

# PERAMALAN TUNGGAKAN TAGIHAN PELANGGAN MENGUNAKAN MODEL AUTOREGRESSIVE FRACTIONALLY INTEGRATED MOVING AVERAGE (ARFIMA)

Tri Wulanda Fitri<sup>1\*</sup>, Gungum Darmawan<sup>2</sup>, Sri Winarni<sup>3</sup>, Yuyun Hidayat<sup>4</sup>,  
Resa Septiani Pontoh<sup>5</sup>

<sup>1,2,3,4,5</sup>Universitas Padjadjaran  
triwulandafitri@gmail.com\*



e-ISSN: 2987-811X

**MARAS: Jurnal Penelitian Multidisiplin**

<https://ejournal.lumbungpare.org/index.php/maras>

Vol. 1 No. 3 Desember 2023

Page: 614-625

## Article History:

Received: 27-12-2023

Accepted: 30-12-2023

**Abstrak** : Kenaikan jumlah pelanggan pada perusahaan X mengakibatkan kenaikan yang sama pada besar tunggakan tagihan pelanggan. Data tunggakan tersebut cenderung memiliki pola tren naik dan memiliki fluktuasi yang besar, akibatnya perusahaan memerlukan adanya peramalan guna perencanaan dan acuan dalam pengambilan keputusan serta evaluasi kebijakan yang tepat untuk mencegah peningkatan jumlah tunggakan tagihan di masa mendatang dan mengurangi risiko kerugian. Metode peramalan yang akan digunakan pada penelitian ini adalah Autoregressive Fractionally Integrated Moving Average (ARFIMA). Metode ARFIMA dipilih karena data memiliki sifat long memory, dilihat dari lag pada plot ACF data yang menurun secara hiperbolik menuju nilai nol. Artinya observasi yang terpisah jauh masih memiliki perilaku berkelanjutan atau ketergantungan yang kuat. Ciri lain dari metode ARFIMA adalah nilai differencing-nya. Berbeda dengan metode ARIMA yang memiliki parameter  $d$  berupa bilangan asli, metode ARFIMA memiliki parameter  $d$  berupa bilangan pecahan. Pada penelitian digunakan differencing sebesar 0,697 dengan model ARFIMA  $(1, d, 0)$  untuk melakukan peramalan 6 periode kedepan. Diperoleh nilai MAPE untuk peramalan tersebut sebesar 9,28%.

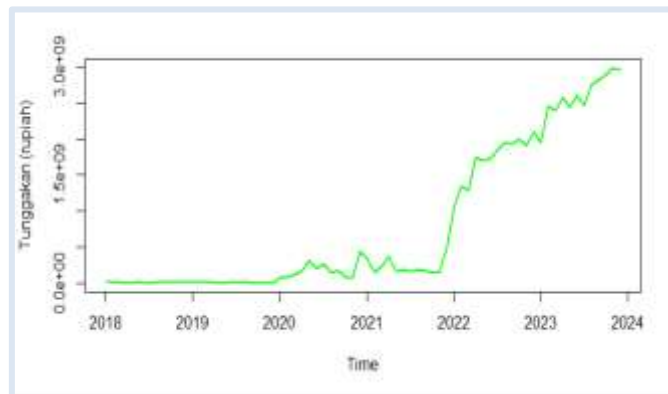
**Kata Kunci** : ARFIMA, Long Memory, Differencing, Tunggakan Tagihan

## PENDAHULUAN

Perusahaan X adalah perusahaan yang bergerak di bidang jasa layanan teknologi informasi komunikasi (TIK) dan jaringan telekomunikasi di Indonesia. Pada perjalanan bisnisnya, perusahaan X memiliki unit yang bergerak khusus untuk melayani para pelanggan bisnis dan korporat yang mana pelanggan unit ini mengalami peningkatan yang signifikan seiring berjalannya waktu. Tercatat bahwa jumlah

pelanggan unit tersebut pada Juli 2021 adalah sebanyak 493 pelanggan, dan pada Juli 2022 sebanyak 577 pelanggan. Hal ini menunjukkan adanya pertumbuhan jumlah pelanggan sebanyak 84 pelanggan atau tumbuh 17,04% dibanding periode yang sama tahun lalu.

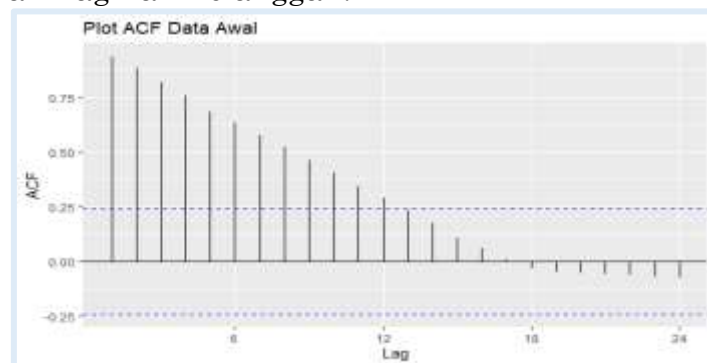
Pertumbuhan jumlah pelanggan tidak hanya berdampak positif bagi perusahaan. Di sisi lain, pertumbuhan jumlah pelanggan juga memberikan masalah baru bagi perusahaan, terutama dalam hal keuangan. Salah satu masalah yang ditimbulkan oleh bertumbuhnya jumlah pelanggan adalah tunggakan tagihan pelanggan yang cenderung meningkat seiring berjalannya waktu. Kenaikan pada tunggakan ini dapat dilihat dari grafik data bulanan Tunggakan Tagihan periode Januari 2018 sampai Desember 2023 berikut.



**Gambar 1.** Grafik Tunggakan Tagihan Pelanggan

Grafik data tunggakan di atas terlihat memiliki pola *trend* naik dan fluktuasi yang sangat besar. Data tersebut juga diduga memiliki sifat *long memory* yang artinya data deret waktu memiliki perilaku berkelanjutan atau ketergantungan yang kuat antara data saat ini dan pengamatan yang terjadi jauh di masa lalu. Hal tersebut diduga karena banyaknya tunggakan yang terjadi pada observasi terakhir disebabkan oleh kurangnya tindakan dari internal perusahaan dalam menanggulangi isu tunggakan pelanggan yang sudah berlangsung cukup lama yang mengakibatkan tunggakan pelanggan semakin naik seiring berjalannya waktu.

Hal selanjutnya yang menjadi alasan bahwa data penelitian ini memiliki sifat *long memory* terletak pada plot ACF-nya. Diketahui bahwa apabila nilai autokorelasi pada plot ACF menurun secara hiperbolik lambat menuju nilai nol, maka data diidentifikasi sebagai data deret waktu *long memory* (Wei, 2006). Berikut adalah plot ACF data Tunggakan Tagihan Pelanggan.



**Gambar 2.** Plot ACF Data Tunggakan Pelanggan

Berdasarkan Gambar 2 di atas dapat dilihat bahwa *lag* pada plot ACF tersebut sudah turun secara lambat menuju nilai nol. Hal ini membenarkan dugaan bahwa data deret waktu pada penelitian ini memiliki sifat *long memory*.

Berdasarkan penjelasan di atas, penulis tertarik untuk membahas model dan peramalan pada data *long memory*. Diketahui metode yang paling tepat untuk pemodelan dan peramalan data *long memory* adalah model ARFIMA (*Autoregressive Fractionally Integrated Moving Average*), sehingga penulis akan menggunakan metode analisis ARFIMA pada penelitian ini.

## METODE PENELITIAN

### Data Penelitian

Data yang digunakan berupa data tunggakan tagihan jasa pelanggan perusahaan X periode Januari 2018 sampai dengan Desember 2023. Data tersebut akan dibagi menjadi data pelatihan dan data uji. Data pelatihan digunakan dalam pembentukan model dan data uji digunakan untuk memeriksa keakuratan model yang dibangun. Data yang digunakan dalam penelitian ini berjumlah 72 observasi. Sebanyak 66 data digunakan sebagai data pelatihan yaitu observasi periode Januari 2018 sampai dengan Juni 2023, dan 6 observasi lainnya akan digunakan sebagai data uji.

### Metode Analisis ARFIMA

#### 1. Identifikasi *Long Memory*

Tahapan pertama untuk mengetahui apakah sebuah data deret waktu memiliki ciri *long memory* adalah dengan melihat plot ACF data tersebut. Jika nilai autokorelasi pada plot data ACF menurun secara hiperbolik menuju nilai nol, maka data tersebut dapat dianggap sebagai data *long memory*. Kemudian untuk memastikan data bersifat *long memory* secara statistik dilakukan pengujian menggunakan statistik uji *Hurst* (Hurst, 1951) dengan langkah-langkah sebagai berikut:

- Menentukan rata-rata, *adjusted mean*, dan simpangan baku dari data deret waktu dengan persamaan sebagai berikut:

$$\bar{Y} = \frac{1}{T} \sum_{i=1}^T Y_t \quad (7)$$

$$Y_t^{adj} = Y_t - \bar{Y} \quad (8)$$

$$S_t = \sqrt{\frac{1}{T} \sum_{i=1}^T (Y_t^{adj})^2} \quad (9)$$

dengan  $t = 1, 2, \dots, T$

- Menentukan deviasi kumulatif dan rentang dari deviasi kumulatif.

$$Y_t^* = \sum_{i=1}^T Y_t^{adj} \quad (10)$$

$$R_t = \text{Max}(Y_1^{adj}, Y_2^{adj}, \dots, Y_t^{adj}) - \text{Min}(Y_1^{adj}, Y_2^{adj}, \dots, Y_t^{adj}) \quad (11)$$

dengan  $t = 1, 2, \dots, T$

- Menentukan nilai eksponensial *Hurst* ( $H$ ) melalui statistik R/S dari data deret waktu.

$$H = \frac{\log\left(\frac{R}{S}\right)_t}{\log(t)} \quad (12)$$

Kriteria uji statistik *Hurst*:

- $0 < H < 0,5$  menunjukkan sifat *short memory*
- $0,5 < H < 1$  menunjukkan sifat *long memory*

## 2. Estimasi Parameter $d$ dengan Metode GPH

Estimasi GPH adalah salah satu teknik pengujian paling umum digunakan untuk fraksi integrasi  $d$ . Kelebihan dari metode Geweke dan Porter-Hudak adalah parameter  $d$  dapat langsung diestimasi tanpa terlebih dahulu mengetahui nilai orde AR ( $p$ ) dan MA ( $q$ ). Parameter  $d$  dapat ditaksir melalui persamaan (Geweke dan Porter-Hudak, 1983):

$$\hat{d} = \frac{\sum_{j=1}^m (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^m (x_j - \bar{x})^2}, m = \lceil T^{1/2} \rceil \quad (13)$$

dengan,

$$x_j = \ln\left(\frac{1}{4\sin^2(\lambda_j/2)}\right), \lambda_j = \frac{2\pi j}{T}$$

dimana,  $T$  = banyaknya pengamatan,  $j = 1, 2, \dots, m$

$$y_j = \ln I(\lambda_j), I(\lambda_j) = \frac{1}{2\pi} \left\{ \gamma_0 + 2 \sum_{t=1}^{T-1} \gamma_t \cos(t \lambda_j) \right\}, \lambda_j \in (-\pi, \pi)$$

## 3. Pemodelan

Proses pemodelan pada metode ARFIMA hampir sama dengan pemodelan pada model ARIMA yaitu, dengan mengidentifikasi orde AR( $p$ ) dan MA( $q$ ) pada plot ACF (*autocorrelation function*) dan PACF (*partial autocorrelation function*) untuk data yang sudah stasioner terhadap varians dan rata-rata. Langkah selanjutnya adalah melihat plot ACF dan PACF dari *correlogram cut off* di *lag* berapa. Dari plot ACF dan PACF tersebut bisa diidentifikasi beberapa kemungkinan model yang cocok untuk dijadikan model.

## 4. Uji Signifikansi Parameter

Uji signifikansi parameter dilakukan untuk melihat apakah model yang didapatkan dari proses sebelumnya cukup baik atau tidak. Berikut hipotesis yang digunakan dalam uji signifikansi parameter:

- Hipotesis:
  - $H_0 : \theta = 0$  (parameter tidak menjelaskan variabel respon)
  - $H_1 : \theta \neq 0$  (parameter menjelaskan variabel respon)
- Statistik Uji:

$$t_{hitung} = \frac{\hat{\theta}}{SE(\hat{\theta})} \quad (14)$$

dengan,

$\hat{\theta}$  : nilai estimasi parameter

$SE(\hat{\theta})$  : standar *error* dari nilai estimasi parameter ( $\hat{\theta}$ )

- Kriteria Uji:

Tolak  $H_0$  jika nilai  $|t_{hitung}| > t_{\frac{\alpha}{2},(T-p)}$ , dengan nilai  $n$  sebagai banyaknya observasi dan  $p$  adalah jumlah parameter yang ditaksir.

## 5. Uji Diagnostik

Tahap ini menguji apakah residual data memenuhi asumsi *white noise*, yaitu apakah residual data bersifat acak dan terdistribusi normal. Dalam pengujian diagnostik dilakukan prosedur sebagai berikut:

### a. Asumsi nilai residual berdistribusi normal

Pengujian diagnostik model yang pertama adalah melakukan uji normalitas untuk melihat kenormalan dari residual data. Untuk menguji normalitas residual dapat digunakan salah satunya adalah dengan menggunakan uji *Kolmogorov-Smirnov*.

- Hipotesis:

$H_0 : a_t \sim N(0, \sigma_{a_t}^2)$  (residual berdistribusi normal)

$H_1 : a_t \neq N(0, \sigma_{a_t}^2)$  (residual tidak berdistribusi normal)

- Statistik Uji :

$$D = \max |F_0(a_t) - S(a_t)| \quad (15)$$

dengan,

$F_0(a_t)$  : fungsi peluang kumulatif dari distribusi normal

$S(a_t)$  : fungsi peluang kumulatif data

- Kriteria Uji:

$H_0$  ditolak jika nilai maksimum  $|F_0(a_t) - S(a_t)| >$  nilai pada tabel *Kolmogorov-Smirnov*, sebaliknya apabila nilai maksimum  $|F_0(a_t) - S(a_t)| <$  nilai pada tabel *Kolmogorov-Smirnov* maka  $H_0$  diterima. Cara lain adalah dengan melihat nilai *p-value* dari analisis uji *Kolmogorov-Smirnov* yaitu, tolak  $H_0$  apabila *p-value*  $<$   $\alpha$ .

### b. Asumsi nilai residual bersifat acak

Pengujian diagnostik model yang kedua dilakukan adalah menguji keacakan (korelasi) terhadap residual data dengan menggunakan statistik uji *Box – Pierce* dan *Ljung – Box*.

- Hipotesis:

$H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0$  (residual bersifat acak (*uncorrelated*))

$H_1$  : minimal terdapat satu nilai  $\rho_k \neq 0$  (terdapat korelasi pada residual)

- Statistik Uji:

$$Q = n(n+2) \sum_{k=1}^m \frac{r_k^2}{n-k} \quad (16)$$

dimana  $r_k^2$  adalah autokorelasi residual *lag*- $k$ , dan  $m$  adalah *lag* maksimum (*lag* maksimum ditentukan sembarang, tetapi untuk *lag* yang cukup besar).

Autokorelasi residual :

$$r_k^2 = \frac{\sum_{t=k+1}^n e_t e_{t-k}}{\sum_{t=1}^n e_t^2} \quad (17)$$

- Kriteria Uji:

$Q \leq X_{(a),(k-p-q)}^2$   $H_0$  diterima,

$Q > X_{(a),(k-p-q)}^2$   $H_0$  ditolak.

## 6. Pemilihan Model Terbaik

Model yang telah memenuhi asumsi akan dilihat nilai *Akaike Information Criterion* (AIC). Model dengan nilai AIC terkecil yang akan menjadi model terbaik untuk peramalan. Berikut persamaan untuk menghitung nilai AIC (Wei, 2006):

$$AIC = n \ln \hat{\sigma}_a^2 + 2M \quad (18)$$

Nilai *Mean Absolute Percentage Error* (MAPE) digunakan untuk mengukur keakuratan nilai model yang diestimasi dan dinyatakan dalam bentuk rata-rata persentase *absolute* kesalahan. MAPE juga menunjukkan seberapa besar kesalahan prediksi ketika membandingkan nilai sebenarnya dari data dengan hasil prediksi. Model berkinerja sangat baik ketika MAPE kurang dari 10% dan berkinerja baik ketika MAPE antara 10% dan 20% (Lewis, 1982). Berikut persamaan MAPE:

$$MAPE = \frac{\sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t}}{n} \times 100\% \quad (19)$$

dengan,

$y_t$ : nilai aktual waktu  $t$

$\hat{y}_t$ : nilai peramalan  $y_t$

$n$ : jumlah data historis

## HASIL DAN PEMBAHASAN

### 1. Identifikasi Long Memory

Tahapan pertama untuk mengidentifikasi data deret waktu *long memory* adalah dengan melihat plot ACF data. Plot ACF data penelitian ini dapat dilihat pada Gambar 2. Pada gambar tersebut terlihat bahwa data deret waktu telah memiliki sifat *long memory* dengan ciri lag pada plot ACF-nya yang turun secara lambat menuju nilai nol (Wei, 2006).

Langkah selanjutnya untuk mengidentifikasi sifat *long memory* pada data deret waktu adalah dengan uji statistik *Hurst*. Didapatkan hasil uji *Hurst* menggunakan bantuan *software* R untuk data ini yaitu sebesar 0,766 yang mana nilai tersebut sudah memenuhi kriteria  $0,5 < H < 1$  yang menunjukkan bahwa suatu data deret waktu mengandung sifat *long memory*.

### 2. Uji Stasioneritas

#### a. Stasioneritas Varians

Pemeriksaan stasioneritas varians pada data dapat dilihat dari nilai lambda. Apabila nilai lambda mendekati 1 maka, data sudah stasioner dalam varians. Data yang belum stasioner dalam varians harus melewati tahapan transformasi *Box-Cox* terlebih dahulu dengan rumus  $Y_t^\lambda$ . Berikut adalah nilai estimasi lambda sebelum dan sesudah transformasi.

**Tabel 1.** Estimasi Nilai  $\lambda$

Transformasi $Y_t^\lambda$	Sebelum Transformasi	Sesudah Transformasi
Estimasi Nilai $\lambda$	0,21	1,19
Keterangan	Belum Stasioner	Sudah Stasioner

Dari tabel di atas dapat dilihat bahwa data sudah stasioner dalam varians setelah dilakukannya 1 kali transformasi dengan nilai akhir lambda 1,194.

#### b. Stasioneritas Rata-rata

Stasioneritas rata-rata pada data dapat dilihat melalui pengujian ADF (*Augmented Dickey Fuller*). Jika nilai  $p\text{-value} > \alpha = 5\%$  maka, data tidak stasioner

dalam rata-rata, sebaliknya nilai  $p\text{-value} < \alpha = 5\%$  menyatakan data deret waktu memiliki rata-rata yang stasioner. Didapatkan hasil analisis  $p\text{-value}$  dengan menggunakan bantuan *software* R adalah sebesar 0,37 yang artinya data belum stasioner dalam rata-rata. Apabila rata-rata data belum stasioner maka, diperlukan adanya proses *differencing*. Pada data *long memory* proses *differencing* dilakukan dengan parameter bernilai pecahan. Hal ini *differencing* yang dilakukan dengan nilai pecahan membantu menghilangkan komponen tren atau pola sekuensial dari data dengan cara yang lebih lembut daripada *differencing* dengan bilangan bulat (J. Pan et al, 2008).

Metode yang sering digunakan untuk memperoleh parameter  $d$  untuk data *long memory* adalah dengan metode Geweke Poter-Hudak (GPH). Berikut hasil pengujian menggunakan bantuan *software* R:

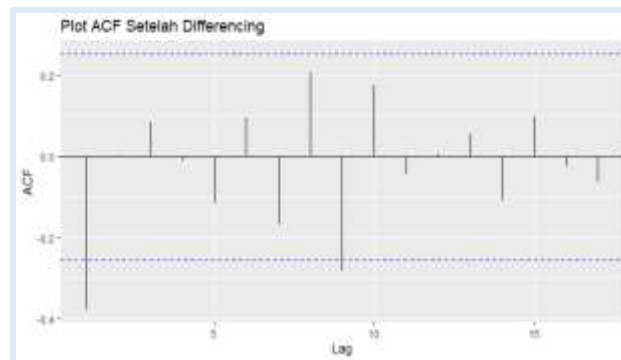
**Tabel 2.** Hasil *Differencing*

$p\text{-value}$ awal	$\alpha$	Keterangan	<i>Differencing</i>	$p\text{-value}$ setelah <i>differencing</i>	Keterangan
0,37	0,05	Belum Stasioner	0,69	0,01	Sudah Stasioner

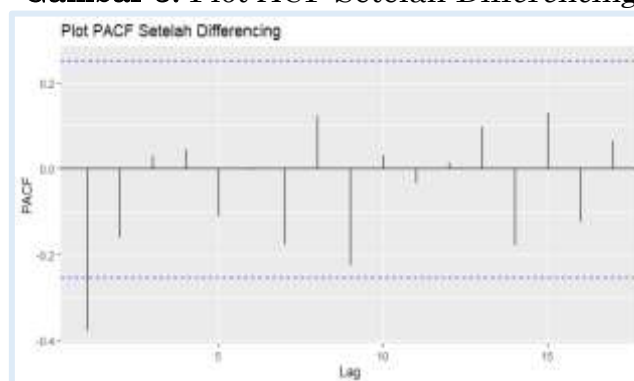
Dari tabel di atas dapat dilihat bahwa data telah stasioner dalam rata-rata setelah dilakukan *differencing* sebesar 0,69.

### 3. Identifikasi Model

Langkah selanjutnya adalah menentukan orde  $AR(p)$  dan  $MA(q)$  untuk mengidentifikasi kemungkinan model peramalan terbaik dengan melihat plot ACF dan PACF data deret waktu setelah stasioner dalam varians dan rata-rata. Berikut disajikan plot ACF dan PACF data setelah stasioner.



**Gambar 3.** Plot ACF Setelah *Differencing*



**Gambar 4.** Plot PACF Setelah *Differencing*

Dapat dilihat *cut off* pada plot ACF dan PACF di atas sama-sama terjadi setelah *lag* pertama, dengan parameter  $d$  sebesar 0,69 sehingga estimasi model ARFIMA yang akan digunakan adalah sebagai berikut:

**Tabel 3.** Model ARFIMA

Model ARFIMA
<i>ARFIMA (1, d, 0)</i>
<i>ARFIMA (1, d, 1)</i>
<i>ARFIMA (0, d, 1)</i>

#### 4. Uji Signifikansi Parameter

Estimasi model yang telah didapatkan perlu dilakukan uji signifikansi untuk melihat apakah parameter berpengaruh secara signifikan terhadap model. Kriteria uji untuk parameter yang berpengaruh signifikan terhadap model adalah ketika nilai  $|t_{hitung}| > t_{\frac{\alpha}{2}, (T-p)}$ . Berikut hasil analisis signifikansi parameter pada model-model estimasi sebelumnya:

**Tabel 4.** Parameter Model ARFIMA (1, d, 0)

	$\phi_1$
Estimasi Parameter	0,96
<i>Standard Error</i>	0,03
$ t_{hitung} $	27,61
$t_{\frac{\alpha}{2}, (T-p)}$	2,00
Keterangan	Signifikan

**Tabel 5.** Parameter Model ARFIMA (1, d, 1)

	$\phi_1$	$\theta_1$
Estimasi Parameter	0,96	0,37
<i>Standard Error</i>	0,04	0,25
$ t_{hitung} $	23,47	1,45
$t_{\frac{\alpha}{2}, (T-p)}$	2,00	2,00
Keterangan	Signifikan	Tidak Signifikan

Berdasarkan tabel di atas, dapat dilihat bahwa kedua model estimasi memperoleh nilai  $p\text{-value} > \alpha = 5\%$  artinya asumsi nilai residual bersifat normal pada kedua model estimasi terpenuhi.

#### 5. Uji Diagnostik

##### a. Uji Normalitas Residual

Uji normalitas pada nilai sisa bertujuan untuk melihat apakah residual berdistribusi normal atau tidak. Pengujian dilakukan menggunakan uji *Kolmogorov-Smirnov* dengan bantuan *software* R. Berikut ditampilkan rangkuman *output* pengujian normalitas residual menggunakan uji *Kolmogorov-Smirnov* pada *software* R:

**Tabel 6.** Hasil Analisis Uji Normalitas Residual

Model ARFIMA ( $d = 0,69$ )	$p\text{-value}$	Keterangan
(1, d, 0)	0,07	$H_0$ diterima
(0, d, 1)	0,06	$H_0$ diterima



### b. Uji Non-autokorelasi Residual

Pengujian non-autokorelasi pada nilai sisa menggunakan statistik uji *Box-Ljung*. Berikut rangkuman *output* uji non-autokorelasi residual menggunakan statistik uji *Box-Ljung* menggunakan *software R*:

**Tabel 7.** Hasil Analisis Uji Non-Autokorelasi Residual

Model ARFIMA ( $d = 0,69$ )	$p$ -value	Keterangan
(1, $d$ ,0)	0,75	$H_0$ diterima
(0, $d$ ,1)	0,02	$H_0$ ditolak

Berdasarkan tabel di atas dapat dilihat bahwa model estimasi yang memenuhi asumsi non-autokorelasi hanyalah model ARFIMA (1, $d$ ,0), artinya model tersebut memiliki residual data yang bersifat acak atau *uncorrelated*.

### 6. Model Terbaik

Dari analisis sebelumnya, diketahui bahwa hanya terdapat 1 model yang signifikan dan memenuhi asumsi residual yaitu model ARFIMA (1, $d$ ,0) dengan parameter  $d = 0,69$ . Selanjutnya akan dilihat nilai AIC dan MAPE dari model tersebut.

**Tabel 8.** Nilai AIC dan MAPE Model Terpilih

Model ARFIMA ( $d = 0,69$ )	AIC	MAPE (%)
(1, $d$ ,0)	2256,88	9,28

Nilai MAPE didapat dari data testing sebanyak 6 observasi yang diambil dari periode Juli 2023 sampai Desember 2023. Sebelumnya telah dilakukan peramalan dan pemodelan pada data *training* periode Januari 2018 sampai Juni 2023 menggunakan model ARFIMA (1, $d$ ,0) dengan  $d = 0,69$ . Berikut adalah tabel analisis nilai MAPE untuk data *testing* pada penelitian ini.

**Tabel 9.** Analisis MAPE Model Terpilih

Periode	Data Aktual	<i>Forecast</i>	Persentase (%)
Juli 2023	2464640158	2604807195	5,68
Agustus 2023	2745750660	2605284602	5,11
September 2023	2825728957	2601503245	7,93
Oktober 2023	2904822571	2595024804	10,66
November 2023	2983229008	2586610332	13,29
Desember 2023	2961108837	2576716156	12,98
		<b>MAPE</b>	<b>9,28</b>

Model ARFIMA (1;  $d = 0,69$ ; 0) sudah dapat dikategorikan baik karena memiliki nilai persentase MAPE sebesar 9,28% yang mana MAPE dengan persentase antara 0% - 20% menunjukkan bahwa model memiliki akurasi peramalan yang baik (Meng Dun et al, 2020). Apabila model ARFIMA (1;  $d = 0,69$ ; 0) dimasukkan ke dalam persamaan (1) dan (6) maka, dapat dituliskan sebagai berikut:

$$\begin{aligned} \phi_p(B)(1-B)^d Y_t &= \theta_q(B) a_t \\ 0,96_1(B)(1-B)^{0,69} Y_t &= \theta_q(B) a_t \\ (1-B)^{0,69}(1-0,96) Y_t &= 0 \quad (20) \end{aligned}$$

dengan,

$$(1 - B)^d = 1 - dB - \frac{1}{2}(1 - d)dB^2 - \frac{1}{6}(1 - d)(2 - d)dB^3 + \dots$$

$$(1 - B)^{0,69} = 1 - (0,69)(B) - \frac{1}{2}(1 - 0,69)(0,69)B^2 - \frac{1}{6}(1 - 0,69)(2 - 0,69)(0,69)B^3 + \dots$$

$$(1 - B)^{0,69} = 1 - 0,69B - 0,10B^2 - 0,04B^3 + \dots$$

sehingga model menjadi,

$$(1 - 0,69B - 0,10B^2 - 0,04B^3 + \dots)(1 - 0,96)Y_t = 0$$

$$(1 - 0,69B - 0,10B^2 - 0,04B^3 + \dots)(0,03)Y_t = 0$$

$$(0,03Y_t - 0,02BY_t - 0,0031B^2Y_t - 0,0013B^3Y_t + \dots) = 0$$

dengan  $BY_t = Y_{t-1}$ , maka:

$$(0,0301Y_t - 0,021Y_{t-1} - 0,00317Y_{t-2} - 0,001378Y_{t-3} + \dots) = 0$$

diperoleh persamaan ARFIMA (1;  $d = 0,69$ ; 0) sebagai berikut:

$$Y_t = \frac{0,021Y_{t-1} + 0,00317Y_{t-2} + 0,001378Y_{t-3} - \dots}{0,0301} \quad (21)$$

## 7. Peramalan

Peramalan data tunggakan akan dilakukan untuk 6 observasi kedepan yaitu untuk tunggakan periode bulan Januari 2024 sampai Juni 2024. Peramalan dilakukan dengan model yang telah terpilih sebelumnya yaitu, model ARFIMA (1;  $d = 0,69$ ; 0) dan MAPE sebesar 9,28% menggunakan persamaan :

$$Y_t = \frac{0,021Y_{t-1} + 0,00317Y_{t-2} + 0,001378Y_{t-3} - \dots}{0,0301}$$

Berikut hasil peramalannya :

**Tabel 10.** Hasil Peramalan ARFIMA (1;  $d = 0,69$ ; 0)

Periode	Peramalan Tunggakan
Januari 2024	Rp2.961.997.399,00
Februari 2024	Rp2.960.660.303,00
Maret 2024	Rp2.957.341.131,00
April 2024	Rp2.952.430.036,00
Mei 2024	Rp2.946.227.012,00
Juni 2024	Rp2.938.956.800,00

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada dosen-dosen Prodi Statistika Universitas Padjadjaran

1. Dr. Gungum Darmawan, S.Si., M.Si.
2. Sri Winarni, S.Si., M.Si.
3. Prof. Yuyun Hidayat, MS., Ph.D.
4. Resa Septiani Pontoh, S.Si., M.A.B., M.Stat.Sci.

Atas bimbingan dan telaah keilmiah yang dilakukan guna meningkatkan kualitas artikel ilmiah ini. Semoga ilmu yang diberikan dapat bermanfaat bagi bidang keilmuan terkait.

## KESIMPULAN DAN SARAN

Penelitian ini membahas pemodelan dan peramalan menggunakan ARFIMA untuk data tunggakan tagihan pelanggan observasi bulan Januari 2018 sampai Desember 2023 dengan peramalan untuk periode Januari 2024 sampai Juni 2024.

Model terbaik yang digunakan untuk penelitian ini adalah model ARFIMA (1;  $d = 0,69$ ; 0) dengan persamaan:

$$Y_t = \frac{0,021Y_{t-1} + 0,00317Y_{t-2} + 0,001378Y_{t-3} - \dots}{0,0301}$$

Model dipilih karena bersifat signifikan dan telah memenuhi asumsi residual normalitas dan non-autokorelasi. Model terpilih juga sudah memiliki akurasi peramalan yang cukup baik dengan MAPE sebesar 9,28%.

Saran yang dapat diberikan untuk penelitian selanjutnya adalah dalam mengestimasi parameter  $d$  dapat menggunakan metode lainnya seperti SPR, GPHT<sub>r</sub>, GPHT<sub>a</sub>, dan MGPH.

## DAFTAR PUSTAKA

- [1] Agustiana, Ferdian., dkk. (2016). *Perbandingan Penaksiran Parameter Pembeda (d) ARFIMA Nonstasioner GPH dengan Modifikasi GPH (Geweke and Porter-Hudak)*. BIAStatistics, 10(2), 34-44.
- [2] Assaf, A. (2006). *Persistence and Long Memory Dependence In The Emerging Stock Market of Kuwait*, The Middle East Business and Economic Review, 18(1).
- [3] Baillie, R., & Chung, S. (2002). *Modeling and forecasting from trend-stationary long memory models with applications to climatology*. International Journal of Forecasting, 18, 215-226. [https://doi.org/10.1016/S0169-2070\(01\)00154-6](https://doi.org/10.1016/S0169-2070(01)00154-6).
- [4] Breidert, Christoph. (2006). *Estimation of Willingness-to-Pay*. Wiesbaden: Gabler Edition Wissenschaft.
- [5] Caraka, Rezzy Eko, dkk. (2016). *Model Long Memory Dalam Prediksi Suhu*. Prosiding 2016 Anomali Cuaca dan Iklim Sebagai Peluang dan Tantangan, 222-237.
- [6] Darmawan, Gungum. (2009). *Perbandingan Metode Peramalan ARIMA dan ARFIMA pada Data Long Memory*. Jurnal Unisba, 9(2), 109-113.
- [7] Dun, M., Xu, Z., Chen, Y., & Wu, L. (2020). *Short-Term Air Quality Prediction Based on Fractional Grey Linear Regression and Support Vector Machine*. Mathematical Problems in Engineering, 2020, 1-13. <https://doi.org/10.1155/2020/8914501>.
- [8] Geweke, J dan Porter-Hudak, S. 1983. *The Estimation and Application of Long Memory Time Series Models*. *Journal of Time Series Analysis*. 4(4), 221-238.
- [9] Haba, A., dkk. (2021). *Prediksi Jumlah Tunggakan Tagihan Rekening Air Menggunakan Metode Least Square*. Jurnal Nasional Cosphi, 5(2), 31-35.
- [10] Hartini, Dwi. (2016). *Penerapan Model ARFIMA (Autoregressive Fractionally Integrated Moving Average) dalam Prakiraan Data Suku Bunga PUAB (Pasar Uang Antar Bank)*.
- [11] Hurst, HE. (1951). *The Problem of Long Term Storage in Reservoirs: An Experimental Study*. International Association of Scientific Hydrology, Bulletin. 116: 770-799.

- [12] Hutabarat, Magdaniar. (2017). *Pemodelan Hubungan Antara IHSG, Nilai Tukar Dolar Amerika Serikat Terhadap Rupiah (Kurs) dan Inflasi dengan Vector Error Correction Model (VECM)*.
- [13] Lewis, C. (1982). *Industrial and Business Forecasting Methods*. London: Butterworth.
- [14] Montgomery, Douglas C., C. L. Jennings., M. Kulahci. (2008). *Introduction to Time Series Analysis and Forecasting*. New Jersey: John Wiley & Sons, Inc.
- [15] Palma, Wilfredo. (2007). *Long-Memory Time Series*. New Jersey: John Wiley & Sons, Inc.
- [16] Panjaitan, Helmi, A. Prahutama, dan Sudarno. (2018). *Peramalan Jumlah Penumpang Kereta Api Menggunakan Metode ARIMA, Intervensi, dan ARFIMA*. Jurnal Gaussian, 7(1), 96-109.
- [17] Pan, J., & Chen, S. (2008). *Monitoring long-memory air quality data using ARFIMA model*. Environmetrics, 19. <https://doi.org/10.1002/ENV.882>.
- [18] Paridi. (2018). *Perbandingan Metode ARIMA (Box Jenkins) dan ARFIMA dalam Peramalan Jumlah Kasus Demam Berdarah Dengue di Rumah Sakit Hasan Sadikin Bandung*. Jurnal Pendidikan Mandala, 3(3), 338-348.
- [19] Petriella, Yanita. (2022). *Tagihan Indihome Naik Mulai April 2022, Imbas PPN 11 Persen*. Diakses pada Agustus 2022, dari <https://ekonomi.bisnis.com/>.
- [20] Prasetya, B. D., F. S. Pamungkas, I. Kharisudin. (2020). *Pemodelan dan Peramalan Data Saham dengan Analisis Time Series menggunakan Python*. PRISMA. Prosiding Seminar Nasional Matematika 3, 714-718.
- [21] Pudiastuti, V. E., F. Hasun, M. Dellaroswati. (2015). *Evaluasi Besaran Tarif Telkom Speedy dengan Metode Value Based Pricing*. Jurnal Tugas Akhir Fakultas Rekayasa Industri, 2(1), 870-877.
- [22] Sitepu, F. T. Br., V. A. P. Sirait., R. Yunis. (2021). *Analisis Deret Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Prophet Facebook*. Jurnal Paradigma, 23(1), 99-105.
- [23] Telkom Indonesia. (2022). *Tentang Telkomgroup*. Diakses pada Agustus 2022, dari <https://www.telkom.co.id/>.
- [24] Wei, William W.S. (2006). *Time Series Analysis: Univariate and Multivariate Methods, 2nd edition*. New York: Pearson Education, Inc.