

Peramalan Data Runtun Waktu menggunakan Model *Hybrid Time Series Regression – Autoregressive Integrated Moving Average*

Melisa Arumsari¹, Andrea Tri Rian Dani^{2*}

¹Program Studi Statistika, Jurusan Matematika, Fakultas MIPA, Universitas Mulawarman, 75119, Indonesia

²Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember Surabaya, 60111, Indonesia

*Email korespondensi: andreatriandani98@gmail.com

Abstrak

Peramalan merupakan metode yang digunakan untuk memperkirakan atau memprediksi suatu nilai di masa yang akan datang dengan menggunakan data dari masa lampau. Semakin berkembangnya metode dalam analisis data runtun waktu, dikembangkan metode yang bersifat *hybrid* dimana dilakukan kombinasi beberapa model dengan tujuan untuk menghasilkan peramalan yang lebih akurat. Tujuan dari penelitian ini adalah untuk mengetahui apakah metode *hybrid* TSR-ARIMA memiliki tingkat akurasi lebih tinggi dibandingkan dengan metode TSR secara individu sehingga diperoleh hasil peramalan yang lebih akurat. Data pada penelitian ini adalah data bulanan jumlah penumpang maskapai penerbangan Amerika Serikat periode Januari Tahun 1949 hingga Desember Tahun 1960. Berdasarkan hasil analisis, metode *hybrid* TSR-ARIMA menghasilkan MAPE sebesar 3,061% dan metode TSR menghasilkan MAPE sebesar 7,902%.

Kata kunci: *Autoregressive Integrated Moving Average (ARIMA), Hybrid, Time Series Regression (TSR).*

Abstract

Forecasting is a method used to estimate or predict a value in the future using data from the past. With the development of methods in time series data analysis, a hybrid method was developed in which a combination of several models was carried out in order to produce a more accurate forecast. The purpose of this study was to determine whether the TSR-ARIMA hybrid method has a better level of accuracy than the individual TSR method so that more accurate forecasting results are obtained. The data in this study are monthly data on the number of passengers on American airlines for the period January 1949 to December 1960. Based on the analysis, the TSR-ARIMA hybrid method produces a MAPE of 3,061% and the TSR method produces an MAPE of 7,902%.

Keywords: *Autoregressive Integrated Moving Average (ARIMA), hybrid, Time Series Regression (TSR).*

1. Pendahuluan

Semakin berkembangnya teknologi, muncul perkembangan baru dari peramalan runtun waktu yaitu peramalan yang bersifat *hybrid*, yang berarti kombinasi antar beberapa model. Hal ini tentunya dilakukan karena ingin memanfaatkan kelebihan antara kedua metode dan meminimalisasi adanya kekurangan pada masing-masing metode [1]. Dua metode atau lebih dikombinasikan dengan tujuan untuk menghasilkan peramalan yang lebih akurat dibandingkan jika metode tersebut digunakan secara individu [2]. Salah satu metode *hybrid* yang dapat digunakan adalah *Time Series Regression (TSR)* dan *Autoregressive Integrated Moving Average (ARIMA)*.

Pada TSR, komponen deterministik dimodelkan dengan menguraikannya menjadi komponen regresi linier berganda sebagai fungsi dari variabel yang tersedia ditambah komponen siklus yang menghitung periodisitas tahunan dan tren linier [3]. Sedangkan metode ARIMA cukup fleksibel dalam memodelkan sebagian besar pola data runtun waktu karena dapat menganalisis situasi yang acak, tren, musiman bahkan bersifat siklis dari data yang akan di analisis. Meskipun metode ARIMA cukup fleksibel dalam memodelkan sebagian besar pola runtun waktu, metode ARIMA kurang cocok untuk sebagian besar permasalahan yang bersifat nonlinier [4]. Oleh karena itu, perlu dilakukan kombinasi metode TSR dan ARIMA untuk menghasilkan tingkat akurasi yang lebih tinggi.

Secara umum, metode TSR dapat digunakan pada data yang mengandung pola tren maupun musiman, salah satunya adalah pada data sektor transportasi. Dalam kehidupan masyarakat, transportasi sangatlah penting [5]. Transportasi memiliki peran strategis bagi masyarakat, terutama jasa transportasi udara [6]. Transportasi udara merupakan salah satu moda transportasi untuk perjalanan jarak menengah maupun jarak jauh, salah satunya adalah pesawat [7]. Peningkatan jumlah penumpang pesawat dalam maskapai penerbangan di periode mendatang perlu dianalisis menggunakan analisis runtun waktu (*time series analysis*) dengan tujuan untuk memprediksi jumlah penumpang maskapai penerbangan beberapa periode ke depan [8].

Beberapa penelitian telah membahas metode *hybrid*. Suhartono, dkk. [9] melakukan peramalan kebutuhan air dengan *hybrid* SSA-TSR-ARIMA dan diperoleh hasil bahwa peramalan kebutuhan air bulanan di Kabupaten Wonogiri dengan metode SSA-TSR-ARIMA sebagai kombinasi beberapa metode individual memberikan peramalan yang lebih akurat dibandingkan *hybrid* SSA-ARIMA sebagai metode yang lebih sederhana dan ARIMA sebagai metode individu. Sulistyowati, dkk. [10] melakukan peramalan model *hybrid* untuk memprediksi penumpang udara dan kargo di Indonesia dan diperoleh hasil bahwa metode ARIMAX-NN dan TSR-NN menghasilkan kinerja peramalan yang lebih baik daripada TSR-SVR dan ARIMAX-SVR. Sehingga penggunaan NN secara model *hybrid* memiliki kinerja yang lebih baik dibandingkan dengan SVR untuk data penumpang dan kargo yang di prediksi di tiga bandara internasional terbesar di Indonesia.

Berdasarkan uraian tersebut, penulis akan melakukan peramalan jumlah penumpang maskapai penerbangan Amerika Serikat dengan menggunakan model *hybrid* TSR-ARIMA. Pada penelitian ini, dilakukan peramalan menggunakan metode TSR terlebih dahulu, kemudian dilakukan peramalan kembali pada data residual menggunakan metode ARIMA. Untuk memperoleh hasil akhir peramalan metode *hybrid* TSR-ARIMA, dilakukan penjumlahan dari peramalan metode TSR dan ARIMA.

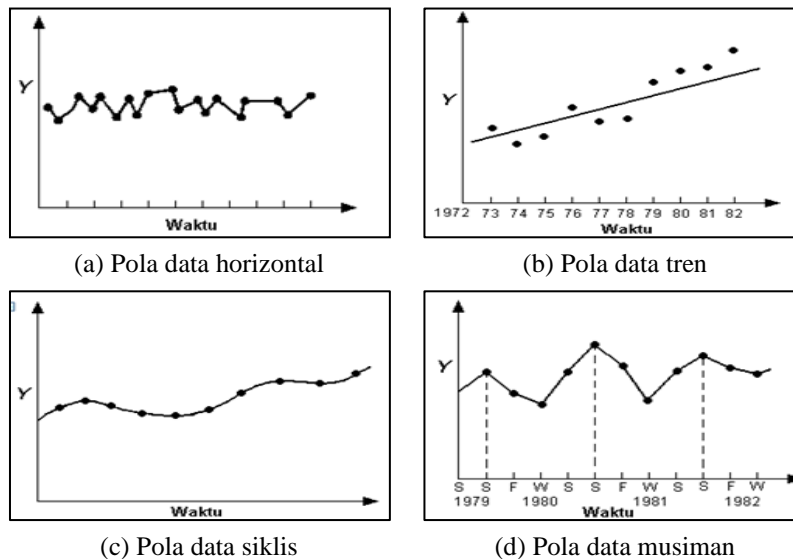
2. Tinjauan Pustaka

2.1 Peramalan

Peramalan merupakan metode yang berguna untuk memperkirakan suatu nilai di periode yang akan datang dengan menggunakan data historis [11]. Metode peramalan merupakan suatu pendekatan untuk memperkirakan sesuatu yang akan terjadi pada beberapa periode ke depan berdasarkan data histori [12]

2.2 Analisis Runtun Waktu

Data runtun waktu merupakan sekumpulan data berdasarkan waktu secara terurut dengan interval waktu yang tetap [13]. Sedangkan analisis runtun waktu merupakan analisis sekumpulan data dengan tujuan untuk mengetahui atau meramalkan kondisi di masa depan [14]. Pola data runtun waktu dapat dibedakan menjadi empat jenis, yaitu seperti pada Gambar 1 berikut [15].



Gambar 1. Pola data runtun waktu

2.3 Time Series Regression

Time Series Regression (TSR) pada dasarnya sama dengan regresi, terutama regresi dengan variabel *dummy*. Dalam penelitian ini model TSR merupakan model penanganan tren dan komponen musiman secara terpisah. Secara umum tren diartikan sebagai arah jangka panjang yang terus menerus naik atau turun, dan musiman adalah

pola yang berulang dengan periode yang sama, misalnya 12 bulan pertahun. Model TSR untuk pola tren dapat digambarkan sebagai regresi polinomial sebagai berikut:

$$\hat{T}_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_m t^m \tag{1}$$

Sedangkan model TSR pola musiman secara umum ditulis sebagai berikut:

$$\hat{S}_t = \beta_1 D_1 + \beta_2 D_2 + \dots + \beta_s D_s \tag{2}$$

dimana D_j (untuk $j = 1, 2, \dots, s$) adalah variabel *dummy* untuk komponen musiman [9].

2.4 Kestasioneran Data

Stasioneritas memiliki makna yaitu fluktuasi data berada disekitar nilai rata-rata yang tetap, tidak bergantung pada waktu dan variansi itu sendiri [15].

a. Stasioner dalam variansi

Apabila kondisi data tidak stasioner dalam variansi, perlu dilakukan transformasi agar data tersebut stasioner. Salah satu transformasi yang dapat digunakan adalah transformasi pangkat (*power transformation*) yang diperkenalkan oleh Box dan Cox pada tahun 1964 dengan ketentuan sebagai berikut:

$$Z'_t = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda}, \lambda \neq 0 \\ \ln(Z_t), \lambda = 0 \end{cases} \tag{3}$$

dengan Z_t adalah data deret waktu periode ke- t dan λ adalah parameter transformasi [16].

b. Stasioner dalam rata-rata

Jika diperoleh data deret waktu yang belum stasioner dalam rata-rata, data tidak dapat digunakan untuk mendapatkan model ARIMA terbaik, tetapi terlebih dahulu distasionerkan. Cara yang dapat dilakukan untuk menstasionerkan data yang tidak stasioner dalam rata-rata yaitu dengan menggunakan metode *differencing* (pembedaan).

Metode *differencing* diperoleh dengan cara mengurangi nilai pengamatan pada waktu t dengan nilai pengamatan pada waktu sebelumnya dan membentuk suatu data baru. Jika hasil *differencing* tersebut disimbolkan dengan W_t , maka secara umum *differencing* orde 1 dapat ditulis sebagai berikut [16]:

$$W_t = Z_t - Z_{t-1} \tag{4}$$

$$W_t = (1 - B)Z_t \tag{5}$$

dimana B adalah operator *backshift* (operator mundur).

2.5 Autocorrelation Function dan Partial Autocorrelation Function

Nilai *autocorrelation function* (ACF) ρ_k dapat diduga dengan persamaan sebagai berikut:

$$\rho_k = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2} \tag{6}$$

dimana n adalah jumlah observasi, k adalah selisih waktu (*lag*), Z_t adalah data pada waktu ke- t , Z_{t+k} adalah data pada waktu ke- $t+k$, dan \bar{Z} adalah rata-rata dari Z_t . Grafik ACF dapat digunakan untuk mendeteksi kestasioneran data. Jika grafik ACF cenderung turun lambat (*dies down*), maka data tersebut dikatakan belum stasioner dalam rata-rata [15].

Nilai *partial autocorrelation function* (PACF) untuk *lag* k dan $j = 1, 2, \dots, k$ adalah sebagai berikut [15]:

$$\phi_{kk} = \begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \\ \hline 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix} \tag{7}$$

2.6 Seasonal Autoregressive Integrated Moving Average

Seasonal Autoregressive Integrated Moving Average atau SARIMA mengandung model yang cocok pada data dengan pola musiman dan merupakan pengembangan dari model ARIMA. Model SARIMA ini dinotasikan dengan ARIMA $(p,d,q)(P,D,Q)^S$. Bentuk dari model SARIMA adalah sebagai berikut [17]:

$$\Phi_p(B^S)\phi_p(B)(1-B)^d(1-B^S)^D Z_t = \theta_q(B)\Theta_Q(B^S)a_t \quad (8)$$

dimana,

$$\begin{aligned} \phi_p(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\ \Phi_p(B^S) &= 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_p B^{pS} \\ \theta_q(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \\ \Theta_Q(B^S) &= 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS} \end{aligned}$$

2.7 Pengujian Signifikansi Parameter

Pengujian ini dilakukan memeriksa apakah suatu parameter model layak masuk dalam model ARIMA atau tidak [16]. Sebagai contoh, pengujian signifikansi parameter model ARIMA untuk parameter model AR (p) sebagai berikut:

Hipotesis

$$H_0 : \phi = 0 \text{ (Parameter model tidak signifikan)}$$

$$H_1 : \phi \neq 0 \text{ (Parameter cukup signifikan)}$$

Statistik uji

$$t_{hit} = \frac{\hat{\phi}}{SE(\hat{\phi})} \quad (9)$$

Daerah kritis

$$\text{Menolak } H_0 \text{ pada taraf signifikansi } \alpha \text{ jika nilai } |t_{hit}| > t_{(\alpha/2; db)} \text{ atau } p\text{-value} < \alpha \text{ dengan } (db) = n - n_p,$$

dimana n adalah banyaknya data dan n_p adalah banyaknya parameter dalam model.

2.8 Pemeriksaan Diagnostik

Pemeriksaan diagnostik meliputi uji independensi residual dan uji residual berdistribusi normal [16].

a. Pengujian independensi residual

Uji ini digunakan untuk mengetahui apakah residual antar *lag* saling berkorelasi atau tidak. Hipotesis yang digunakan pada uji independensi residual adalah sebagai berikut:

Hipotesis

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0 \text{ (Tidak terdapat korelasi antar lag)}$$

$$H_1 : \text{Minimal ada satu } \rho_j \neq 0, j = 1, 2, \dots, k \text{ (Terdapat korelasi antar lag)}$$

Statistik uji

$$Q^* = n(n+2) \sum_{j=1}^k \frac{\hat{\rho}_j^2}{(n-j)} \quad (10)$$

Daerah Kritis

Menolak H_0 pada taraf signifikansi α jika nilai $Q^* > \chi^2_{(\alpha, k-(p+q))}$, $k > (p+q)$ atau $p\text{-value} < \alpha$, nilai p dan q menunjukkan orde dari model ARMA(p,q).

b. Pengujian residual berdistribusi normal

Salah satu cara yang dapat ditempuh untuk melakukan uji asumsi kenormalan residual ini adalah uji *Kolmogorov-Smirnov*. Tahapan dalam pengujian kenormalan residual adalah sebagai berikut:

Hipotesis

$$H_0 : \text{Residual berdistribusi normal}$$

$$H_1 : \text{Residual tidak berdistribusi normal}$$

Statistik uji

$$D_{hit} = \sup\{|F_0(Z) - F_n(Z)|, -\infty \leq Z \leq \infty\} \quad (11)$$

Daerah Kritis

H_0 pada taraf signifikansi α jika nilai $D_{hit} > D_{(\alpha;n)}$ atau $p\text{-value} < \alpha$.

2.9 Hybrid TSR-ARIMA

Secara umum struktur *hybrid* pada deret waktu terdiri dari dua komponen sebagai berikut [18]:

$$y_t = L_t + N_t \quad (12)$$

Berdasarkan model pada Persamaan (12) digunakan model statistik untuk memperoleh estimasi model komponen L_t yang dilambangkan dengan \hat{L}_t . Selanjutnya diperoleh pemodelan sisa atau residual menggunakan $y_t - \hat{L}_t$.

Dari residual tersebut akan diperoleh estimasi model komponen N_t dan dilambangkan dengan \hat{N}_t . Berdasarkan model pada Persamaan (12) terbentuk model estimasi sebagai berikut:

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

2.10 Mean Absolute Percentage Error

Mean Absolute Percentage Error (MAPE) menghitung perbedaan dari data aktual dan data hasil peramalan yang telah di absolutkan. Suatu metode dikatakan berkinerja dengan baik apabila menghasilkan nilai MAPE kurang dari 10%, atau dengan kata lain, semakin kecil MAPE maka semakin baik kinerja dari metode yang digunakan. Berikut ini adalah rumus perhitungan MAPE [19].

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - x_i}{x_i} \right| \quad (14)$$

3. Metodologi Penelitian

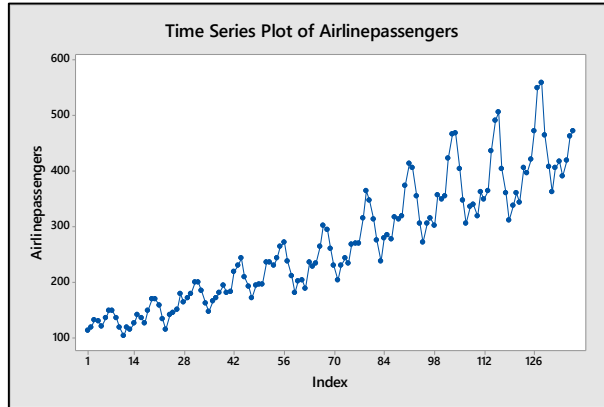
Sampel yang digunakan dalam penelitian ini adalah data bulanan jumlah penumpang maskapai penerbangan Amerika Serikat periode Januari Tahun 1949 hingga Desember Tahun 1960 yang diperoleh dari R dataset sebanyak 144 data. Adapun langkah-langkah analisis data menggunakan model *hybrid* TSR-ARIMA adalah sebagai berikut:

- 1) Membagi sampel menjadi dua yaitu *in sample* (sebanyak 137 data yaitu periode Januari Tahun 1949 hingga Mei Tahun 1960) dan *out sample* (sebanyak 7 data yaitu periode Juni hingga Desember Tahun 1960).
- 2) Melakukan analisis statistika deskriptif menggunakan grafik runtun waktu.
- 3) Melakukan peramalan dengan metode TSR melalui tahapan sebagai berikut:
 - a. Melakukan identifikasi variabel *dummy* sebanyak 12 variabel, yaitu Bulan Januari hingga Bulan Desember.
 - b. Melakukan peramalan dengan model TSR.
 - c. Membuat grafik perbandingan antara data aktual dengan hasil peramalan menggunakan model TSR.
 - d. Menghitung nilai residual dari model TSR.
 - e. Menghitung MAPE dari peramalan dengan model TSR.
- 4) Memeriksa apakah residual dari model TSR telah memenuhi asumsi *white noise* atau tidak.
- 5) Melakukan peramalan dengan metode ARIMA melalui tahapan sebagai berikut:
 - a. Melakukan pemeriksaan stasioneritas data residual dari model TSR.
 - b. Penetapan model ARIMA sementara.
 - c. Mengestimasi parameter model ARIMA sementara.
 - d. Pemeriksaan diagnostik model ARIMA sementara.
 - e. Pemilihan model ARIMA yang memenuhi seluruh pengujian pada pemeriksaan diagnostik.
 - f. Menentukan model terbaik ARIMA berdasarkan MAPE terkecil.
 - g. Melakukan peramalan dengan model ARIMA terbaik.
- 6) Melakukan peramalan dengan model *hybrid* TSR-ARIMA.
- 7) Membuat grafik perbandingan antara data aktual dengan hasil peramalan menggunakan model *hybrid* TSR-ARIMA.
- 8) Menghitung MAPE dari peramalan dengan model *hybrid* TSR-ARIMA.

4. Hasil dan Pembahasan

4.1 Grafik Runtun Waktu

Langkah awal yang perlu dilakukan adalah menampilkan statistika deskriptif berupa grafik runtun waktu yang ditunjukkan pada Gambar 2.



Gambar 2. Grafik runtun waktu data jumlah penumpang maskapai penerbangan Amerika Serikat periode Januari Tahun 1949 hingga Mei Tahun 1960

Berdasarkan grafik runtun waktu pada Gambar 2, dapat diketahui bahwa data bulanan jumlah penumpang maskapai penerbangan Amerika Serikat periode Januari Tahun 1949 hingga Mei Tahun 1960 menunjukkan pola musiman dan tren naik dikarenakan terjadi perubahan yang berulang dengan sendirinya dan mengalami peningkatan dari waktu ke waktu.

4.2 *Time Series Regression (TSR)*

Berdasarkan pola data yang diperoleh, proses identifikasi variabel *dummy* yang digunakan dapat dilihat pada Tabel 1.

Tabel 1 Identifikasi Variabel *Dummy*

Bulan	Variabel <i>Dummy</i>	Bulan	Variabel <i>Dummy</i>
Januari	D_1	Juli	D_7
Februari	D_2	Agustus	D_8
Maret	D_3	September	D_9
April	D_4	Oktober	D_{10}
Mei	D_5	November	D_{11}
Juni	D_6	Desember	D_{12}

Dengan menggunakan bantuan *software* Minitab, model TSR untuk masing-masing variabel *dummy* secara berturut-turut diberikan oleh:

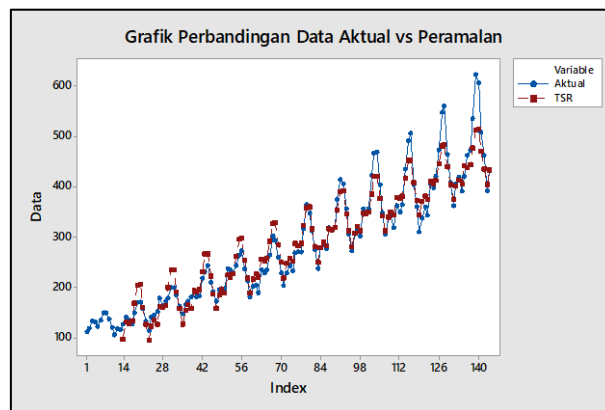
$$\begin{aligned}
 D_1 &= 69,08 + 2,5771_t & D_7 &= 154,06 + 2,5771_t \\
 D_2 &= 59,76 + 2,5771_t & D_8 &= 152,67 + 2,5771_t \\
 D_3 &= 92,35 + 2,5771_t & D_9 &= 105,91 + 2,5771_t \\
 D_4 &= 86,69 + 2,5771_t & D_{10} &= 68,51 + 2,5771_t \\
 D_5 &= 88,86 + 2,5771_t & D_{11} &= 35,57 + 2,5771_t \\
 D_6 &= 121,27 + 2,5771_t & D_{12} &= 60,81 + 2,5771_t
 \end{aligned}$$

Berdasarkan model-model tersebut diperoleh peramalan 7 periode ke depan seperti pada Tabel 2.

Tabel 2 Peramalan dengan Model TSR

Tahun	Bulan	Jumlah Penumpang
1960	Juni	477
	Juli	513
	Agustus	514
	September	470
	Oktober	435
	November	405
	Desember	432

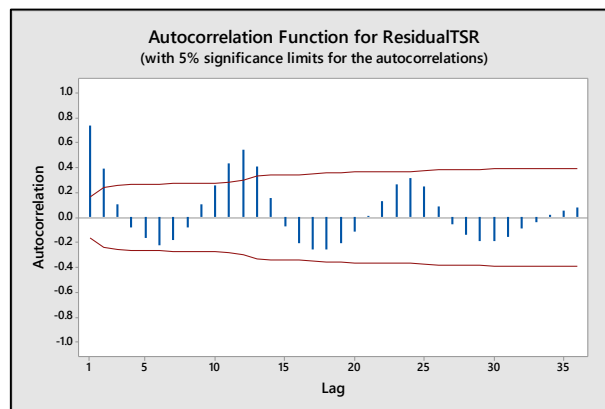
Selanjutnya berdasarkan hasil prediksi menggunakan model TSR, maka dapat dibentuk grafik perbandingan antara data bulanan jumlah penumpang maskapai penerbangan Amerika Serikat dan peramalannya menggunakan model TSR periode Januari Tahun 1949 hingga Desember Tahun 1960 seperti pada Gambar 3.



Gambar 3. Grafik perbandingan data aktual dan peramalan dengan model TSR

Berdasarkan Gambar 3, dapat dilihat bahwa pola data peramalan dengan model TSR hampir menyerupai data aktualnya. Sehingga dapat dilakukan perhitungan akurasi peramalan menggunakan MAPE untuk melihat tingkat akurasi dari peramalan tersebut. Dengan menggunakan Persamaan (14) diperoleh MAPE untuk peramalan dengan model TSR sebesar 7,902%.

Langkah selanjutnya adalah melakukan pemeriksaan apakah data residual telah memenuhi asumsi *white noise* atau tidak dengan cara menampilkan grafik ACF seperti pada Gambar 4.

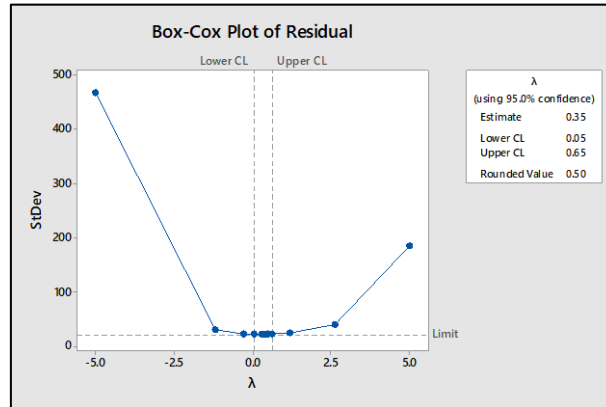


Gambar 4. Grafik ACF data residual

Berdasarkan Gambar 4, dapat dilihat bahwa data residual dari model TSR belum memenuhi asumsi *white noise* dikarenakan masih ada beberapa *lag* yang keluar dari batas signifikan. Sehingga untuk mengatasi permasalahan tersebut, dapat dilakukan pemodelan *hybrid* TSR-ARIMA dengan data yang di analisis menggunakan metode ARIMA adalah data residual dari pemodelan TSR.

4.3 *Autoregressive Integrated Moving Average (ARIMA)*

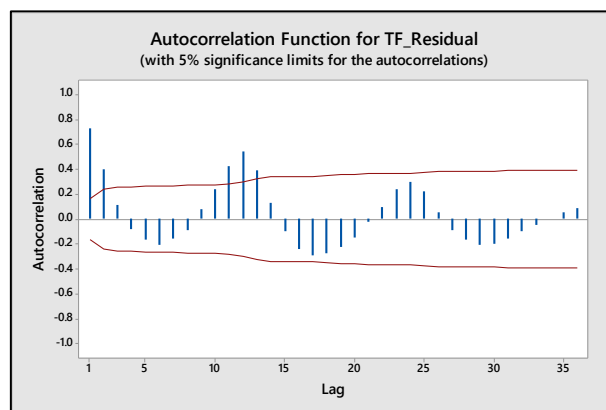
Berdasarkan peramalan menggunakan metode TSR, diperoleh data residual sebanyak 137 data yang selanjutnya diramalkan menggunakan metode ARIMA. Langkah awal metode ARIMA adalah mengidentifikasi model ARIMA. Untuk membuktikan apakah data residual telah stasioner dalam variansi, maka akan dilakukan pengecekan data untuk melihat nilai estimasi (λ) seperti pada Gambar 5.



Gambar 5. Box-Cox *plot* data residual

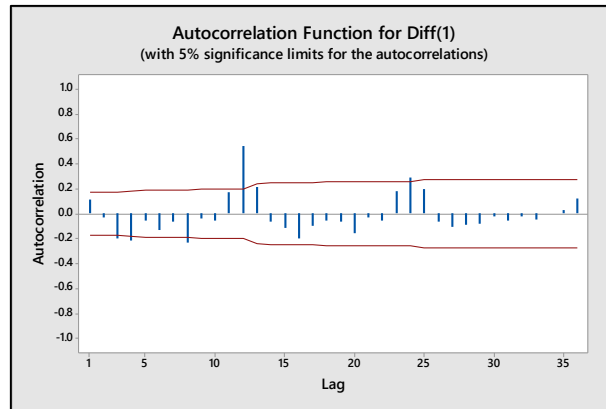
Berdasarkan Gambar 5 dapat dilihat bahwa data residual memiliki nilai λ sebesar 0,35. Karena nilai λ harus mendekati angka 1, selanjutnya perlu dilakukan transformasi pangkat pada data tersebut ($Z_t^{0,35}$). Setelah dilakukan transformasi pangkat, maka dilakukan pengecekan kembali dan diperoleh nilai λ sebesar 1,01. Sehingga dapat disimpulkan bahwa data telah stasioner dalam variansi.

Langkah selanjutnya adalah memeriksa apakah data hasil transformasi telah stasioner dalam rata-rata atau tidak. Untuk membuktikan apakah data residual setelah di transformasi stasioner dalam rata-rata, maka akan ditampilkan grafik ACF sebagai berikut:



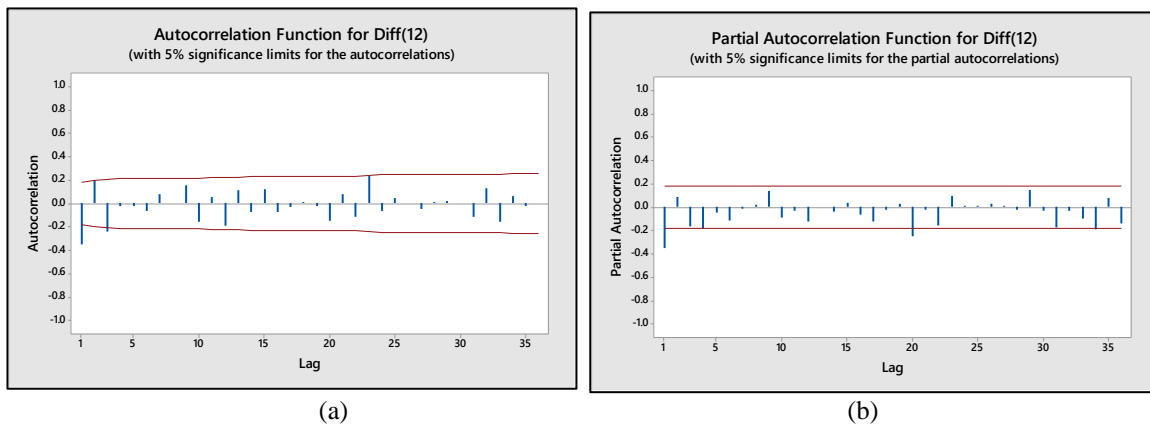
Gambar 6. Grafik ACF data residual setelah transformasi

Berdasarkan Gambar 6, dapat dilihat bahwa pada grafik ACF, nilai *lag* cenderung turun lambat sedemikian sehingga diindikasikan bahwa data residual belum stasioner dalam rata-rata. Sehingga, perlu dilakukan *differencing* orde 1 pada data setelah di transformasi. Berikut grafik ACF dari data setelah dilakukan *differencing* orde 1.



Gambar 7. Grafik ACF data hasil differencing orde 1

Berdasarkan Gambar 7, dapat dilihat bahwa pada grafik ACF, nilai lag-lag musiman cenderung turun lambat sedemikian sehingga diindikasikan bahwa data residual belum stasioner dalam rata-rata. Sehingga, perlu dilakukan differencing orde musiman yaitu 12 pada data setelah di differencing orde 1. Berikut grafik ACF dari data setelah dilakukan differencing orde 12.



Gambar 8. Grafik ACF dan PACF data hasil differencing orde 12

Berdasarkan Gambar 8(a), dapat dilihat bahwa pada grafik ACF, nilai lag untuk orde non musiman mengalami pola cut-off pada lag 1, 2, 3 dan tidak ada nilai lag yang cut-off untuk orde musiman. Sedangkan berdasarkan Gambar 8(b), dapat dilihat bahwa pada grafik PACF, nilai lag untuk orde non musiman mengalami pola cut-off pada lag 1 dan tidak ada nilai lag yang cut-off untuk orde musiman. Karena pada grafik ACF nilai lag telah mengalami pola cut-off, maka dapat diindikasikan bahwa data residual telah stasioner dalam rata-rata. Sehingga dari grafik ACF dan PACF dapat dibentuk model dugaan ARIMA musiman yaitu ARIMA (0,1,1)(0,1,0)¹², ARIMA (0,1,2)(0,1,0)¹², ARIMA (0,1,3)(0,1,0)¹², ARIMA (1,1,0)(0,1,0)¹², ARIMA (1,1,1)(0,1,0)¹², ARIMA (1,1,2)(0,1,0)¹² dan ARIMA (1,1,3)(0,1,0)¹².

Untuk mengetahui model terbaik di antara 7 model dugaan tersebut, maka perlu dilakukan pengujian hipotesis terhadap model dugaan tersebut. Hasil penaksiran parameter dari 7 model dugaan awal dapat dilihat pada Tabel 3.

Tabel 3 Estimasi Parameter model ARIMA

Model	Koefisien	p-value	Keputusan	Kesimpulan
ARIMA (0,1,1)(0,1,0) ¹²	0,3271	0,000	Tolak H ₀	Signifikan
ARIMA (0,1,2)(0,1,0) ¹²	0,3272	0,000	Gagal	Tidak
ARIMA (0,1,3)(0,1,0) ¹²	0,0018	0,984	Tolak H ₀	Signifikan
ARIMA (1,1,0)(0,1,0) ¹²	0,3192	0,000	Gagal	Tidak
ARIMA (1,1,1)(0,1,0) ¹²	-0,1290	0,180	Tolak H ₀	Signifikan
ARIMA (1,1,2)(0,1,0) ¹²	0,3263	0,001	Tolak H ₀	Signifikan
ARIMA (1,1,3)(0,1,0) ¹²	-0,3154	0,000	Tolak H ₀	Signifikan
ARIMA (1,1,1)(0,1,0) ¹²	-0,5766	0,024	Gagal	Tidak
ARIMA (1,1,1)(0,1,0) ¹²	-0,2924	0,309	Tolak H ₀	Signifikan

ARIMA	0,8140	0,000		
(1,1,2)(0,1,0) ¹²	1,1456	0,000	Tolak H_0	Signifikan
	-0,1703	0,000		
	-0,0003	0,999		
ARIMA	0,3189	0,271	Gagal	Tidak
(1,1,3)(0,1,0) ¹²	-0,1290	0,322	Tolak H_0	Signifikan
	0,3264	0,001		

Berdasarkan Tabel 3 dapat dilihat bahwa parameter model ARIMA (0,1,1)(0,1,0)¹², ARIMA (1,1,0)(0,1,0)¹² dan ARIMA (1,1,2)(0,1,0)¹² telah signifikan. Sehingga berdasarkan pengujian signifikan parameter model ARIMA sementara, maka model ARIMA (0,1,1)(0,1,0)¹², ARIMA (1,1,0)(0,1,0)¹² dan ARIMA (1,1,2)(0,1,0)¹² adalah model yang dipilih dan layak untuk digunakan. Selanjutnya akan dilakukan pemeriksaan diagnostik pada model tersebut.

Selanjutnya Uji independensi residual dilakukan dengan melihat nilai *p-value* yang lebih besar dari 0,05 pada lag 12, 24, 36, dan 48. Apabila terdapat 1 lag yang tidak signifikan, model dugaan tersebut tidak dapat digunakan. Adapun pengujian independensi residual dengan *Ljung-Box* dapat dilihat pada Tabel 4.

Tabel 4 Uji Independensi Residual

Model	Lag	<i>p-value</i>	Keputusan	Kesimpulan
ARIMA (0,1,1)(0,1,0) ¹²	12	0,275	Gagal Tolak H_0	Tidak terdapat korelasi antar lag
	24	0,111		
	36	0,315		
	48	0,282		
ARIMA (1,1,0)(0,1,0) ¹²	12	0,299	Gagal Tolak H_0	Tidak terdapat korelasi antar lag
	24	0,116		
	36	0,346		
	48	0,313		
ARIMA (1,1,2)(0,1,0) ¹²	12	0,113	Gagal Tolak H_0	Tidak terdapat korelasi antar lag
	24	0,105		
	36	0,271		
	48	0,346		

Berdasarkan Tabel 4 dapat dilihat bahwa seluruh model dengan lag yang dicobakan adalah 12, 24, 36 dan 48 memiliki *p-value* > 0,05 sehingga dapat disimpulkan model ARIMA yang terbaik telah memenuhi asumsi independensi residual.

Setelah melakukan pengujian independensi residual, selanjutnya adalah melakukan pengujian kenormalan residual dengan melihat nilai *p-value*, apabila nilai *p-value* > 0,05 maka dapat disimpulkan jika data residual berdistribusi normal. Kenormalan residual diuji dengan menggunakan statistik uji *Kolmogorov-Smirnov* dan dapat di sajikan pada Tabel 5.

Tabel 5 Uji Normalitas Residual

Model	D_{hit}	<i>p-value</i>	Keputusan	Kesimpulan
ARIMA (0,1,1)(0,1,0) ¹²	0,070	0,141	Gagal Tolak H_0	Residual berdistribusi normal
ARIMA (1,1,0)(0,1,0) ¹²	0,061	0,150	Gagal Tolak H_0	Residual berdistribusi normal
ARIMA (1,1,2)(0,1,0) ¹²	0,070	0,139	Gagal Tolak H_0	Residual berdistribusi normal

Berdasarkan Tabel 5 dapat dilihat bahwa seluruh model memiliki nilai *p-value* > 0,05, sehingga dapat disimpulkan bahwa model ARIMA yang terbaik telah memenuhi asumsi normalitas residual.

Berdasarkan tiga model ARIMA yang lolos tahap pemeriksaan diagnostik, maka selanjutnya dapat ditentukan model yang terbaik dengan menggunakan kriteria MAPE. Model ARIMA (1,1,2)(0,1,0)¹² merupakan model ARIMA yang memiliki nilai MAPE terkecil, sehingga terpilih menjadi model terbaik dan akan digunakan pada proses selanjutnya.

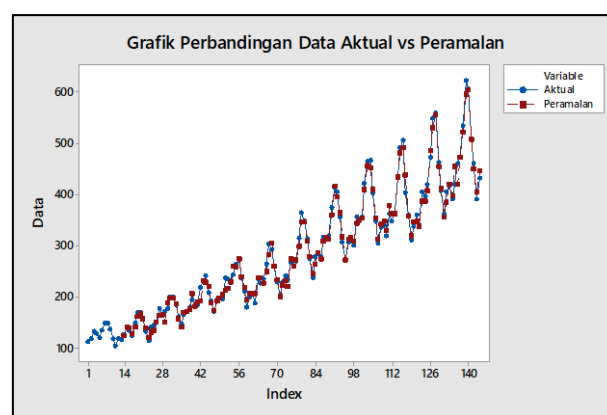
4.4 Hybrid TSR-ARIMA

Dengan bantuan *software* Minitab, diperoleh hasil peramalan model *hybrid* TSR-ARIMA tujuh periode ke depan dengan cara menjumlahkan hasil peramalan dari metode TSR dan hasil peramalan dari metode ARIMA sehingga diperoleh hasil peramalan model *hybrid* TSR-ARIMA pada Tabel 6 sebagai berikut:

Tabel 6 Hasil Peramalan dengan Model Hybrid TSR-ARIMA

Tahun	Bulan	Jumlah Penumpang
1960	Juni	522
	Juli	596
	Agustus	605
	September	508
	Oktober	451
	November	405
	Desember	447

Berdasarkan hasil prediksi dan peramalan dengan model *hybrid* TSR-ARIMA, perlu di visualisasikan perbandingan data aktual dengan data hasil peramalan yaitu dengan menggunakan grafik yang disajikan pada Gambar 9. Berdasarkan Gambar 9 dapat dilihat bahwa pola data peramalan hampir menyerupai pola data aktual.



Gambar 9. Grafik perbandingan data aktual dengan data hasil peramalan model *hybrid* TSR-ARIMA

4.5 Pembahasan

TSR secara langsung meramalkan data bulanan jumlah penumpang maskapai penerbangan Amerika Serikat periode Januari Tahun 1949 hingga Mei Tahun 1960 sedangkan model ARIMA terbaik dipilih berdasarkan kriteria MAPE terkecil yaitu model ARIMA (1,1,2)(0,1,0)¹². Hasil peramalan kedua metode tersebut selanjutnya dijumlahkan dan hasil penjumlahan tersebut merupakan hasil peramalan dari metode *hybrid* TSR-ARIMA (1,1,2)(0,1,0)¹². Pengukuran akurasi peramalan menggunakan MAPE dapat dilihat pada Tabel 7.

Tabel 7 MAPE dari Metode TSR dan Hybrid TSR-ARIMA

Metode	MAPE (%)
TSR	7,902
<i>Hybrid</i> TSR-ARIMA (1,1,2)(0,1,0) ¹²	3,061

Berdasarkan Tabel 7 dapat dilihat bahwa metode *hybrid* TSR-ARIMA (1,1,2)(0,1,0)¹² menghasilkan MAPE yang lebih kecil dibandingkan dengan metode TSR. Hal ini menunjukkan bahwa kinerja peramalan dengan metode *hybrid* TSR-ARIMA lebih baik dan akurat dibandingkan dengan metode TSR secara individu.

5. Kesimpulan

Berdasarkan pembahasan sebelumnya dapat disimpulkan bahwa kinerja peramalan dengan metode *hybrid* TSR-ARIMA lebih baik dan akurat dibandingkan dengan metode TSR dalam meramalkan jumlah penumpang maskapai penerbangan Amerika Serikat periode Januari Tahun 1949 hingga Mei Tahun 1960, hal ini berdasarkan bahwa *hybrid* TSR-ARIMA menghasilkan MAPE lebih kecil dibandingkan dengan metode TSR.

Daftar Pustaka:

- [1] Muslim, A. 2017. Peramalan Ekspor dengan Hibrida ARIMA-ANFIS. *Kajian Ekonomi & Keuangan*, 1(2), 128-142.
- [2] Gunaryati, A., Fauziah, Andryana, S. 2019. Hybrid Exponential Smoothing Neural Network Untuk Peramalan Data Pengguna Pita Lebar di Indonesia. *Jurnal Sistem Komputer dan Kecerdasan Buatan*, 2(2), 2-7.
- [3] Hejase, H. A. N, dan Assi, A. H. 2012. Time-Series Regression Model for Prediction of Mean Daily Global Solar Radiation in Al-Ain, UAE. *ISRN Renewable Energy*, 2012, 1-11.
- [4] Kashei, M., dan Bijari, M. 2010. An Artificial Neural Network (p,d,q) Model For Timeseries Forecasting. *Expert Syst. Appl.*, 37(1), 479-489.
- [5] Marine, M., Ginting, K. B., & Ariyanto. 2019. Peramalan Jumlah Penumpang Pesawat dengan Menggunakan Metode Dekomposisi (Studi Kasus: Unit Penyelenggara Bandar Udara (UPBU) Kelas II Frans Seda Maumere). *Jurnal Diferensial*, 1(1), 1-11.
- [6] Iqbalullah, J., & Winahju, W. S. 2014. Peramalan Jumlah Penumpang Pesawat Terbang di Pintu Kedatangan Bandar Udara Internasional Lombok dengan Metode ARIMA Box-Jenkins, ARIMAX, dan Regresi Time Series. *Jurnal Sains dan Seni Pomits*, 3(2), 212-217.
- [7] Shulfi, Z., & Syahnur, S. 2017. Analisis Kausalitas Transportasi Udara dan Pertumbuhan Ekonomi di Indonesia. *Jurnal Ilmiah Mahasiswa*, 2(4), 626-636.
- [8] Durrah, F. I., Yulia, Parhusip, T. P., & Rusyana, A. 2018. Peramalan Jumlah Penumpang Pesawat di Bandara Sultan Iskandar Muda dengan Metode SARIMA (Seasonal Autoregressive Integrated Moving Average). *Journal of Data Analysis*, 1(1), 01-11.
- [9] Suhartono, Isnawati, S., Salehah, N. A., Prastyo, D. D., Kuswanto, H., & Lee, M. H. 2018. Hybrid SSA-TSR-ARIMA for Water Demand Forecasting. *International Journal of Advances in Intelligent Informatics*, 4(3), 238-250.
- [10] Sulistyowati, R., Suhartono, Kuswanto, H., & Setiawan. 2018. Hybrid Forecasting Model to Predict Air Passenger and Cargo in Indonesia. *International Conference on Information and Communications Technology (ICOIACT)*, 442-447.
- [11] Wardah, S., dan Iskandar. 2016. Analisis Peramalan Penjualan Produk Keripik Pisang Kemasan Bungkus (Studi Kasus: Home Industry Arwana Foos Tembilihan). *Jurnal Teknik Industri*, 11(3), 135-142.
- [12] Ahmad, F. 2020. Penentuan Metode Peramalan Pada Produksi Part New Granada Bowl ST di PT.X. *Jurnal Integrasi Sistem Industri*, 7(1), 31-39.
- [13] Dani, A. T. R., Wahyuningsih, S., & Rizki, N. A. 2019. Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu. *Jambura Journal of Mathematics*, 1(2), 64-78.
- [14] Soejoeti, Z. 1987. *Analisis Runtun Waktu*. Jakarta: Karunika.
- [15] Makridakis, S., Wheelwright, S. C., dan McGee, V. E. 1999. *Metode dan Aplikasi Peramalan Jilid 1*. Jakarta: Binarupa Aksara.
- [16] Aswi dan Sukarna. 2006. *Analisis Deret Waktu: Teori dan Aplikasi*. Makassar: Andira Publisher.
- [17] Lubis, D.A., Johra, M.B., & Darmawan, G. 2017. Peramalan Indeks Harga Konsumen dengan Metode Singular Spectrum Analysis (SSA) dan Seasonal Autoregressive Integrated Moving Average (SARIMA). *MANTIK*, 3(2), 76-82.
- [18] Waeto, S., Chuarkham, K., & Intarasit, A. 2017. Forecasting Time Series Movement Direction with Hybrid Methodology. *Journal of Probability and Statistics*, 2017, 2-8.
- [19] Listiowarni, I., Dewi, N. P., & Hapantenda, A. K. W. 2020. Perbandingan Metode Double Exponential Smoothing dan Double Moving Average untuk Peramalan Harga Beras Eceran di Kabupaten Pamekasan. *Jurnal Komputer Terapan*, 6(2), 158-169