

Perbandingan Akurasi Peramalan Wisatawan Mancanegara di Provinsi Bali Menggunakan Model *Hybrid Time Series Regression-Autoregressive Integrated Moving Average* dan Model *Autoregressive Integrated Moving Average*

Safira Nanda Pradhina¹, Etik Zukhronah², Irwan Susanto³

Universitas Sebelas Maret^{1,2,3}

safiranananda@student.uns.ac.id

Abstrak—Ancaman resesi ekonomi menyebabkan Indonesia perlu memperkuat cadangan devisa negara yang ada pada sektor pariwisata Provinsi Bali. Perkembangan pariwisata di Bali sangat ditentukan oleh jumlah wisatawan mancanegara. Salah satu cara yang dapat dilakukan untuk memprediksi jumlah wisatawan tersebut di masa mendatang, yaitu dengan melakukan peramalan. Semakin berkembangnya model dalam peramalan, model *hybrid* dibangun untuk menghasilkan peramalan yang lebih akurat dengan menggabungkan beberapa model. Tujuan dari penelitian ini adalah untuk membandingkan akurasi peramalan menggunakan model *hybrid* TSR-ARIMA dan model ARIMA pada data wisatawan mancanegara di Provinsi Bali. Data yang digunakan adalah data bulanan jumlah wisatawan mancanegara di Provinsi Bali selama enam tahun. Model *hybrid* TSR-ARIMA diawali dengan memodelkan TSR untuk dicari nilai residu, kemudian residu dimodelkan menggunakan ARIMA. Selanjutnya dilakukan identifikasi dan penetapan model ARIMA berdasarkan orde plot ACF dan PACF dari data yang telah stasioner. Setelah itu, model *hybrid* sementara dilakukan uji signifikansi parameter, uji diagnostik, dan perhitungan nilai MAPE. Model ARIMA diperoleh dengan mengidentifikasi orde plot ACF dan PACF dari data yang telah stasioner. Setelah itu, model ARIMA sementara dilakukan uji signifikansi parameter, uji diagnostik, dan perhitungan nilai MAPE. Berdasarkan hasil analisis, model ARIMA menghasilkan nilai MAPE sebesar 13,68%, sedangkan model *hybrid* menghasilkan nilai MAPE sebesar 5,77%. Dapat disimpulkan bahwa model *hybrid* TSR-ARIMA merupakan model yang lebih baik dibandingkan model ARIMA karena mempunyai nilai MAPE yang lebih kecil.

Kata kunci: Peramalan, *Hybrid*, *TSR-ARIMA*, *ARIMA*

I. PENDAHULUAN

Perekonomian dunia terancam mengalami resesi ekonomi pada tahun 2023, tak terkecuali negara Indonesia. Oleh sebab itu, Indonesia perlu mengantisipasi dengan memperkuat cadangan devisa negara yang ada pada sektor pariwisata Provinsi Bali. Meningkatnya jumlah wisatawan mancanegara dapat menambah devisa bagi negara yang ditinggali [1]. Salah satu cara yang dapat dilakukan untuk mengetahui jumlah wisatawan mancanegara di masa mendatang yaitu dengan peramalan. Data jumlah wisatawan mancanegara merupakan data *time series*. Data *time series* merupakan pengamatan yang dikumpulkan dalam suatu urutan waktu dan dalam kurun waktu tertentu [2]. Analisis *time series* adalah prosedur statistik untuk memprediksi kejadian di masa mendatang yang digunakan sebagai pertimbangan mengambil keputusan. Peramalan dilakukan dengan memprediksi nilai masa depan berdasarkan pengamatan sebelumnya. Terdapat berbagai model analisis *time series* antara lain, yaitu *Time Series Regression* (TSR) dan *Autoregressive Integrated Moving Average* (ARIMA).

Model TSR merupakan komponen deterministik yang dimodelkan dengan menguraikan data menjadi komponen regresi linier berganda sebagai fungsi dari variabel yang tersedia dan tren linier [3]. Model ARIMA relatif fleksibel dalam memodelkan sebagian besar pola data yang acak, tren, musiman dari data yang akan di analisis. Model ARIMA cukup fleksibel dalam memodelkan sebagian besar pola runtun waktu, tetapi model ini kurang cocok untuk sebagian besar permasalahan yang bersifat nonlinier [4]. Oleh karena itu, perlu dilakukan kombinasi model TSR-ARIMA untuk menghasilkan nilai MAPE yang terkecil.

Penelitian mengenai peramalan menggunakan model TSR sudah banyak dilakukan seperti penelitian [5] yang meramalkan kebutuhan energi listrik sektor industri di Jawa Timur dan diperoleh nilai MAPE sebesar 3,75%. Penelitian [6] yang meramalkan penjualan properti pada PT XYZ menggunakan model TSR linier menghasilkan nilai MAPE sebesar 1,52%. Selain itu, terdapat beberapa peramalan menggunakan

model ARIMA. Penelitian [7] meramalkan jumlah kasus virus Covid-19 di Provinsi Jawa Tengah diperoleh nilai MAPE sebesar 4,89% dengan model ARIMA. Penelitian [8] meramalkan jumlah penduduk Kabupaten Sleman dan diperoleh nilai MAPE sebesar 3,62%.

Di sisi lain, penelitian menggunakan model *hybrid* TSR-ARIMA sudah banyak dilakukan oleh beberapa peneliti. Penelitian [9] meramalkan harga saham PT. Unilever Indonesia Tbk. menggunakan model *hybrid* TSR linier-ARIMA menghasilkan nilai MAPE yaitu sebesar 1,35%. Penelitian [10] meramalkan jumlah penumpang maskapai penerbangan Amerika Serikat dan diperoleh nilai MAPE model *hybrid* TSR-ARIMA sebesar 3,061% dan nilai MAPE model TSR sebesar 7,902%. Peramalan harga saham PT. Telkom Indonesia Tbk. menggunakan model TSR dan *hybrid* TSR-ARIMA [11] diperoleh model *hybrid* TSR-ARIMA memiliki nilai MAPE lebih kecil yaitu sebesar 1,88%.

Penelitian untuk memprediksi jumlah wisatawan pernah dilakukan oleh beberapa peneliti sebelumnya. Peramalan kunjungan wisatawan domestik di Taman Wisata Alam Kawah Ijen [12] menggunakan model *Exponential Smoothing* menghasilkan MAPE 46,63%. Penelitian [13] yang meramalkan wisatawan mancanegara ke Indonesia menggunakan model *Fuzzy Time Series* menghasilkan nilai MAPE sebesar 18,05%. Penelitian [14] yang meramalkan wisatawan mancanegara di Kota Pontianak menggunakan model *Deseasonalized* menghasilkan nilai MAPE sebesar 20,11%. Penelitian sebelumnya mengenai peramalan wisatawan tidak menggunakan model *hybrid* dan hanya menggunakan satu model analisis *time series*. Selain itu, penelitian tersebut menghasilkan nilai MAPE yang kurang baik dimana menurut [15] peramalan dikatakan akurat apabila nilai MAPE yang diperoleh kurang dari 10%.

Berdasarkan latar belakang yang telah diuraikan, model *hybrid* TSR-ARIMA menghasilkan nilai MAPE yang lebih kecil dibandingkan dengan model yang lain. Oleh karena itu, penelitian ini melakukan peramalan wisatawan mancanegara di Provinsi Bali menggunakan model ARIMA dan model *hybrid* TSR-ARIMA.

II. METODE PENELITIAN

A. Data Penelitian

Data yang digunakan dalam penelitian ini adalah data jumlah wisatawan mancanegara yang datang ke Provinsi Bali melalui pintu masuk darat, air, dan udara bulan Januari 2017 hingga November 2022. Data tersebut diambil dari Badan Pusat Statistik Provinsi Bali. Data tersebut dibagi menjadi dua yaitu data *training* dimulai dari bulan Januari 2017 hingga Mei 2022 dan data *testing* dimulai dari bulan Juni hingga November 2022.

B. Model Time Series Regression (TSR)

Model *time series regression* merupakan fungsi antara satu variabel respon dengan satu atau lebih variabel prediktor dimana kedua variabel tersebut bersyarat pada waktu. Secara umum model TSR dapat dituliskan dalam Persamaan 1 [16]

$$Y_t = T_t + S_t + \varepsilon_t \quad (1)$$

dengan Y_t adalah variabel respon periode ke- t , T_t merupakan komponen tren periode ke- t , S_t merupakan komponen musiman periode ke- t , dan ε_t adalah eror yang diasumsikan berdistribusi Normal dan independen dengan rata-rata 0 dan variansi σ_ε^2 .

Apabila data mengandung tren, maka Persamaan 1 dapat dituliskan menjadi Persamaan 2

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_t t^m + \varepsilon_t \quad (2)$$

dengan β_0 adalah parameter konstan, β_t adalah parameter ke- t , t^m merupakan komponen tren polinomial. Apabila mengandung tren linier, maka Persamaan 2 dituliskan menjadi Persamaan 3 [16]

$$Y_t = T_t + \varepsilon_t \quad (3)$$

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t \quad (4)$$

dengan β_1 adalah parameter ke-1 dan t merupakan komponen tren linier. Model TSR dengan pola musiman terdapat beberapa variabel *dummy* dengan asumsi bahwa S merupakan periode musiman, maka modelnya dapat dituliskan dalam Persamaan 5

$$Y_t = \beta_1 S_{1,t} + \beta_2 S_{2,t} + \dots + \beta_t S_{s,t} + \varepsilon_t \quad (5)$$

dengan $S_{s,t}$ merupakan variabel *dummy* untuk pola data musiman.

C. Model ARIMA

Model ARIMA dibangun dari data yang telah stasioner. Stasioneritas dibagi menjadi dua, yaitu stasioner dalam variansi dan stasioner dalam rata-rata [2]. Data *time series* dikatakan stasioner dalam variansi apabila

struktur data dari waktu ke waktu mempunyai fluktuasi data yang tetap atau konstan. Stasioner dalam rata-rata adalah fluktuasi data berada pada sekitar nilai rata-rata yang konstan dan tidak bergantung pada waktu.

Setelah data telah stasioner, tahap selanjutnya yaitu identifikasi model ARIMA sementara. Notasi model ARIMA dapat dituliskan menjadi ARIMA(p,d,q) dimana p merupakan notasi AR yang diidentifikasi dari plot PACF, d merupakan notasi untuk *differencing*, dan q merupakan notasi MA yang diidentifikasi dari plot ACF. Model ARIMA dapat dinyatakan dalam Persamaan 6.

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)\varepsilon_t \quad (6)$$

dengan ϕ_p adalah parameter MA ke-p, θ_q adalah parameter AR ke-q, B merupakan operator langkah mundur atau *backshift*, Y_t adalah data aktual pada waktu ke-t, dan ε_t adalah eror pada waktu ke-t yang diasumsikan berdistribusi Normal dan independen dengan rata-rata 0 dan variansi σ_ε^2 . Identifikasi model ARIMA dapat dilakukan dengan melihat plot ACF untuk orde MA dan PACF untuk orde AR. Karakteristik ACF dan PACF teoretis untuk model ARIMA terdapat pada Tabel 1.

TABEL 1. KARAKTERISTIK ACF DAN PACF

Model	ACF	PACF
AR (p)	Meluruh menuju nol secara eksponensial (<i>dies down</i>)	Cut off setelah lag ke-p
MA (q)	Cut off setelah lag ke-q	Meluruh menuju nol secara eksponensial (<i>dies down</i>)
ARMA (p, q)	<i>Dies down</i> setelah lag ke- (q-p)	<i>Dies down</i> setelah lag ke- (p-q)

Model ARIMA sementara yang diperoleh harus memenuhi uji signifikansi parameter dan uji diagnostik residu. Uji signifikansi parameter bertujuan untuk mengetahui apakah parameter yang diduga pada model signifikan atau tidak. Pengujian signifikansi parameter dilakukan menggunakan uji *t-student*. Pengujian signifikansi parameter bertujuan untuk mengetahui apakah parameter yang diduga pada model signifikan atau tidak.

Hipotesis :

$H_0 : \beta = 0$ (parameter model tidak signifikan)

$H_1 : \beta \neq 0$ (parameter model signifikan)

dengan taraf signifikansi sebesar α , maka H_0 ditolak jika $|t| > t_{\frac{\alpha}{2}, n-p}$ atau jika *p-value* $< \alpha$ [2].

Statistik uji :

$$t_{hitung} = \frac{\hat{\beta}}{SE(\hat{\beta})}, \text{ untuk } SE(\hat{\beta}) \neq 0 \quad (7)$$

dengan $SE(\hat{\beta})$ adalah standar eror dari nilai estimasi parameter, n adalah banyaknya observasi, dan p adalah jumlah parameter yang ditaksir.

Setelah pengujian signifikansi parameter model ARIMA, dilakukan pengujian diagnostik residu untuk mengetahui apakah model sudah layak atau belum untuk digunakan. Pemeriksaan diagnostik residu meliputi uji asumsi independensi dan uji asumsi normalitas. Asumsi independensi dapat diuji menggunakan uji Ljung-Box. Selain itu, terdapat uji asumsi residu berdistribusi normal yang diuji menggunakan Kolmogorov-Smirnov.

i. Asumsi Residu Independensi

Hipotesis :

$H_0 : \rho_1 = \rho_2 = \dots = \rho_k$ (residu bersifat *random*)

H_1 : Minimal ada satu nilai $\rho_i \neq 0; i = 1, 2, \dots, k$ (residu tidak bersifat *random*)

dengan taraf signifikansi sebesar α , maka H_0 ditolak jika $Q > \chi_{(\alpha, K-p-q)}^2$ atau jika *p-value* $< \alpha$.

Statistik uji :

$$Q = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-k}, n > K \quad (8)$$

dengan K adalah lag maksimum, n adalah jumlah data, dan $\hat{\rho}_k^2$ adalah autokorelasi residu untuk lag ke-k.

ii. Asumsi Residu Berdistribusi Normal

Hipotesis :

$H_0 : F(x) = F_0(x)$ untuk semua x (residu berdistribusi normal)

$H_1 : F(x) \neq F_0(x)$ untuk semua x (residu tidak berdistribusi normal) dengan taraf signifikansi sebesar α , maka H_0 ditolak jika $D_{hitung} > D_{\alpha,n}$ atau jika $p\text{-value} < \alpha$. Statistik uji :

$$D_{hitung} = \sup |S(x) - F_0(x)| \quad (9)$$

dengan D_{hitung} adalah deviasi maksimum, \sup adalah nilai supremum (maksimum) untuk semua x dari selisih mutlak $S(x)$ dan $F_0(x)$, $S(x)$ adalah fungsi distribusi kumulatif dari data sampel, $F_0(x)$ adalah fungsi peluang kumulatif berdistribusi normal.

Setelah diperoleh model ARIMA terbaik, kemudian dilakukan perhitungan nilai MAPE yang terdapat pada Persamaan 10.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|}{n} \times 100\% \quad (10)$$

dengan Y_t adalah data aktual pada periode ke- t , \hat{Y}_t adalah data prediksi pada periode ke- t , dan n adalah banyaknya data pengamatan. Semakin kecil nilai persentase MAPE, maka semakin akurat hasil peramalan tersebut.

D. Model Hybrid TSR-ARIMA

Model *hybrid* TSR-ARIMA adalah gabungan model TSR dan model ARIMA. Penggabungan model dilakukan dengan cara memodelkan data menggunakan model TSR, selanjutnya residu dari model TSR dimodelkan oleh model ARIMA, kemudian hasil peramalan dari kedua model ditambahkan. Model *hybrid* dapat dinyatakan dalam Persamaan 11 [17].

$$Y_t = L_t + N_t + \varepsilon_t \quad (11)$$

dengan Y_t merupakan data runtun waktu ke- t dan L_t merupakan komponen TSR dan N_t merupakan komponen ARIMA. Setelah menentukan model TSR, kemudian diperoleh residu yang dituliskan dalam Persamaan 12.

$$\varepsilon_t = Y_t - \hat{L}_t \quad (12)$$

dengan ε_t merupakan residu TSR dan \hat{L}_t merupakan nilai estimasi dari model TSR dalam waktu ke- t . Selanjutnya, nilai residu dimodelkan dengan ARIMA. Ketika model TSR dan ARIMA telah diperoleh, maka hasil prediksi dengan model *hybrid* TSR-ARIMA dituliskan dalam Persamaan 13.

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t \quad (13)$$

dengan \hat{N}_t adalah nilai estimasi dari model ARIMA dalam waktu ke- t menggunakan residu TSR dan \hat{Y}_t adalah nilai estimasi dari model *hybrid* TSR-ARIMA.

E. Langkah-Langkah Penelitian

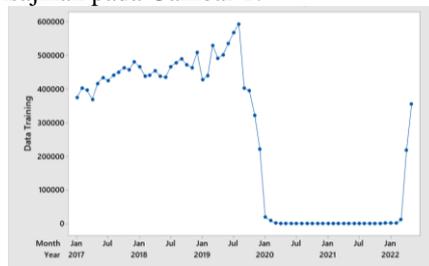
Penelitian dilakukan menggunakan pemodelan *hybrid* TSR-ARIMA dan ARIMA dengan tahapan sebagai berikut.

1. Memodelkan data *training* dengan model *hybrid* TSR-ARIMA melalui tahapan sebagai berikut:
 - a. Melakukan identifikasi variabel *dummy* berdasarkan pola *time series* yang terbentuk
 Model TSR umumnya mengandung komponen tren dan musiman, untuk menentukan variabel *dummy* dapat dilakukan dengan melihat kecenderungan pola data yang ada pada plot *time series*.
 - b. Melakukan pemodelan menggunakan TSR dan menghitung nilai residu dari model TSR.
 Pemodelan TSR dilakukan menggunakan bantuan *software* R. Selanjutnya diperoleh persamaan estimasi Y_t untuk menghitung nilai residu. Perhitungan residu dapat dituliskan dalam Persamaan 8.
 - c. Melakukan pemeriksaan stasioneritas data residu dari model TSR.
 Pemeriksaan stasioneritas dalam variansi dilihat melalui nilai λ . Apabila nilai λ yang diperoleh tidak sama dengan 1, maka perlu dilakukan transformasi Box-Cox[13]. Pengujian stasioneritas dalam rata-rata dapat dilihat dari plot *Autocorrelation Function* [2]. Jika plot ACF menunjukkan adanya lag yang menurun secara perlahan, maka perlu dilakukan *differencing* antar pengamatan.
 - d. Mengidentifikasi dan menetapkan model *hybrid* TSR-ARIMA dengan melihat plot ACF dan PACF.
 Acuan untuk mengidentifikasi dan menetapkan model ARIMA berdasarkan nilai residu TSR dapat dilihat pada Tabel 1.
 - e. Menguji signifikansi model *hybrid* TSR-ARIMA menggunakan uji *t-student*.
 Pengujian ini bertujuan untuk mengetahui apakah parameter yang diduga pada model signifikan atau tidak. Pengujian dilakukan menggunakan rumus yang tertulis pada Persamaan 7.
 - f. Melakukan pemeriksaan diagnostik residu model *hybrid*.

- Pengujian asumsi residu mencakup dua asumsi, yaitu asumsi independensi dengan uji Ljung-Box dan asumsi normalitas dengan uji Kolmogorov-Smirnov.
- g. Memilih model TSR-ARIMA yang memenuhi seluruh pengujian dan nilai MAPE terkecil. *Mean Absolute Percentage Error* (MAPE) digunakan untuk mengevaluasi ketepatan ramalan. Perhitungan nilai MAPE dapat dilakukan dengan menggunakan Persamaan 10.
2. Memodelkan data *training* dengan model ARIMA melalui tahapan sebagai berikut:
 - a. Melakukan pemeriksaan stasioneritas data *training* baik dalam variansi maupun dalam rata-rata. Pemeriksaan stasioneritas dalam variansi dilihat melalui nilai λ . Apabila nilai λ yang diperoleh tidak sama dengan 1, maka perlu dilakukan transformasi Box-Cox[13]. Pengujian stasioneritas dalam rata-rata dapat dilihat dari plot *Autocorrelation Function* [2]. Jika plot ACF menunjukkan adanya lag yang menurun secara perlahan, maka perlu dilakukan *differencing* antar pengamatan.
 - b. Mengidentifikasi dan menetapkan model ARIMA dengan melihat plot ACF untuk orde MA dan plot PACF untuk orde AR.
 Acuan untuk mengidentifikasi dan menetapkan model ARIMA berdasarkan nilai residu TSR dapat dilihat pada Tabel 1.
 - c. Menguji signifikansi parameter model ARIMA menggunakan uji *t-student*.
 Pengujian signifikansi parameter bertujuan untuk mengetahui apakah parameter yang diduga pada model signifikan atau tidak. Pengujian dilakukan menggunakan rumus pada Persamaan 7.
 - d. Melakukan pemeriksaan diagnostik residu model ARIMA.
 Pengujian asumsi residu mencakup dua asumsi, yaitu asumsi independensi dengan uji Ljung-Box dan asumsi normalitas dengan uji Kolmogorov-Smirnov. Pengujian dilakukan menggunakan rumus yang tertulis pada Persamaan 8 dan 9.
 - e. Memilih model ARIMA yang memenuhi seluruh pengujian dan nilai MAPE terkecil
Mean Absolute Percentage Error (MAPE) digunakan untuk mengevaluasi ketepatan ramalan. Perhitungan nilai MAPE dapat dilakukan dengan menggunakan Persamaan 10.
 3. Membuat perbandingan nilai MAPE dari model ARIMA dan model *hybrid* TSR-ARIMA.
 Perbandingan ini bertujuan untuk melihat tingkat keakuratan peramalan dari kedua model yang telah dibangun menggunakan nilai MAPE. Menurut [17] suatu peramalan dikatakan memiliki keakuratan yang baik apabila nilai MAPE yang diperoleh kurang dari 10%.

III. PEMBAHASAN

Langkah awal dalam melakukan peramalan adalah mengidentifikasi pola data. Plot *time series* dibuat dengan tujuan untuk melihat apakah data mempunyai pola musiman dan atau tren atau tidak memiliki kedua pola tersebut. Berikut plot *time series* untuk data jumlah wisatawan mancanegara di Provinsi Bali bulan Januari 2017 hingga Mei 2022 disajikan pada Gambar 1.

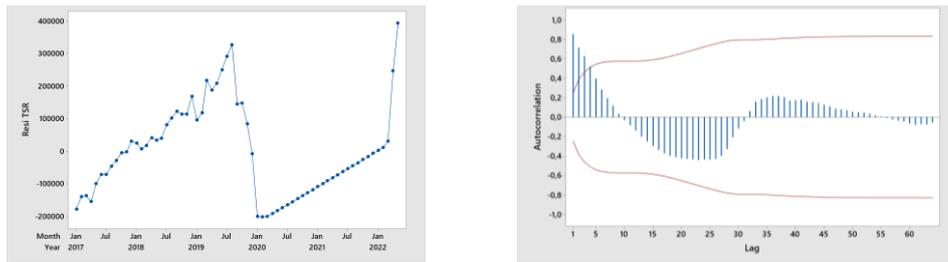


GAMBAR 1. PLOT *TIME SERIES* DATA AKTUAL

Data jumlah wisatawan mancanegara di Provinsi Bali tidak membentuk pola data stasioner karena terdapat pola tren linier naik dan turun. Di sisi lain, Gambar 1 juga menunjukkan bahwa tidak ada pola musiman. Oleh karena itu, digunakan variabel *dummy* berupa waktu (t). Selanjutnya dilakukan pemodelan TSR sehingga diperoleh persamaan yang dituliskan dalam Persamaan 14.

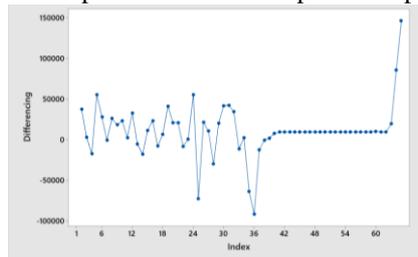
$$TSR : \hat{Y}_t = 561.099 - 9.218t \quad (14)$$

Persamaan 14 menunjukkan bahwa terdapat pengaruh tren pada model TSR. Setiap penambahan 1% tren maka jumlah wisatawan akan menurun sebesar 9.218 jiwa. Nilai residu dari model TSR akan digunakan untuk membentuk model ARIMA. Plot *time series* dan plot ACF residu model TSR disajikan Gambar 2.

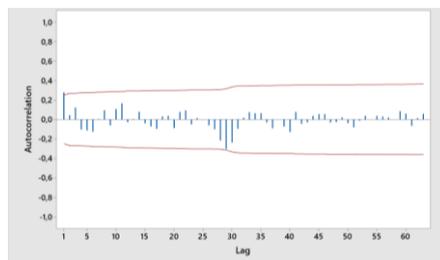


GAMBAR 2. PLOT *TIME SERIES* DAN PLOT ACF RESIDU TSR

Plot ACF pada Gambar 2 menunjukkan adanya lag yang menurun secara perlahan sehingga mengindikasikan adanya kecenderungan pola tren. Hal ini menunjukkan bahwa data tidak stasioner terhadap rata-rata sehingga perlu dilakukan *differencing* non-musiman. Hasil *differencing* lag pertama divisualisasikan dengan plot *time series* pada Gambar 3 dan plot ACF pada Gambar 4.

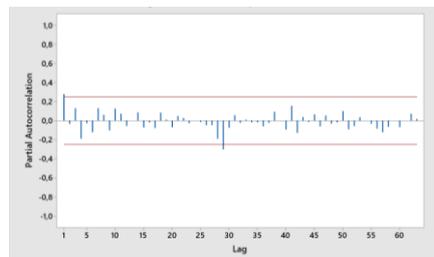


GAMBAR 3. PLOT *TIME SERIES* RESIDU TSR SESUDAH *DIFFERENCING* NON-MUSIMAN



GAMBAR 4. PLOT ACF RESIDU TSR SESUDAH *DIFFERENCING* NON-MUSIMAN

Setelah data di-*differencing* non-musiman, plot *time series* menunjukkan rentang fluktuasi data berada pada sekitar nilai rata-rata yang konstan dan tidak terdapat pola tren naik atau turun lambat. Selain itu, Gambar 4 menunjukkan bahwa lag pertama keluar dari batas pita konfidensi dan *cut off* pada lag setelahnya. Jadi, dapat disimpulkan bahwa data residu TSR telah stasioner. Setelah data stasioner, selanjutnya dilakukan identifikasi model *hybrid* TSR-ARIMA berdasarkan plot ACF pada Gambar 4 dan plot PACF pada Gambar 5.



GAMBAR 5. PLOT PACF RESIDU TSR SESUDAH *DIFFERENCING* NON-MUSIMAN

Plot ACF pada Gambar 4 menunjukkan lag pertama keluar dari pita konfidensi dan *cut off* setelah lag pertama sehingga orde q yang mungkin adalah 0 dan 1. Plot PACF pada Gambar 5 menunjukkan lag pertama keluar dari pita konfidensi dan *cut off* setelah lag pertama sehingga orde p yang mungkin adalah 0 dan 1. Jadi, diperoleh model *hybrid* sementara, yaitu TSR-ARIMA(1,1,0); TSR-ARIMA(1,1,1); TSR-ARIMA (0,1,1).

Model sementara yang diperoleh dilakukan uji signifikansi serta uji diagnostik model menggunakan

asumsi normalitas dan asumsi independensi, hasil pengujian tersebut disajikan pada Tabel 2.

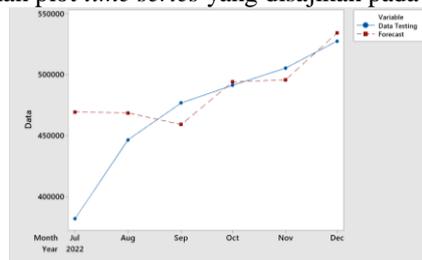
TABEL 2. UJI SIGNIFIKANSI DAN DIAGNOSTIK MODEL HYBRID TSR-ARIMA

Model	<i>p-value</i>	Asumsi Normalitas	Asumsi Independensi
TSR-ARIMA(1,1,0) AR 1	0,001	Memenuhi	Memenuhi
TSR-ARIMA(1,1,1) AR 1 MA 1	0,005 0,345	Tidak memenuhi	Tidak memenuhi
TSR-ARIMA(0,1,1) AR 1	0,000	Memenuhi	Memenuhi

Tabel 2 menunjukkan bahwa model TSR-ARIMA(1,1,0) dan TSR-ARIMA(0,1,1) yang menghasilkan parameter signifikan dan memenuhi asumsi normalitas maupun asumsi independensi. Namun, model TSR-ARIMA(1,1,0) mempunyai nilai MAPE terkecil yaitu sebesar 5,77% sehingga dipilih menjadi model *hybrid* terbaik yang dapat dituliskan ke dalam Persamaan 15.

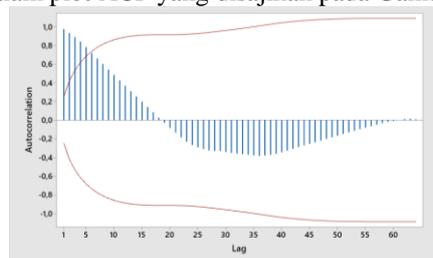
$$TSR - ARIMA(1,1,0) : \widehat{H}_t = 561.099 - 9.218t + 1,487y_{t-1} - 0,487y_{t-2} \quad (15)$$

Setelah diperoleh model *hybrid* TSR-ARIMA terbaik, model yang sudah dibangun diuji performanya menggunakan data *testing* dan diperoleh nilai MAPE sebesar 6,68%. Plot data *testing* dan hasil peramalan Juli-November 2022 menggunakan plot *time series* yang disajikan pada Gambar 6.



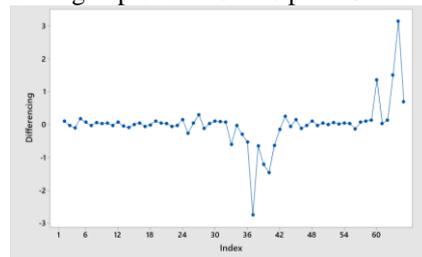
GAMBAR 6. PERBANDINGAN DATA TESTING DAN HASIL PERAMALAN PERIODE JUNI-NOVEMBER 2022

Di sisi lain, terdapat model ARIMA yang umumnya menjadi model standar dalam membandingkan akurasi peramalan model *hybrid* menggunakan nilai MAPE. Langkah pertama dalam memodelkan ARIMA adalah memastikan bahwa data sudah stasioner. Hasil pemeriksaan kestasioneritasan terhadap variansi diperoleh nilai *rounded value* sebesar 0,17. Nilai lambda tersebut menunjukkan bahwa data belum stasioner terhadap variansi. Oleh karena itu, data perlu ditransformasi menggunakan $\ln Y_t$. Data hasil transformasi kemudian divisualisasikan ke dalam plot ACF yang disajikan pada Gambar 7.

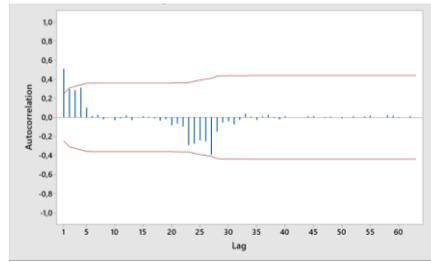


GAMBAR 7. PLOT ACF DATA HASIL TRANSFORMASI

Plot ACF pada Gambar 7 menunjukkan adanya lag menurun secara perlahan. Hal ini mengindikasikan bahwa data tidak stasioner terhadap rata-rata sehingga perlu dilakukan *differencing*. Hasil *differencing* lag pertama kemudian divisualisasikan dengan plot *time series* pada Gambar 8 dan plot ACF pada Gambar 9.

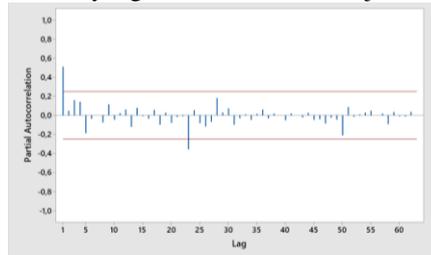


GAMBAR 8. PLOT TIME SERIES DATA SESUDAH DIFFERENCING NON-MUSIMAN



GAMBAR 9. PLOT ACF DATA SESUDAH DIFFERENCING NON-MUSIMAN

Plot *time series* pada Gambar 8 menunjukkan rentang fluktuasi data berada pada sekitar nilai rata-rata yang konstan dan tidak terdapat pola tren naik atau turun lambat, sedangkan Gambar 9 menunjukkan bahwa data lag 1 keluar dari batas pita konfidensi dan *cut off* pada lag setelahnya. Jadi, dapat disimpulkan bahwa data *training* telah stasioner. Plot PACF data yang telah stasioner ditunjukkan pada Gambar 10.



GAMBAR 10. PLOT PACF DATA SESUDAH DIFFERENCING NON-MUSIMAN

Gambar 9 menunjukkan lag pertama keluar dari pita konfidensi dan *cut off* setelah lag pertama sehingga orde q yang mungkin adalah 0 dan 1. Plot PACF pada Gambar 10 menunjukkan lag pertama keluar dari pita konfidensi dan *cut off* setelah lag pertama sehingga orde p yang mungkin adalah 0 dan 1. Jadi, diperoleh model ARIMA sementara, yaitu ARIMA(1,1,0); ARIMA(1,1,1); ARIMA(0,1,1) yang selanjutnya dilakukan uji signifikansi serta uji diagnostik model. Hasil pengujian tersebut disajikan pada Tabel 3.

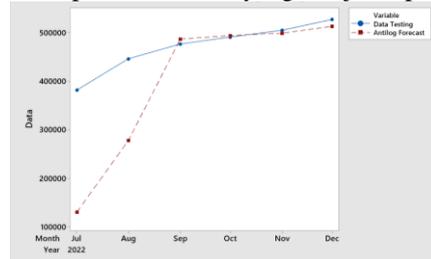
TABEL 3. UJI SIGNIFIKANSI DAN DIAGNOSTIK MODEL ARIMA

Model	<i>p-value</i>	Asumsi Normalitas	Asumsi Independensi
ARIMA(1,1,0) AR 1	0,000	Memenuhi	Memenuhi
ARIMA(1,1,1) AR 1 MA 1	0,000 0,020	Tidak memenuhi	Tidak memenuhi
ARIMA(0,1,1) AR 1	0,001	Memenuhi	Memenuhi

Tabel 3 menunjukkan bahwa model ARIMA(1,1,0) dan ARIMA(0,1,1) menghasilkan parameter yang signifikan dan memenuhi asumsi normalitas maupun asumsi independensi. Namun, model ARIMA (1,1,0) mempunyai nilai MAPE terkecil yaitu sebesar 13,68% sehingga menjadi model ARIMA yang lebih baik. Model tersebut dapat dituliskan ke dalam Persamaan 16.

$$ARIMA(1,1,0) : \ln \hat{y}_t = 1,519 \ln y_{t-1} - 0,519 \ln y_{t-2} \quad (16)$$

Setelah diperoleh model ARIMA terbaik, model yang sudah dibangun diuji performanya menggunakan data *testing* dan diperoleh nilai MAPE sebesar 18,39%. Plot data *testing* dan hasil peramalan bulan Juni hingga November 2022 menggunakan plot *time series* yang disajikan pada Gambar 11.



GAMBAR 11. PERBANDINGAN DATA TESTING DAN HASIL PERAMALAN PERIODE JUNI-NOVEMBER 2022

Gambar 11 menunjukkan bahwa jumlah wisatawan mancanegara mengalami kenaikan. Hal ini dikarenakan adanya penurunan kasus Covid-19 di Indonesia pada akhir tahun 2022. Ringkasan perhitungan nilai MAPE dari model ARIMA dan model *hybrid* TSR-ARIMA terdapat dalam Tabel 4.

TABEL 4. PERBANDINGAN NILAI MAPE

No.	Model	Data	Nilai MAPE
1	ARIMA	Data <i>Training</i>	13,68 %
		Data <i>Testing</i>	18,39 %
2	<i>Hybrid</i> TSR-ARIMA	Data <i>Training</i>	5,77 %
		Data <i>Testing</i>	6,68 %

Tabel 4 menunjukkan bahwa peramalan menggunakan model *hybrid* TSR-ARIMA memiliki nilai MAPE untuk data *training* dan data *testing* yang lebih kecil yaitu sebesar 5,77% dan 6,68% dibandingkan nilai MAPE model ARIMA, yaitu sebesar 13,68% dan 18,39%. Dapat disimpulkan peramalan menggunakan model *hybrid* memiliki nilai MAPE yang kecil. Seperti halnya, pada penelitian [13] yang menggunakan TSR linier sebagai pembandingan model *hybrid* TSR-ARIMA. Penelitian [13] menyimpulkan bahwa model *hybrid* lebih baik dibandingkan model TSR linier karena model *hybrid* nilai MAPE yang diperoleh model *hybrid* lebih kecil.

IV. SIMPULAN DAN SARAN

A. Kesimpulan

Peramalan wisatawan mancanegara di Provinsi Bali dengan model ARIMA menghasilkan nilai MAPE untuk pemodelan data *training* dan pengujian menggunakan data *testing* secara berurutan yaitu 13,68% dan 18,39%. Pada peramalan dengan model *hybrid* menghasilkan nilai MAPE untuk pemodelan data *training* dan pengujian menggunakan data *testing* secara berurutan yaitu 5,77% dan 6,68%. Model *hybrid* menghasilkan peramalan yang lebih akurat dibandingkan dengan model ARIMA berdasarkan nilai MAPE yang diperoleh. Hal ini menunjukkan bahwa peramalan menggunakan model *hybrid* akan menghasilkan nilai MAPE yang lebih kecil.

B. Saran

Saran yang dapat dilakukan untuk penelitian selanjutnya, yaitu diharapkan menggunakan model peramalan dengan analisis intervensi. Analisis intervensi digunakan untuk mengatasi data selama pandemi Covid-19 pada periode tahun 2019 hingga 2021. Adanya pandemi Covid-19 menyebabkan penurunan drastis pada jumlah wisatawan mancanegara di Provinsi Bali.

DAFTAR PUSTAKA

- [1] Z. Shella, M. Said dan M. Nasir, "Faktor-faktor yang mempengaruhi penerimaan daerah sektor pariwisata kota Banda Aceh", *Jurnal Ilmu Ekonomi*, vol. 2, hal. 39-48, 2014.
- [2] W. W. S. Wei, *Time Series Analysis : Univariate and Multivariate*, Addison Wesley Publishing Company: America, 2006.
- [3] H. A. N. Hejase dan A. H. Assi, "Time Series Regression Model for Prediction of Mean Daily Global Solar Radiation in Al-Ain, UAE", *ISRN Renewable Energy*, hal. 1-11, 2014.
- [4] M. Kashei dan M. Bijari, "An Artificial Neural Network (p,d,q) Model For Timeseries Forecasting", *Expert Syst. Appl.*, vol. 37, hal. 479-489.
- [5] R. Duanaputri, Sulistyowati, dan P. A. Insan, "Analisis peramalan kebutuhan energi listrik sektor industri di Jawa Timur dengan model regresi linier," *Jurnal Eltek*, vol. 20, hal. 50-60, 2022.
- [6] G. N. Ayuni dan D. Fitriana, "Penerapan model regresi linier untuk prediksi penjualan properti pada PT XYZ," *Jurnal Telematika*, vol. 14, hal. 79-86, 2019.
- [7] K. Rachmawati dan S. D. Miasary, "Peramalan penyebaran jumlah kasus virus covid-19 Provinsi Jawa Tengah dengan model ARIMA," *J. Zeta-Math*, vol. 6, hal. 11-16, 2021.
- [8] R. Yuliyanti dan E. Arliani, "Peramalan jumlah penduduk menggunakan model ARIMA," *Jurnal Kajian dan Terapan Matematika*, vol. 8, hal. 114-128, 2022.
- [9] Desi, S. W. Rizki, dan Yundari, "Combined Model Time Series Regression-ARIMA on Stocks Prices," *Pure Applied Mathematics Journal*, vol. 3, hal. 65-72, 2022.
- [10] M. Arumsari dan A. T. R. Dani, "Peramalan data runtun waktu menggunakan model hybrid time series regression-autoregressive integrated moving average," *Jurnal Siger Matematika*, vol. 2, hal. 1-12, 2021.
- [11] Ramadani, S. Wahyuningsih, M. N. Hayati, "Forecasting stock price PT. Telkom using hybrid time series regression linier-autoregressive integrated moving average model," *Jurnal Matematika, Statistika dan Komputasi*, vol. 18, hal. 293-307, 2022.
- [12] F. R. Kustiawan dan Hudori, "Forecasting Jumlah Wisatawan di Taman Wisata Alam Kawah Ijen dengan Model Exponential Smoothing", *Jurnal Pendidikan Matematika & Matematika*, vol. 1, hal. 36-49, 2017.
- [13] J. Thira, N. A. Mayangky, dan D. N. Kholifah, "Peramalan Data Kunjungan Wisatawan Mancanegara ke Indonesia menggunakan Fuzzy Time Series", *Jurnal Edukasi d& Penelitian Informatika*, vol 5, hal. 18-23, 2019.
- [14] Masrudin, N. Satyahadewi, N. Imroah, "Peramalan Jumlah Wisatawan Mancanegara di Kota Pontianak dengan Model Deseasonalized", *Jurnal Bimaster*, vol. 3, hal. 159-168, 2018.
- [15] Zhang, Taili, dan K. Wang, "Modeling and Analyzing the Transmission Dynamics of HBV Epidemic in Xinjiang, China." *PLoS ONE J.*, vol. 10, hal. 1-14, 2015.
- [16] B. L. Bowerman dan D. O'Connell, *Forecasting and time series: an applied approach*, Duxbury Press: California, 1993.
- [17] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, hal. 159-173, 2003.