International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2021; 6(5): 08-11

www.mathsjournal.com
Received: 04-06-2021
Accepted: 11-07-2021

David Peter Oaya
Department of Statistics, Nasarawa State University, Keffi, Nasarawa, Nigeria

Nweze Obini Nwanze
Department of Statistics, Nasarawa State University, Keffi, Nasarawa, Nigeria

Abubakar Muhammad Auwal
Department of Statistics, Nasarawa State University, Keffi, Nasarawa, Nigeria

Corresponding Author:
David Peter Oaya
Department of Statistics, Nasarawa State University, Keffi, Nasarawa, Nigeria

Comparing SARIMA model results of a bootstrap resampled and actual time series inflation data (2014-2018)

David Peter Oaya, Nweze Obini Nwanze and Abubakar Muhammad Auwal

Abstract
The researched on the efficiency of establishing of variance obtained from Seasonal Autoregressive Integrated Moving Averages (SARIMA) model for resampled and actual inflation time series data. The study used the monthly inflation data of 2014 to 2018. It conducting stationarity test on the time series data using Augmented Dickey-Fuller (ADF) unit roots test in order to avoid wrong forecast. Fitting a model on resampled and non-resampled time series data was important to get accurate forecasts and predictions, as well as, check the precision between the variance of these models and outcomes. Using the Akaike Information Criterion (AIC), most efficient forecast was obtained. The Seasonal Autoregressive integrated moving averages (SARIMA) was applied to both the stationary data for bootstrap and non-bootstrap samples. Results show that the models fit on resampled inflation series produced higher AIC value of 689.05 of SARIMA (3,0,1) (3,1,0)12 compared to the actual inflation data having an AIC value of 269.9 of SARIMA (2,0,0) (2,0,0)12. The findings reveal that the SARIMA model of the actual inflation data was more perfect for forecasting compared to that of the resampled data. The inflation data became stationary at the sixth difference while that of resampled data became stationary at second difference. The study recommends that it is ideal and better to use actual data of seasonal time series event for forecasting.

Keywords: Bootstrapping, inflation, resampled, stationarity, SARIMA

Introduction

One of the major challenges faced by the developing economy of unindustrialized as well as developing countries has been the rise and fall of the stock market and inflation generally. Inflation is said to be the continual increase in the prices of goods and services that has resulted to increase in the volume of money in circulation in the economy of a country as well as affecting the exchange rate of few goods and services. The instability of the foreign exchange rate (international exchange rate), increment in the price of crude oil just after the removal of oil subsidy in conjunction with the skyrocketed prices of foreign goods had resulted to an increase or high price of importation of goods into the country. The bite of inflation that was witnessed by the undulating rise and fall of stock market and crude oil has grossly affected the economic indices and policies. Most of these measures according to Wiri and Isaac (2018) [12] taken by developing countries to check the problem of inflation are in the form of the use of central bank instruments of credit control. That these measures were to reduce to the barest minimal the volume of money in circulation and its sustenance of ensuring low cost of living (Wiri and Isaac, 2018) [12]. Model selection is an important aspect of statistics because it allows researchers to test the relative quality of one model over another. In this research, the Seasonal Auto Regressive Integrated Moving Averages (SARIMA) model will not be compared to a different model but to the same of its kind with the difference in the type of data used to fit in the model. The study will focus on resampled inflation and non-resampled inflation time series secondary data obtained from the Central Bank of Nigeria Statistical Bulletin 2018 publication, its uses by applying the selection method of Akaike Information

Criterion specifically, considering the period of 2014 – 2018 on resampled and actual time series data for better forecast.

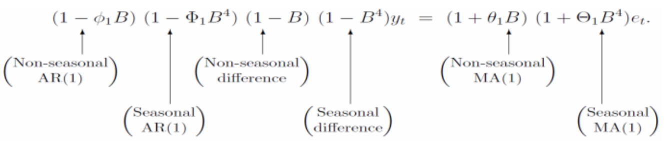
Materials and Methods

Data collection

The data used for this work is monthly data on resampled and actual inflation rate of the years ranging from 2014 to 2018. These data were obtained from Central Bank of Nigeria Statistical Bulletin 2018. The Bootstrap resampling method was used to get resampled inflation time series. ADF was used to make the data stationary for forecasting. SARIMA model was used on the resampled and actual data for forecasting future data points. The R-console statistical package (Version 4.0.3) was used for the analysis.

Method of data analysis

Model of the study



The SARIMA Process

The first order SARMA model with the missing capital "I" is given as

$$x_t = \Phi x_{t-12} + w_t + \Theta w_{t-12}.$$
$$(1 - \Phi B^{12})x_t = (1 + \Theta B^{12})w_t$$

This implies that the model exhibits autocorrelation at past lags of multiple of 12 (which would be defined as the seasonal period) for both auto regression and moving average components. Thus, this model would be of SARMA (P, Q), where P and Q =1.

Multiplicative SARMA model

Combining both the seasonal and non-seasonal ARMA models would be simply expressed as ARMA (p,q) x (P,Q)s, where P,Q,s represent the components and seasonal operators related to parameters of the seasonal model.

$$\Phi_P(B^s)\phi(B)x_t = \Theta_Q(B^s)\theta(B)w_t$$

Seasonal Differencing

Some time series may depict ACF that decays slowly at the lags of the multiples of the seasonal period. The average monthly temperature can be modelled as the following:

$$x_t = S_t + \omega_t$$
$$S_t = S_{t-12} + v_t$$

S_t represents the seasonal component that varies from one year to the next, and represents a random walk expression. Note that w_t and v_t are uncorrelated white noise processes. Applying a seasonal difference to the time series expression will result in a MA (1) with an ACF peak at lag = 12.

$$(1 - B^{12})x_t = x_t - x_{t-12} = v_t + w_t - w_{t-12}$$

Thus, the multiplicative SARIMA model can be expressed in an expression similar to the ARIMA model as shown in the following:

$$\Phi_P(B^s)\phi(B)\nabla_s^D\nabla^d x_t = \delta + \Theta_Q(B^s)\theta(B)w_t$$

Results and Discussion

Using the r studio, we plotted the Auto Correlation Function (ACF) of the inflation rates with five lags to show the seasonality in the data. The result is shown below.

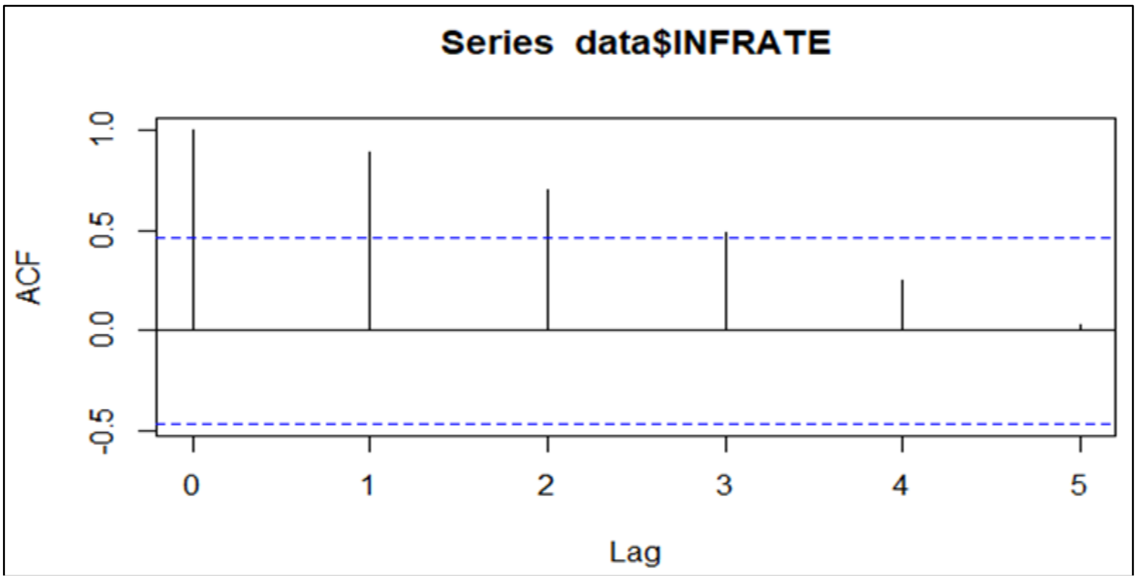


Fig 1: A Graph showing the ACF of Inflation Rate

Testing for Stationarity using the Augmented Dickey–Fuller (ADF), the test below shows the p-values at 1st and 2nd differencing. However, by practice if the p-value of an ADF test is less than 0.05 then the data is stationary at that point. So, we shall take the second difference with value p-value = 0.01 of our time series data for our analysis since from the result of the first differencing ADF test in R-console software, the p-value was 0.5836 greater than 0.05 meaning that the data is not stationary. Regarding the SARIMA Model Estimation, using the R studio, SARIMA model was fit on the stationary data. The SARIMA code at different values for the (p, d, q) (P, D, Q)m was used in R-console, The model that returns the best AIC (Akaike Information Criterion) with the least value was considered best model for forecasting.

Table 1: A Table showing AICs of different SARIMA models

Model	AIC
SARIMA (1,0,1) (1,0,1) 12	286.65
SARIMA (1,1,1) (1,1,1) 12	279.42
SARIMA (1,0,0) (1,0,0) 12	290.11
SARIMA (2,0,1) (2,1,0) 12	271.57
SARIMA (3,0,1) (3,1,0) 12	273.21
SARIMA (2,1,1) (2,1,1) 12	272.5
SARIMA (2,0,0) (2,0,0) 12	269.6

Consequently, this implies that the best model for these data is a SARIMA (2,0,0) (2,0,0) 12 with AIC of 269.9.

Bootstrap Resampling for Sarima Modeling
After obtaining the SARIMA model and forecasts of the actual inflation data, we resampled the data and carried out similar steps on the data in order to test it stationarity. The resampled data was subjected to differencing till it became stationary and fit to carry out the SARIMA model analysis. The AICs at different orders of (p,d,q) (P,D,Q) m were noted and compared with the AIC of the actual time series data.

Resampling the Inflation Data
The resampling was done in excel with the command
= Index (Sample 1, Rand between (1, Rows (sample 1), 1) where index is the column with the actual time series data. The result of the resampling was obtained in the next column as shown in the figure below;

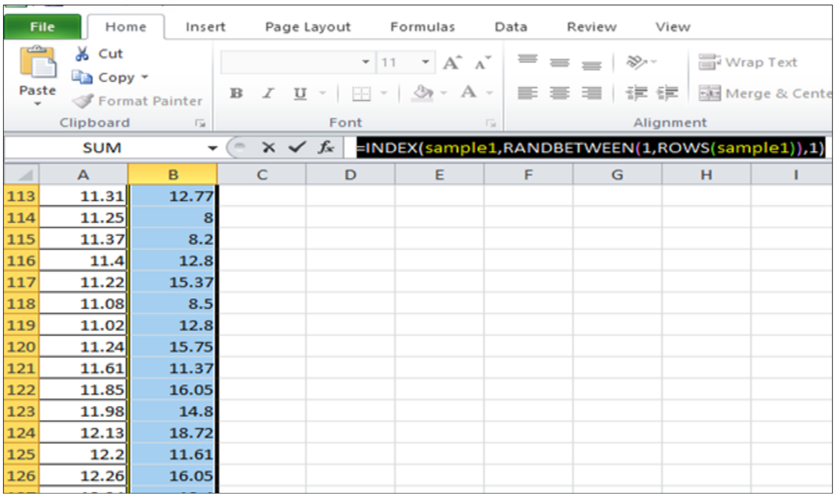


Fig 2: An excel format of resampling

Test for Stationarity using the ADF of resampled data: The Augmented Dickey–Fuller Test was used to make the data stationary at the first differencing. The SARIMA model was applied on the differenced resampled time series data.

Table 2: A Table showing AICs of different SARIMA models of resampled data

Model	AIC
SARIMA (1,0,1) (1,0,1) 12	689.05
SARIMA (1,1,1) (1,1,1) 12	748.47
SARIMA (1,0,0) (1,0,0) 12	747.41
SARIMA (2,0,1) (2,1,0) 12	689.13
SARIMA (3,0,1) (3,1,0) 12	690.1
SARIMA (2,1,1) (2,1,1) 12	722.07
SARIMA (2,0,0) (2,0,0) 12	720.06

From the above table, SARIMA (1,0,1) (1,0,1) 12 returned a lower AIC of 689.05 less than the rest. However, this result is still higher than the lowest AIC obtain from the SARIMA model using the actual time series data.

Conclusion
This study examined that the inflation data became stationary at the second difference while that of resampled data became stationary at first difference. The Seasonal Autoregressive integrated moving averages (SARIMA) was applied to both the stationary data for bootstrap and non-bootstrap samples, our finding shows that the SARIMA model of the actual inflation data is more perfect for forecasting. The values of the AICs obtained from SARIMA of different specification revealed that actual data value of AIC is less than those from resampled inflation data. Therefore, we conclude that the forecast from the actual data is adopted for this study and also recommend the use of actual data for forecasting time series.

Reference
1. Adnan E, Haitham AA, Mohammed FA, Ahmed E. Heavy metal monitoring analysis and prediction in lakes and rivers: state of the art. Environmental Science and pollution Research 2017;24:12104-12117.
2. Barzegar R, Moghaddam AA, Adamowski J, Ozga-zielinski B. Multi-step water quality forecasting using a

boosting ensemble multi-wavelet extreme learning machine model. Stochastic Environmental Research and Risk Assessment 2017;32(3):799-813.
3. Belayneh A, Adamowski J. Drought forecasting using new machine learning methods/Prognozowanie suszy z wykorzystaniem automatycznych samouczących się metod. Journal of Water and Land Development 2013;18:3-12.
4. Box GEP, Jenkin GM. Time series analysis, forecasting and control. Holden day, San Francisco 1976.
5. Certral Bank of Nigeria. Statistical bulletin, http://www.cenbank.org 2018.
6. Etuk EH, Azubuike SA, Bartholomew AU. A forecasting model for monthly Nigeria Treasury Bill rate by box-Jenkins techniques, from January 2006–December 2014. International Journal of life sciences and Engineering 2015;1(1):20-25
7. Fisher RA. Design of Experiments (1st Edition). Edinburgh: Oliver and Boyd 1935.
8. Gautam R, Sinha AK. Time series analysis of reference crop evapotranspiration for Bokaro District, Jharkhand, India. Journal of Water and Land Development 2016;30:51-56.
9. Gerretsadikan A, Sharma MK. Modelling and forecasting of rainfall data of Mekele for Tigray region (Ethiopia). Statistics and Applications 2011;9(1–2):31-53.
10. Osarumwense OI. Applicability of Box Jenkins SARIMA model in rainfall forecasting: A case study of Port-Harcourt south Nigeria. Canadian Journal on Computing in Mathematics, Natural Sciences, Engineering and Medicine 2013;4(1):1-4.
11. Tiwari M, Adamowski J, Adamowski K. Water demand forecasting using extreme learning machines. Journal of Water and Land Development 2016;28:37-52.
12. Wiri, Leneenadogo & Isaac, Didi Essi. Seasonal Autoregressive Integrated Moving Average (SARIMA) Modelling and Forecasting of Inflation Rates in Nigerian (2003- 2016). International Journal of Applied Science and Mathematical Theory 2018;4(1):1-14.
13. Yadav B, CH S, Mathur S, Adamowski J. Assessing the suitability of extreme learning machines (ELM) for groundwater level prediction. Journal of Water and Land Development 2017;32:103-112.
14. Yang T, Asanjan AA, Welles E, Gao X, Sorooshian S, Liu X. Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information. Water Resources Research 2017;53(4):2786-2812.