diharapkan dapat menjadi dasar yang akurat bagi Toko Obat Basaria dalam mengoptimalkan pengelolaan stok, mengurangi kerugian akibat obat kadaluwarsa, dan meningkatkan ketersediaan produk demi kepuasan pelanggan.

Kata kunci: ARIMA, SARIMA, prediksi penjualan, prediksi stok obat

Abstract

Toko Obat Basaria faces significant challenges in inventory management, primarily due to the accumulation of expired drugs and stock shortages caused by inaccurate demand forecasting. This situation negatively impacts both operational efficiency and customer satisfaction. This research aims to evaluate the performance of Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) models in predicting drug sales. Model selection is based on the decomposition analysis of the dataset components: ARIMA is applied when the dataset does not exhibit a seasonal pattern, while SARIMA is considered if a seasonal pattern is identified. Furthermore, this study analyzes the accuracy of both models using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) metrics to determine the optimal predictive model for projecting future sales. The study utilizes time series sales data for three distinct drug types: Bodrex, Vitamin, and Paramex. The research findings indicate that the optimal model performance varies across each drug type. For the Bodrex sales dataset, the ARIMA (1,0,0) model

yielded the best accuracy, with an MAE(18.14), RMSE(22.16), and MAPE(27.07%). Conversely, the Vitamin sales dataset was most accurately predicted by the SARIMA (1,0,2),(3,1,2)12 model, achieving an MAE(18.92), RMSE(25.37), and MAPE(47.68%). For Paramex sales, the SARIMA (2,0,2),(1,1,2)12 model proved superior, demonstrating an accuracy with an MAE(15.57), RMSE(26.13), and MAPE(12.61%). These sales predictions are expected to provide Toko Obat Basaria with an accurate foundation for optimizing stock management, thereby minimizing losses from expired inventory and enhancing product availability to improve overall customer satisfaction.

Keywords: ARIMA, SARIMA, sales forecasting, medicine stock prediction

1. PENDAHULUAN

Toko Obat Basaria merupakan salah satu tempat usaha yang bergerak di bidang penjualan berbagai jenis obat, khususnya obat bebas dan obat bebas terbatas yang tidak memerlukan resep dokter. Dalam operasionalnya, Toko Obat Basaria menghadapi permasalahan dalam pengelolaan persediaan stok obat. Terdapat beberapa jenis obat yang tidak terjual dan menumpuk di gudang hingga melewati masa kedaluwarsa, sementara beberapa obat lainnya justru mengalami kekurangan stok. Permasalahan ini timbul karena pihak toko mengalami kesulitan dalam memperkirakan kebutuhan stok yang tepat untuk setiap jenis obat. Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah dengan menerapkan metode peramalan berbasis data historis. Metode ini memungkinkan pelaku usaha untuk menganalisis pola penjualan dari waktu ke waktu dan menggunakannya sebagai dasar untuk memperkirakan kebutuhan stok di masa mendatang. Dalam penelitian ini, digunakan metode *Autoregressive Integrated Moving Average* (ARIMA) dan *Seasonal ARIMA*.

Untuk mengevaluasi tingkat akurasi prediksi yang dihasilkan oleh model ARIMA atau SARIMA, digunakan 3 metrik evaluasi yaitu MAPE (*Mean Absolute Percentage Error*), MAE (*Mean Absolute Error*), RMSE (*Root Mean Squared Error*). Metrik-metrik ini digunakan untuk mengukur tingkat kesalahan antara nilai aktual dan nilai prediksi, sehingga dapat menilai kinerja model secara kuantitatif (Hyndman & Athanasopoulos, 2018).

2. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Prediksi kebutuhan stok obat merupakan aspek penting dalam rantai pasok apotek untuk mencegah kekurangan maupun kelebihan stok. Untuk mendukung peramalan ini, metode statistik seperti ARIMA dan SARIMA telah banyak digunakan dalam berbagai penelitian.

Tempati et al. (2022) mengembangkan model kombinasi *Artificial Neural Network* (ANN) dan ARIMA untuk memprediksi stok obat di Apotik Total Life Clinic. Hasilnya menunjukkan tingkat akurasi prediksi yang tinggi dengan variasi MAPE antar obat. Penelitian ini menunjukkan bahwa kombinasi model statistik dan *machine learning* dapat meningkatkan performa prediksi stok.

Suseno dan Wibowo (2023) membandingkan model ARIMA dan SARIMA dalam meramalkan penjualan telur ayam. SARIMA menunjukkan performa lebih baik berdasarkan nilai MSE yang lebih rendah, mengindikasikan keunggulan model musiman dalam menangani data dengan pola berulang.

2.2 Model ARIMA dan SARIMA

Model *Autoregressive Integrated Moving Average* (ARIMA) merupakan salah satu metode statistik yang paling umum digunakan dalam analisis data deret waktu, khususnya untuk data non-stationer. Model ini diperkenalkan oleh Box dan Jenkins pada tahun 1976, dan dikenal juga

sebagai model *Box-Jenkins* [1]. ARIMA menggabungkan tiga komponen utama, yaitu *autoregressive* (AR), *integrated* (I), dan *moving average* (MA), yang secara umum dinotasikan sebagai ARIMA (p, d, q). Komponen AR merepresentasikan pengaruh nilai masa lalu terhadap nilai saat ini, komponen I menunjukkan jumlah *differencing* yang diperlukan agar data menjadi stationer, dan komponen MA menggambarkan pengaruh kesalahan masa lalu terhadap nilai observasi sekarang [2].

ARIMA digunakan untuk membangun model peramalan dengan mempertimbangkan hubungan linier antara nilai saat ini dengan nilai sebelumnya serta nilai residualnya. Model ini dinilai efektif dalam memodelkan tren dan fluktuasi acak jangka pendek, dan telah banyak diterapkan dalam berbagai bidang, termasuk peramalan penjualan dan pengelolaan persediaan [2].

SARIMA (*Seasonal ARIMA*) merupakan perluasan dari ARIMA untuk menangani data musiman, dengan menambahkan komponen musiman (P, D, Q) ke dalam struktur model [3]. SARIMA lebih unggul dalam menangkap pola musiman dan tren periodik, menjadikannya pilihan yang tepat dalam peramalan stok dengan pola berulang [4].

2.3 Stationeritas dan Transformasi Data

Model ARIMA mengharuskan data bersifat stationer, yaitu memiliki rata-rata dan varians yang konstan dari waktu ke waktu [1]. Uji *Augmented Dickey-Fuller* (ADF) digunakan untuk mendeteksi unit root sebagai indikator ketidakstationeran [2]. Jika data tidak stationer, transformasi melalui *differencing* dilakukan untuk menghilangkan tren. Umumnya cukup dilakukan sekali (*first-order*), sedangkan *second-order* hanya jika tren nonlinier masih terlihat. *Differencing* sebaiknya dihindari karena dapat mengaburkan pola musiman [3].

2.4 Identifikasi Parameter Model

Parameter ARIMA terdiri dari p (*autoregressive*), d (*differencing*), dan q (*moving average*), sedangkan SARIMA merupakan parameter musiman berupa P, D, Q, dan periode musiman s. penentuan nilai d atau D dilakukannya berdasarkan hasil uji stationeritas, sementara nilai p, q, P, dan Q diidentifikasi melalui pola yang muncul pada grafik *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). ACF digunakan untuk mengidentifikasi orde dari komponen q dan Q, sedangkan PACF untuk komponen p dan P [1].

2.5 Kriteria Pemilihan Model

Pemilihan model ARIMA dan SARIMA optimal mempertimbangkan keseimbangan antara akurasi prediksi dan kompleksitas model. Dalam penelitian ini digunakan *Akaike Information Criterion* (AIC) dan *Bayesian Information Criterion* (BIC) untuk memilih kombinasi parameter model terbaik. AIC menghukum model dengan banyak parameter untuk mencegah *overfitting*, sementara BIC memberikan penalti lebih besar terhadap kompleksitas sehingga cenderung memilih model yang lebih sederhana. Nilai AIC dan BIC yang lebih rendah menunjukkan model lebih efisien secara statistik dan prediktif [2].

2.6 Evaluasi Performa Model

Setelah model ARIMA dan SARIMA dibangun, analisis residual dilakukan untuk mengevaluasi apakah model telah menangkap seluruh informasi dalam data. Penelitian ini menggunakan uji Ljung-Box guna mendeteksi autokorelasi residual secara keseluruhan. Statistik uji ini mengikuti distribusi *chi-kuadrat*, dan jika nilai $p > 0,05$, residual dianggap acak atau *white noise*, menandakan kecukupan model. Sebaliknya, adanya autokorelasi menunjukkan bahwa model belum cukup baik dalam merepresentasikan data [1].

3. METODE PENELITIAN

Tahapan penelitian ini adalah pengumpulan data, pra-pemrosesan data, membangun dan melatih model peramalan, dan prediksi kebutuhan di masa mendatang. Penelitian ini menggunakan data historis penjualan obat yang dikumpulkan secara manual dari catatan

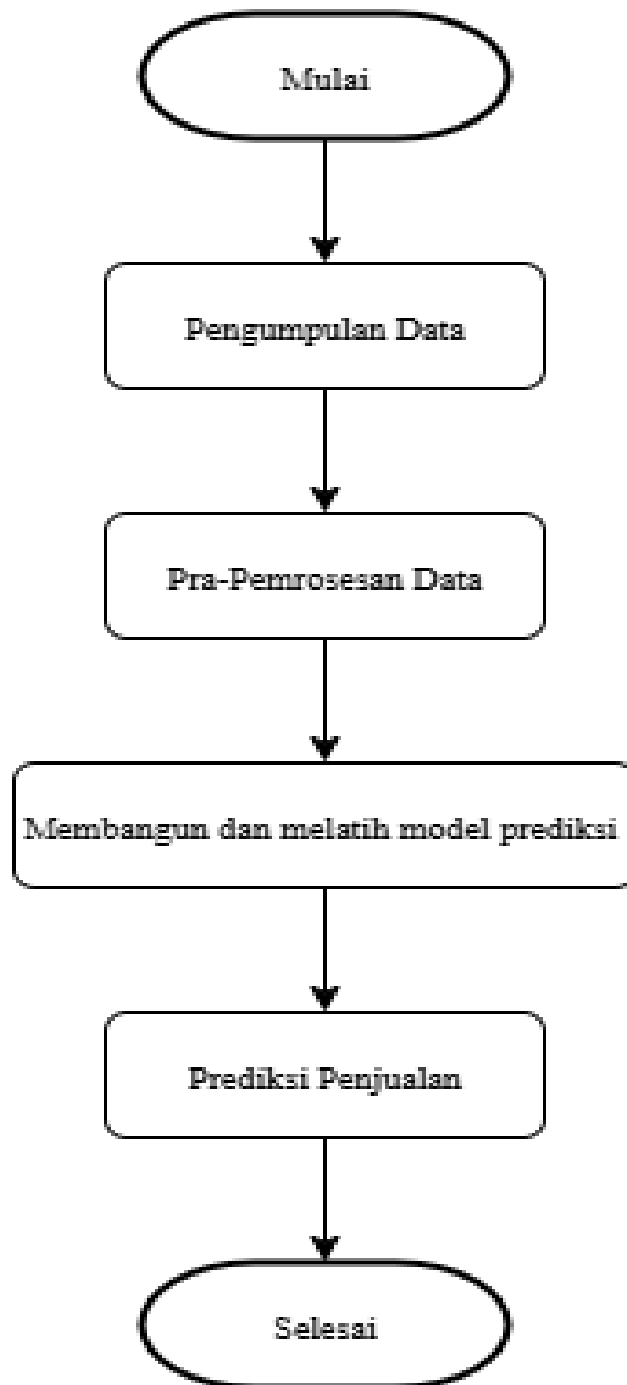
transaksi selama 70 bulan (sejak Januari 2019 hingga Oktober 2024) di Toko Obat Basaria. Dataset tersebut mencakup tanggal dan total unit obat terjual perbulannya untuk 3 jenis obat yaitu Bodrex, Vitamin, dan Paramex. Model peramalan yang digunakan adalah ARIMA untuk dataset yang tidak mengandung pola musiman, dan SARIMA untuk dataset yang mengandung pola musiman. Analisis data dan implementasi model dilakukan menggunakan bahasa pemrograman *Python* pada platform *Google Colaboratory*.

Dalam analisis data deret waktu menggunakan model ARIMA dan SARIMA, pra-pemrosesan data adalah tahapan fundamental untuk menjamin validitas inferensi statistik dan akurasi prediksi. Proses ini meliputi identifikasi dan penanganan *outlier* menggunakan metode *Interquartile Range (IQR)* yang kemudian diatasi dengan *winsorisasi*. Selanjutnya, stasioneritas data diuji dengan melakukan pengujian *Augmented Dickey-Fuller (ADF)*; jika data belum stasioner, diterapkan *differencing*, dimulai dengan *seasonal differencing* jika ada pola musiman, diikuti *non-seasonal differencing*, hingga data mencapai stasioneritas berdasarkan *p-value* hasil uji ADF. Namun jika data tidak memiliki pola musiman, dilakukan *non-seasonal differencing*.

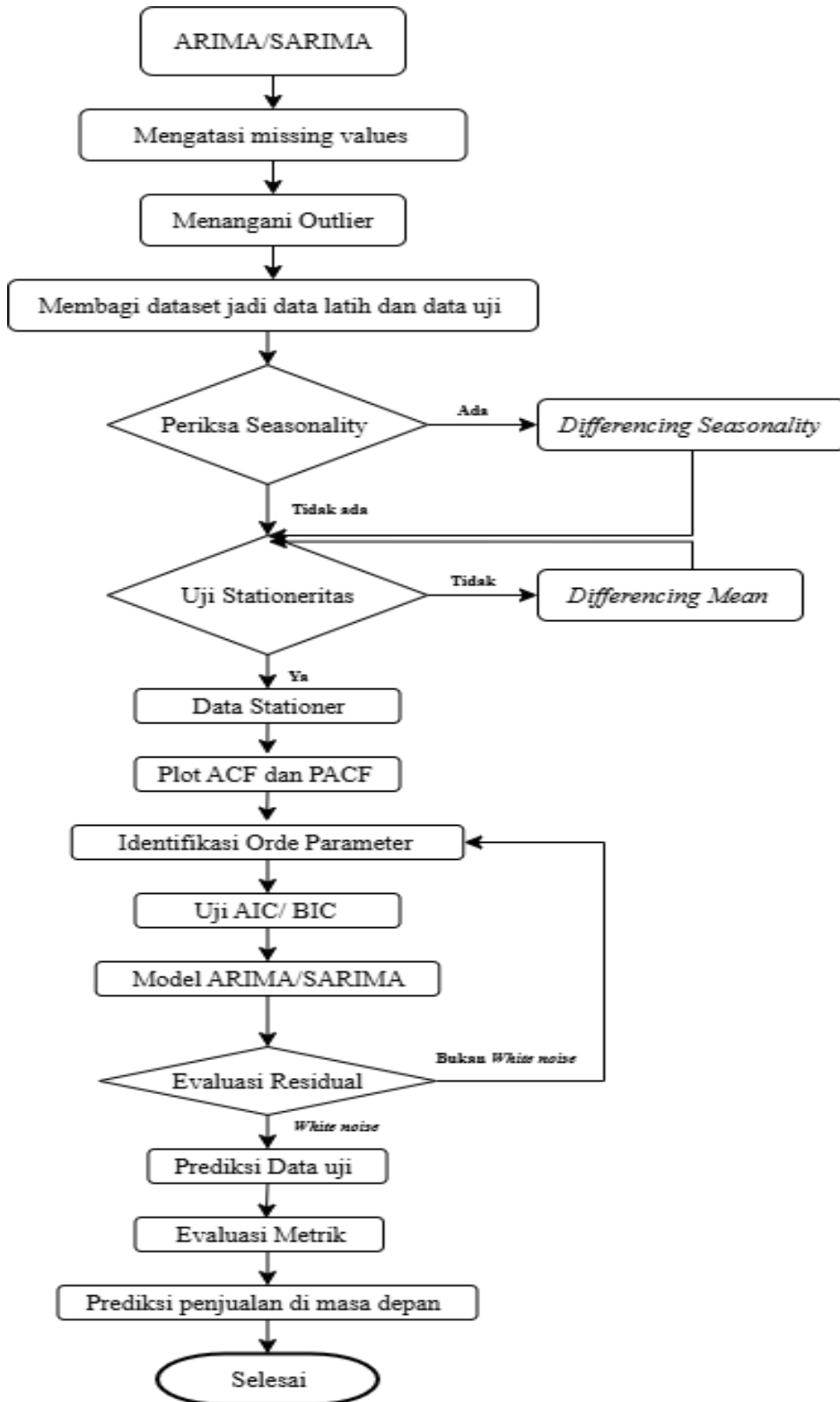
Setelah data mencapai stasioneritas melalui pra-pemrosesan, pembangunan model ARIMA (*Autoregressive Integrated Moving Average*) atau SARIMA (*Seasonal Autoregressive Integrated Moving Average*) dimulai. Tahap ini diawali dengan analisis plot *Autocorrelation Function (ACF)* dan *Partial Autocorrelation Function (PACF)* dari data yang telah stasioner untuk mengidentifikasi orde komponen non-musiman (yaitu parameter *p* untuk AR dan *q* untuk MA) dan musiman (yaitu parameter *P* untuk AR dan *Q* untuk MA). Berdasarkan deteksi pola musiman dari dekomposisi STL (*Seasonal-Trend decomposition using Loess*), model awal dipilih: ARIMA(*p,d,q*) jika tidak ada pola musiman, atau SARIMA(*p,d,q*)(*P, D, Q*)*s* jika pola musiman teridentifikasi (di mana *d* dan *D* masing-masing menunjukkan orde *differencing* non-musiman dan musiman, serta *s* adalah panjang musim). Setiap kombinasi parameter potensial kemudian dievaluasi menggunakan Kriteria Informasi Akaike (AIC) dan Kriteria Informasi Bayesian (BIC) untuk menemukan keseimbangan optimal antara kecocokan model dan kompleksitasnya. Model dengan nilai AIC dan BIC terbaik selanjutnya menjalani evaluasi residual, termasuk Uji Ljung-Box untuk memastikan bahwa residualnya bersifat *white noise*. Jika hasil evaluasi residual memadai, model tersebut digunakan untuk prediksi dan dievaluasi lebih lanjut dengan metrik kinerja untuk menentukan model final yang paling akurat, misalnya dalam memprediksi kebutuhan stok obat di masa depan.

Tahapan selanjutnya dalam proses ini adalah prediksi kebutuhan di masa depan, di mana model prognostik terbaik—yang telah divalidasi melalui serangkaian pengujian dan menunjukkan nilai metrik evaluasi terendah seperti MAE (*Mean Absolute Error*), RMSE (*Root Mean Squared Error*), dan MAPE (*Mean Absolute Percentage Error*)—akan diaplikasikan untuk memproyeksikan data penjualan obat pada periode mendatang. Hasil proyeksi penjualan ini kemudian berfungsi sebagai basis fundamental untuk mengestimasi secara akurat potensi kebutuhan stok obat yang optimal, memungkinkan pihak pengelola toko obat membuat keputusan yang lebih tepat dan efisien dalam perencanaan pengadaan serta pengelolaan inventaris.

Berikut adalah gambar alur peramalan persediaan stok obat pada penelitian ini. Gambar 1 menjelaskan alur tahapan penelitian yang dilakukan. Gambar 2 menjelaskan alur prediksi menggunakan metode ARIMA dan SARIMA.



Gambar 1 Diagram alur tahapan penelitian



Gambar 2 Diagram langkah ARIMA dan SARIMA

4. PEMBAHASAN

4.1 Dataset Penjualan Bodrex

Model peramalan terbaik untuk memprediksi penjualan Bodrex adalah model ARIMA (1,0,0), dengan nilai error MAE (18.14), RMSE (22.16), dan MAPE (27.07%). Berikut ini adalah

prediksi penjualan obat Bodrex setiap bulannya, untuk 12 bulan berikutnya setelah Oktober 2024 menggunakan model ARIMA (1, 0, 0):

Tabel 1 Prediksi penjualan bodrex 12 bulan selanjutnya

Tanggal	Prediksi Penjualan
30-11-2024	92
31-12-2024	103
31-01-2025	89
28-02-2025	103
31-03-2025	110
30-04-2025	73
31-05-2025	77
30-06-2025	75
31-07-2025	75
31-08-2025	76
30-09-2025	76
31-10-2025	88

Berdasarkan Tabel 1, estimasi penjualan bodrex di bulan november 2024 sebanyak 92 unit, bulan Desember 2024 sebanyak 103 unit, bulan Januari 2025 sebanyak 89 unit, bulan Februari 2025 sebanyak 103 unit, bulan Maret 2025 sebanyak 110 unit, bulan April 2025 sebanyak 73 unit, bulan Mei 2025 sebanyak 77 unit, bulan Juni 2025 sebanyak 75 unit, bulan Juli 2025 sebanyak 75 unit, bulan Agustus 2025 sebanyak 76 unit, bulan September 2025 sebanyak 76 unit, dan bulan Oktober 2025 sebanyak 88 unit. Secara keseluruhan estimasi jumlah stok yang harus disediakan ± 1037 unit obat.

4.2 Dataset Penjualan Vitamin

Model peramalan terbaik untuk memprediksi penjualan Vitamin adalah model SARIMA (1,0,2)(3,1,2)₁₂ dengan nilai error MAE (18.92), RMSE (25.37), dan MAPE (47.68%). Berikut ini adalah prediksi penjualan vitamin 12 bulan selanjutnya menggunakan model SARIMA (1,0,2)(3,1,2)₁₂:

Tabel 2 Prediksi penjualan bodrex 12 bulan selanjutnya

Tanggal	Prediksi Penjualan
30-11-2024	55
31-12-2024	33
31-01-2025	3
28-02-2025	32
31-03-2025	107
30-04-2025	115
31-05-2025	164
30-06-2025	166
31-07-2025	76
31-08-2025	51
30-09-2025	64
31-10-2025	46

Berdasarkan Tabel 2, estimasi penjualan bodrex di bulan november 2024 sebanyak 55 unit, bulan Desember 2024 sebanyak 33 unit, bulan Januari 2025 sebanyak 3 unit, bulan Februari 2025 sebanyak 32 unit, bulan Maret 2025 sebanyak 107 unit, bulan April 2025 sebanyak 115 unit, bulan Mei 2025 sebanyak 164 unit, bulan Juni 2025 sebanyak 166 unit, bulan Juli 2025 sebanyak 76 unit, bulan Agustus 2025 sebanyak 51 unit, bulan September 2025 sebanyak 64 unit, dan bulan Oktober 2025 sebanyak 64 unit. Secara keseluruhan estimasi jumlah stok yang harus disediakan ± 912 unit obat.

4.3 Dataset Penjualan Paramex

Model peramalan terbaik untuk memprediksi penjualan Paramex adalah model SARIMA (2,0,2)(1,1,2)₁₂ dengan nilai error MAE (15.57), RMSE (26.13), dan MAPE (12.61%). Berikut

ini adalah prediksi penjualan vitamin 12 bulan selanjutnya menggunakan model SARIMA (2,0,2)(1,1,2)₁₂:

Tabel 3 Prediksi penjualan obat paramex 12 bulan berikutnya

Tanggal	Prediksi Penjualan
30-11-2024	162
31-12-2024	160
31-01-2025	137
28-02-2025	146
31-03-2025	164
30-04-2025	162
31-05-2025	156
30-06-2025	145
31-07-2025	168
31-08-2025	173
30-09-2025	150
31-10-2025	167

Berdasarkan Tabel 3, estimasi penjualan bodrex di bulan november 2024 sebanyak 162 unit, bulan Desember 2024 sebanyak 160 unit, bulan Januari 2025 sebanyak 137 unit, bulan Februari 2025 sebanyak 146 unit, bulan Maret 2025 sebanyak 164 unit, bulan April 2025 sebanyak 162 unit, bulan Mei 2025 sebanyak 156 unit, bulan Juni 2025 sebanyak 145 unit, bulan Juli 2025 sebanyak 168 unit, bulan Agustus 2025 sebanyak 173 unit, bulan September 2025 sebanyak 150 unit, dan bulan Oktober 2025 sebanyak 167 unit. Secara keseluruhan estimasi jumlah stok yang harus disediakan ± 1890 unit obat.

5. KESIMPULAN

Performa dan akurasi model ARIMA dan SARIMA dalam memprediksi kebutuhan stok obat untuk tiga jenis dataset penjualan obat yaitu: Untuk dataset penjualan Bodrex hasil evaluasi model ARIMA (1,0,0) memiliki nilai MAE (18.14), RMSE (22.16), dan MAPE (27.07%), sedangkan model SARIMA (1,0,1),(1,1,2)₁₂ dengan nilai MAE (23.00), RMSE (28.42), dan MAPE (32.55%). Untuk dataset penjualan Vitamin hasil evaluasi model SARIMA (1,0,2),(3,1,2)₁₂, memiliki nilai MAE (18.92), RMSE (25.37), dan MAPE (47.68%), sedangkan model ARIMA (2,0,2) dengan nilai MAE (29.5), RMSE (37.76), dan MAPE (55.26%). Untuk dataset penjualan Paramex hasil evaluasi model SARIMA (2,0,2),(1,1,2)₁₂ memiliki nilai MAE (15.57), RMSE (26.13), dan MAPE (12.61%), sedangkan model ARIMA (1,0,2) dengan nilai MAE (21.14), RMSE (26.21), dan MAPE (14.54%).

Model prediksi yang menunjukkan akurasi tertinggi dan paling layak direkomendasikan untuk memprediksi kebutuhan stok obat di masa depan, berdasarkan perbandingan metrik evaluasi untuk ketiga dataset penjualan obat adalah model prediksi ARIMA (1,0,0) untuk dataset penjualan Bodrex. Model prediksi SARIMA (1,0,2)(3,1,2)₁₂, untuk dataset penjualan Vitamin. Model prediksi SARIMA (2,0,2)(1,1,2)₁₂, untuk dataset penjualan Paramex.

6. DAFTAR PUSTAKA

- [1] 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: Wiley, 2015.
- [2] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne, Australia: OTexts, 2021.
- [3] C. Chatfield, *The Analysis of Time Series: An Introduction*, 6th ed. Boca Raton, FL, USA: Chapman & Hall/CRC, 2003.
- [4] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, 2nd ed. Hoboken, NJ, USA: Wiley, 2015.