

Forecasting the Ambon City Consumer Price Index Using Arima Box-Jenkins

Ronald John Djami^{1*}, Jefri Esna Thomas Radjabaycolle², Gabriella Haumahu³

¹ Universitas Pattimura, Jalan Ir. M. Putuhena, Kampus Poka, Kota Ambon, Indonesia.

*Email: ronalddjami@gmail.com

Manuscript submitted : October 2021.

Accepted for publication : November 2021

doi : <https://doi.org/10.30598/tensorvol2iss2pp87-96>

Abstract: The Consumer Price Index is an index number that measures the average price of goods and services consumed by households. The index number is the price comparison in a certain month against the previous month, in which case the price in the previous month is the price in the base year in the CPI calculation. CPI is time series data, so CPI data in the next period can be known by forecasting through time series analysis. Arima is a technique for finding the most suitable pattern from a group of data (curve fitting). Based on the results of the analysis, the best ARIMA model used in forecasting CPI in Ambon city for the period January 2007 to December 2020 is the ARIMA model (1,1,1), namely $\alpha_1 = 0.9000$ and $\beta_1 = 0.9933$.

2010 Mathematical Subject Classification: 62M10

Keywords: Arima, Forecasting, Consumer Price Index

1. Pendahuluan

Indeks Harga Konsumen (*Consumer Price Index*) merupakan nomor indeks yang mengukur harga rata-rata dari barang dan jasa yang dikonsumsi oleh rumah tangga. Nomor indeks tersebut merupakan perbandingan harga pada suatu bulan tertentu terhadap bulan sebelumnya, di mana dalam hal ini harga pada bulan sebelumnya merupakan harga pada tahun dasar pada perhitungan IHK. IHK sering digunakan untuk mengukur tingkat kenaikan secara umum harga barang dan jasa yang merupakan kebutuhan pokok masyarakat suatu negara dan juga sebagai pertimbangan untuk penyesuaian gaji, upah, uang pensiun, dan kontrak lainnya [5].

Data mengenai IHK secara resmi dikeluarkan oleh pemerintah melalui Badan Pusat Statistik (BPS). BPS melakukan perhitungan IHK melalui analisis pada perhitungan jumlah jenis barang/jasa yang tercakup dalam komoditas IHK, nilai konsumen dari suatu jenis barang/jasa, maupun perbandingan harga suatu jenis barang/jasa dari bulan yang sedang berjalan terhadap bulan sebelumnya. Dengan demikian, data-data IHK merupakan data yang saling berhubungan. Hal ini dapat dilihat dari IHK pada periode sekarang dipengaruhi

oleh periode sebelumnya. Karena data-data tersebut memiliki keterkaitan antar periode yang tetap maka data IHK dapat dikatakan sebagai data runtun waktu. Oleh karena IHK merupakan data runtun waktu maka data IHK pada periode berikutnya dapat diketahui dengan melakukan peramalan melalui analisis runtun waktu. Salah satu metode statistika untuk peramalan data runtun waktu adalah Arima Box-Jenkins.

Arima adalah teknik mencari pola yang paling cocok dari sekelompok data (curve fitting). Model ARIMA adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. ARIMA cocok jika observasi dari runtun waktu (time series) secara statistik berhubungan satu sama lain (Dependent). Model ARIMA adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. ARIMA cocok jika observasi dari runtun waktu (time series) secara statistik berhubungan satu sama lain (Dependent)[1].

ARIMA terdiri dari model *Autoregressive* (AR) yang menggunakan hubungan data dengan data masa lalu (data lag). [2] dalam bukunya menjelaskan bahwa model AR(p) mengikuti persamaan umum model sebagai berikut

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_p$$

yang mana ε_p mengikuti *white noise* (berdistribusi normal dengan rata-rata 0 dan varians σ^2 dan tidak berkorelasi dengan y_s untuk setiap $s < t$)

Integrated (I) merupakan proses *differencing* untuk menjadikan data dari tidak stasioner menjadi data yang stasioner. Stasioner adalah kondisi data dengan rata-rata dan variansi yang konstan sepanjang waktu.

Sedangkan model *Moving Average* (MA) menggunakan hubungan data dengan residual yang diaplikasikan pada data masa lalu. Persamaan umum model MA(q) mengikuti

$$y_t = \varepsilon_p - \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_p \varepsilon_{t-p}$$

Dengan $\theta_1 \dots \theta_p$ adalah koefisien eror masa lalu. Untuk mendapatkan nilai p dan q bisa dilakukan dengan mengamati pola korelogram *Autocorellation* (AC) dan *Partial Corellation* (PAC), yaitu dengan mengamati garis bar yang melewati garis putus-putus.

Penelitian terkait IHK dengan metode ARIMA pernah dilakukan oleh [3]. Data IHK yang digunakan adalah data Indonesia periode September 2008 hingga oktober 2012. Hasil penelitiannya menemukan bahwa pemodelan IHK Indonesia mengikuti model ARIMA(3,1,2) dengan nilai MSE 0,14966.

Peramalan IHK juga dilakukan oleh [4]. Data IHK yang digunakan adalah data Indonesia periode Januari 2009 hingga Maret 2015. Hasil penelitiannya menemukan bahwa pemodelan IHK Indonesia mengikuti model ARIMA(2,1,3) dengan orde intervensi $b = 0$, $s = 2$, dan $r = 0$.

Berdasarkan uraian di atas, maka penulis akan membuat sebuah peramalan IHK dengan metode ARIMA Box - Jenkins dikota Ambon periode januari 2007 hingga desember 2020.

2. Metodologi

1.1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder tentang data Indeks Harga Konsumen (IHK) bulanan kota Ambon periode januari 2007 sampai desember 2020 yang diperoleh dari Badan Pusat Statistik Indonesia melalui situs <https://www.bps.go.id/indicator/3/2/1/indeks-harga-konsumen-umum->

[.htmi](#). Pengolahan data, khususnya pada pendugaan parameter model, dilakukan dengan menggunakan software MINITAB 16.

1.2. Tahapan Penelitian

Tahapan analisis yang digunakan dalam penelitian ini seperti yang ada pada gambar 1. Tahapan tersebut meliputi 3 tahap besar, yaitu identifikasi, estimasi pengujian, serta tahap aplikasi.

1) Identifikasi Model

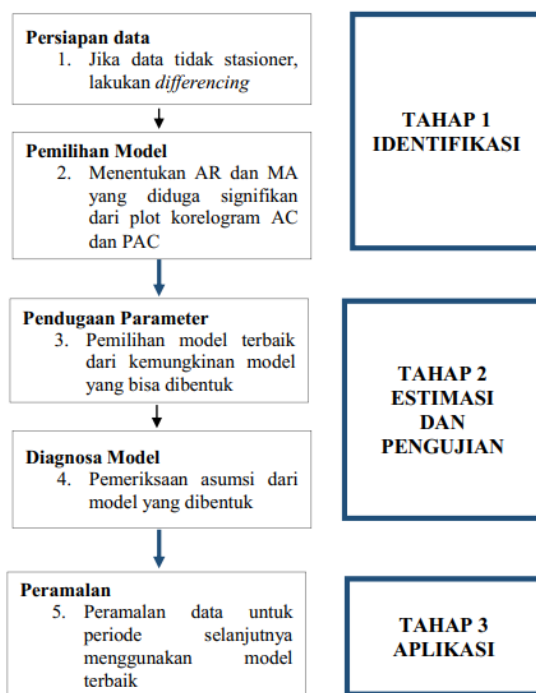
Tahap identifikasi model meliputi proses pengecekan kestasioneran data. Jika data sudah stasioner, langkah selanjutnya adalah melakukan penentuan ordo AR dan MA melalui bar korelogram.

2) Estimasi Pengujian Parameter

Tahap estimasi dan pengujian Parameter melalui pemilihan model yang akan saling diperbandingkan dari semua kemungkinan model yang bisa dibentuk. Model terbaik akan digunakan untuk menduga nilai parameter. Langkah selanjutnya adalah diagnosa model terbaik melalui pemeriksaan asumsi-asumsi yang harus dipenuhi oleh model terbaik.

3) Peramalan

Tahap ketiga adalah peramalan atau *forecasting* terhadap data periode selanjutnya. Peramalan ini dilakukan dengan menggunakan model terbaik yang didapatkan.

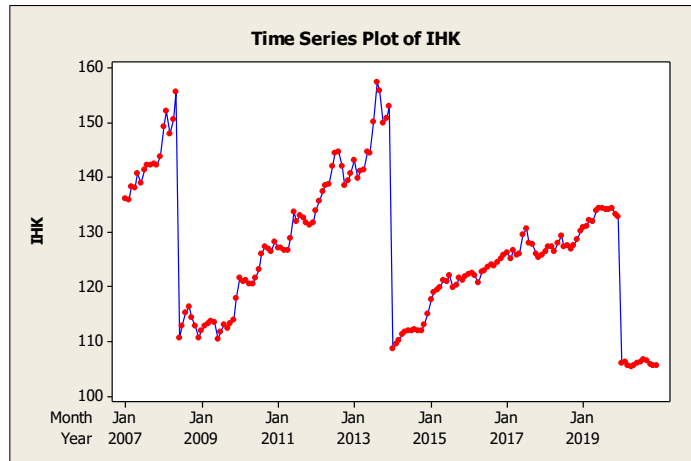


Gambar 1. Tahapan Analisis

3. Hasil dan Diskusi

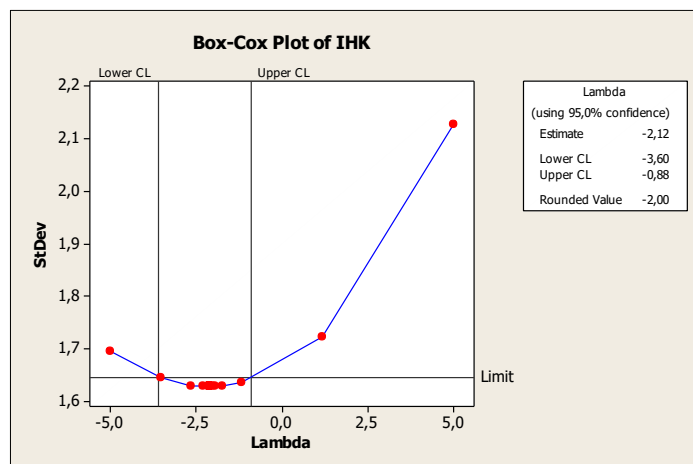
Dalam penelitian ini model time series akan dibuat berdasarkan data IHK periode januari 2007 sampai dengan Desember 2020 dengan menggunakan model ARIMA Box - Jenkins. Adapun langkah-langkah pada analisis model ARIMA menggunakan bantuan Software Minitab.

Langkah awal dalam ARIMA adalah membuat plot data peramalan IHK dalam bentuk plot data untuk mengetahui gerakan perubahan IHK terhadap waktu. Berikut ini adalah data plot data IHK di Kota Ambon periode Januari 2007 hingga Desember 2020.



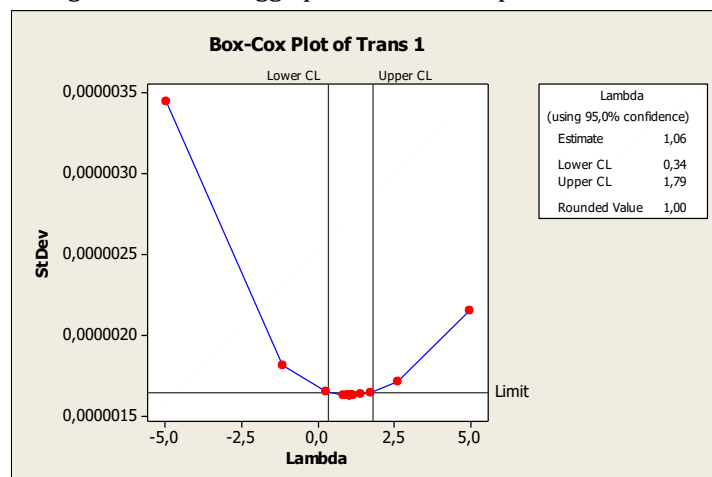
Gambar 2. Plot IHK

Berdasarkan Gambar 2, dapat diketahui bahwa *plot* yang terjadi pada data IHK adalah data dengan *trend* acak sehingga data tersebut merupakan data non musiman, sehingga metode yang tepat untuk melakukan peramalan adalah metode ARIMA Box-Jenkins. Setelah melihat plot data IHK, selanjutnya akan dilihat stasioner terhadap varian dan dapat dilihat pada Gambar 3. Box-cox plot of IHK.



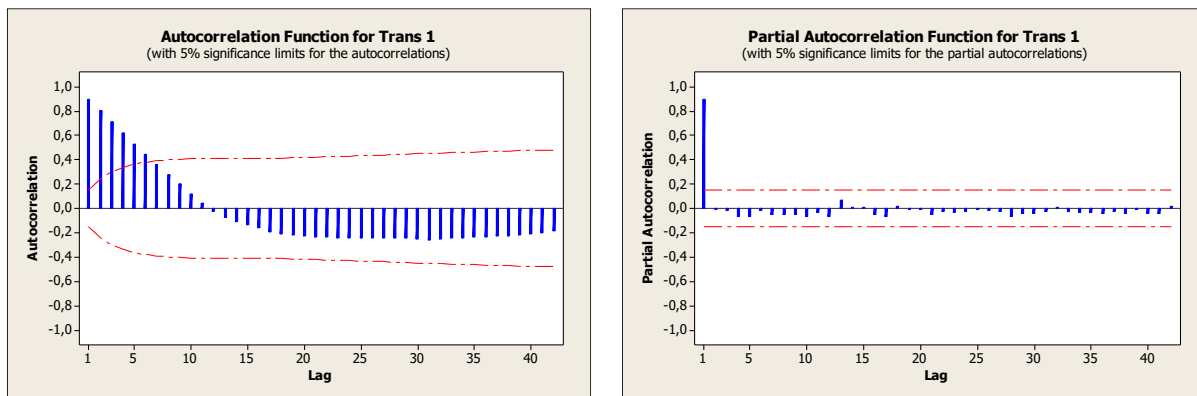
Gambar 3. Box-cox plot of IHK.

Gambar 3. Box-Cox di atas dapat diketahui bahwa data IHK di kota Ambon dalam periode bulanan mulai dari Januari 2007 sampai Desember 2020 belum stasioner terhadap varian, hal ini dilihat berdasarkan nilai rounded value yang sama dengan -2,00, sehingga perlu dilakukan proses transformasi data.



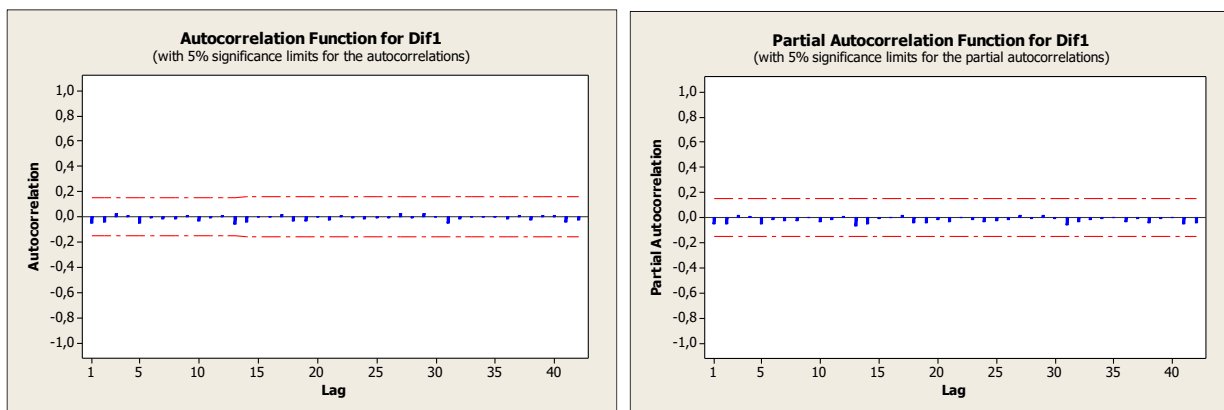
Gambar 4. Box-cox Plot Transformasi

Setelah melakukan transformasi data dengan menggunakan Box-Cox transformasi maka dihasilkan nilai *Rounded Value* sama dengan 1 itu artinya data IHK di Kota Ambon periode Januari 2007 sampai dengan Desember 2020 telah stasioner dalam varian, hal ini dapat dilihat pada Gambar 4. Kestasioneran data tidak hanya dilihat berdasarkan varian tapi juga dilihat berdasarkan rata-rata. Untuk mengetahui data IHK sudah stasioner terhadap rata-rata harus dilihat berdasarkan plot ACF dan PACF.



Gambar 5. Plot ACF dan PACF data transformasi

Pada Gambar 5 menunjukkan bahwa data transformasi IHK belum stasioner terhadap rata-rata, hal ini dapat dilihat pada plot ACF yaitu lag yang keluar dari garis stasioner lebih dari 3. Untuk membuat data menjadi stasioner terhadap rata-rata maka kita harus melakukan proses differencing.



Gambar 6. Plot ACF dan PACF hasil Differencing

Setelah melakukan proses differencing pada Gambar 6 dapat dilihat bahwa lag yang keluar dari garis stasioner pada plot ACF dan PACF tidak lebih dari 3, sehingga dapat dikatakan data IHK hasil differencing di kota Ambon pada periode Januari 2007 sampai Desember 2020 telah stasioner terhadap rata-rata.

3.1. Estimasi Parameter

Berdasarkan plot ACF dan PACF data yang telah stasioner dari hasil differencing maka dilakukan identifikasi model awal sehingga model awal yang diperoleh adalah ARIMA (0,1,1), (1,1,0) dan (1,1,1). Dari ketiga model tersebut akan ditentukan satu model terbaik yang akan dipakai sebagai untuk melakukan peramalan. Model terbaik dapat diketahui berdasarkan nilai p-Value dan nilai MSE dari tiap-tiap model. Hal ini dapat dilihat pada tabel 1.

Tabel 1. Pengujian Model Awal

Model ARIMA	P-Value			Nilai MSE	Parameter Signifikan
	AR	MA	Constant		
ARIMA (0.1.1)	-	0.322	0.650	31.79	Tidak Signifikan
ARIMA (1.1.0)	0.379	-	0.655	31.81	Tidak Signifikan
ARIMA (1.1.1)	0.000	0.000	0.074	30.78	Signifikan

Pada Tabel 3 di atas dapat dilihat bahwa model ARIMA (1,1,1) yang dipakai sebagai model awal merupakan model terbaik . hal ini dapat dilihat berdasarkan nilai *P-Value* yang lebih kecil dari 0.05 (α) yang artinya model awal ini signifikan dan merupakan model terbaik yang dapat digunakan dalam metode peramalan. Setelah model diperoleh, tahap selanjutnya yaitu mengestimasi parameter dalam model ARIMA (1,1,1).

Berdasarkan hasil output MINITAB diatas dapat dilihat bahwa nilai koefisien parameter untuk ARIMA (1,1,1) adalah

Tabel 2. Estimasi Parameter Model ARIMA (1,1,1)

Parameter	Koefisien	P-Value
AR ₁ (φ_1)	0,9000	0.000
MA ₁ (θ_1)	0,9933	0.000

Tabel 2, menunjukkan hasil estimasi parameter dalam model ARIMA (1,1,1) yaitu $\varphi_1 = 0,7880$ dan $\theta_1 = 0,9855$. Selanjutnya dilakukan uji signifikansi parameter tersebut dengan menggunakan nilai P-value.

1. Uji signifikansi parameter AR₁ (φ_1) yaitu = 0,9000

Hipotesis :

H0 : parameter AR(1) tidak signifikan dalam model

H1 : parameter AR(1) signifikan dalam model

Parameter AR (1) mempunyai nilai P-value sebesar 0.000, dengan level toleransi 5% berarti P-value < α yaitu 0.000 < 0.05. Sehingga dapat disimpulkan untuk menolak H0, yang berarti $\varphi_1 = 0,9000$ signifikan dalam model.

2. Uji signifikansi parameter MA₁ (θ_1) yaitu = 0,9933

Hipotesis :

H0: parameter MA(1) musiman tidak signifikan dalam model

H1: parameter MA(1) musiman signifikan dalam model

Parameter MA (1) musiman mempunyai nilai P-value sebesar 0.000, dengan level toleransi 5% berarti P-value < α yaitu 0.000 < 0.05. Sehingga dapat disimpulkan untuk menolak H0, yang berarti $\theta_1 = 0,9933$ signifikan dalam model.

Berdasarkan hasil yang diperoleh pada tahap estimasi parameter, maka parameter-parameter hasil estimasi yang signifikan dalam model ARIMA (1,1,1) yaitu $\varphi_1 = 0,7880$ dan $\theta_1 = 0,9855$. Berdasarkan persamaan (1) maka model ARIMA (1,1,1) sebagai berikut:

$$Z_t = (1 + \varphi_1)Z_{t-1} - \varphi_1 Z_{t-2} + \alpha_t + \theta_1 Y_{t-1}$$

$$Z_t = (1 + 0,9000)Z_{t-1} - 0,9000Z_{t-2} + \alpha_t + 0,9933Y_{t-1}$$

$$Z_t = 1,9000Z_{t-1} - 0,9000Z_{t-2} + \alpha_t + 0,9933Y_{t-1}$$

3.2 Diagnosis Model

1. Uji Independensi

Residual Uji dilakukan untuk mendeteksi independensi residual antar lag. Dua lag dikatakan tidak berkorelasi jika antar lag tidak ada korelasi yang cukup berarti.

a. Hipotesis

$H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0$ (Tidak ada korelasi antar lag/ independent)

$H_1 : \rho_i \neq 0$ (Ada korelasi antar lag/ dependent)

b. Daerah Penolakan

Statistik Ljung-Box-Pierce $> \chi^2_{(\alpha, df)}$ Dengan $df = K - k$. Dalam hal ini K berarti pada lag K dan k adalah jumlah parameter model. Daerah penolakan dalam bentuk grafik menggambarkan $\alpha = 5\%$ pada grafik statistik χ^2 untuk $df = 18 - 3 = 15$

c. Interpretasi Output Diagnosis Model

Deteksi independensi antar lag dilakukan pada tiap lag, sebagai berikut:

Tabel 3. Modified Box-Pierce (Ljung-Box) Chi-Square Statistik

Lag (K)	$df(K - k)$	$\chi^2_{(\alpha, df)}$	P -value
12	9	3.2	0.958
24	21	5.1	1.000
36	33	6.7	1.000
48	45	9.4	1.000

Dari Tabel 4, dapat disimpulkan bahwa Statistik Ljung-Box-Pierce dapat dilihat berdasarkan nilai p -value yaitu pada lag 12 sampai lag 48 semua nilai p -value $> \alpha = 0.05$. Sehingga dapat disimpulkan bahwa residual telah memenuhi asumsi *independent*.

2. Uji Kenormalan Residual

Uji kesesuaian model untuk membuktikan model sementara yang telah ditetapkan cukup memadai dengan menggunakan analisis galat untuk memenuhi asumsi kenormalan model. Uji kenormalan model dilakukan dengan uji *Kolmogorov Smirnov*.

Hipotesis:

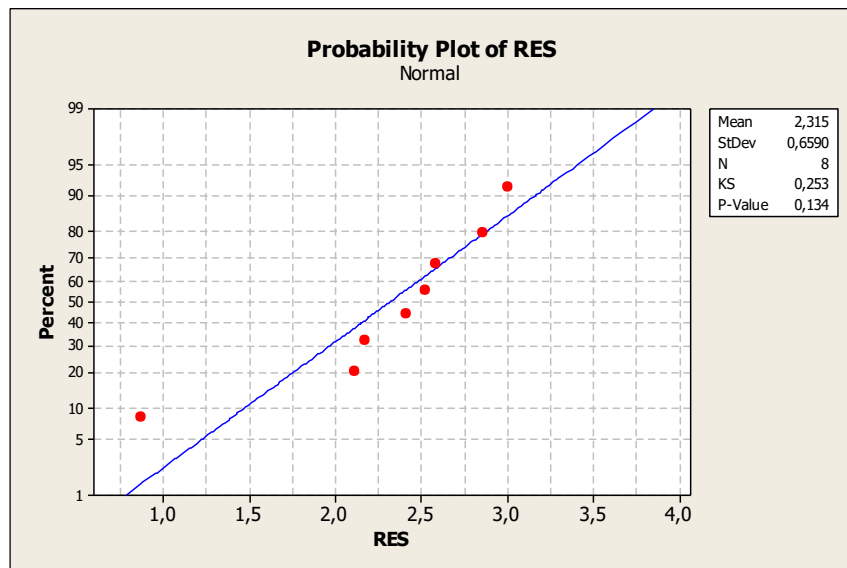
H_0 : sampel berasal dari populasi berdistribusi normal

H_1 : sampel tidak berada dari populasi berdistribusi normal

Uji normalitas dilakukan menggunakan *software* MINITAB 16.

Kriteria keputusan: tolak H_0 jika nilai signifikansi $< \alpha$.

Selain melakukan uji Kolmogorov Smirnov, dilakukan uji *white noise* untuk memenuhi asumsi tidak ada autokorelasi residual dengan menggunakan statistik uji Ljung Box.



Gambar 8. Plot Kolmogorov Smirnov

Berdasarkan Gambar 9, P-value = 0.134 > α (0.05) maka model ARIMA (1,1,1) memenuhi normalitas residual. Karena uji independensi sudah terbukti signifikan (tidak ada hubungan antar data satu dengan data yang lain), sehingga dapat melakukan peramalan dengan model ARIMA (1,1,1).

3.3. Peramalan

Hasil pemodelan data apabila telah signifikan dan memenuhi asumsi yang disyaratkan, berarti ias diandalkan. Model ARIMA (1,1,1) cukup memuaskan berarti ias dibuat peramalan data ke depan. Pada output sebelumnya menunjukkan hasil peramalan, sebagai berikut :

Tabel 5. Hasil Peramalan

Period	Forecast	Aktual
169	106.41	105.54
170	107.2	105.09
171	107.899	105.49
172	108.516	105.66
173	109.061	106.48
174	109.539	107.37
175	109.958	107.44
176	110.324	107.32

Output di atas menunjukkan periode peramalan dilakukan mulai periode 169 sampai periode 176. Ini berarti peramalan dilakukan untuk 8 bulan periode kedepan yaitu peramalan untuk tahun 2021. Maka kita dapat melihat IHK untuk 8 bulan depan yaitu bulan Januari 2021 sebesar 106.41 (lihat forecast periode 169) sampai bulan Agustus 2021 (lihat forecast periode 176). Pada tabel diatas juga dapat dilihat bahwa data hasil peramalan (forcast) dan data aktual mempunyai perbedaan nilai yang kecil hal ini berarti hasil peramalan yang dilakukan sangat baik karena data hasil peramalan tidak jauh dari data aktualnya.

4. KESIMPULAN

Berdasarkan hasil analisis pada pembahasan dapat diperoleh simpulan sebagai berikut:

1. Model ARIMA terbaik yang digunakan dalam melakukan peramalan IHK di kota Ambon periode januari 2007 hingga Desember 2020 adalah model ARIMA (1,1,1) yaitu $\varphi_1 = 0,9000$ dan $\theta_1 = 0,9933$. Berdasarkan persamaan (1) maka model ARIMA (1,1,1) adalah

$$Z_t = (1 + \varphi_1)Z_{t-1} - \varphi_1 Z_{t-2} + \alpha_t + \theta_1 Y_{t-1}$$

$$Z_t = (1 + 0,9000)Z_{t-1} - 0,9000Z_{t-2} + \alpha_t + 0,9933Y_{t-1}$$

$$Z_t = 1,9000Z_{t-1} - 0,9000Z_{t-2} + \alpha_t + 0,9933Y_{t-1}$$

2. Hasil peramalan (forcast) dan data aktual mempunyai perbedaan nilai yang kecil hal ini berarti hasil peramalan yang dilakukan sangat baik karena data hasil peramalan tidak jauh dari data aktualnya.

References

- [1] John, D. R., & Latupeirissa, S. J. (2020). "Peramalan Harga Emas Di Indonesia Tahun 2014-2019 Dengan Metode Arima Box-Jenkins", *VARIANCE: Journal of Statistics and Its Applications*, 2(2), 53-62.
- [2] Montgomery, D.C., et.al. (2008), "Forecasting Time Series Analysis 2nd. John Wiley
- [3] Pimpi, L. (2013), "Penerapan metode ARIMA dalam meramalkan indeks harga konsumen (IHK) Indonesia Tahun 2013". *Jurnal paradigma*, 17(2).
- [4] SARI, D. R. (2015), "Peramalan Indeks Harga Konsumen Menggunakan Model Intervensi Fungsi Step" (Doctoral dissertation, FSM Universitas Diponegoro).
- [5] https://id.wikipedia.org/wiki/Indeks_harga_konsumen

