International Journal of Economics and Financial Issues

ISSN: 2146-4138

available at http: www.econjournals.com

International Journal of Economics and Financial Issues, 2019, 9(3), 268-276.

# Forecasting of India VIX as a Measure of Sentiment

**Arindam Banerjee***

Birla Institute of Management Technology, Greater Noida, India. *Email: arindam20011@gmail.com

**Received:** 16 March 2019 **Accepted:** 21 May 2019 **DOI:** https://doi.org/10.32479/ijefi.8046

**ABSTRACT**

The India VIX represents the sentiment of traders in the Indian market, so by forecasting the future value of India VIX, we get a feel for investor sentiment in future. The objective of this study is to fit a forecasting model on India VIX using auto regressive integrated moving average (ARIMA). The model would be useful in having a glimpse of investor mood in near future. This is probably the first of its kind study based on Indian market. The motivation of this study lies not only on the pervasive agreement that the VIX is a barograph of the general marketplace sentiment as to what concerns investors' risk appetite, but also on the fact that there are many trading strategies that depend on the VIX index for speculative and hedging determinations. The study found ARIMA (1-0-2) forecasting model on VIX produces better forecasting result. We also validated the model with an out-of-sample dataset and found the model reliable.

**Keywords:** VIX, India, Sentiment, Forecasting, Auto Regressive Integrated Moving Average
**JEL Classifications:** C53, G17

## 1. INTRODUCTION

Volatility index or VIX captures the investors' expectation about volatility. Often termed as "investor fear gauge," VIX is always been considered as a strong indicator of investors' fear and emotions (Durand et al., 2011; Whaley, 2009). India Volatility Index i.e., India VIX was launched by National Stock Exchange (NSE) of India in 2009. It measures investors' view of the market's volatility in the immediate term. The India VIX is a good pointer of whether the market players are feeling complacent or fearful about near future. It reflects the behaviour of traders from the representativeness, affect, and extrapolation bias concepts of behavioural finance (Hibbert, et al. 2008). Thus, the VIX determines investors' expected returns since its changes are reflected in the time-varying systematically priced risk premia (Durand et al., 2011). The objective of this study is to offer a unique and simple method of forecasting India VIX. Our argument is, forecasting of India VIX may help the market participants in gauging the sentiment of the market, and may lead to better investment decisions. This paper develops a analytical model for forecasting of India VIX in the auto-regressive integrated moving average (ARIMA) framework for the period March 2009 until October 2016. In this context, our first contribution to the literature is methodology. To prove the robustness of our model, it is validated by using daily data from November 2016 to October 2017.

The rest of the paper is organized as follows. Section "about India VIX" provides a brief introduction about the volatility index. Section "literature review on predictability of VIX and its behavioural explanation" describes the past research works in the area of predictability of VIX and also its behavioural explanation. Section "objective" describes the motivation behind this study. Section "period of study and data" underlines the period covered under this study for fitting the model and also for validating the same. The section "description of methodology" presents the description about methodology used in this paper. The section "empirical results and analysis" shows the data analysis along with the estimated results. The section "evaluation of forecasts" presents the validation of the model. And finally, the paper concludes with the section "conclusion and future scope," which also indicates towards few areas for further research.

This study uses India VIX, which is traded at the NSE of India platform. India VIX is a volatility index launched by NSE of India, which measures the market expectations of near-term volatility. The value of India VIX is computed on the basis of order book of NIFTY options. The best bid-ask quotes of near and next month NIFTY option contracts which are traded on the NSE platform are used for computation of India VIX. The methodology for computation may be accessed at https://www.nseindia.com/content/indices/white_paper_IndiaVIX.pdf.

The motivation of this study lies not only on the widespread agreement that the VIX is a barograph of the overall marketplace sentiment as to what concerns investors' risk appetite, but also on the fact that there are many trading strategies that rely on the VIX index for speculative and hedging purposes (Fernandes et al., 2014).

Since the India VIX represents the sentiment of traders in the Indian market, so by forecasting the future value of India VIX, we get a feel for investor sentiment in future. The objective of this study is to fit a forecasting model on India VIX using ARIMA. The model would be useful in having a glimpse of investor mood in near future.

## 2. LITERATURE REVIEW ON PREDICTABILITY OF VIX AND ITS BEHAVIORAL EXPLANATION

VIX was first introduced by Chicago Board Option Exchange in 1993. It is widely known as "investor fear gauge." VIX reflects the expected market volatility of the market index over the upcoming days based on the implied volatility in the prices of options on the market index. In case of "India VIX," the index is Nifty 50.

Katja (2006) models the implied volatility of the S&P 500 index, with the aim of producing useful forecasts for option traders. His results indicate that an ARIMA(1,1,1) model enhanced with exogenous regressors has predictive power regarding the directional change in the VIX index.

Whaley (2009) provided a history of the VIX. He said "the VIX has been dubbed the "investor fear gauge"... (because) the S and P 500 index option market has become dominated by hedgers who buy index puts when they are concerned about a potential drop in the stock market... (the) VIX is an indicator that reflects the price of portfolio insurance."

Durand et al. (2011) showed that VIX that captures the market expectations of the investors, also affect the market return.

Chandra and Thenmozhi (2015) examined the asymmetric relationship between India volatility index (India VIX) and stock market returns, and demonstrates that Nifty returns are negatively related to the changes in India VIX levels, but in case of highly ascending movements in the market, the returns on the two indices incline to move independently.

## 3. DATA AND METHODOLOGY

This section describes data used and the methodology used for designing the India VIX model that forecasts the future direction. For this purpose, the ARIMA model has been used. But before that the unit root test has been performed to check the stationarity of the dataset.

The period of study under consideration is from March 2009 until October 2016 for fitting the model. We also used daily data from November 2016 to October 2017 for validating the ARIMA model. The daily closing value of India VIX is downloaded from the website of the NSE of India (www.nseindia.com).

### 3.1. Unit Root Analysis

As many a times, the time series variables suffer from the non-stationary problem, we have tested for unit root under augmented Dickey-Fuller (ADF) test. Section 6.1 shows the result and analysis of ADF test.

### 3.2. ARIMA

The basic idea behind ARIMA or the Box-Jenkins (BJ) methodology for forecasting is to analyse the probabilistic or stochastic properties of economic time series on their own under the philosophy "let the data speak for themselves." This concept is very different from traditional regression models, in which the dependent variable $Y_t$ is explained by k explanatory variables $X_1$, $X_2$,...........$X_n$, the BJ time series models allow $Y_t$ to be explained by the past, or lagged, values of $Y_t$ itself and the current and lagged values of $u_t$, which is an uncorrelated random error with zero mean and constant variance $\sigma^2$ – that is, a white noise error term.

The BJ methodology is based on the assumption that the time series under consideration is stationary.

#### *3.2.1. The AR model*

Consider the following model:

$$Y_t=B_0+B_1Y_{t-1}+B_2Y_{t-2}+\ldots\ldots+B_pY_{t-p}+u_t$$

Where $u_t$ is a white noise error term.

This model is termed as an AR model of order p, AR (p), for it involves regressing Y at time t on its values lagged p periods into the past, the value of p being determined empirically using some criterion, such as the Akaike information criterion.

#### *3.2.2. The MA model*

The AR process is not the only mechanism that may have generated $Y_t$. In some situation, it might be possible to capture the process of generation of $Y_t$ series by following model.

$$Y_t=u_t+\Theta u_{t-1}$$

Where, as before, $u_t$ is a white error term. The model implies that $Y_t$ is determined as a MA of the current and immediate past values of the error term. This model is called the first-order MA or MA(1) model.

The general form of the MA model is an MA(q) model of the form

$$Y_t=u_t+\Theta_1 u_{t-1}+\Theta_2 u_{t-2}+\ldots\ldots+\Theta_q u_{t-q}$$

It appears that a MA process is simply a linear combination of white noise processes, so that $Y_t$ depends on the current and previous values of a white noise error term. Further, as long as q is finite, the MA(q) process is stationary as it is an average of q stationary white noise error terms which are stationary.

### *3.2.3. The ARMA model*

If we suppose that $Y_t$ has characteristics of both AR and MA, then its is called ARMA process. For example, an ARMA (1,1) model may be written as

$$Y_t=\Phi Y_{t-1}+u_t+\Theta u_{t-1}$$

In general, an ARMA (p,q) process will have p AR and q MA terms. It is written as

$$Y_t=\Phi_1 Y_{t-1}+\Phi_2 Y_{t-2}+\ldots\ldots+\Phi_p Y_{t-p}+u_t+\Theta_1 u_{t-1}+\Theta_2 u_{t-2}+\ldots\ldots+\Theta_q u_{t-q}$$

### *3.2.4. The ARIMA model*

If a time series is integrated of order d and we apply ARMA (p,q) model to it, then we say that the original time series is ARIMA (p,d,q), i.e., it is an ARIMA time series. Clearly, if a time series is ARIMA (2,1,2), it has to be differenced once to make it stationary and the stationary time series can be modelled as ARMA (2,2) process, i.e., it will have two AR and two MA terms. Similarly, an ARIMA (p,0,p) series is same as ARMA (p,q) when the time series is stationary at the beginning. On the other hand, ARIMA (p,0,0) and ARIMA (0,0,q) series represent AR (p) and MA (q) stationary processes, respectively. Thus, given the values of p,d, and q, one can say what process is being modelled.

## 4. EMPIRICAL RESULTS AND ANALYSIS

This study fits a forecasting model based on India VIX, the fear and emotion gauge of Indian market. The study considers log of daily VIX (LVIX) value. The idea behind plotting the log of VIX and not the VIX itself is that changes in the log of a variable represents a relative change (or rate of return), whereas a change in the log of a variable itself represents an absolute change. Returns are unit-free and they are more comparable (Gujarati, 2015). The total number of observations are n = 1898.

### 4.1. Test of Stationarity

First test the stationarity of the time series data is tested. To test stationarity, the ADF test is being used. The test is performed by using the following form:

$$\Delta LVIXt = B_1 + B_2 t + B_3\ LVIX_{t-1} + \sum_{i=1}^{m} \alpha_i\ LVIX_{t-I} + \varepsilon_t \quad (1)$$

In each case, the null hypothesis is $B_3=0$ (i.e., unit root exists) and the alternative hypothesis is that $B_3 < 0$ (i.e., no unit root). The result of the unit root test of VIX with intercept is shown in Table 1.

As the $R^2$ (0.014328) is less than Durbin-Watson stat (2.0125553), therefore the regression is not spurious (Bhowmick, 2015).

**Table 1: Unit root test of VIX with intercept**

| | | | | |
|---|---|---|---|---|
| **Null hypothesis: LVIX has a unit root** | | | | |
| **Exogenous: Constant, linear trend** | | | | |
| **Lag length: 0 (Automatic - based on SIC, Maxlag=25)** | | | | |
| | | | **t-statistic** | **Prob.*** |
| Augmented dickey-fuller test statistic | | | −5.23489 | 0.0001 |
| Test critical values: | 1% level | | −3.96288 | |
| | 5% level | | −3.41218 | |
| | 10% level | | −3.12801 | |
| ***MacKinnon (1996) one-sided *P*-values** | | | | |
| **Augmented dickey-fuller test equation** | | | | |
| **Dependent variable: D (LVIX)** | | | | |
| **Method: Least squares** | | | | |
| **Sample (adjusted): March 02, 2009 October 30, 2016** | | | | |
| **Included observations: 1898 after adjustments** | | | | |
| **Variable** | **Coefficient** | **Std. error** | **t-statistic** | **Prob.** |
| LVIX(−1) | −0.027603 | 0.005273 | −5.23489 | 0 |
| C | 0.091081 | 0.01781 | 5.113952 | 0 |
| R-squared | 0.014328 | Mean dependent var | | −0.00054 |
| Adjusted R-squared | 0.013288 | S.D. dependent var | | 0.0522 |
| S.E. of regression | 0.051852 | Akaike info criterion | | -3.07927 |
| Sum squared resid | 5.094955 | Schwarz criterion | | −3.0705 |
| Log likelihood | 2925.224 | Hannan-quinn criter. | | −3.07604 |
| F-statistic | 13.77294 | Durbin-watson stat | | 2.012553 |
| Prob (F-statistic) | 0.000001 | | | |

Source: Author's own computation. SD: Standard deviation

Table 1 shows the results of ADF test. The LVIX lagged one period. The ADF test statistic is (−5.234888). However, the DF critical values are: −3.962884 (1% level), −3.412178 (5% level), and −3.128012 (10% level). In absolute terms, 5.234888 is greater than any of DF critical t values in absolute terms. Hence, the conclusion is that the VIX time series is stationary. (Gujarati, 2015). To confirm the Stationarity, also plotted the graph of LVIX over time (Figure 1). The graph confirms the Stationarity of LVIX.

## 4.2. Determination of p, q, and d

As the LVIX is stationary time series with level unit root, therefore we consider the value of d=0. We already have showed that the level order time series LVIX is stationary. So, we work with LVIX only here.

To see, which ARIMA model fits LVIX, and following the BJ methodology, we computed the correlogram of this series up to 80 lags. For determining the number of lags, we followed the rule of thumb suggested by Schwert (1989). Due to space constraint, we show the correlogram up to 15 lags in Table 2 below. The complete correlogram is given in annexure A1 at the end of this paper.

Table 2 produces two types of correlation coefficients: Autocorrelation (AC) and partial AC (PAC). The AC function (ACF) shows correlation of current LVIX with its values with various lags. The PAC function (PACF) shows the correlation between observations that are k periods apart after controlling for the effects of intermediate lags. The BJ methodology uses both these correlation coefficients to identify the type of ARMA model that is appropriate for this case.

**Figure 1:** LVIX time series


**Table 2: AC function and PAC function of LVIX**

| AC | PAC | | AC | PAC | Q-stat | Prob. |
|---|---|---|---|---|---|---|
| \|******* | \|******* | 1 | 0.983 | 0.983 | 1837.4 | 0 |
| \|******* | \|\| | 2 | 0.966 | 0.004 | 3614 | 0 |
| \|******* | \|\| | 3 | 0.951 | 0.034 | 5335.8 | 0 |
| \|******* | \|\| | 4 | 0.937 | 0.021 | 7007.2 | 0 |
| \|******* | \|\| | 5 | 0.924 | 0.032 | 8633.2 | 0 |
| \|******* | \|\| | 6 | 0.91 | −0.007 | 10214 | 0 |
| \|******\| | \|\| | 7 | 0.898 | 0.028 | 11754 | 0 |
| \|******\| | \|\| | 8 | 0.887 | 0.032 | 13257 | 0 |
| \|******\| | \|\| | 9 | 0.876 | −0.003 | 14724 | 0 |
| \|******\| | \|\| | 10 | 0.864 | −0.036 | 16151 | 0 |
| \|******\| | \|\| | 11 | 0.851 | −0.017 | 17537 | 0 |
| \|******\| | \|\| | 12 | 0.84 | 0.019 | 18886 | 0 |
| \|******\| | \|\| | 13 | 0.827 | −0.04 | 20195 | 0 |
| \|******\| | \|\| | 14 | 0.815 | 0.017 | 21466 | 0 |
| \|******\| | \|\| | 15 | 0.803 | 0.013 | 22702 | 0 |

Source: Author's own computation. AC: Autocorrelation, PAC: Partial autocorrelation

Table 2 shows gradual decline in AC and changes in positive and negative signs for PAC. However, it does not show any sign of exponential decay for any sustained period.

To see, which correlations are statistically significant, we calculate the standard error of sample correlation coefficients given by $\sqrt{1/n} = \sqrt{1/1899} = 0.022948$, where n is the sample size. Therefore, the 95% confidence interval for the true correlation coefficients is about 0 ± 1.96*(0.022948) = (−0.044977 to 0.044977). Correlation coefficients lying outside these bounds are statistically significant at 5% level. On this basis, it seems that PACF correlations at lag(s) 1, 2, 40, and 49 are statistically significant.

Since we do not have any clear-cut pattern of the ACF and PACF, we will proceed by trial and error.

First, we fit an AR model at lags 1, 2, 40, and 49.

Then we fit an MA model at lags 1, 2, 40, and 49. The result of AR (1,2,40.49) is shown in Table 3.

Since the AR(2), AR(40), and AR(49) coefficients are not significant, we can drop these from consideration and re-estimate the model with AR(1). The result is shown in Table 4 below. The model is significant at AR(1).

In the next stage, we fit the MA model. Again, we go through the trial and error method for MA(1), MA(2), MA(40) and MA(49). The result of MA model is given in Table 5 below:

The model is significant at MA(1), MA(2), MA(40), and MA(49).

Thus we use ARIMA(1,0,1), ARIMA(1,0,2), ARIMA(1,0,40), and ARIMA(1,0,49) now. Table 6 shows ARIMA(1,0,1) model fit.

**Table 3: AR model fit at lags 1, 2, 40, and 49**

| | | | | |
|---|---|---|---|---|
| **Dependent variable: LVIX** | | | | |
| **Method: ARMA maximum likelihood (OPG - BHHH)** | | | | |
| **Sample: March 02, 2009 October 30, 2016** | | | | |
| **Included observations: 1899** | | | | |
| **Convergence achieved after 19 iterations** | | | | |
| **Coefficient covariance computed using outer product of gradients** | | | | |
| **Variable** | **Coefficient** | **Std. error** | **t-statistic** | **Prob.** |
| C | 3.046914 | 0.127597 | 23.8792 | 0 |
| AR (1) | 0.969734 | 0.017887 | 54.21501 | 0 |
| AR (2) | 0.007231 | 0.018351 | 0.39404 | 0.6936 |
| AR (40) | 0.011858 | 0.009928 | 1.194428 | 0.2325 |
| AR (49) | 0.002974 | 0.009621 | 0.30908 | 0.7573 |
| SIGMASQ | 0.002697 | 4.39E-05 | 61.46501 | 0 |
| R-squared | 0.969851 | Mean dependent var | | 3.007222 |
| Adjusted R-squared | 0.969771 | S.D. dependent var | | 0.299151 |
| S.E. of regression | 0.052012 | Akaike info criterion | | −3.06954 |
| Sum squared resid | 5.120982 | Schwarz criterion | | −3.052 |
| Log likelihood | 2920.525 | Hannan-quinn criter. | | −3.06308 |
| F-statistic | 12178.97 | Durbin-watson stat | | 2.000455 |
| Prob (F-statistic) | 0 | | | |

In Table 6, the AR(1) is significant, but the MA(1) is not significant. So, we will not consider ARIMA(1,0,1) for fitting the model.

Next, we try ARIMA(1,0,2) model. Table 7 shows the ARIMA(1,0,2) model fit.

The result from Table 7 shows that both AR(1), and MA(2) are statistically significant. So, we accept the ARIMA(1,0,2) model for LVIX.

**Table 4: An AR (1) model for LVIX**

| | | | | |
|---|---|---|---|---|
| **Dependent variable: LVIX** | | | | |
| **Method: ARMA maximum likelihood (OPG - BHHH)** | | | | |
| **Sample: March 02, 2009 October 30, 2016** | | | | |
| **Included observations: 1899** | | | | |
| **Convergence achieved after 13 iterations** | | | | |
| **Coefficient covariance computed using outer product of gradients** | | | | |
| **Variable** | **Coefficient** | **Std. error** | **t-statistic** | **Prob.** |
| C | 3.024242 | 0.082506 | 36.65501 | 0 |
| AR (1) | 0.986062 | 0.00374 | 263.6586 | 0 |
| SIGMASQ | 0.002706 | 4.33E-05 | 62.46023 | 0 |
| R-squared | 0.969743 | Mean dependent var | | 3.007222 |
| Adjusted R-squared | 0.969711 | S.D. dependent var | | 0.299151 |
| S.E. of regression | 0.052063 | Akaike info criterion | | −3.06925 |
| Sum squared resid | 5.139273 | Schwarz criterion | | −3.06048 |
| Log likelihood | 2917.249 | Hannan-quinn criter. | | −3.06602 |
| F-statistic | 30383.78 | Durbin-watson stat | | 2.024454 |
| Prob (F-statistic) | 0 | | | |
| Inverted AR roots | 0.99 | | | |

Source: Author's own computation

**Table 5: MA model fit at lags 1, 2, 40, and 49**

| | | | | |
|---|---|---|---|---|
| **Dependent variable: LVIX** | | | | |
| **Method: ARMA maximum likelihood (OPG - BHHH)** | | | | |
| **Sample: March 02, 2009 October 30, 2016** | | | | |
| **Included observations: 1899** | | | | |
| **Convergence achieved after 35 iterations** | | | | |
| **Coefficient covariance computed using outer product of gradients** | | | | |
| **Variable** | **Coefficient** | **Std. error** | **t-statistic** | **Prob.** |
| C | 3.007858 | 0.009301 | 323.4038 | 0 |
| MA (1) | 1.25166 | 0.014879 | 84.12337 | 0 |
| MA (2) | 0.728046 | 0.015508 | 46.94592 | 0 |
| MA (40) | 0.063456 | 0.013094 | 4.846213 | 0 |
| MA (49) | 0.054395 | 0.012574 | 4.326118 | 0 |
| SIGMASQ | 0.01337 | 0.000441 | 30.30454 | 0 |
| R-squared | 0.850521 | Mean dependent var | | 3.007222 |
| Adjusted R-squared | 0.850126 | S.D. dependent var | | 0.299151 |
| S.E. of regression | 0.115812 | Akaike info criterion | | −1.46877 |
| Sum squared resid | 25.38971 | Schwarz criterion | | −1.45123 |
| Log likelihood | 1400.593 | Hannan-quinn criter | | −1.46231 |
| F-statistic | 2154.202 | Durbin-watson stat | | 0.966532 |
| Prob (F-statistic) | 0 | | | |

SD: Standard deviation

Next, we try ARIMA(1,0,40) model fit. The results are shown in Table 8 below:

In Table 8, AR(1) is statistically significant, but MA(40) is not. Therefore, we reject the ARIMA (1,0,40) model.

**Table 6: ARIMA (1,0,1) model fit**

| | | | | |
|---|---|---|---|---|
| **Dependent variable: LVIX** | | | | |
| **Method: ARMA maximum likelihood (OPG - BHHH)** | | | | |
| **Sample: March 02, 2009 October 30, 2016** | | | | |
| **Included observations: 1899** | | | | |
| **Convergence achieved after 20 iterations** | | | | |
| **Coefficient covariance computed using outer product of gradients** | | | | |
| **Variable** | **Coefficient** | **Std. error** | **t-statistic** | **Prob.** |
| C | 3.024706 | 0.083767 | 36.10841 | 0 |
| AR (1) | 0.986472 | 0.003756 | 262.6262 | 0 |
| MA (1) | −0.013534 | 0.018029 | −0.750681 | 0.4529 |
| SIGMASQ | 0.002706 | 4.35E-05 | 62.22056 | 0 |
| R-squared | 0.969748 | Mean dependent var | | 3.007222 |
| Adjusted R-squared | 0.9697 | S.D. dependent var | | 0.299151 |
| S.E. of regression | 0.052073 | Akaike info criterion | | −3.06836 |
| Sum squared resid | 5.138424 | Schwarz criterion | | −3.05667 |
| Log likelihood | 2917.404 | Hannan-quinn criter | | −3.06405 |
| F-statistic | 20248.62 | Durbin-watson stat | | 1.999234 |
| Prob (F-statistic) | 0 | | | |
| Inverted AR Roots | 0.99 | | | |
| Inverted MA Roots | 0.01 | | | |

**Table 7: ARIMA (1,0,2) model fit for LVIX**

| | | | | |
|---|---|---|---|---|
| **Dependent variable: LVIX** | | | | |
| **Method: ARMA maximum likelihood (OPG - BHHH)** | | | | |
| **Sample: March 02, 2009 October 30, 2016** | | | | |
| **Included observations: 1899** | | | | |
| **Convergence achieved after 14 iterations** | | | | |
| **Coefficient covariance computed using outer product of gradients** | | | | |
| **Variable** | **Coefficient** | **Std. error** | **t-statistic** | **Prob.** |
| C | 3.025538 | 0.086049 | 35.16046 | 0 |
| AR (1) | 0.987248 | 0.003622 | 272.5759 | 0 |
| MA (2) | −0.039814 | 0.019388 | −2.053567 | 0.0402 |
| SIGMASQ | 0.002702 | 4.32E-05 | 62.52751 | 0 |
| R-squared | 0.969787 | Mean dependent var | | 3.007222 |
| Adjusted R-squared | 0.96974 | S.D. dependent var | | 0.299151 |
| S.E. of regression | 0.052039 | Akaike info criterion | | −3.06965 |
| Sum squared resid | 5.131749 | Schwarz criterion | | −3.05796 |
| Log likelihood | 2918.634 | Hannan-quinn criter. | | −3.06535 |
| F-statistic | 20275.78 | Durbin-watson stat | | 2.029595 |
| Prob (F-statistic) | 0 | | | |
| Inverted AR Roots | 0.99 | | | |
| Inverted MA Roots | 0.2 | | −0.2 | |

SD: Standard deviation

**Table 8: ARIMA (1,0,40) for LVIX**

| Dependent variable: LVIX | | | | |
|---|---|---|---|---|
| Method: ARMA maximum likelihood (OPG - BHHH) | | | | |
| Sample: March 02, 2009 October 30, 2016 | | | | |
| Included observations: 1899 | | | | |
| Convergence achieved after 16 iterations | | | | |
| Coefficient covariance computed using outer product of gradients | | | | |
| Variable | Coefficient | Std. error | t-statistic | Prob. |
| C | 3.02412 | 0.082288 | 36.75045 | 0 |
| AR (1) | 0.985905 | 0.003779 | 260.9204 | 0 |
| MA (40) | 0.007426 | 0.024453 | 0.303692 | 0.7614 |
| SIGMASQ | 0.002706 | 4.55E-05 | 59.42377 | 0 |
| R-squared | 0.969745 | Mean dependent var | | 3.007222 |
| Adjusted R-squared | 0.969697 | S.D. dependent var | | 0.299151 |
| S.E. of regression | 0.052076 | Akaike info criterion | | −3.06825 |
| Sum squared resid | 5.13899 | Schwarz criterion | | −3.05656 |
| Log likelihood | 2917.301 | Hannan-quinn criter | | −3.06395 |
| F-statistic | 20246.32 | Durbin-watson stat | | 2.024428 |
| Prob (F-statistic) | 0 | | | |

SD: Standard deviation

**Table 9: ARIMA (1,0,49) for LVIX**

| Dependent variable: LVIX | | | | |
|---|---|---|---|---|
| Method: ARMA maximum likelihood (OPG - BHHH) | | | | |
| Sample: March 02, 2009 October 30, 2016 | | | | |
| Included observations: 1899 | | | | |
| Convergence achieved after 15 iterations | | | | |
| Coefficient covariance computed using outer product of gradients | | | | |
| Variable | Coefficient | Std. error | t-Statistic | Prob. |
| C | 3.02447 | 0.08283 | 36.51429 | 0 |
| AR (1) | 0.986338 | 0.003752 | 262.8917 | 0 |
| MA (49) | −0.013802 | 0.025166 | −0.548437 | 0.5835 |
| SIGMASQ | 0.002706 | 4.37E-05 | 61.94302 | 0 |
| R-squared | 0.969749 | Mean dependent var | | 3.007222 |
| Adjusted R-squared | 0.969701 | S.D. dependent var | | 0.299151 |
| S.E. of regression | 0.052072 | Akaike info criterion | | −3.06838 |
| Sum squared resid | 5.138258 | Schwarz criterion | | −3.05669 |
| Log likelihood | 2917.429 | Hannan-quinn criter. | | −3.06408 |
| F-statistic | 20249.29 | Durbin-watson stat | | 2.023808 |
| Prob (F-statistic) | 0 | | | |

Next, we consider the ARIMA(1,0,49). The result is given below in Table 9:

Table 9 shows AR(1) is significant, but MA(49) is not significant. Therefore, we reject ARIMA(1,0,49).

Finally, after going through the abovementioned tests, and under consideration of principles of parsimony, finally, we select the ARIMA(1,0,2) or ARMA(1,2) as a fit model for forecasting of India VIX. This model may be used as an estimator for predicting the future values of India VIX.

**Table 10: Inverse roots of AR and MA**

| Inverse Roots of AR/MA Polynomial (s) | |
|---|---|
| Specification: LGVIX C AR (1) MA (2) | |
| Sample: March 02, 2009 October 30, 2016 | |
| Included observations: 1899 | |
| AR root (s) | Modulus |
| 0.987248 | 0.987248 |
| No root lies outside the unit circle | |
| ARMA model is stationary | |
| MA Root (s) | Modulus |
| −0.199534 | 0.199534 |
| 0.199534 | 0.199534 |
| No root lies outside the unit circle | |
| ARMA model is invertible | |

**Table 11: Forecasted values of log of India VIX**

| Date | Observed value | Model value |
|---|---|---|
| November 30th, 2016 | 2.824351 | 2.874933 |
| December 31st, 2016 | 2.738579 | 2.729323 |
| January 31st, 2017 | 2.822866 | 2.81418 |
| February 28th, 2017 | 2.624125 | 2.621679 |
| March 31st, 2017 | 2.519308 | 2.504164 |
| April 30th, 2017 | 2.385086 | 2.410127 |
| May 31st, 2017 | 2.464917 | 2.469022 |
| June 30th, 2017 | 2.462363 | 2.441327 |
| July 31st, 2017 | 2.476538 | 2.417994 |
| August 31st, 2017 | 2.556258 | 2.610972 |
| September 31st, 2017 | 2.524528 | 2.581056 |
| October 30th, 2017 | 2.494238 | 2.448128 |

The generalized ARIMA(1,0,2) model may be written as (Chatfield, 2003):

$$x_t=\mu\ (1-\alpha)+\alpha\ (x_{t-1})+\beta_1 e_{t-1}+\beta_2 e_{t-2}$$

$$\text{or, } x_t=\mu\ (1-\alpha)+\alpha\ (x_{t-1})+\beta_1\ (x_{t-1}-x_{t-2})+\beta_2(x_{t-2}-x_{t-3}) \quad (2)$$

Now, we may put the values of ARIMA(1,0,2) from Table 7 into equation 2, where,
$\mu$=3.025538
$\alpha$=0.987248
$\beta_1$=0, and
$\beta_2$=−0.039814.

By putting the abovementioned values in equation 2, the model becomes,

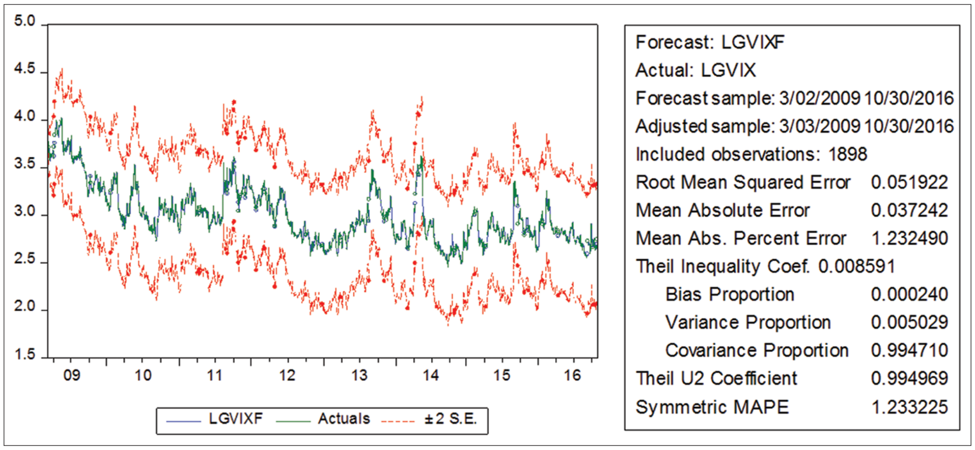
$$x_t=3.025538\ (1-0.987248)+0.987248\ (x_{t-1})-0.039814e_{t-2} \quad (3)$$

We compute the difference between observed value and computed model value for the sample data using equation 3, and find root mean squared errors (RMSE) equals to 0.0517. Which again established the appropriateness of model.

### 4.3. ARIMA Forecasting

We now use ARIMA(1,0,2) model for forecasting India VIX. Figure 2 shows the static forecast of VIX. This figure shows the actual and forecast values of logs of closing India VIX, as well as the confidence interval of forecast. The accompanying table gives the same measures of the quality of the forecast, namely,

**Figure 2:** Actual and forecast VIX




RMSE, mean absolute error (MAE), mean absolute percent error, and Theil inequality coefficient. The Theil coefficient is very low (0.008591), suggesting that the fitted model is quite good. This is also clearly shown in Figure 2, which demonstrates how closely the actual and forecast values track each other.

## 5. EVALUATION OF FORECASTS

The forecast of VIX appears to be very reliable on the basis of the following criterion:

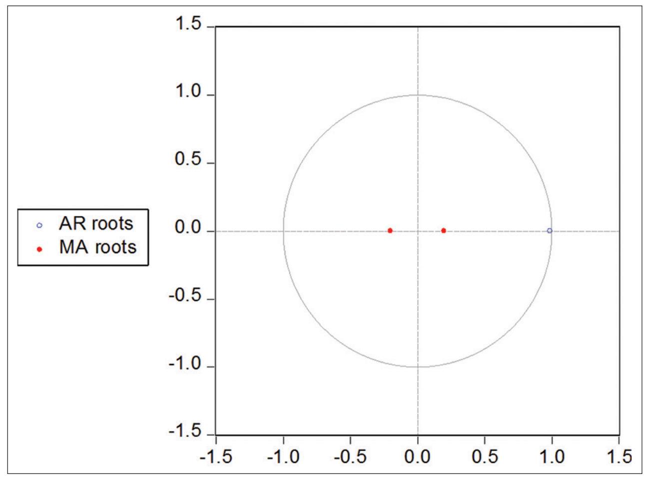
i. The estimated coefficients of both AR(1) and MA(2) terms are statistically significant (Table 7).
ii. The value of RMSE for the estimated ARIMA(1,0,2) model is 0.051922 (Figure 2), which is pretty low.
iii. The values of "bias proportion," "variance proportion," and "covariance proportion" are 0.000240, 0.005029, and 0.994710 (Figure 2) respectively. Since the values of bias and variance proportions are low, and that of covariance proportion is high, therefore the forecast may be considered satisfactory.
iv. All inverted AR and MA roots are within the unit circle (Figure 3), which implies that the chosen ARIMA model is stationary and the model has been correctly specified.

We also presented the invert roots of AR and MR in Table 10 above. No root lies outside the unit circle. The ARMA model is invertible.

For validating the model, we again considered the daily India VIX daily data from November 1st, 2016 to October 30th, 2017. Table 11 shows the month-end estimated value and the observed value of log of India VIX by using equation 3. Though all the computations are based upon daily closing quotes of India VIX, in Table 11, we are only showing the month-end data for 12 month due to limited space.

To assess the validity of observed value and model value again, we computed the RMSE of daily India VIX quotes, which we found at 0.045814, which is pretty low. The computed MAE 0.029349 and mean absolute percent error (MAPE) at 1.149486[1]. This again validates the appropriateness of the model described in equation 3.

**Figure 3:** Inverse roots of autoregressive/moving average polynomials of VIX




## 6. CONCLUSION AND FUTURE SCOPE

Our objective was to fit a forecasting model for India VIX. Based on literature (Durand et al., 2011; Whaley, 2009), we considered the India VIX as a measure of investor sentiment. We find ARIMA(1,0,2) is the fittest model to forecast future India VIX values. The evaluation of forecasting ARIMA model is also found to be reliable.

A reliable forecast of India VIX may prove to be very useful in predicting how the investor sentiment may turn in coming days. Investors may find this extremely useful in taking investment decisions. They will be able to gauge, whether the market participants are in happy mood or feeling cautious.

This study may be extended by linking INDIA VIX with the Index (NIFTY) returns. A derivative trader may be able to take a better decision by considering the forecasted values of VIX when of taking a position in derivative contracts.

## REFERENCES

Bhowmick, S. (2015). Principles of Econometrics. New Delhi: Oxford University Press.

1 The computed value of RMSE, MAE, and MAPE is slightly different from Figure 7.2 due to rounding off. This does not change the conclusions.

Chandra, M., Thenmozhi, M. (2015), On asymetric relationship of India volatility index (India VIX) with stock market return and risk management. Decision, 42, 33-55.

Chatfield, C. (2003), The Analysis of Time Series. USA: Chapman and Hall.

Durand, R.B., Lim, D., Zumwalt, J.K. (2011), Fear and Fama-French factors. Financial Management, 40(2), 409-426.

Fernandes, M., Medeiros, M.C., Scharth, M. (2014), Modeling and predicting the CBOE market volatility index. Journal of Banking and Finance, 40, 1-10.

Gujarati, N. (2015), Econometrics by Example. USA: Palgrave Macmillan.

Hibbert, A.M., Daigler, R.T., Dupoyet, B (2008), A behavioral explanation for the negative asymmetric return-volatility relation. Journal of Banking and Finance, 32(10), 2254-2266.

Katja, A. (2006), Modeling and Forecasting Implied Volatility an Econometric Analysis of the VIX Index (HECER-Helsinki Center of Economic Research Discussion Paper No. 129. University of Helsinki. Helsinki Center of Economic Research Website. Available from: https://www.helda.helsinki.fi/bitstream/handle/10138/16698/modeling.pdf?sequence=1.

Kengatharan, L., Kengatharan, N. (2014), The influence of behavioral factors in making investment decisions and performance: Study on investors of Colombo stock exchange, Srilanka. Asian Journal of Finance and Accounting, 6(1), 1-23.

Schwert, G.W. (1989), Why does stock market volatility change over time? The Journal of Finance, 44(5), 1115-1153.

Whaley, R.J. (2009), Understanding the VIX. The Journal of Portfolio Management, 35, 98-105.

## ANNEXURE

**A1: Correlogram**

<table>
<tr><th>AC</th><th>PAC</th><th></th><th>AC</th><th>PAC</th><th>Q-stat</th><th>Prob.</th></tr>
<tr><td>|*******</td><td>|*******</td><td>1</td><td>0.983</td><td>0.983</td><td>1837.4</td><td>0</td></tr>
<tr><td>|*******</td><td>| |</td><td>2</td><td>0.966</td><td>0.004</td><td>3614</td><td>0</td></tr>
<tr><td>|*******</td><td>| |</td><td>3</td><td>0.951</td><td>0.034</td><td>5335.8</td><td>0</td></tr>
<tr><td>|*******</td><td>| |</td><td>4</td><td>0.937</td><td>0.021</td><td>7007.2</td><td>0</td></tr>
<tr><td>|*******</td><td>| |</td><td>5</td><td>0.924</td><td>0.032</td><td>8633.2</td><td>0</td></tr>
<tr><td>|*******</td><td>| |</td><td>6</td><td>0.91</td><td>-0.007</td><td>10214</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>7</td><td>0.898</td><td>0.028</td><td>11754</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>8</td><td>0.887</td><td>0.032</td><td>13257</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>9</td><td>0.876</td><td>−0.003</td><td>14724</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>10</td><td>0.864</td><td>−0.036</td><td>16151</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>11</td><td>0.851</td><td>−0.017</td><td>17537</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>12</td><td>0.84</td><td>0.019</td><td>18886</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>13</td><td>0.827</td><td>−0.04</td><td>20195</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>14</td><td>0.815</td><td>0.017</td><td>21466</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>15</td><td>0.803</td><td>0.013</td><td>22702</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>16</td><td>0.792</td><td>−0.009</td><td>23904</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>17</td><td>0.78</td><td>−0.002</td><td>25072</td><td>0</td></tr>
<tr><td>|******|</td><td>| |</td><td>18</td><td>0.769</td><td>0</td><td>26207</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>19</td><td>0.758</td><td>−0.006</td><td>27310</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>20</td><td>0.747</td><td>0.002</td><td>28381</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>21</td><td>0.736</td><td>−0.006</td><td>29422</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>22</td><td>0.725</td><td>0.024</td><td>30434</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>23</td><td>0.715</td><td>−0.016</td><td>31417</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>24</td><td>0.705</td><td>0.01</td><td>32374</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>25</td><td>0.695</td><td>0</td><td>33304</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>26</td><td>0.685</td><td>−0.006</td><td>34208</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>27</td><td>0.676</td><td>0.022</td><td>35089</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>28</td><td>0.668</td><td>0.024</td><td>35948</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>29</td><td>0.659</td><td>−0.005</td><td>36787</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>30</td><td>0.651</td><td>0.018</td><td>37607</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>31</td><td>0.644</td><td>0.015</td><td>38408</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>32</td><td>0.638</td><td>0.027</td><td>39195</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>33</td><td>0.632</td><td>0.025</td><td>39969</td><td>0</td></tr>
<tr><td>|*****|</td><td>| |</td><td>34</td><td>0.627</td><td>0.015</td><td>40730</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>35</td><td>0.623</td><td>0.026</td><td>41481</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>36</td><td>0.619</td><td>0.033</td><td>42224</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>37</td><td>0.617</td><td>0.034</td><td>42961</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>38</td><td>0.613</td><td>−0.034</td><td>43690</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>39</td><td>0.608</td><td>−0.012</td><td>44408</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>40</td><td>0.602</td><td>−0.045</td><td>45113</td><td>0</td></tr>
</table>

**A1: Correlogram**

<table>
<tr><th>AC</th><th>PAC</th><th></th><th>AC</th><th>PAC</th><th>Q-stat</th><th>Prob.</th></tr>
<tr><td>|****|</td><td>| |</td><td>41</td><td>0.596</td><td>−0.002</td><td>45804</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>42</td><td>0.592</td><td>0.025</td><td>46485</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>43</td><td>0.587</td><td>0.016</td><td>47156</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>44</td><td>0.582</td><td>−0.025</td><td>47816</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>45</td><td>0.577</td><td>−0.009</td><td>48465</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>46</td><td>0.573</td><td>0.023</td><td>49104</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>47</td><td>0.569</td><td>0.003</td><td>49735</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>48</td><td>0.564</td><td>−0.015</td><td>50357</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>49</td><td>0.558</td><td>−0.051</td><td>50964</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>50</td><td>0.552</td><td>0.036</td><td>51560</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>51</td><td>0.548</td><td>0.021</td><td>52147</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>52</td><td>0.545</td><td>0.03</td><td>52727</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>53</td><td>0.542</td><td>0.024</td><td>53302</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>54</td><td>0.539</td><td>0.006</td><td>53871</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>55</td><td>0.537</td><td>0.023</td><td>54437</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>56</td><td>0.535</td><td>−0.006</td><td>54997</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>57</td><td>0.531</td><td>−0.027</td><td>55551</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>58</td><td>0.527</td><td>−0.006</td><td>56096</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>59</td><td>0.524</td><td>0.019</td><td>56634</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>60</td><td>0.52</td><td>−0.015</td><td>57165</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>61</td><td>0.516</td><td>−0.01</td><td>57687</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>62</td><td>0.511</td><td>−0.016</td><td>58200</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>63</td><td>0.507</td><td>0.023</td><td>58706</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>64</td><td>0.502</td><td>−0.027</td><td>59202</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>65</td><td>0.498</td><td>−0.001</td><td>59689</td><td>0</td></tr>
<tr><td>|****|</td><td>| |</td><td>66</td><td>0.493</td><td>0</td><td>60168</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>67</td><td>0.488</td><td>−0.01</td><td>60637</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>68</td><td>0.483</td><td>−0.009</td><td>61096</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>69</td><td>0.477</td><td>−0.005</td><td>61545</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>70</td><td>0.471</td><td>−0.043</td><td>61983</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>71</td><td>0.465</td><td>0.016</td><td>62410</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>72</td><td>0.459</td><td>−0.025</td><td>62825</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>73</td><td>0.452</td><td>−0.011</td><td>63229</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>74</td><td>0.446</td><td>0.011</td><td>63623</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>75</td><td>0.44</td><td>−0.004</td><td>64006</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>76</td><td>0.435</td><td>0.016</td><td>64381</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>77</td><td>0.431</td><td>0.037</td><td>64749</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>78</td><td>0.427</td><td>0.016</td><td>65111</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>79</td><td>0.423</td><td>−0.013</td><td>65466</td><td>0</td></tr>
<tr><td>|**** |</td><td>| |</td><td>80</td><td>0.418</td><td>−0.005</td><td>65813</td><td>0</td></tr>
</table>

AC: Autocorrelation, PAC: Partial correlation