International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No.01, February 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

# A Study on Whole Sale Price Index based Inflation rate in India

**J.Krishnan[1], R.Vijayaragavan[2], K.M.Sakthivel [3]**

[1]*Research Scholar, Department of Statistics, Bharathiar University, Coimbatore, Tamil Nadu*

[2]*Professor, Department of Statistics, Bharathiar University, Coimbatore, Tamil Nadu*

[3]*Assistant Professor, Department of Statistics, Bharathiar University, Coimbatore, Tamil Nadu*

**ABSTRACT**

*Inflation affects many developing nations; India is also a nation which is affected by Inflation. As food prices have led much of the inflation, it causes social instability as the poor are impacted more by price increases of necessary commodities. Inflation is a very unpopular happening in an economy. An important social cost of inflation, especially in developing countries is its bad effect on long-run economic growth. Opinion surveys conducted in India, The U.S.A. and other countries reveal that inflation is the most important concern of the people as it badly affects their standard of living. There are many methods used to model and forecast time series data such as trend, regression, moving average etc. In this paper, ARIMA method is considered to predict inflation for upcoming years. ARIMA techniques are used to analyze time series data and have been mainly used for loading forecast due to its accuracy, mathematical soundness and flexible due to inclusion of AR and MA terms over a regression analysis. A detailed explanation of the above is presented and summarized in tables and figures using MINITAB software.*

***Keywords: Time series, Inflation, Forecast, AR, MA, ARIMA.***

## I. INTRODUCTION

In economics, inflation is a persistent increase in the general price level of goods and services over a period of time. Each unit of currency purchases smaller quantity of goods and services when the general price level rises. Therefore, inflation returns a reduction in the purchasing power per unit of money – a loss of real value in the medium of exchange and unit of account within the economy. The most important measure of price inflation is the inflation rate, the annualized percentage change in a general price index (normally the consumer price index) over time. Inflation influences an economy in a variety of ways, both positively and negatively. Generally economists believe that high rates of inflation and hyperinflation are caused by an extreme growth of the money supply.

### 1. Definition of Inflation

Inflation is the percentage change in the value of the Wholesale Price Index (WPI) on a year basis. It effectively measures the change in the prices of a basket of goods and services in a year.

$$Inflation = \frac{WPI\ of\ end\ of\ year - WPI\ of\ beginning\ of\ year}{WPI\ of\ beginning\ of\ year\ X\ 100}$$

### 2. Time Series Analysis

Time series is a series of data points indexed or graphed in the order of time. It is often plotted using line charts. Time series are used in statistics, signal processing, pattern recognition, econometrics, finance, forecasting of weather, earthquake prediction, control engineering, astronomy, and mostly in any area of applied science and engineering. Time series analysis consists of techniques for analyzing time series data in order to dig out meaningful characteristics of the data. Time series analysis can be useful to real-valued, continuous data, discrete numeric data, or discrete symbolic data. Time series forecasting is the technique used in a model to forecast future values based on previously observed values.

There are many different notations used for time-series analysis. A common notation specifying a time series $X$ that is indexed by the natural numbers is written as

$$X = \{X_1, X_2, ...\}.$$

Another common notation is

$$Y = \{Y_t: t \in T\},$$

where $T$ is the index set.

## II. ARIMA MODEL

An autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model in time series analysis. These models are fitted to time series data for understanding the data better or to forecast future points in the series. ARIMA models are applied to the data that shows evidence of non-stationarity, where an initial differencing step can be applied one or more times to remove the non-stationarity. The AR part of ARIMA shows that the variable of interest is regressed on its own lagged values. The MA part shows that the regression error is actually a linear combination of error terms whose values occurred contemporaneously and at various times in the past. ARIMA models can be estimated following the Box–Jenkins approach.

### 1. Autocorrelation

Autocorrelation is the correlation between a variable lagged one or more periods and itself. Autocorrelation coefficients for different time lags of a variable are used to identify time series data patterns. The formula for computing the lag $k$ autocorrelation coefficient between $Y_t$ and $Y_{t-k}$, which are k periods apart, is given by

$$r(k) = \frac{\sum_{t:k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t:1}^{n} (Y_t - \bar{Y})^2}, \quad for \;\; k = 0, 1, 2, \ldots$$

where $\bar{Y}$ the mean value of the values of the series is, $Y_t$ is the observation in time period $t$ and $Y_{t-k}$ is the observation $k$ time periods earlier or at time period $t$-$k$.

### 2. Partial Autocorrelation

A partial autocorrelation at time lag $k$ is the correlation between $Y_t$ and $Y_{t-k}$, after adjusting for the effects of the intervening values $Y_{t-1}, Y_{t-2,}, \ldots, Y_{t-k+1}$. In time series analysis, the partial autocorrelation function (PACF) plays an important role in data analyses aimed at identifying the extent of the lag in an autoregressive model. The use of this function was introduced as part of the Box-Jenkins approach to time series modeling, where by plotting the partial auto correlative functions one could determine the appropriate lag $p$ in an $AR(p)$ model or in an extended *ARIMA* ($p$, $d$, $q$) model.

### 3. Models

Box Jenkins Model assumes that a time series is a linear function of past actual values and error terms. That is,

$$Y_t = b_t + \varepsilon_t$$

The error terms are distributed as normally and independently distributed, having no pattern, with a mean of zero and an error variance that is lower than the variance of $Y_t$. With these assumptions the Box-Jenkins models are classified as Auto Regressive (AR), Moving Average (MA), or a combination of the two or Auto Regressive Integrated Moving Averages (ARIMA) models. The standard notation $p$ identifies the order of autoregressive, $d$ identifies integration or differencing and $q$ identifies the moving averages.

### 4. Auto Regressive model

It is one in which the current value of the variable is a function of previous values and an error term. The reason that this is called an auto regressive model is because $Y_t$ is being regressed on itself as

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \ldots + \Phi_p Y_{t-p} + \varepsilon_t,$$

where $Y_t$ is the dependent variable, $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$ are the independent variables based on the dependent variable lagged ($p$) specific time periods, $\Phi_0, \Phi_1, \ldots, \Phi_p$ are the computed regression coefficients and $\varepsilon_t$ is the random error term measured in time $t$.

### 5. Moving Averages model

Link the current values of the time series to random errors that have occurred in previous time periods. A moving average model is as follows

$$Y_t = \theta_0 - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t,$$

where $Y_t$ is the dependent variable, $\theta_0$ is the mean about which the series fluctuates, $\theta_0, \theta_1, \theta_2 \ldots, \theta_q$ are the moving average parameters to be estimated, $\varepsilon_{t-1}, \varepsilon_{t-2}, .. \varepsilon_{t-q}$ are the error terms and $\varepsilon_t$ is the random error term measured in time $t$. The highest order of the model is called $q$ and refers to the number of lagged time periods of the model.

### 6. ARIMA model

It is the combination of the AR and MA models. Thus, the model is given as

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \ldots + \Phi_p Y_{t-p} - \theta_0 - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

When using the ARIMA model, it is able to use a combination of past values and past errors. The order of the model is commonly written as *(p, d, q)*. To select an appropriate model for forecasting, one would depend on the autocorrelation (*AC*) and partial autocorrelation (*PAC*) statistics of time series.

## III. ANALYSIS OF TIME SERIES DATA ON AVERAGE INFLATION

This section presents the analysis of time series data on Inflation data. Appropriate time series models for making forecasts for future period are constructed using Box –Jenkins Methodology.

In order to forecast the average inflation for future periods the time series analysis is carried out for the given data. The scatter plot of the time series data against time periods is given in fig 1. From this figure, it is observed that there is downward trend in the observed time series data. Hence, the observed trend has to be removed. In order to verify whether the data is stationary time series plot has plotted and it is shown in the fig 2. It is observed that the time series data is non stationary and did not vary about a fixed level. Hence, it is necessary to transform the non-stationary data into stationary. For eliminating the trend and create a stationary series, the first differences are found and are presented in table 2. The first difference series is also plotted and is displayed in fig 3. The plot of the first differences against the lagged time variables indicates the stationary.

The values of autocorrelation and partial autocorrelation functions are given in table 3 and the plots of those values are displayed in fig 4 and 5 respectively. The fig 4 indicates that the auto-correlations appear to cut off after lag 1 which indicates a MA (1) behavior. At the same time fig 5 reveals that the partial autocorrelations have two spikes by which it is inferred that an auto-regressive of order one is appropriate for the time series data. As differencing is done one times for converting the non-stationary data into stationary data it is now possible to define an auto-regressive integrated moving average (ARIMA) model as an appropriate model in studying the given time series data. As the moving average model of order one and the autoregressive model of order two are identified on the basis of the pattern exhibited by ACs and PACs with the first order differencing, the ARIMA model is defined with reference to the parameters (*p, d, q*) = (2, 1, 1). This model can be represented as

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} - \theta_0 - \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

For the given time series data, the ARIMA (2, 1, 1) model is fitted using Minitab statistical software (Version 16). The model parameters are estimated as

$\Phi_0 - \theta_0 = 0.5287,\ \ \Phi_1 = -1.1163\ \Phi_2 = -0.1746$ and $\theta_1 = -1.0181.$ Thus, the fitted time series model for average inflation is given by

$$Y_t = 0.5287 - 1.1163 Y_{t-1} - 0.1746 Y_{t-2} + 1.0181 \varepsilon_{t-1} + \varepsilon_t$$

The estimates of the model parameters are provided in table 4 along with their standard errors. It can be observed from the results that the estimate of MA1 model is highly significant that the P-values corresponding to such estimates are equal to zero. From these observations of times series analysis it is inferred that the constructed ARIMA model is adequate. The forecasts for the period from 2015 – 2019 are generated based on the model and are displayed in table 6. The fig 6 displays residual plots of residuals. The straight line like behaviour of the points in the normal probability plot suggests a good fit between the data and the normal distribution. When the residuals are plotted against the fitted values, it is observed that the error variability is constant and the underlying relationship $Y_t$ and $t$ appears to be linear. The fig 7 shows that the time series plot of average inflation against various time points along with the forecasted values are also plotted and it is observed from the plot that the forecast values are quite similar to the actual values.

International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No.01, February 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

## IV. FIGURES AND TABLES

Table 1: Average inflation data during 1970-2014

| Year | Average inflation | Year | Average inflation | Year | Average inflation | Year | Average inflation | Year | Average inflation |
|---|---|---|---|---|---|---|---|---|---|
| 1970 | 5.09 | 1979 | 6.20 | 1988 | 9.40 | 1997 | 7.24 | 2006 | 6.01 |
| 1971 | 3.07 | 1980 | 11.39 | 1989 | 7.12 | 1998 | 13.16 | 2007 | 4.91 |
| 1972 | 6.44 | 1981 | 13.10 | 1990 | 8.95 | 1999 | 4.75 | 2008 | 8.66 |
| 1973 | 16.77 | 1982 | 7.92 | 1991 | 13.88 | 2000 | 3.92 | 2009 | 2.39 |
| 1974 | 28.52 | 1983 | 11.82 | 1992 | 11.88 | 2001 | 3.75 | 2010 | 9.57 |
| 1975 | 6.60 | 1984 | 8.43 | 1993 | 6.31 | 2002 | 4.22 | 2011 | 9.47 |
| 1976 | -7.57 | 1985 | 5.55 | 1994 | 10.24 | 2003 | 3.64 | 2012 | 7.46 |
| 1977 | 8.32 | 1986 | 8.71 | 1995 | 10.21 | 2004 | 3.97 | 2013 | 6.19 |
| 1978 | 2.53 | 1987 | 8.78 | 1996 | 8.97 | 2005 | 3.97 | 2014 | 4.18 |

Table 2: First Order Differences for Checking Stationary in Average inflation data

| Year | Average inflation | Fod | Year | Average inflation | Fod | Year | Average inflation | Fod |
|---|---|---|---|---|---|---|---|---|
| 1970 | 5.09 | * | 1985 | 5.55 | -2.88 | 2000 | 3.92 | -0.83 |
| 1971 | 3.07 | -2.02 | 1986 | 8.71 | 3.16 | 2001 | 3.75 | -0.17 |
| 1972 | 6.44 | 3.37 | 1987 | 8.78 | 0.07 | 2002 | 4.22 | 0.47 |
| 1973 | 16.77 | 10.33 | 1988 | 9.4 | 0.62 | 2003 | 3.64 | -0.58 |
| 1974 | 28.52 | 11.75 | 1989 | 7.12 | -2.28 | 2004 | 3.97 | 0.33 |
| 1975 | 6.6 | -21.92 | 1990 | 8.95 | 1.83 | 2005 | 3.97 | 0 |
| 1976 | -7.57 | -14.17 | 1991 | 13.88 | 4.93 | 2006 | 6.01 | 2.04 |
| 1977 | 8.32 | 15.89 | 1992 | 11.88 | -2 | 2007 | 4.91 | -1.1 |
| 1978 | 2.53 | -5.79 | 1993 | 6.31 | -5.57 | 2008 | 8.66 | 3.75 |
| 1979 | 6.2 | 3.67 | 1994 | 10.24 | 3.93 | 2009 | 2.39 | -6.27 |
| 1980 | 11.39 | 5.19 | 1995 | 10.21 | -0.03 | 2010 | 9.57 | 7.18 |
| 1981 | 13.1 | 1.71 | 1996 | 8.97 | -1.24 | 2011 | 9.47 | -0.1 |
| 1982 | 7.92 | -5.18 | 1997 | 7.24 | -1.73 | 2012 | 7.46 | -2.01 |
| 1983 | 11.82 | 3.9 | 1998 | 13.16 | 5.92 | 2013 | 6.19 | -1.27 |
| 1984 | 8.43 | -3.39 | 1999 | 4.75 | -8.41 | 2014 | 4.18 | -2.01 |

Table 3: Values of ACF and PACF Coefficients along With the Values of T-statistics

| Lag | ACF | T-Statistics | LBQ | PACF | T-Statistics |
|---|---|---|---|---|---|
| 01 | -0.17384 | -1.153098482 | 1.422401419 | -0.17384 | -1.153098482 |
| 02 | -0.38995 | -2.511844049 | 8.750295272 | -0.43326 | -2.873933571 |

| | | | | | |
|---|---|---|---|---|---|
| 03 | 0.065635 | 0.372707728 | 8.962963903 | -0.13563 | -0.899645261 |
| 04 | 0.029035 | 0.164358216 | 9.005622533 | -0.20234 | -1.342170719 |
| 05 | -0.15845 | -0.896384608 | 10.30861283 | -0.30234 | -2.005507457 |
| 06 | 0.092711 | 0.515155421 | 10.76642665 | -0.14684 | -0.974029882 |
| 07 | 0.089147 | 0.492389826 | 11.20115692 | -0.15694 | -1.041002299 |
| 08 | -0.05757 | -0.316215142 | 11.38746264 | -0.14535 | -0.96412328 |
| 09 | 0.026882 | 0.147333805 | 11.42925238 | -0.06776 | -0.449437223 |
| 10 | 0.040861 | 0.223837121 | 11.52864323 | -0.04786 | -0.317447684 |
| 11 | -0.07519 | -0.411410867 | 11.8753694 | -0.06787 | -0.450217194 |

Table 4: Model Parameters

| Type | Coef | SE Coef | T | P |
|---|---|---|---|---|
| AR1 | -1.1163 | 0.1587 | -7.03 | 0 |
| AR2 | -0.1746 | 0.1625 | -1.07 | 0.289 |
| MA1 | -1.0181 | 0.0151 | -67.36 | 0 |
| Constant | 0.5287 | 0.8639 | 0.61 | 0.544 |

Table 5: Modified Box-Pierce (Ljung-Box) Chi-Square statistic

| Lag | 12 | 24 | 36 | 48 |
|---|---|---|---|---|
| Chi-square | 11.1 | 20.4 | 28.1 | * |
| DF | 8 | 20 | 32 | * |
| P-value | 0.196 | 0.434 | 0.666 | * |

Table 6: Forecast Values for Average inflation

| Period | Forecasts | Actual value |
|---|---|---|
| 2015 | 5.91 | 5.88 |
| 2016 | 4.85 | 4.97 |
| 2017 | 6.26 | 4.88 |
| 2018 | 5.41 | - |
| 2019 | 6.64 | - |

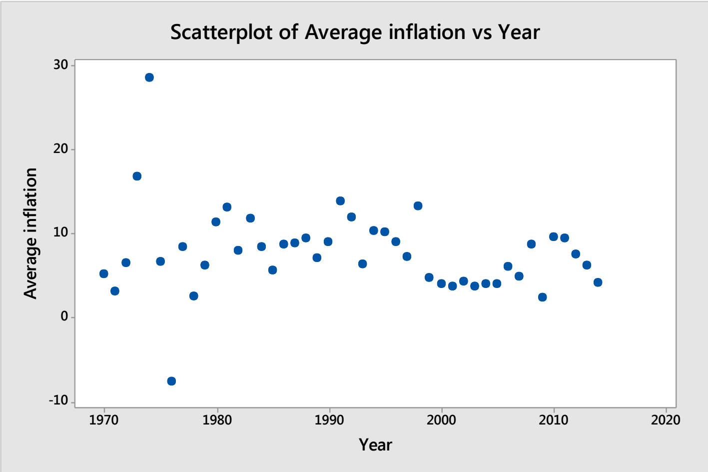


Fig 1: scatter plot for average inflation against time

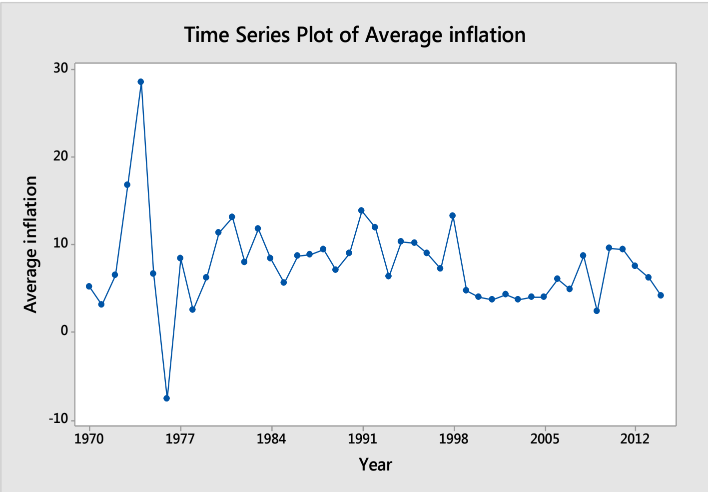


Fig 2: Time series plot of average inflation against time

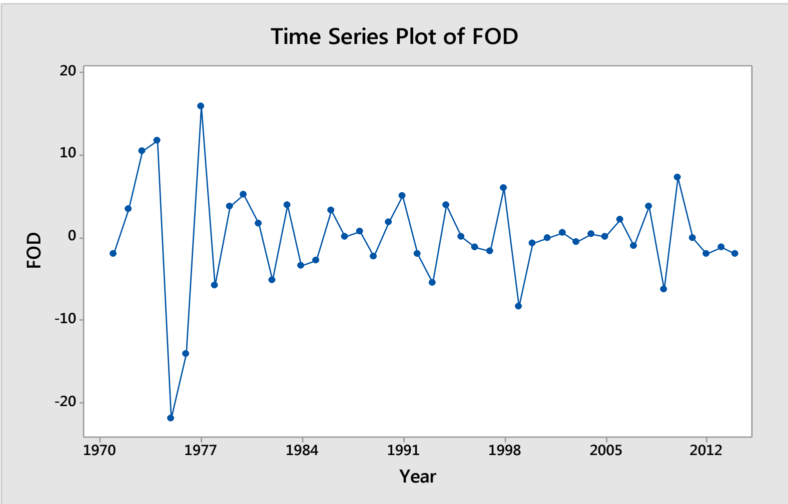


Fig 3: plot of first order differences against time lag

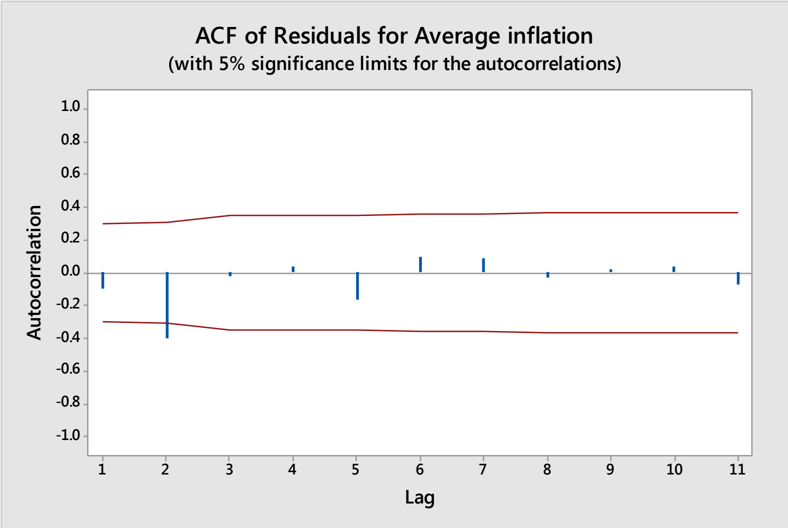


Fig 4: plot of Auto correlation against time lag

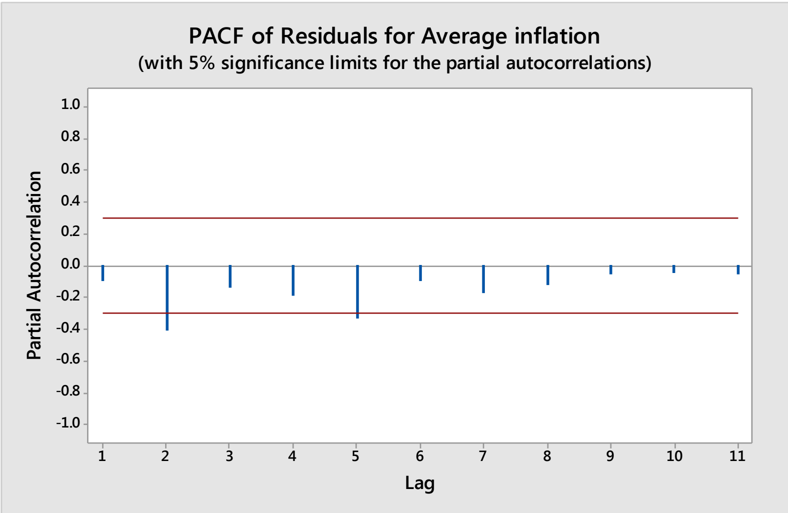


Fig 5: plot of Partial Auto correlation against time lag

IJARSE
ISSN: 2319-8354

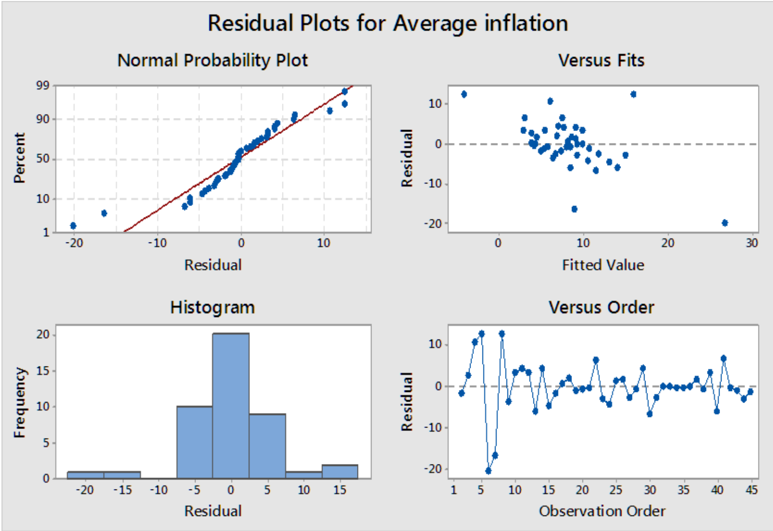


Fig 6: Residuals plots for average inflation

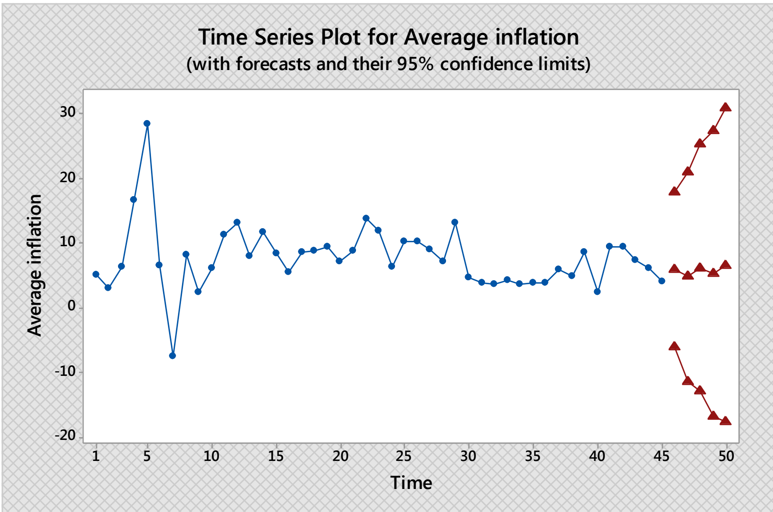


Fig 7: Time series plots for average inflation with forecasting values

## V. CONCLUSION

The data is analyzed in terms of time series plot and ARIMA model for forecasting the future values. Especially the ARIMA model is considered to be an appropriate to predict the future values in terms of time series data. In the above said result, it has been observed that the inflation rate went down rapidly in the year of 1974 to 1977 which refers to 21-months period during 1975–77 when Prime Minister Indira Gandhi unilaterally had a state of emergency declared across the country. It is suggested to a common man that, average inflation is normal. Inflation was controlled with mild variation except 1974-77 and maintaining the average inflation for the next few years. It is further informed that the future inflation rate will not make a hazardous situation to the common man in the upcoming five years. It is obvious that using ARIMA models, the forecasted values are quite similar to the actual values.

## VI. ACKNOWLEDGEMENTS

I sincerely thank all the professors of Department of Statistics, Bharathiar University who rendered their support in the success of my work especially Dr.R.Vijayaragavan and Dr.K.M.Sakthivel.

## REFERENCES

[1] S.C Gupta and V.C Kapoor, *fundamentals of applied statistics 4th edition* sultan Chand & Son's 2007

[2] Amodar N. Gujarati, *basic econometrics, 3rd Edition* McGraw-Hill International Editions 1995

[3] O. D. ANDERSON, *time series analysis and forecasting, the box-jenkins approach* butterworths, london, 1975

[4] O. D. ANDERSON, *Identifying Box-Jenkins Processes, Computer Applications* vol. 2, 1974, p. 275-283.

[6] John E.Hanke, Dean Wichern and Arthur G Reitsch; *Business Forecasting, 7th Edition* Prentice Hall Publication 2001

[7] Paul Newbold, *ARIMA model building and the time series approach to forecasting* Journal of forecasting, vol.2, 23-35, 1983

[8] T. AMEMIYA and R. Y. WU, *The Effect of Aggregation on Prediction in the Autoregressive Model* J.A.S.A., vol. 67, 1972, p. 628-632.

[9] O. D. ANDERSON, *an inequality with a time series application* J. Econom.,vol. 2, 1974, p. 189-193