

Penerapan Analisis Deret Waktu Arima dalam Meramalkan Wisatawan Mancanegara Asal Malaysia Tahun 2024

Rido Julistio^{*}, Ilham Faishal Mahdy

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

ridojulistio99@gmail.com, Ilhamfaishalmahdy@unisba.ac.id

Abstract. Forecasting is an activity aimed at predicting future events, covering short-term, medium-term, and long-term periods using time series data. Time series data refers to a collection of data arranged continuously over a specific period. Forecasting techniques in time series data analysis involve statistical modeling methods, including exponential smoothing, moving average, regression, and autoregressive integrated moving average (ARIMA). The ARIMA model is categorized into three groups: autoregressive (AR), moving average (MA), and the combined ARMA (autoregressive moving average) model, which incorporates characteristics of the first two models. Stationarity checks on data can be conducted using time series plots, the Augmented Dickey-Fuller (ADF) test, and the autocorrelation function (ACF) test. The Box-Cox transformation is applied when time series data exhibit non-stationarity in variance, requiring an assessment of stationarity issues through the ADF test and ACF test. Differencing is performed after the data achieve stationarity in variance. If the data remain non-stationary in the mean, further differencing is applied. International tourists are defined as individuals traveling to a country outside their place of residence. Using the Akaike Information Criterion (AIC) test, the best-fitting model was determined to be ARIMA (1,2,1) based on the smallest AIC value. This study concludes that the ARIMA time series analysis can be utilized to forecast the number of Malaysian international tourists in 2024.

Keywords: *Time Series Analysis, ARIMA, Box-Cox Transformation, Differencing, ACF test, AIC test, Travelers.*

Abstrak. Peramalan adalah kegiatan untuk memprediksi kejadian yang akan datang yang meliputi jangka pendek, menengah dan panjang menggunakan data deret waktu. Data deret waktu merupakan kumpulan data yang disusun pada waktu tertentu secara terus menerus. Teknik peramalan analisis data deret waktu dengan model peramalan berdasarkan pemodelan statistik yaitu (exponential smoothing, moving average, regresi dan autoregressive integrated moving average (ARIMA)). Model ARIMA dibagi ke dalam 3 kelompok, yaitu: model autoregressive (AR), moving average (MA), dan model campuran ARMA (autoregressive moving average) yang mempunyai karakteristik dari dua model pertama. Pegecekan stasioner data bisa dilakukan dengan plot time series, uji ADF dan uji ACF. Transformasi Box Cox digunakan ketika data deret waktunya mengandung masalah adanya ketidakstasioneran data dalam varians, sehingga harus ada proses pemeriksaan masalah kestasioneran yang dapat diperiksa menggunakan statistik ADF dan uji ACF. Proses pembedaan (differencing) dilakukan setelah data stasioner dalam varians. Proses pembedaan dilakukan jika data tidak stasioner dalam rata-rata. wisatawan secara internasional yakni setiap orang yang mengunjungi suatu negara di luar tempat tinggalnya. Dengan menggunakan Uji AIC dengan nilai terkecil didapatkan model terbaik ARIMA (1, 2, 1). Berdasarkan penelitian ini diketahui bahwa Analisis Deret Waktu ARIMA dapat digunakan untuk melakukan peramalan jumlah wisatawan mancanegara asal Malaysia tahun 2024.

Kata Kunci: *Analisis deret waktu, ARIMA, Transformasi Box-Cox, Differencing, Uji ACF, Uji AIC, Wisatawan.*

A. Pendahuluan

Pariwisata telah menjadi sektor prioritas dalam pembangunan perekonomian Indonesia. Pariwisata dibutuhkan bisa menjadi salah satu penggerak utama pada meningkatkan kecepatan pertumbuhan ekonomi pada Indonesia melalui penciptaan lapangan pekerjaan serta kesempatan berusaha, penerimaan devisa, dan pembangunan infrastruktur (Widyawati, 2013). Selain itu, pariwisata pula dapat digunakan buat memperkenalkan identitas serta kebudayaan nasional. Sehingga, pengembangan pariwisata akan terus dilanjutkan serta ditingkatkan melalui ekspansi serta pemanfaatan sumber serta potensi pariwisata nasional (Firmansyah & Suwanda, 2024).

Sebagai salah satu komoditi ekspor yang tidak dapat dilihat secara nyata, peran pariwisata semakin meningkat dalam perekonomian Indonesia. Dalam mengembangkan pariwisata internasional sangat diperlukan program yang terarah dan tepat dalam rangka meningkatkan jumlah kedatangan wisatawan mancanegara (wisman) (Ajeng Mega Pratiwi & Mutaqin, 2021). Hal ini dapat dilakukan dengan meningkatkan kegiatan pemasaran dan perbaikan dari berbagai fasilitas yang diperlukan wisman, seperti pelayanan imigrasi, fasilitas angkutan, perbankan, akomodasi, restoran, biro perjalanan, dan sebagainya (Wicaksana, 2021).

Untuk meningkatkan kegiatan pemasaran, diperlukan perencanaan yang sesuai berdasarkan informasi kuantitatif dan kualitatif tentang wisman pada masa-masa sebelumnya (Khalis Syahril Suryana et al., 2024). Dengan adanya pandemi Covid-19 yang melanda Indonesia sejak awal tahun 2020, informasi tentang wisman semakin diperlukan sebagai bahan evaluasi dan perencanaan pembangunan di masa yang akan datang. Tidak tersedianya data pada masa lalu akan menyulitkan dalam membuat rencana yang cermat dan terarah (Muhammad Farhan Praja Utama & Teti Sofia Yanti, 2024). Data statistik yang disajikan dalam publikasi ini dimaksudkan untuk memenuhi salah satu kebutuhan tersebut (Muhammad Rifqi Naufal & Nur Azizah Komara Rifai, 2024). Pada masa ini telah dikembangkan sejumlah metode peramalan dengan berbagai asumsi mengenai data yang akan diramalkan untuk masa yang akan datang (Fausett, 1994; Wei, 2006). Dengan adanya berbagai metode peramalan diharapkan akan terciptanya suatu aplikasi dan implementasi yang lebih baik untuk terwujudnya diberbagai bidang kehidupan, salah satunya adalah bidang pariwisata (Ma, 2018; Zhu, 2018).

Analisis deret waktu adalah salah satu strategi kuantitatif di mana evaluasi masa depan dibuat dengan mempertimbangkan kualitas masa lalu. Alasan teknik Analisis deret waktu adalah untuk melihat model dalam deret informasi yang dicatat dan memproyeksikan contoh-contoh tersebut ke masa depan (Guo, 2019; Gong, 2020). Ada beberapa teknik dalam melakukan peramalan termasuk moving average, Exponential Smoothing dan ARIMA, namun dalam ulasan ini teknik penentuan yang digunakan adalah ARIMA, karena ARIMA memiliki ketepatan yang sangat baik untuk melakukan estimasi peramalan (Zhang, 2003).

Pengukuran/harapan dikelompokkan menjadi 3-time skyline, yaitu estimasi jangka pendek, estimasi jangka menengah, danantisipasi jangka panjang. Perkiraan eksplorasi ini menggunakan penentuan jangka menengah, pada umumnya pengukuran jangka menengah hanya mencakup bulan hingga tahun, ukuran ini berguna untuk mengatur kesepakatan, membuat rencana pengeluaran, rencana keuangan tunai, dan membedah rencana kerja yang berbeda (Liboschik, 2017). Kajian ini akan mengantisipasi jumlah wisatawan ke Indonesia dengan menggunakan pendekatan ARIMA.

Berdasarkan uraian diatas, maka yang menjadi fokus kajian dalam penelitian ini adalah untuk mengetahui penerapan model Model ARIMA, Model ARIMA dibagi ke dalam 3 kelompok, yaitu: model autoregressive (AR), moving average (MA), dan model campuran ARMA (autoregressive moving average) yang mempunyai karakteristik dari dua model pertama.

B. Metode

Pada penelitian ini, penulis menggunakan data sekunder yang diperoleh dari website resmi Badan Pusat Statistik (BPS) bps.go.id

Model ARIMA

Transformasi Box-Cox adalah salah satu metode untuk proses stasioneritas data dalam varians yang dikenalkan oleh Box dan Cox. Transformasi Box-Cox juga sering disebut dengan transformasi

kuasa. Secara matematis, transformasi Box-Cox dirumuskan sebagai berikut:

$$\phi_p(B)(1 - B)^d Z_t = \theta_q(B)\alpha_t \quad \dots (1)$$

Keterangan :

Z_t = data (obsevasi) θ

Φ_p = parameter *autoregressive*

θ_q = parameter *moving average* ke-q

α_t = nilai residual (*error*) pada waktu ke-t

Model Autoregressive (AR)

Model Autoregressive (AR) secara umum disebut ARIMA (p,0,0). Nilai pengamatan Z_t bergantung pada Z_{t-p} , sedangkan nilai dari koefisien autoregresif α_t mempunyai nilai terbatas (Makridakis, Wheelwright, & McGee, 1999). Fungsi model AR pada orde p dinyatakan dalam bentuk persamaan

$$Z_t = \phi_1 + \dots + \phi_p Z_{t-p} + \alpha_t \quad \dots (2)$$

Keterangan :

ϕ_p = Parameter *autoregressive* ke-p

Z_t = Taksiran variable saat waktu ke-t

α_t = Sisaan saat waktu ke-t

Z_{t-1}, Z_{t-p} = Taksiran variabel saat waktu t-1 dan t-p

Model Moving Avarege (MA)

Model *Moving Average* (MA) secara umum ditulis ARIMA (0,0,q). Nilai pengamatan Z_t bergantung pada nilai kesalahan α_t dan juga kesalahan sebelumnya α_{t-q} , dengan koefisien θ_q (Makridakis, Wheelwright, & McGee, 1999). Fungsi model MA pada orde q dinyatakan dalam bentuk persamaan

$$Z_t = \alpha_t - \theta_1 \alpha_{t-1} - \dots - \theta_q \alpha_{t-q} \quad \dots (3)$$

Keterangan :

θ_q = parameter *moving average* ke-q

α_t = nilai residual (*error*) pada waktu ke-t

Model Autoregressive Moving Avarege (MA)

Model ARMA merupakan model gabungan antara model AR dan model MA atau bisa disebut ARMA (p, q). Fungsi model ARMA pada orde p dan q dinyatakan dalam bentuk persamaan

$$Z_t = \phi_1 Z_t + \dots + \phi_p Z_{t-p} + \alpha_t - \theta_1 \alpha_{t-1} - \dots - \theta_q \alpha_{t-q} \quad \dots (4)$$

Fungsi Autokorelasi (ACF)

Fungsi autokorelasi merupakan hubungan linear pada data deret waktu antara Y_t dengan Y_{t+k} yang terpisahkan oleh waktu lag k. ACF digunakan untuk mengidentifikasi model data deret waktu dan melihat kestasioneran data dalam rata-rata. Fungsi autokorelasi yang dihitung berdasarkan sampel data ditulis pada persamaan berikut (Wei, 2006).

$$\rho_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n-k} (Y_t - \bar{Y})^2} \quad \dots (5)$$

Keterangan :

P_k adalah nilai ACF pada lag k

Y_t adalah nilai dari deret waktu pada waktu t

\bar{Y} adalah rata-rata dari deret waktu

K adalah lag (interval waktu)

n adalah jumlah total pengamatan dalam deret waktu

Partial Autocorrelation Function (PACF)

Fungsi autokorelasi parsial (*Partial Autocorrelation Function*) digunakan untuk menunjukkan besarnya hubungan antara nilai suatu variabel saat ini dengan nilai sebelumnya dari variabel yang sama. Fungsi autokorelasi parsial merupakan korelasi antara Y_t dan Y_{t+k} setelah dependensi linier Y_{t+1} , Y_{t+2} , ..., Y_{t+k} pada variabel dihilangkan. Fungsi autokorelasi parsial dari sampel dapat dihitung dengan persamaan matematis sebagai berikut (Wei, 2006).

$$\hat{\phi}_{k+1,k+1} = \frac{\sum_{j=1}^{k-1} \hat{\phi}_{k,j} \hat{P}_{k+1-j}}{1 - \sum_{j=1}^k \hat{\phi}_{k,j} \hat{P}_{k+1-j}} \quad \dots (6)$$

Keterangan :

$\hat{\phi}$ adalah koefisien dari model AR (p) pada lag j

\hat{P}_{k+1-j} adalah nilai PACF pada lag $K+1-j$.

Tranformasi Box-cox

Transformasi Box-Cox adalah salah satu metode untuk proses stasioneritas data dalam varians yang dikenalkan oleh Box dan Cox. Transformasi Box-Cox juga sering disebut dengan transformasi kuasa. Secara matematis, transformasi Box-Cox dirumuskan sebagai berikut:

$$T(Z_t) \begin{cases} Z_t^\lambda & \lambda = 1 \\ \ln Z_t & \lambda \neq 1 \end{cases} \quad \dots (7)$$

Notasi λ melambangkan parameter transformasi. Setiap nilai λ mempunyai rumus transformasi yang berbeda. Transformasi dilakukan jika belum diperoleh nilai $\lambda = 1$ yang artinya data telah stasioner dalam varians

Differencing

Proses pembedaan (*differencing*) dilakukan setelah data stasioner dalam varians. Proses pembedaan dilakukan jika data tidak stasioner dalam rata-rata. Pembedaan dapat dilakukan untuk beberapa periode sampai data stasioner. Proses pembedaan dilakukan dengan cara mengurangi suatu data dengan data sebelumnya. Notasi B (operator backshift) digunakan dalam proses pembedaan.

$$Z_t^2 = (1 - B)^d Z_t \dots \dots \dots \quad \dots (8)$$

Uji Normalitas

Uji normalitas residual digunakan untuk melihat apakah suatu residual memiliki distribusi normal atau tidak. Statistik uji yang digunakan adalah uji Shapiro-Wilk dengan rumus sebagai berikut:

$$W = \frac{(\sum_{i=1}^n \alpha_i x_{(i)})^2}{\sum_{i=1}^n 1 (x_i - \bar{x})^2} \quad \dots (9)$$

Rumusan hipotesis yang digunakan adalah sebagai berikut:

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Taraf signifikan atau α yang digunakan sebesar 5%

Kriteria keputusan tolak H_0 jika $p - value < \alpha$ atau $Whitung < W\alpha, n$ yang berarti residual tidak berdistribusi normal.

Akaike Information Criterion (AIC)

Dalam menguji ketepatan suatu model, seorang ahli statistik yaitu Profesor Hirotu Akaike mengusulkan suatu metode yaitu Akaike Information Criterion (AIC). Nilai AIC dapat dihitung sebagai berikut:

$$AIC = \log \left(\frac{\sum(e_i)^2}{n} + \frac{2k}{n} \right) \quad \dots(10)$$

Keterangan :

e_i^2 : residual kuadrat

k : jumlah parameter

n : banyak pengamatan

Untuk memilih model ARMA terbaik, digunakan hasil perhitungan nilai AIC yang didapatkan untuk tiap modelnya. Nilai AIC terkecil merupakan aspek untuk menentukan model ARMA terbaik untuk peramalan.

C. Hasil Penelitian dan Pembahasan**Data****Tabel 1.** Data Wisatawan Mancanegara Asal Malaysia Januari 2017- Maret 2024.

Bulan	Tahun							
	2017	2018	2019	2020	2021	2022	2023	2024
Januari	149488	207123	231021	207523	46246	55214	140071	155213
Februari	142045	205944	255407	164043	34425	46894	150078	218057
Meret	172097	252966	260448	114453	37095	54006	152447	160269
April	171536	214399	245963	61496	35871	63079	147248	
Mei	176147	180063	25378	66414	45322	95099	169185	
Juni	176276	225707	287051	62033	34345	95277	168759	
Juli	174367	202645	23276	58052	41913	124849	154967	
Agustus	194468	210158	258845	57864	37530	112314	174788	
September	174748	19762	229173	53370	37946	124445	152210	
Oktober	17878	17380	238656	45332	44113	131169	138597	
November	18388	19359	247866	43292	39434	121035	141746	
Desember	228056	239329	239783	46246	48728	189193	211146	

Pemilihan model terbaik**Tabel 2.** Model ARIMA

Model	AIC
ARIMA (1,2,2)	1969.995
ARIMA (1,2,1)	1968.005
ARIMA (0,2,2)	1970.922

Perhitungan menggunakan software R Studio, Berdasarkan hasil dari AIC, nilai model terbaik adalah ARIMA (1 2 1).

Peramalan

Setelah didapatkan model terbaik yang akan digunakan untuk melakukan peramalan. Langkah selanjutnya adalah mencari nilai yang akan diramalkan pada data wisatawan mancanegara asal Malaysia ke Indonesia untuk 9 bulan kedepan.

Dengan bantuan menggunakan software R studio dapatkan hasil sebagai berikut:

Tabel 3. Hasil Peramalan

Periode	Hasil Ramalan	Lo 80	Hi 80	Lo 95	Hi 95
April 2024	188472	156605	220339	139736	237208
Mei 2024	179942	140866	219018	120180	239703
Juni 2024	187103	137269	236937	110888	263317
Juli 2024	187561	128479	246643	97203	277919
Agustus 2024	190883	121926	259939	85270	296495
September 2024	192981	113967	271994	72140	313821
Oktober 2024	195602	106294	284910	59017	332187
November 2024	197999	98171	297828	45325	350674
Desember 2024	200492	89855	311129	31288	369697

D. Kesimpulan

Berdasarkan hasil penelitian menggunakan metode Analisis deret waktu ARIMA yang telah dibahas pada bab sebelumnya dengan membandingkan data sebelum dan sesudah di transformasi Box-Cox, didapatkan kesimpulan sebagai berikut:

1. Didapat model terbaik yaitu ARIMA (1,2,1) dengan nilai AIC terkecil sebesar 1968.005.
2. Penerapan metode ARIMA terbukti dapat digunakan untuk meramalkan jumlah wisatawan mancanegara asal Malaysia ke Indonesia pada tahun 2024
3. Dari hasil peramalan dapat diketahui bahwa jumlah wisatawan Mancanegara asal Malaysia ke Indonesia tahun 2024 bergerak stabil dari tahun sebelumnya walaupun pada periode mei 2024 mengalami sedikit penurunan dan Jumlah wisatawan tertinggi terdapat pada bulan desember 2024 sebanyak 200492 orang sementara yang terendah pada bulan Mei 2024 sebanyak 17994 orang.

Ucapan Terimakasih

Alhamdulillah, for all the blessings that Allah SWT has given. I dedicate this thesis as a sign of my responsibility to my parents who always pray for, give advice, encourage and give support. Don't forget to say thank you to my mentor and the people i love and love me for their help and support to me. May Allah will give you with a lot of goodness.

Daftar Pustaka

Ajeng Mega Pratiwi, & Mutaqin, A. K. (2021). Penerapan Algoritma Naïve Bayes Classifier dalam Memprediksi Status Keberlanjutan Polis Nasabah Asuransi PT.X. *Jurnal Riset Statistika*, 1(2), 117–126. <https://doi.org/10.29313/jrs.v1i2.435>

- Octiana, Vivin. (2014). MODEL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE. (pp. 4-6). Malang: Academia.edu.
- Deviana.S, Nusyirwan, Azis.D , & Pandri Ferdias. (2021). Analisis Model Autoregressive Integrated Moving Average Data Deret Waktu Dengan metode momen sebagai estimasi Parameter . (pp .59). Lampung: Academia.edu.
- Maulana.C . (2023). *Penerapan Tranformasi Box-cox untuk mengatasi ketidakstasioneran dan pola periodik dalam data deret waktu pada ekspor bidang pertanian di Indonesia*. Skripsi tidak dipublikasikan. Bandung: Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Univeritas Islam Bandung.
- BPS. (2024).jumlah kunjungan wisatawan mancanegara per bulan menurut kebangsaan (kunjungan)
- Firmansyah, E. N. R., & Suwanda. (2024). Diagram Kendali Multivariate Exponentially Weighted Moving Variance pada Produksi Kertas Koran. *Jurnal Riset Statistika*, 1–10. <https://doi.org/10.29313/jrs.v4i1.3834>
- Khalis Syahril Suryana, Syahla Anisah, & Aceng Komarudin Mutaqin. (2024). Uji Dua Rata-Rata Waktu Belajar Mandiri Antara Mahasiswa Laki-Laki dan Perempuan. *Jurnal Riset Statistika*, 103–110. <https://doi.org/10.29313/jrs.v4i2.5002>
- Muhammad Farhan Praja Utama, & Teti Sofia Yanti. (2024). Diagram Kendali Adaptive Exponentially Weighted Moving Average Bayesian dalam Pengendalian Penyaluran Air Minum. *Jurnal Riset Statistika*, 83–92. <https://doi.org/10.29313/jrs.v4i2.5014>
- Muhammad Rifqi Naufal, & Nur Azizah Komara Rifai. (2024). Analisis Perbandingan Jamaah Umrah Berdasarkan Jenis Kelamin Menggunakan Independent Sample T-Test. *Jurnal Riset Statistika*, 111–118. <https://doi.org/10.29313/jrs.v4i2.5020>
- Ismayanti. (2010). *pengantar pariwisata*. Tangerang: Universitas Terbuka.