

APPLICATION OF ARIMA AND ARIMAX METHODS TO PREDICT THE NUMBER OF VISITORS TO HOTEL XYZ PEKANBARU

PENERAPAN METODE *ARIMA* DAN *ARIMAX* UNTUK MEMPREDIKSI JUMLAH PENGUNJUNG HOTEL XYZ PEKANBARU

Julio Vernando¹, Fitri Insani^{2*}, Okfalisa³, Fitra Kurnia⁴

^{1, 2, 3, 4} Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Riau

email: 12050116440@students.uin-suska.ac.id¹, fitri.insani@uin-suska.ac.id², okfalisa@gmail.com³,
fitra.k@uin-suska.ac.id⁴

Abstract - Predicting the number of visitors to Hotel XYZ is one of the steps that can be taken by the hotel to find out how many visitors will increase in each upcoming holiday season. The purpose of this study is to forecast the number of visitors to Hotel XYZ from June 2023 to July 2024 using the ARIMA and ARIMAX comparison methods. The research methodology encompasses problem identification, data collection, data processing, and ARIMA and ARIMAX analysis, which involves testing the parameters (p , d , q) selected based on the ACF and PACF using the AIC Model. Based on the test results, ARIMAX (5, 0, 3) has the lowest AIC, which is 3495.2, followed by ARIMAX (3, 0, 5), which has a slightly higher AIC. The results showed that the ARIMAX (5, 0, 3) model is the most accurate model for predicting data (eg the number of hotel guests, room demand, or income), with an RMSE value of 15.80% and a MAPE of 18.90%. Therefore, research that applies the ARIMAX model can provide real benefits in supporting operational efficiency, resource management, and hotel business strategy, ultimately increasing the competitiveness and profitability of the hotel.

Keywords - Prediction, Number of Visitors, Hotel, ARIMA Method, ARIMAX Method.

Abstrak - Memprediksi jumlah pengunjung Hotel XYZ merupakan salah satu langkah yang dapat dilakukan oleh pihak hotel untuk mengetahui seberapa besar peningkatan pengunjung setiap musim liburan mendatang. Penelitian ini bertujuan untuk memprediksi jumlah pengunjung Hotel XYZ dari bulan juni 2023 sampai dengan Juli 2024 menggunakan metode perbandingan ARIMA dan ARIMAX. Metodologi penelitian meliputi, identifikasi masalah, pengumpulan data, pengolahan data, analisis ARIMA dan ARIMAX dengan menguji parameter (p , d , q) yang dipilih berdasarkan ACF dan PACF menggunakan Model AIC. Berdasarkan hasil pengujian, ARIMAX (5, 0, 3) memiliki AIC terendah, yaitu 3495.2, diikuti oleh ARIMAX (3, 0, 5) yang memiliki AIC sedikit lebih tinggi. Hasil penelitian menunjukkan bahwa model ARIMAX (5, 0, 3) merupakan model yang paling akurat untuk memprediksi data (misalnya jumlah tamu hotel, permintaan kamar, atau pendapatan), dengan nilai RMSE sebesar 15,80% dan MAPE sebesar 18,90%. Oleh karena itu penelitian dengan penerapan Model ARIMAX ini dapat memberikan manfaat nyata dalam mengatur operasional, termasuk penjadwalan karyawan, pengadaan logistik, dan pengelolaan sumber daya secara efisien.

Kata Kunci – Prediksi, Jumlah Pengunjung, Hotel, Metode ARIMA, Metode ARIMAX.

I. PENDAHULUAN

Kemajuan teknologi yang cepat dapat memicu berbagai perubahan signifikan. Salah satu sektor yang terdampak adalah industri pariwisata. Pengembangan sektor ini berpotensi meningkatkan pendapatan daerah serta menarik minat wisatawan, baik dari dalam maupun luar negeri. Salah satu bentuk akomodasi yang berkembang dalam industri ini adalah hotel. Hotel merupakan jenis akomodasi yang secara komersial menyediakan layanan penginapan, makanan, minuman, dan berbagai fasilitas lainnya bagi para tamu[1]. Hotel XYZ Pekanbaru adalah salah satu hotel berbintang empat yang senantiasa berusaha meningkatkan mutu layanannya guna memenuhi ekspektasi para tamu. Meski demikian, hotel ini menghadapi masalah dengan perubahan jumlah pengunjung harian yang signifikan, khususnya saat musim libur atau event tertentu. Perubahan jumlah pengunjung yang signifikan ini menyulitkan pihak manajemen dalam mengatur operasional, termasuk penjadwalan karyawan, pengadaan logistik, dan pengelolaan sumber daya secara efisien[2].

Berdasarkan data atau variabel yang memengaruhi kondisi saat ini, prediksi merupakan proses untuk memprediksi apa yang akan terjadi di masa mendatang. Memprediksi jumlah pengunjung menjadi langkah penting bagi pihak hotel untuk mengetahui seberapa besar peningkatan pengunjung di setiap musim liburan[3]. Penelitian ini menggunakan data harian jumlah pengunjung Hotel XYZ dari Juni 2023 hingga Juli 2024. Dengan demikian, dibutuhkan pendekatan dan metode yang tepat, karena metode yang lebih akurat akan menghasilkan prediksi dengan tingkat ketepatan yang lebih baik. Salah satu metode statistik paling populer yang digunakan untuk memodelkan dan memprediksi data deret waktu adalah model *Autoregressive Integrated Moving Average* (ARIMA), yang terdiri dari tiga komponen utama: *autoregressive* (AR), *integrated* (I), dan *moving average* (MA), yang memungkinkan model untuk menangani data stasioner melalui perbedaan[3]. Meskipun ARIMA sangat efektif, namun ARIMA memiliki beberapa keterbatasan di antaranya membutuhkan data stasioner. Metode ini juga tidak dapat menangani pola non-linier[4]. Sensitif juga terhadap outlier dan kesalahan data, serta kesulitan pada data dengan perubahan struktur mendadak. Agar prediksi menjadi lebih akurat dan relevan, pada data harian yang biasanya memiliki pola musiman dan pengaruh faktor eksternal (seperti hari libur), model ini harus diubah, misalnya menambahkan Metode ARIMAX dengan menambahkan variabel eksogen. ARIMAX, memperkenalkan variabel eksogen sebagai input tambahan untuk menangkap pengaruh faktor eksternal yang dapat memengaruhi pola deret waktu utama[5].

Menurut[6] Peramalan Jumlah Kedatangan Jalur Udara di Bandara Depati Amir dengan menggunakan Model ARIMA, akurasi diperoleh dengan MAE sebesar 46,197, MAPE sebesar 46,197, dan RMSE sebesar 13905,436, bersama dengan hasil prediksi yang menunjukkan kedatangan. Metode ARIMAX efek variasi kalender libur akhir tahun digunakan untuk meramalkan jumlah wisatawan asing yang akan datang ke Indonesia. Dari hasil analisis jumlah wisatawan, Hasil dari peramalan jumlah wisatawan yang datang ke Indonesia dengan metode ARIMAX pada bulan Januari 2016 dan februari 2016 menunjukkan bahwa variasi ini terjadi pada saat ke-t sebesar 814052,96 dan 858423,27 orang[7]. Selanjutnya Peramalan Tingkat Penghunian Hotel di Provinsi Riau Menggunakan ARIMA dan ARIMAX dimana penelitian ini mendapatkan akurasi yang lebih baik dibandingkan ARIMA, dengan nilai MAPE sebesar 9,64 persen dan RMSE sebesar 4,21 persen. Penelitian ini menyimpulkan jika tidak ada perubahan kebijakan pemerintah terkait pembatasan sosial hingga akhir tahun, model ARIMAX memprediksi Tingkat Penghunian Hotel Desember 2021 sebesar 38,59 persen[8]. Berdasarkan penjelasan latar belakang, penelitian ini bermaksud untuk melakukan Penerapan Metode ARIMA dan ARIMAX dalam memprediksi jumlah pengunjung pada Hotel XYZ Pekanbaru. Tujuan dilakukan penelitian ini untuk mengetahui seberapa meningkat jumlah pengunjung pada saat musim liburan, serta memberikan rekomendasi yang dapat menjadi rujukan dalam mengatur operasional, termasuk penjadwalan karyawan, pengadaan logistik, dan pengelolaan sumber daya secara efisien.

II. SIGNIFIKASI STUDI

A. Prediksi

Prediksi dapat didefinisikan sebagai aktivitas terstruktur untuk memprediksi apa yang mungkin terjadi di masa mendatang dengan menggunakan informasi masa lalu dan masa kini yang ada untuk mengurangi kesalahan. Prediksi juga dapat dilakukan dengan membuat gambaran masa depan menggunakan teknik statistik. [1].

B. Pengunjung hotel

Pengunjung dan tamu hotel adalah bagian penting dari bisnis hotel, jadi mereka berhak untuk dilayani dengan baik. Ada dua kategori pengunjung hotel diantaranya:

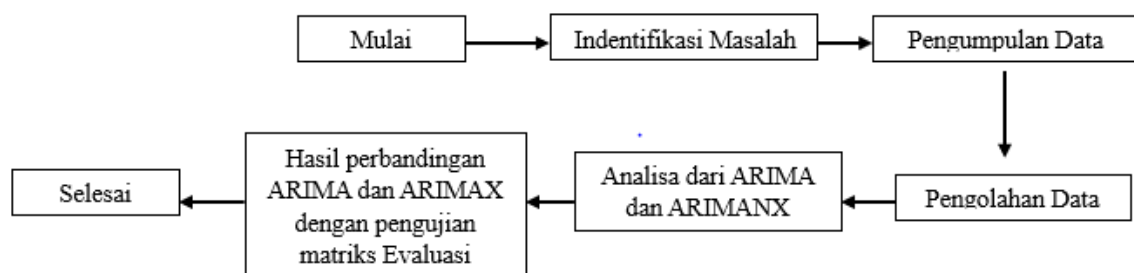
1. Pengunjung berdasarkan kebutuhan, yaitu pengunjung hotel yang datang untuk keperluan tertentu seperti rekreasi, olah raga, bisnis, transaksi dagang, konferensi, kegiatan seremonial, dan lain sebagainya.
2. Pengunjung berdasarkan lama tinggal, yaitu pengunjung yang tinggal dalam jangka waktu lama (umumnya 3 bulan). Oleh karena itu, hotel tertentu menyediakan fasilitas akomodasi bagi mereka seperti rumah sewa apartemen/hotel hunian.

C. Model ARIMA dan ARIMAX

Model ARIMA merupakan gabungan dari tiga komponen utama AR (autoregressive), I (integrated), dan MA (*moving average*) yang digunakan untuk menganalisis data stasioner melalui proses diferensiasi. Setelah diferensiasi dilakukan sebanyak d kali, model ARMA menjadi ARIMA [9]. ARIMAX adalah pengembangan ARIMA yang menambahkan variabel eksogen, sehingga tidak hanya mempertimbangkan nilai masa lalu variabel Y , tetapi juga pengaruh variabel independen pada waktu yang sama. [10].

D. Metodologi Penelitian

Penelitian ini melewati beberapa tahapan metodologi yang dirancang untuk memperoleh data yang valid dan menyusun analisis yang jelas tentang jumlah pengunjung Hotel XYZ di Pekanbaru. Gambar 1 menunjukkan alur kerja dari penelitian ini.



Gambar 1. Metodologi Penelitian

1. Identifikasi Masalah

Identifikasi masalah dilakukan untuk mengetahui permasalahan yang terjadi pada studi kasus penelitian. Hotel XYZ di Pekanbaru menghadapi fluktuasi jumlah pengunjung, terutama selama musim liburan, lonjakan tamu yang tidak terduga sering menyebabkan beban kerja berlebih bagi para karyawan. Perubahan jumlah pengunjung yang signifikan ini menyulitkan pihak manajemen dalam mengatur operasional, termasuk penjadwalan karyawan, pengadaan logistik, dan pengelolaan sumber daya secara efisien. Untuk itu diperlukannya metode prediksi yang tepat guna membantu manajemen dalam pengambilan keputusan strategis. Metode ARIMA dan ARIMAX dipilih karena mampu menghasilkan estimasi yang lebih akurat dengan memanfaatkan data historis serta variabel eksternal, seperti hari libur[11].

2. Pengumpulan Data

Data juga dapat didefinisikan sebagai bagian atau fakta yang belum tersusun yang memiliki makna dan terkait dengan peristiwa yang sebenarnya terjadi. Peneliti mengumpulkan data setelah masalah diidentifikasi. Data yang digunakan dalam penelitian ini berjumlah 365 data dari Juli 2023 hingga Juni 2024. Selanjutnya, data akan diolah untuk keberlanjutan penelitian [12].

3. Pengolahan Data

Pengolahan data akan melalui proses pembersihan data (*cleaning*) untuk memastikan tidak ada kesalahan pencatatan, data *noise*, atau outlier yang dapat mempengaruhi hasil analisis. Pada tahap ini juga data kemudian dianalisis untuk menentukan apakah data bersifat stasioner atau memerlukan proses transformasi seperti, *differencing* atau *log transformation*, agar sesuai dengan asumsi model ARIMA [13]. Data juga akan dikumpulkan untuk digunakan dalam model ARIMAX jika ada variabel eksternal yang relevan, seperti promosi hotel, musim liburan, atau kondisi ekonomi lokal. Pada pengolahan data dilakukan nya uji stasioner dimana Uji stasioner merupakan tahap penting dalam analisis deret waktu untuk memastikan bahwa data yang digunakan dalam pemodelan tidak memiliki tren atau variabilitas yang berubah seiring Waktu. Salah satu metode yang umum digunakan untuk menguji stasioneritas suatu data adalah Uji *Augmented Dickey-Fuller* (ADF), jika nilai statistik ADF lebih kecil dari nilai kritis atau p-value < 0,05, maka hipotesis nol ditolak dan data dianggap stasioner [14]. Persamaan uji ADF dapat dilihat pada persamaan (1):

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

4. Analisa dari Metode ARIMA dan ARIMAX

Ada beberapa tahapan yang dilakukan pada analisa dari metode ARIMA dan ARIMAX ini, Tahap pertama dilakukannya analisis ACF dan PACF dimana, *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) digunakan untuk menentukan nilai parameter dalam model ARIMA [15]. ACF mengukur sejauh mana data saat ini berkorelasi dengan lag sebelumnya, sedangkan PACF mengisolasi hubungan langsung antara lag tertentu dengan data saat ini, mengabaikan efek lag di antaranya. Rumus ACF dapat dilihat pada persamaan (2) dan rumus PACF dapat dilihat pada persamaan (3).

$$\rho(k) = \frac{\sum_{t=1}^{N-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^N (X_t - \bar{X})^2} \quad (2) \quad \alpha_{kk} = \frac{p_k - \sum_{j=1}^{k-1} \alpha_{j,k-1} p_{k-1}}{1 - \sum_{j=1}^{k-1} \alpha_{j,k-1} p_{k-1}} \quad (3)$$

Tahap Kedua dilakukannya analisis metode ARIMA dan ARIMAX menggunakan Metode AIC. *Akaike Information Criterion* (AIC) merupakan suatu ukuran relatif dari model statistika. Karena AIC bertujuan sebagai peramalan dan dapat menjelaskan kecocokan model dengan data saat ini dan nilai yang akan datang (peramalan di luar sampel), syarat utama model dianggap baik jika memiliki nilai AIC terkecil. Rumus AIC dapat dilihat pada persamaan (4).

$$AIC = 2k - 2 \ln(\mathcal{L}) \quad (4)$$

5. Hasil Perbandingan ARIMA dan ARIMAX dengan Pengujian Matriks Evaluasi

Setelah melakukan analisa kedua metode selanjutnya dilakukan nya perbandingan ARIMA dan ARIMAX dengan pengujian Matriks Evaluasi. Metrik evaluasi adalah ukuran yang digunakan untuk menilai performa suatu model prediksi atau klasifikasi dengan membandingkan hasil prediksi model dengan nilai aktual. Berikut adalah jenis dan fungsi dari beberapa metrik evaluasi: *Root Mean Squared Error* (RMSE), metrik yang digunakan untuk mengukur seberapa jauh nilai prediksi model dari nilai sebenarnya. RMSE menghitung akar dari rata-rata kuadrat dari semua selisih antara nilai prediksi dan nilai aktual. Rumus persamaan RMSE dapat dilihat pada persamaan (5). *Mean Absolute Percentage Error* (MAPE), metrik yang menunjukkan rata-rata kesalahan

absolut sebagai persentase antara nilai aktual dan prediksi. Rumus persamaan MAPE dapat dilihat pada Persamaan (6).

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5) \quad \text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (6)$$

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Data Jumlah Pengunjung Hotel XYZ Pekanbaru dari Juli 2023 hingga Juni 2024 dapat dilihat pada tabel 1. Tabel 1 ini merupakan data deret waktu harian yang mencakup pengamatan dari tanggal 1 Juli 2023 hingga 30 Juni 2024, dengan jumlah total 366 hari yang mencerminkan tahun kabisat. Tabel ini terdiri dari tiga kolom utama, yaitu "Tanggal," "Jumlah," dan "Hari_Libur." Kolom "Jumlah" merepresentasikan nilai harian dari suatu aktivitas atau kejadian tertentu dan dapat dilihat pada tabel tersebut bahwa data tertinggi yaitu pada tanggal 5 juli 2023 sebanyak 132 pengunjung dan data terendah pada tanggal 30 Juni 2024 sebanyak 46 pengunjung. Sedangkan kolom "Hari_Libur" menunjukkan status hari libur, hari libur dapat dilihat pada tanggal 2 juli 2023 dan 30 juni 2024 dimana nilai 1 mengindikasikan hari libur. Angka 0 mengindikasikan hari biasa yang terjadi di tanggal 1,3,4,5 Juli 2023 sampai dengan tanggal 30 juni 2024.

Tabel 1. Data Historis Pengunjung Hotel XYZ

No	Tanggal	Jumlah	Hari_Libur
1	1-Jul-23	69	0
2	2-Jul-23	64	1
3	3-Jul-23	90	0
4	4-Jul-23	102	0
5	5-Jul-23	132	0
.....			
.....			
.....			
362	26-Jun-24	115	0
363	27-Jun-24	83	0
364	28-Jun-24	128	0
365	29-Jun-24	79	0
366	30-Jun-24	46	1

B. Pengolahan Data

Pada tahap ini, dilakukan uji stasioner dengan uji Augmented Dickey-Fuller (ADF) Hasil Pengujian dari stasioner dapat dilihat pada tabel 2.

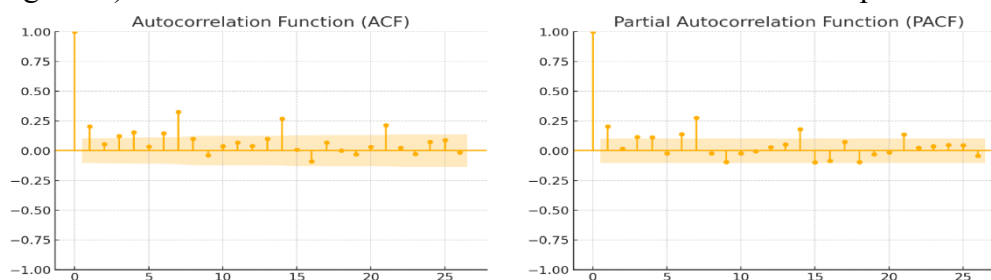
Tabel 2. Uji Stasioner dengan ADF

Hasil Uji	Nilai
ADF Statistic	-3,45062
p-value	0,00935
Critical Value 1%	-3,44928
Critical Value 5%	-2,86988
Critical Value 10%	-2,57121

Pada tabel 2 Menunjukkan bahwa hasil uji yang dilakukan adalah uji *Augmented Dickey-Fuller* (ADF). Statistik ADF yang diperoleh adalah -3,450615, yang menunjukkan nilai yang lebih negatif dibandingkan dengan nilai kritis pada tingkat signifikansi 1%, yaitu -3,449282. Ini berarti hipotesis nol, yang mengasumsikan bahwa data deret waktu tersebut tidak stasioner, dapat ditolak. Selain itu, *p-value* yang sebesar 0,009350, yang lebih kecil dari 0,05, juga memberikan bukti kuat untuk menolak hipotesis nol pada tingkat signifikansi 5%. Nilai kritis pada 5% adalah -2,869881, dan pada 10% adalah -2,571214, yang semuanya lebih besar daripada statistik ADF yang diperoleh. Dengan demikian, dapat disimpulkan bahwa data deret waktu tersebut stasioner, karena statistik ADF lebih negatif dari nilai kritis pada tingkat signifikansi yang relevan.

C. Analisis Metode RIMA dan ARIMAX

Ada beberapa tahapan yang dilakukan pada tahap ini, yang pertama melakukan analisis ACF dan PACF. Analisis ini penting dalam pemilihan parameter *p* (autoregressive/AR) dan *q* (moving average/MA) dalam model ARIMA. Grafik analisa ACF dan PACF dapat dilihat pada gambar 2.



Gambar 2. Grafik Analisa ACF dan PACF

Menunjukkan dua grafik yaitu fungsi autokorelasi (ACF) di sisi kiri dan fungsi autokorelasi parsial (PACF) di sisi kanan. Kedua grafik ini digunakan untuk menganalisis keterkaitan lag dalam suatu deret waktu. Nilai autokorelasi pada lag 0 pada grafik ACF adalah 1.00, seperti yang diharapkan. Beberapa lag lainnya memiliki nilai positif atau negatif kecil, dan sebagian besar titik berada dalam rentang $\pm 0,25$. Area *orange* menunjukkan interval kepercayaan dan sebagian besar titik yang berada di rentang ini, menunjukkan tidak banyak lag yang memiliki autokorelasi yang signifikan. Kedua grafik PACF di sisi kanan menunjukkan pola yang serupa, dengan nilai autokorelasi parsial terbesar terjadi pada lag 0 (1.00), dan beberapa lag lainnya memiliki nilai sekitar $\pm 0,25$ atau lebih rendah, dengan sebagian besar tetap dalam batas interval kepercayaan ini mungkin menunjukkan bahwa tidak ada komponen musiman atau tren yang signifikan dalam data tersebut. Interpretasi Grafik ACF yang dihasilkan menunjukkan perlambatan (*decay*) secara bertahap tanpa adanya cutoff yang jelas. Pola ini mengindikasikan bahwa terdapat komponen *moving average* (MA) dalam data, yang berarti hubungan antara nilai saat ini dengan nilai masa lalu tidak langsung terputus setelah beberapa lag, tetapi melemah secara bertahap.

Setelah dilakukannya analisis ACF dan PACF tahap selanjutnya adalah analisis metode ARIMA dan ARIMAX menggunakan Metode AIC. Pengujian Model ARIMA dan ARIMAX dengan AIC dapat dilihat pada tabel 3.

Tabel 3. Pengujian model dengan menggunakan AIC

Model	Model (p, d, q)	AIC
ARIMA	(5, 0, 5)	3526.5
ARIMA	(3, 0, 5)	3548.2
ARIMA	(3, 0, 4)	3548.7
ARIMAX	(5, 0, 3)	3495.2
ARIMAX	(3, 0, 5)	3495.6
ARIMAX	(3, 0, 4)	3496.8

Hasil pengujian menunjukkan bahwa ARIMA (5, 0, 5) memiliki AIC terendah, yaitu 3526.52, diikuti oleh ARIMA (3, 0, 5) dan ARIMA (3, 0, 4) dengan nilai AIC sedikit lebih tinggi. Namun, model-model ini masih memiliki kemungkinan untuk meningkatkan kemampuan dalam menangkap pola data. Model ARIMAX diuji dengan menambahkan variabel eksogen hari libur untuk melihat apakah faktor eksternal dapat meningkatkan akurasi prediksi. Berdasarkan hasil pengujian, ARIMAX (5, 0, 3) memiliki AIC terendah, yaitu 3495.2, diikuti oleh ARIMAX (3, 0, 5) yang memiliki AIC sedikit lebih tinggi. Model ARIMAX ini lebih baik dalam menangkap pola data, berkat pemanfaatan variabel eksternal yang relevan. Dari hasil pengujian, model ARIMAX (5, 0, 3) dipilih sebagai model terbaik karena memiliki AIC terendah. Langkah selanjutnya adalah melakukan prediksi jumlah pengunjung untuk periode mendatang, serta membandingkan hasil prediksi dengan data aktual guna mengukur akurasi model yang dipilih. Grafik perbandingan ARIMA dan ARIMAX menggunakan model terbaik AIC dapat dilihat pada gambar 4.2 dimana ARIMA (3, 0, 4) memiliki tingkat AIC tertinggi dengan nilai 3548,7 dan nilai ARIMAX (5, 0, 3) memiliki AIC terendah, yaitu 3495.2.

D. Hasil Perbandingan Metode ARIMA dan ARIMAX dengan Matriks Evaluasi

Pada tahap ini dilakukannya pengujian evaluasi kinerja model dengan menggunakan dua metrik kesalahan yang sering digunakan dalam analisis deret waktu, yaitu *Root Mean Squared Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE) berdasarkan persamaan rumus (3) dan (4). Hasil evaluasi RMSE dan MAPE dapat dilihat tabel 4.

Tabel 4. Hasil Evaluasi RMSE dan MAPE		
Model	RMSE	MAPE (%)
ARIMA (5, 0, 5)	21.82	28.41
ARIMA (3, 0, 5)	22.43	28.9
ARIMA (3, 0, 4)	22.81	29.69
ARIMAX (5, 0, 3)	15.8	18.9
ARIMAX (3, 0, 5)	15.83	18.94
ARIMAX (3, 0, 4)	16.11	19.4

Tabel 4 menunjukkan bahwa model ARIMAX (5, 0, 3) yang menggunakan variabel eksternal hari libur sebagai faktor tambahan memiliki kinerja terbaik di antara semua model yang diuji. Dengan RMSE sebesar 15.80 dan MAPE 18.90%, model ini mampu memberikan prediksi yang lebih akurat dibandingkan model-model lainnya. Dalam hal ini, pemanfaatan variabel eksternal terbukti meningkatkan kemampuan model dalam menangkap pola data yang lebih kompleks. Sementara itu, model ARIMA (5, 0, 5) juga memberikan hasil yang baik dengan RMSE 21.82 dan MAPE 28.41%, tetapi tidak sebaik model ARIMAX (5, 0, 3). Model ARIMA (3, 0, 5) dan ARIMA (3, 0, 4) menunjukkan akurasi yang lebih rendah, dengan RMSE dan MAPE masing-masing lebih tinggi, yaitu 22.43 dan 28.90% untuk ARIMA (3, 0, 5), serta 22.81 dan 29.69% untuk ARIMA (3, 0, 4).

Temuan ini sejalan dengan studi sebelumnya, seperti oleh Rizalde et al. (2022) di Provinsi Riau, yang juga menunjukkan bahwa model ARIMAX lebih akurat dalam memprediksi tingkat hunian hotel dengan mempertimbangkan hari libur. Penelitian lain oleh Riestiansyah et al. (2022) di Bali bahkan menunjukkan peningkatan prediksi lebih signifikan saat menyertakan variabel harga dan promosi. Meskipun penelitian ini baru menggunakan satu variabel eksogen, hasilnya telah membuktikan bahwa penambahan faktor eksternal memberikan kontribusi positif terhadap akurasi prediksi. Ini menegaskan bahwa pola kunjungan hotel di Pekanbaru menunjukkan kesamaan tren dengan wilayah lain, namun juga membuka peluang pengembangan model lebih lanjut dengan menambahkan variabel-variabel lainnya.

Selanjutnya hasil model ARIMAX (5,0,3) digunakan untuk memprediksi jumlah pengunjung selama 30 hari ke depan. Dengan mempertimbangkan variabel eksternal berupa hari libur, model ini mampu menangkap pola musiman atau tren yang mungkin tidak terdeteksi hanya dari data historis. Hasil prediksi dari model ARIMA (5,0,5) dan ARIMAX (5,0,3) ditampilkan dalam tabel 5 dan divisualisasikan pada gambar 3.

Tabel 5. Hasil Prediksi Jumlah Pengunjung ARIMA dan ARIMAX

Tanggal	ARIMA(5,0,5)	ARIMAX(5,0,3)	Tanggal	ARIMA(5,0,5)	ARIMAX(5,0,3)
7/1/2024	70.66	54.88	7/16/2024	75.57	58.42
7/2/2024	74.2	56.78	7/17/2024	96.32	51.12
7/3/2024	96.37	85.09	7/18/2024	88.1	93.44
7/4/2024	89.73	95.38	7/19/2024	76.05	88.45
7/5/2024	76.74	86.76	7/20/2024	90.44	90.57
7/6/2024	90.1	92	7/21/2024	92.06	91.66
7/7/2024	92.07	90.73	7/22/2024	71.47	87.73
7/8/2024	70.97	87.91	7/23/2024	76.21	58.1
7/9/2024	74.9	58.67	7/24/2024	96.29	87.16
7/10/2024	96.35	85.99	7/25/2024	87.36	92.74
7/11/2024	88.89	94.22	7/26/2024	75.78	89.13
7/12/2024	76.36	87.65	7/27/2024	90.62	90.07
7/13/2024	90.26	91.23	7/28/2024	91.99	91.88
7/14/2024	92.08	91.29	7/29/2024	71.66	87.81
7/15/2024	71.24	87.75	7/30/2024	76.83	57.73

Hasil prediksi ini menunjukkan estimasi jumlah pengunjung berdasarkan model-model yang telah diuji. Berdasarkan evaluasi RMSE dan MAPE, model ARIMAX (5, 0, 3) lebih akurat dalam memprediksi jumlah pengunjung, karena model ini mempertimbangkan faktor eksternal, yaitu Hari Libur. Dengan RMSE 15.80 dan MAPE 18.90%, ARIMAX (5, 0, 3) memberikan hasil yang lebih baik dibandingkan model lainnya. Meskipun model ARIMA (5, 0, 5) juga memberikan hasil yang baik dengan RMSE 21.82 dan MAPE 28.41%, ARIMAX (5, 0, 3) yang menggunakan informasi tambahan terbukti lebih efektif dalam menangkap pola data. Oleh karena itu, ARIMAX (5, 0, 3) dipilih sebagai model terbaik untuk prediksi jumlah pengunjung di masa mendatang.

Model yang digunakan dalam penelitian ini memiliki sejumlah keterbatasan yang perlu dicermati. Secara umum, baik ARIMA maupun ARIMAX mengasumsikan hubungan linear antar variabel dan belum mampu menangkap pola non-linear yang mungkin terdapat dalam data. Selain itu, pengujian terhadap asumsi model seperti normalitas residual dan independensi residual belum dilakukan secara menyeluruh. Faktor outlier yang dapat mempengaruhi kestabilan model juga belum dianalisis secara eksplisit. Penelitian ini hanya melibatkan satu variabel eksogen, yaitu hari libur, padahal terdapat banyak faktor lain yang berpotensi memengaruhi jumlah pengunjung hotel, seperti kondisi cuaca, harga kamar, kegiatan promosi, serta event lokal yang sedang berlangsung. Oleh karena itu, hasil prediksi dapat ditingkatkan lebih lanjut dengan memasukkan variabel tambahan dan mempertimbangkan pendekatan model yang lebih fleksibel.

IV. KESIMPULAN

Dengan mempertimbangkan hasil perbandingan dan prediksi data jumlah pengunjung dari tahun 2023 sampai dengan 2024 dengan menggunakan Metode ARIMA dan ARIMAX, dapat menghasilkan kesimpulan sebagai berikut:

1. Model ARIMA (3, 0, 5) dan ARIMA (3, 0, 4) menunjukkan akurasi yang lebih rendah, dengan RMSE dan MAPE masing-masing lebih tinggi, yaitu 22.43 dan 28.90% untuk ARIMA (3, 0, 5), serta 22.81 dan 29.69% untuk ARIMA (3, 0, 4).
2. Hasil Ramalan model ARIMA (5, 0, 5) memberikan hasil yang baik dengan RMSE 21.82 dan MAPE 28.41%, tetapi tidak sebaik model ARIMAX (5, 0, 3).
3. Jumlah pengunjung Hotel XYZ pada tahun 2023 sampai dengan 2024 memiliki variasi yang berbeda beda.
4. Hasil dari mencari prediksi menggunakan error RMSE untuk masing-masing model menunjukkan bahwa semakin rendah nilai RMSE nya semakin baik model tersebut bekerja dalam melakukan prediksi.
5. Manajemen disarankan untuk mengadopsi model ini sebagai dasar dalam perencanaan strategis, khususnya dalam proyeksi permintaan kamar dan pengambilan keputusan operasional.
6. Penyesuaian jadwal kerja staf, pengelolaan logistik, dan penyediaan perlengkapan hotel harus disesuaikan dengan fluktuasi jumlah tamu agar efisiensi operasional tetap terjaga.
7. Selanjutnya perlunya ada pengembangan Metode selain dari metode ARIMA dan ARIMAX, contohnya seperti, SARIMA dan SARIMAX atau penambahan variabel cuaca untuk dapat melanjutkan penelitian ini.

REFERENSI

- [1] C. Chandra and A. A. Fajrin, "Forecasting Konsumsi Barang Barang Pada Storage Hotel Dengan Metode Autoregressive Integrated Moving Average," *J. Comasie*, vol. 03, 2021.
- [2] A. Pramudita, "Prosiding Seminar Edusainstech MEMPERKIRAKAN TINGKAT PENGHUNI HOTEL MENGGUNAKAN ANALISIS ARIMA DENGAN APLIKASI MINITAB," *Pros. Semin. Edusaunstech*, pp. 249–257, 2020.
- [3] F. Fejriani, M. Hendrawansyah, L. Muharni, S. F. Handayani, and Syaharuddin, "Forecasting Peningkatan Jumlah Penduduk Berdasarkan Jenis Kelamin menggunakan Metode Arima," *J. Kajian, Penelit. dan Pengemb. Pendidik.*, vol. 8, no. 1 April, pp. 27–36, 2020, [Online]. Available: <http://journal.ummat.ac.id/index.php/geography/article/view/2261/pdf>
- [4] S. Mizan, B. D. P. R, N. Suriani, R. Mihaimin, N. Laili, and W. Ernita, "Peramalan Data Penduduk Miskin Provinsi Nusa Tenggara Barat (NTB) Model Auto Regressive Integrated Moving Average (ARIMA)," *J. Pemikir. dan Penelit. Pendidik. Mat.*, vol. 2, no. 1, pp. 1–10, 2019.
- [5] L. Y. Chiu, D. J. Arcega Rustia, C. Y. Lu, and T. Te Lin, "Modelling and Forecasting of Greenhouse Whitefly Incidence Using Time-Series and ARIMAX Analysis," *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 196–201, 2019, doi: 10.1016/j.ifacol.2019.12.521.
- [6] J. Fraction, M. Nalda, A. Simanjuntak, R. Tipani, Z. M. Afriyanti, and N. Hidayati, "(Forecasting the Number of Arrivals By Air At Depati Amir Airport Uses," vol. 3, no. 2, pp. 44–52, 2023.
- [7] N. N. Lingga, I. Indwiarti, and A. A. Rohmawati, "Pemodelan dan Peramalan Kedatangan Wisatawan ke Tempat Wisata dengan Google Trends Menggunakan Metode Variasi Kalender ARIMAX," *e-Proceeding Eng.*, vol. 8, no. 2, pp. 3361–3372, 2021.
- [8] F. A. Rizalde, S. Mulyani, and N. Bachtiar, "Forecasting Hotel Occupancy Rate in Riau Province Using ARIMA and ARIMAX," *Proc. Int. Conf. Data Sci. Off. Stat.*, vol. 2021, no. 1, pp. 578–589, 2022, doi: 10.34123/icdsos.v2021i1.199.
- [9] D. Rahman, I. W. SUMARJAYA, and I. K. G. SUKARSA, "Perbandingan Peramalan Hasil Produksi Ikan Menggunakan Metode Permulusan Eksponensial Holt-Winters Dan Arima," *E-Jurnal Mat.*, vol. 7, no. 4, p. 371, 2018, doi: 10.24843/mtk.2018.v07.i04.p227.
- [10] I. F. Amri, W. N. Ramadhan, S. Ainurrofiah, and M. Al Haris, "Pemodelan ARIMA dan ARIMAX untuk Memprediksi Jumlah Produksi Padi di Kota Magelang," *Sq. J. Math. Math. Educ.*, vol. 5, no. 2, pp. 93–105, 2023, doi: 10.21580/square.2023.5.2.17059.
- [11] R. Arianti, S. Sahriman, and L. P. Talangko, "Model ARIMA dengan variabel eksogen dan GARCH pada data kurs rupiah," *Estimasi J. Stat. Its Appl.*, vol. 3, no. 1, pp. 41–48, 2022, doi: 10.20956/ejsa.vi.11603.
- [12] W. Chiu, T. C. M. Fan, S. B. Nam, and P. H. Sun, "Knowledge mapping and sustainable development of esports research: A bibliometric and visualized analysis," *Sustain.*, vol. 13, no. 18, 2021, doi: 10.3390/su131810354.
- [13] F. Riestiansyah, D. Damayanti, M. Reswara, and R. Susetyoko, "Perbandingan metode ARIMA dan ARIMAX dalam Memprediksi Jumlah Wisatawan Nusantara di Pulau Bali," *J. Infomedia*, vol. 7, no. 2, p. 58, 2022, doi: 10.30811/jim.v7i2.3336.
- [14] M. A. Baker Siddique *et al.*, "Forecasting of tilapia (*Oreochromis niloticus*) production in Bangladesh using ARIMA model," *Heliyon*, vol. 10, no. 10, 2024, doi: 10.1016/j.heliyon.2024.e31174.
- [15] S. Soraya, I. Fitriana Aziza, M. Rizky, U. Juanda, G. Primajati, and P. Rahima, "Peramalan Jumlah Kunjungan Wisatawan di Provinsi Nusa Tenggara Barat (NTB) Menggunakan Metode Arima Box-Jenkins," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 6, no. 01, pp. 35–43, 2024, doi: 10.35580/variensiunm150.