Jurnal Matematika Vol. 17, No. 2, November 2018
ISSN: 1412-5056 / 2598-8980
http://ejournal.unisba.ac.id

Diterima: 21/08/2018
Disetujui: 10/11/2018
Publikasi Online: 29/11/2018

Model Peramalan Indeks Harga Konsumen Kota Palangka Raya Menggunakan *Seasonal* ARIMA (SARIMA)

Forecasting Model for Consumer Price Index of Palangka Raya City using Seasonal ARIMA (SARIMA)

Ananto Wibowo

Badan Pusat Statistik, Kotawaringin Barat

ananto.wibowo@bps.go.id

Abstrak. Palangka Raya merupakan salah kota indikator perhitungan Indeks Harga Konsumen (IHK) di Provinsi Kalimantan Tengah. Secara tidak langsung, persentase perubahan IHK, senantiasa dijaga oleh pemerintah agar tetap rendah dan stabil untuk prasyarat pertumbuhan ekonomi berkesinambungan serta mampu memberikan manfaat bagi peningkatan kesejahteraan masyarakat. Oleh karena itu, peramalan data IHK perlu dilakukan untuk membantu pemerintah dalam menyusun suatu kebijakan. Salah satu metode statistik yang paling tepat untuk melakukan peramalan data IHK Kota Palangka Raya dengan menggunakan *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Metode SARIMA sangat cocok untuk diterapkan pada data IHK karena terdapat pola musiman yang terjadi pada waktu tertentu. Data yang digunakan merupakan data sekunder bersumber dari Badan Pusat Statistik (BPS) Provinsi Kalimantan Tengah dengan *series* sebanyak 48 observasi. Berdasarkan hasil *trial and error*, model SARIMA pada saat $p = 1, d = 1, q = 0, P = 3, D = 1, Q = 2$, dan $S = 6$ dianggap sebagai model terbaik untuk peramalan data IHK dengan koefisien determinasi *adjusted* sebesar 84,67%. Selain itu, model peramalan memenuhi seluruh uji asumsi diagnostik dan sangat *powerful* mendekati data aktual.

Kata kunci: SARIMA, Peramalan, IHK

Abstract. Palangka Raya is one of the indicator city for Consumer Price Index (CPI) calculation in Central Kalimantan Province. The percentage of changes in CPI had always been indirectly maintained by government in order to remain low and stable as a preconditions for sustainable economic growth and to provide benefits for improving people's walfare. Therefore, the forecasting of CPI is required to assist the development of government's policy. One of the precise statistical method for forecasting Palangka Raya's CPI is Seasonal Autoregressive Integrated Moving Average (SARIMA). SARIMA is very suitable to be applied in Consumer Price Index's data because of seasonal patterns that occured in several times. The data used in this research are secondary data sourced from Statistics Indonesia of Central Kalimantan Office with 48 observations. Based on trial and error process, SARIMA when $p = 1, d = 1, q = 0, P = 3, D = 1, Q = 2,\ S = 6$ is considered the best statistical model to forecast CPI data with an adjusted coefficent of determination 84.67%. This forecasting model also fulfilled all diagnostic assumptions' tests and powerfully approached the actual data.

Keywords: SARIMA, Forcasting, CPI

1. Pendahuluan

Dalam perekonomian angka Indeks Harga Konsumen (IHK) memainkan peranan penting sebagai dasar perhitungan inflasi suatu wilayah seperti di Kota Palangka Raya. Secara tidak langsung, persentase perubahan IHK senantiasa dijaga oleh pemerintah agar tetap rendah dan stabil untuk prasyarat pertumbuhan ekonomi berkesinambungan serta mampu memberikan manfaat bagi peningkatan kesejahteraan masyarakat [3]. Hal ini juga sangat diperlukan untuk memperhitungkan biaya produksi dengan kenaikan tingkat harga seperti pada waktu yang lalu [8]. Dengan demikian, peramalan data IHK perlu dilakukan untuk membantu pemerintah dalam menyusun suatu kebijakan.

Terdapat beberapa metode peramalan yang digunakan seperti *Moving Average*, Metode *Winter*, *Exponensial Smoothing*, *Autoregressive Integrated Moving Average* (ARIMA), *Seasonal* ARIMA (SARIMA) dan sebagainya. ARIMA dipandang lebih populer dalam melakukan peramalan karena lebih fleksibel dan mampu mewakili banyak variasi data pada deret waktu tertentu [1]. Adapun SARIMA merupakan pengembangan model ARIMA yang memiliki efek musiman [6]. Pola data IHK Kota Palangka Raya yang dirilis setiap bulan sangat dimungkinkan terdapatnya efek musiman terutama pada hari raya lebaran. Oleh karena itu, penelitian ini memiliki tujuan untuk memodelkan data IHK Kota Palangka Raya menggunakan metode SARIMA.

2. Metodologi

2.1. Sumber Data

Data IHK yang digunakan dalam penelitian ini merupakan data sekunder bersumber dari Badan Pusat Statistik Provinsi Kalimantan Tengah dari Bulan Januari tahun 2014 hingga Desember 2017 sebanyak 48 observasi [2]. Pada periode tersebut, tahun 2012 merupakan tahun dasar perhitungan IHK $(2012 = 100)$ dan pemaparan hasil penelitian ini didukung oleh perangkat *Eviews* 7.0.

2.2. ARIMA

Metodologi ARIMA merupakan suatu pendekatan model analisis *time series* yang sangat akurat. Susunan ARIMA terdiri dari model *Autoregressive* (AR) dan *Moving Average* (MA) dimana *Integrated* (I) menunjukkan nilai dari orde *differencing* suatu data hingga dapat dianggap stasioner. Brockwell & Davis [4] menjelaskan bahwa ide dasar pada $AR(p)$ merupakan nilai X_t oleh fungsi masa lalu ke-p, X_{t-1}, $X_{t-2}, \dots, X_{t-p}$ dengan persamaan:

$$X_t = \mu + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \dots + \emptyset_p X_{t-p} + Z_t \quad (1)$$

untuk $\{Z_t\}$ adalah *white noise* yaitu $Z_t \sim N(0, \sigma^2)$ dan tidak berkorelasi dengan X_s untuk setiap $s < t$. Sedangkan model $MA(q)$ menjelaskan nilai X_t dari kesalahan (*error*) prediksi masa lalu dengan persamaan:

$$X_t = Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2} - \cdots - \theta_q Z_{t-q} \quad (2)$$

dengan $\theta_1, \dots, \theta_q$ adalah koefisien eror masa lalu.

Menurut Baumöhl & Lyócsa [5], suatu data dikatakan stasioner jika mean dan variansnya konstan dari waktu ke waktu yang dapat diuji dengan *Augmented Dicky Fuller*. Pada data *time series*, seringkali ditemukan data tidak stasioner pada orde ke nol $(q = 0)$. Sebagai contoh terdapat model $AR(1)$ dengan persamaan:

$$X_t = \emptyset_1 X_{t-1} + Z_t \quad (3)$$

maka data *time series* memiliki *unit root* jika $\emptyset_1 = 1$ dan stasioner jika $|\emptyset_1| < 1$. Apabila data tersebut masih belum stasioner (masih memiliki *unit root*), maka dilakukan *differencing* hingga stasioner pada orde ke-d. Dengan demikian model $ARIMA(p, d, q)$ merupakan kombinasi dari model $ARMA(p, q)$ dengan proses *differencing* hingga orde ke-d. Misalkan $W_t = X_t - X_{t-1}$ atau $W_t = (1 - B)X_t$ maka secara umum model $ARIMA(p, d, q)$ memiliki persamaan:

$$\emptyset_p(B)(1 - B)^d X_t = \theta_q(B) Z_t \quad (4)$$

dengan B merupakan efek dari perubahan waktu ke-t terhadap waktu ke-$t-1$ atau $BX_t = X_{t-1}$ dan $B^2 X_t = X_{t-2}$ dan seterusnya.

Untuk mendapatkan nilai p dan q pada model $ARIMA(p, d, q)$ dilakukan dengan melihat pola *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Nilai Autokorelasi didefinisikan sebagai korelasi antar suatu deret waktu dengan deret waktu itu sendiri pada selisih waktu (lag) $0, 1, 2$ periode atau lebih. Secara matematis ACF untuk lag ke $1, 2, 3, \dots, k$ dapat dinotasikan dengan:

$$r_k = \frac{\sum_{t=1}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2} \tag{5}$$

Adapun PACF r_{kk} digunakan untuk mengukur tingkat keeratan antara X_t dan X_{t-1} apabila pengaruh dari lag ke $1, 2, 3, ...$ dan seterusnya sampai $k-1$ dianggap terpisah [7]. Notasi matematisnya adalah:

$$r_{kk} = \frac{r_k - \sum_{j=1}^{n-k} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j}\, r_{jk}} \tag{6}$$

Model $ARIMA(p, d, q)$ yang terbaik dipilih dengan mempertimbangkan nilai ${R^2}_{adj}$ terbesar, signifikansi masing-masing parameter dan lolos semua uji diasnostik diantaranya kenormalan, nonautokorelasi, dan nonmultikolinieritas. ${R^2}_{adj}$ dihitung dengan formula:

$${R^2}_{adj} = 1 - \frac{\dfrac{\sum_{i=1}^{n}(x_i - \hat{x}_n)^2}{n-p-1}}{\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{7}$$

dengan p adalah jumlah parameter di dalam model.

2.3. SARIMA

SARIMA merupakan teknik time series berupa pengembangan dari ARIMA yang berisi efek musiman secara periodik dan terjadi pengulangan di setiap s observasi [6]. Model $SARIMA(p, d, q)(P, D, Q)_s$ dapat ditulis dengan persamaan:

$$\phi_P(B^s)\emptyset_p(B)(1 - B^s)^D(1 - B)^d X_t = \Theta_Q(B^s)\theta_q(B)Z_t \tag{8}$$

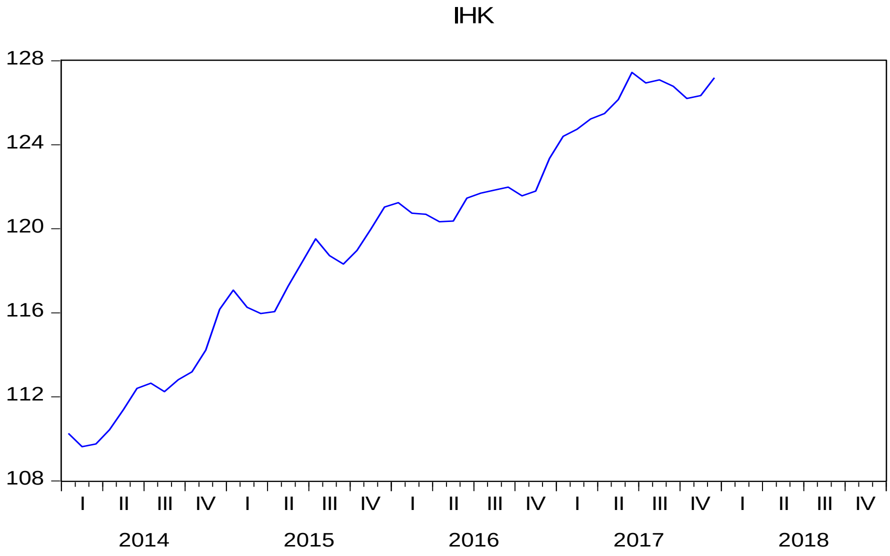
Dimana :

$\phi_P(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - \cdots - \phi_P B^{Ps}$
$\emptyset_p(B) \ = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p$
$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \cdots + \Theta_Q B^{Qs}$
$\theta_q(B) \ = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$

Seperti halnya ARIMA, pemodelan SARIMA juga harus dilakukan pada data stasioner dengan nilai d yang merupakan orde *differencing* untuk data nonmusiman dan D merupakan orde *differencing* untuk data musiman. Sedangkan penentuan nilai p, q, P dan Q dilakukan melalui pola ACF dan PACF [6].

3. Hasil dan Pembahasan

Sebelum pemodelan dilakukan, grafik data IHK perlu ditampilkan untuk melihat pola yang terjadi dan mengetahui keberadaan efek musiman. Beradasarkan Gambar 1, terlihat pengaruh musiman yang ada pada setiap semester dimana data IHK melonjak drastis dibandingkan bulan-bulan sebelumnya. Hal ini sangat dimungkinkan oleh naiknya harga-harga pada hari raya seperti Idul Fitri dan liburan Natal juga tahun baru. Oleh karena itu, pemodelan SARIMA sangatlah cocok digunakan untuk data IHK Kota Palangka Raya.



Gambar 1. Grafik Data IHK Kota Palangka Raya

Tampilan Gambar 1 secara jelas juga menyimpulkan bahwa data memiliki *trend* yang meningkat dari waktu ke waktu. Hal ini dapat dibuktikan dengan hasil uji stasioner IHK Kota Palangka Raya pada orde ke nol melalui uji *Augmented Dicky Fuller*. Berdasarkan hasil yang diperoleh pada Tabel 1, data masih memiliki *unit root* atau belum stasioner dimana nilai *p-value* sebesar 0,4045 yang masih lebih besar dari batas signifikansi 5%.

Tabel 1. Uji *Augmented Dicky Fuller* Data IHK Orde ke-0

		t-Statistik	P-Value
Uji Statistik ADF		-1.739944	0.4045
Nilai Kritis:	Orde 1%	-3.592462	
	Orde 5%	-2.931404	
	Orde 10%	-2.603944	

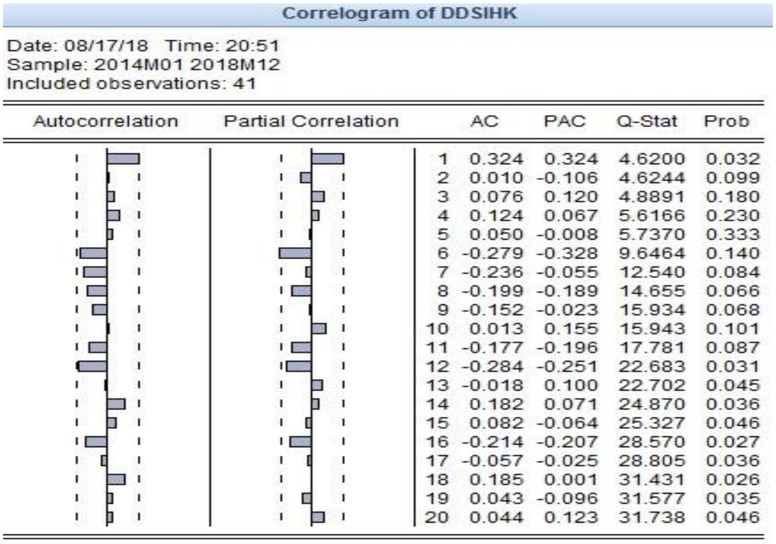
Karena data IHK orde ke-0 masih belum stasioner maka perlu dilakukan *differencing* pada orde pertama dengan pola musiman (semesteran atau Lag 6) dan nonmusiman serta dilakukan kembali uji *Augmented Dicky Fuller*. Hasil perhitungan yang ditunjukkan oleh Tabel 2 menyimpulkan bahwa data IHK telah stasioner di orde pertama baik taraf signifikansi 1%, 5% hingga 10%. Itu berarti rata-rata dan varians telah stasioner pada *differencing* pertama secara musiman dan nonmusiman sepanjang waktu observasi penelitian. Dengan demikian, model $ARIMA(p,d,q)(P,D,Q)_S$ yang sesuai untuk menggambarkan data IHK memiliki nilai $d = 1, D = 1$, dan $S = 6$.

 ISSN: 1412-5056 | 2598-8980

Tabel 2. Uji *Augmented Dicky Fuller* Data IHK Orde ke-1

		t-Statistik	P-Value
Uji Statistik ADF		-4,376440	0,0012
Nilai Kritis:	Orde 1%	-3,605593	
	Orde 5%	-2,936942	
	Orde 10%	-2,606857	

Tahapan berikutnya adalah menentukan nilai p dan q yang tepat melalui pola ACF dan PACF berdasarkan Gambar 2. Penentuan parameter model p dan q dilakukan dengan cara mencoba-coba (*trial and error*) untuk mendapatkan hasil yang terbaik. Dari pola yang terbentuk dapat diduga beberapa nilai diantaranya $p = 1, q = 1, P = 1, Q = 1, P = 2, Q = 2, P = 3$ sebagai dasar untuk kombinasi model terbaik. Nilai $p = 1$ dan $q = 1$ merupakan indikasi dari grafik ACF dan PACF pada bar pertama yang melewati garis putus-putus sedangkan nilai $P = 1, Q = 1, P = 2, Q = 2,$ dan $P = 3$ merupakan dugaan secara musiman yang memungkinkan kombinasi tersebut dapat mencapai model yang terbaik.


Correlogram of DDSIHK

Date: 08/17/18 Time: 20:51
Sample: 2014M01 2018M12
Included observations: 41

	AC	PAC	Q-Stat	Prob
1	0.324	0.324	4.6200	0.032
2	0.010	-0.106	4.6244	0.099
3	0.076	0.120	4.8891	0.180
4	0.124	0.067	5.6166	0.230
5	0.050	-0.008	5.7370	0.333
6	-0.279	-0.328	9.6464	0.140
7	-0.236	-0.055	12.540	0.084
8	-0.199	-0.189	14.655	0.066
9	-0.152	-0.023	15.934	0.068
10	0.013	0.155	15.943	0.101
11	-0.177	-0.196	17.781	0.087
12	-0.284	-0.251	22.683	0.031
13	-0.018	0.100	22.702	0.045
14	0.182	0.071	24.870	0.036
15	0.082	-0.064	25.327	0.046
16	-0.214	-0.207	28.570	0.027
17	-0.057	-0.025	28.805	0.036
18	0.185	0.001	31.431	0.026
19	0.043	-0.096	31.577	0.035
20	0.044	0.123	31.738	0.046

Gambar 2. Plot ACF dan PACF Data *Differencing* IHK Kota Palangka Raya

Kemungkinan nilai p, q, P dan Q untuk mencapai model SARIMA terbaik dengan kriteria $R^2{}_{adj}$ terbesar pada ditunjukkan oleh Tabel 3. Berdasarkan hasil yang diperoleh, peramalan data IHK Kota Palangka Raya yang paling tepat terjadi pada model $SARIMA(1,1,0)(3,1,2)_6$ karena nilai $R^2{}_{adj}$ merupakan yang tertinggi yakni sebesar 84,67 persen. $R^2{}_{adj}$ yang terbentuk sangatlah besar dan mampu melakukan peramalan dengan sangat akurat.

Tabel 3. Kombinasi SARIMA

Kombinasi Model	$R^2{}_{adj}$
$ARIMA(0,1,0)(1,1,1)_6$	0,3013
$ARIMA(0,1,1)(1,1,1)_6$	0,2787
$ARIMA(1,1,0)(1,1,1)_6$	0,2890
$ARIMA(1,1,1)(1,1,1)_6$	0,2745
.	.
.	.
.	.
$ARIMA(0,1,0)(3,1,2)_6$	0,8031
$ARIMA(0,1,1)(3,1,2)_6$	0,7857
$ARIMA(1,1,0)(3,1,2)_6$	0,8467

Secara matematis, model $SARIMA(1{,}1{,}0)(3{,}1{,}2)_6$ yang ditunjukkan oleh Tabel 4 dapat ditulis dengan persamaan berikut:

$$\begin{aligned}(1 + 0{,}7699B^6 + 0{,}610423B^{12} - 0{,}4611B^{18})(1 + 0{,}1551B)(1 - B^6)(1 - B)X_t \\ = -0{,}0024 + (1 - 0{,}8774B^6 + 0{,}9592B^{12})Z_t \qquad (9)\end{aligned}$$

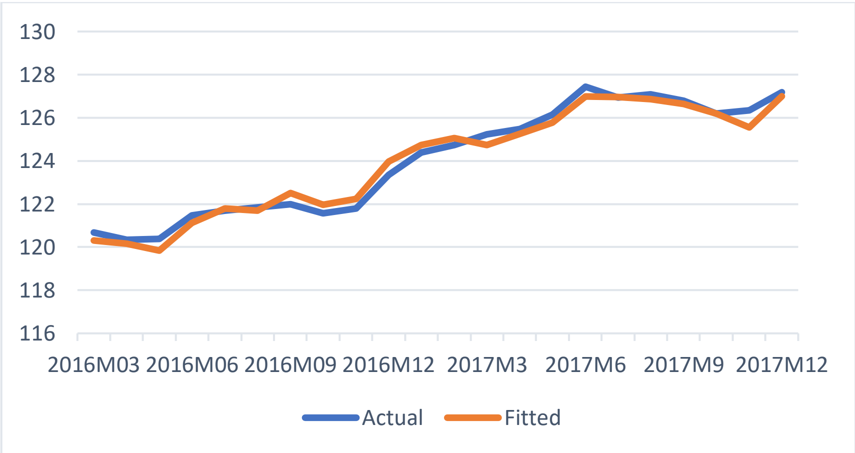
$$\begin{aligned}(1 - 0{,}8449B - 0{,}1551B^2 - 0{,}2301B^6 + 0{,}1944B^7 + 0{,}2745B^8 - 0{,}159477B^{12} \\ + 0{,}13477B^{13} - 0{,}2141B^{14} - 1{,}0715B^{18} + 0{,}9053B^{19} + 0{,}1662B^{20} \\ + 0{,}4611B^{24} - 0{,}5326B^{25} + 0{,}0715B^{26})X_t \\ = -0{,}0024 + (1 - 0{,}8774B^6 + 0{,}9592B^{12})Z_t \qquad (10)\end{aligned}$$

$$\begin{aligned}X_t = -0{,}0024 + 0{,}8449X_{t-1} + 0{,}1551X_{t-2} + 0{,}2301X_{t-6} - 0{,}1944X_{t-7} - 0{,}2745X_{t-8} \\ + 0{,}159477X_{t-12} - 0{,}13477X_{t-13} + 0{,}2141X_{t-14} + 1{,}0715X_{t-18} \\ - 0{,}9053X_{t-19} - 0{,}1662X_{t-20} - 0{,}4611X_{t-24} + 0{,}5326X_{t-25} \\ - 0{,}0715X_{t-26} - 0{,}8774Z_{t-6} + 0{,}9592Z_{t-12} + Z_t \qquad (11)\end{aligned}$$

Tabel 4. Model $SARIMA(1{,}1{,}0)(3{,}1{,}2)_6$ Berdasarkan Pengolahan *Eviews*

Variable	Coefficient	Std. Error
C	-0.002361	0.052755
AR(1)	-0.155130	0.239726
SAR(6)	-0.769887	0.343364
SAR(12)	-0.610423	0.176852
SAR(18)	0.461085	0.238145
MA(6)	-0.877439	0.265568
MA(12)	0.959262	0.025622

Model $SARIMA(1{,}1{,}0)(3{,}1{,}2)_6$ juga memenuhi semua uji asumsi diagnostik. Hasil uji kenormalan dengan $H_0: Z_t \sim N(0, \sigma^2)$ dan statistik uji *Jarque Berra* memberikan kesimpulan bahwa H_0 tidak ditolak atau *error* berdistribusi secara normal. Selain itu, uji autokorelasi menggunakan *Breusch-Godfrey Serial* dengan $H_0: E(Z_i, Z_j) = 0$ atau *error* nonautokorelasi menyimpulkan bahwa H_0 juga tidak ditolak. Adapun, nilai *Variance Inflation Factor* (VIF) menyatakan bahwa tidak ada multikolinieritas pada data $AR(1)$, $SAR(6)$, $SAR(12)$, $SAR(18)$, $MA(6)$, dan $MA(12)$ yang semua nilai VIF-nya kurang dari 5 (Lampiran 1).



Gambar 3. Perbandingan Data Ramalan IHK dengan Data Aktual

Perbandingan data ramalan IHK dengan data asli ditampilkan pada Gambar 3. Berdasarkan hasil yang diperoleh, data peramalan mendekati data aktual dan sangat *powerfull*. Dengan demikian, $SARIMA(0,1,0)(3,1,2)_6$ sangatlah cocok untuk model peramalan IHK Kota Palangka Raya.

4. Kesimpulan

Model $SARIMA(1,1,0)(3,1,2)_6$ merupakan model terbaik untuk peramalan data IHK Kota Palangka Raya dengan nilai $R^2{}_{adj}$ yang besar yakni 84,67% dan sangat *powerfull*. Selain itu, hasil pemodelan juga memenuhi seluruh uji asumsi diagnostik.

Referensi

[1] Ardhikari, A. & Agrawal, R. K., *An Introductory Study on Time Series Modeling and Forecasting*. CoRR, abs/1302.6613, 2013

[2] Badan Pusat Statistik, *Indeks Harga Konsumen Kota Palangka Raya*. [Alamat url diakses 2017]. https://kalteng.bps.go.id/statictable/2017/07/07/443/indeks-harga-konsumen-kota-palangka-raya-1996-2017.html.

[3] Bank Indonesia, *Pengenalan Inflasi*. [diakses 31 Juli 2018]. https://www.bi.go.id/id/moneter/inflasi/pengenalan/Contents/Pentingnya.aspx.

[4] Brockwell J. P. & Davis R. A., *Introdution to Time Series and Forecasting*. New York: Springer. 2002.

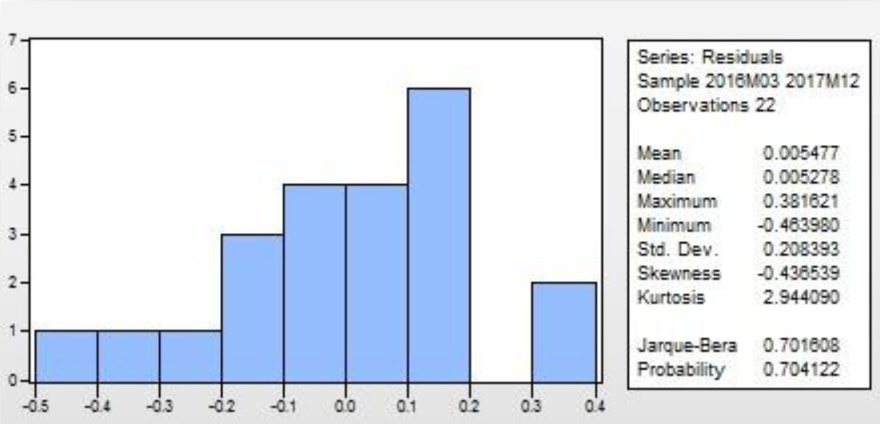
[5] Baumöhl E. & Lyócsa S., *Stationarity of Time Series and the Problem of Spurious Regression*. MPRA Paper 27926, University Library of Munich, Germany. 2009.

[6] Gikungu S. W., Waititu A. G., Kihoro J. M., *Forecasting inflation rate in Kenya using SARIMA model*. American Journal of Theoretical and Applied Statistics. 2015.

[7] Octavia T., Yulia, Lydia, *Peramalan Stok Barang untuk Membantu Pengambilan Keputusan Pembelian Barang Pada Toko Bangunan XYZ dengan Metode Arima*. Seminar Nasional Informatika, 2013.

[8] Suseno & Aisyah, S., *Inflasi*, Jakarta: Bank Indonesia, 2009.

Lampiran 1. Uji Kenormalan, Non Autokorelasi dan Multikolinieritas



Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.834236	Prob. F(2,13)	0.1988
Obs*R-squared	4.829441	Prob. Chi-Square(2)	0.0894

Variance Inflation Factors
Date: 08/21/18 Time: 16:09
Sample: 2014M01 2017M12
Included observations: 22

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.002783	4.308866	NA
AR(1)	0.057469	1.723850	1.484870
SAR(6)	0.117899	4.755130	1.329144
SAR(12)	0.031277	1.347079	1.341305
SAR(18)	0.056713	3.379705	2.385544
MA(6)	0.070526	1.499545	1.486024
MA(12)	0.000656	1.448130	1.404162