Quest Journals
Journal of Research in Applied Mathematics
Volume 9 ~ Issue 3 (2023) pp: 54-60
ISSN(Online) : 2394-0743 ISSN (Print): 2394-0735
www.questjournals.org

Research Paper

Application of Vector Autoregressive Model

Harsha S and Shreya Maheshchandra
DEPARTMENT OF PG STUDIES AND RESEARCH IN STATISTICS
MANGALORE UNIVERSITY MANGALAGANGOTRI -57419

Abstract
Time series analysis is a specific way of analysing a sequence of data points collected over an interval of time. The variables of interest in the time series can broadly classified into univariate or multivariate. If we use the univariate class models to analyse and forecast the data which is having inter-connection, it will mislead the investors. Thus it is very important to handle the relationships between the different stock prices under study than ignoring it and carrying out a univariate time series analysis. In this work, we analysed the daily closing price of shares of State Bank of India (SBI), Axis bank & Industrial Credit and Investment Corporation of India bank (ICICI bank) which are listed under NSE (National Stock exchange) dated from 02-01-2017 to 31-12-2021. The study revealed that multivariate model gives more accurate forecast values than those of univariate model.
Keywords: *VAR, GARCH, forecast accuracy.*

Received 11 Mar., 2023; Revised 23 Mar., 2023; Accepted 25 Mar., 2023 © The author(s) 2023. Published with open access at www.questjournals.org

I. Introduction

Time series analysis is a specific way of analysing a sequence of data points collected over an interval of time. In time series analysis, analysts record data points at consistent intervals over a set of time rather than just recording the data points intermittently or randomly. However, this type of analysis is not merely the act of collecting data over time. What sets time series apart from other data is that the analysis can show how variables change over time. In other words, time is a crucial variable because it shows how the data adjusts over the course of the data points as well as the final results. It provides an additional source of information and order of dependencies between the data. The use of time series data for understanding the past and predicting the future is a fundamental part of business decisions in every sector of the economy and public services. The variables of interest in the time series can broadly classified into univariate or multivariate. If the response variable is influenced by only one factor then we call it univariate time series analysis; whereas if the response variable is influenced by multiple factors then it falls into the class of multivariate. Multivariate time series model is an extension of the univariate case and involves two or more input variables. It does not limit itself to its past information but also incorporate the past of other variables. Thu multivariate processes arise when several related time series are observed simultaneously over time, instead of observing a single series.

In recent decades, the issue of stock investment, stock market and stock trading are treated with more interest. Most people find it difficult to believe that stock is another viable area to invest in especially at this period of economic doldrums. One of the primary factors in accessing global economic and financial situation is that, usually the movement of stock market index is influenced by the moves of other stock market indices around the world or in that region, making it a critical topic to monitor over time. Consequently, the potential to reliably forecast the future value of stock market indices by taking trade relationships into account is critical. Time series models stand tall in addressing these challenges. Thus an extensive work on time series modelling is carried out by many of the researchers. Several models developed to describe the nature of time series data created a bench mark in the literature. Some of the important literatures are reviewed in order to understand the theory behind the modelling technique.

A research article by **Stock and Watson (2001)** critically reviews the use of vector autoregressions (VARs) for four tasks: data description, forecasting, structural inference, and policy analysis. The paper begins with a review of VAR analysis, highlighting the differences between reduced-form VARs, recursive VARs and structural VARs. Three variables VAR model that includes the unemployment rate, price inflation and the short-term interest rate is used to show how VAR method is used for the four tasks. They concluded that VARs have proven to be powerful and reliable tools for data description and forecasting, but have been less useful for structural inference and policy analysis. **Runkle, D.E. (2002)** questioned the statistical significance of variance decompositions and impulse response functions for unrestricted vector autoregressions. It suggests that previous authors have failed to provide confidence intervals for variance decompositions and impulse response functions. He developed two methods of computing such confidence intervals **Ahammad Hossain *et al.* (2015)** carried out a study on Vector Auto Regressive (VAR) models on selected indicators of Dhaka stock exchange (DSE) for the period from June 2004 to July 2013 as the basis on daily scale. The forecast performance of the different VAR models is discussed. **Iberedem and Blessing (2016)** forecasted stocks of the Nigerian banking sector using multivariate time series models. The study involved the stocks from six different banks that were found to be analytically interrelated. They found that the vector autoregressive model of order 1 is more suitable to explain the nature of time series under study. **Jacopo De Stefani (2019)** presented a description of the fundamentals of time series analysis and a review of the state-of-the-art in the domain of multivariate, multiple-step-ahead forecasting. The experimental results show that the proposed strategies by Jacopo De Stefani are a promising alternative to state-of-the-art models, overcoming their limitations in terms of problem size (in case of statistical models) and interpretability (in case of large-scale black-box machine learning models, such as Deep Learning techniques). **Castán-Lascorz *et al.* (2022)** proposed a new algorithm to predict both univariate and multivariate time series based on a combination of clustering, classification and forecasting techniques. The main goal of the proposed algorithm is first to group windows of time series values with similar patterns by applying a clustering process. The new algorithm has been designed using a flexible framework that allows the model to be generated using any combination of approaches within multiple machine learning techniques. To evaluate the model, several experiments are carried out using different configurations of the clustering, classification and forecasting methods that the model consists of. The results are analyzed and compared to classical prediction models, such as autoregressive, integrated, moving average and Holt-Winters models, to very recent forecasting methods, including deep, long short-term memory neural networks, and to well-known methods in the literature, such as k nearest neighbours, classification and regression trees, as well as random forest.

II. Methodology

In this work we used Generalised AutoRegressive Conditional Heteroscedastic (GARCH) model as a representation of the models under univariate time series modelling and Vector Autoregressive Model (VAR) as that of multivariate case. GARCH is a statistical model that can be used to analyse a number of financial data. The representation of GARCH model is given by,

$$Y_t = \sigma_t \varepsilon_t \text{ and } \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-i}^2$$

where $\{\varepsilon_{t1}\}$ is a sequence of white noise, $\alpha_0>0$, $\alpha_i \geq 0$, $\beta_j \geq 0$. Where α_i are ARCH parameters and β_j are GARCH parameters.

In the VAR model, each variable is modelled as a linear combination of past values of itself and the past values of other variables in the system**.** Since we have multiple time series that influence each other, it is modelled as a system of equations with one equation per variable (time series). That is, if you have 5 time series that influence each other, we will have a system of 5 equations. If the series under consideration is stationary, we forecast them by fitting a VAR to the data directly (known as a "VAR in levels"). Otherwise we take differences of the data in order to make them stationary. Then fit a VAR model (known as a "VAR in differences"). In both cases, the models are estimated by equations using the principle of least squares. The system of equations for a VAR (2) model with three time series namely, Y_1, Y_2 and Y_3) is given by,

$$Y_{1,t} = c_1 + \phi_{11,1}Y_{1,t-1} + \phi_{12,1}Y_{2,t-1} + \phi_{13,1}Y_{3,t-1} + \phi_{11,2}Y_{1,t-2} + \phi_{12,2}Y_{2,t-2} + \phi_{13,2}Y_{3,t-2} + \varepsilon_{1,t}$$

$$Y_{2,t} = c_2 + \phi_{21,1}Y_{1,t-1} + \phi_{22,1}Y_{2,t-1} + \phi_{23,1}Y_{3,t-1} + \phi_{21,2}Y_{1,t-2} + \phi_{22,2}Y_{2,t-2} + \phi_{23,2}Y_{3,t-2} + \varepsilon_{2,t}$$

$$Y_{3,t} = c_3 + \phi_{31,1}Y_{1,t-1} + \phi_{32,1}Y_{2,t-1} + \phi_{33,1}Y_{3,t-1} + \phi_{31,2}Y_{1,t-2} + \phi_{32,2}Y_{2,t-2} + \phi_{33,2}Y_{3,t-2} + \varepsilon_{2,t}$$

To check the accuracy of the fitted models we used the mean absolute percentage error (MAPE), also called the mean absolute percentage deviation (MAPD). It measures the accuracy as a percentage by taking the absolute average (mean) of the ratio of the difference between actual values and predicted values to the actual values. They are scale independent and used to compare forecast performance between different time series.

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Diamantis Koutsandreas (2021) noted the main advantage of the MAPE measures is that enables us to evaluate the forecasting accuracy across multiple time series of different scales. Also, it is easy to communicate, especially within businesses and organizations (Kolassa & Martin, 2011). Thus MAPE is the most popular choice to find the accuracy of forecasted series (Fildes & Goodwin, 2007).

III. Analysis and Discussion

The data which we took for analysis contains daily closing price of shares of State Bank of India (SBI), Axis bank & Industrial Credit and Investment Corporation of India bank (ICICI bank) that is listed under NSE (National Stock exchange) dated from 02-01-2017 to 31-12-2021. The data is collected from website https://finance.yahoo.com/ .We observed that the stock prices of SBI and ICICI are positively skewed whereas stock price of AXIS bank is negatively skewed. A stock with negative skewness is one that generates frequent small gains and few extreme or significant losses in the time period considered. On the other hand, a stock with positive skewness is one that generates frequent small losses and few extreme gains.

From Figure 1, we can visualize that the data is not stationary and contains trend component. Further, Mann-Kendall test confirmed the presence of monotonic trend in all the series (with p-vlaue less than 2.2E-16 in all the three cases). The fluctuations in the three data sets are similar over the time period. Thus, the variations in the stock price over the time period is not much effecting on the accuracy of the fitted model. We also confirmed that there is instantaneous causality between SBI, AXIS and ICICI share (with p-value <2.20E-16) using Granger Causality Test which means that there is interconnection between three share prices. From Table 1, we can observe that p-value of Augmented Dickey-Fuller (ADF) test for SBI, AXIS and ICICI bank are greater than significant level (0.05). Thus we can say that the data is having unit root non stationarity. Hence we need to convert the data into stationary by the method of differencing .

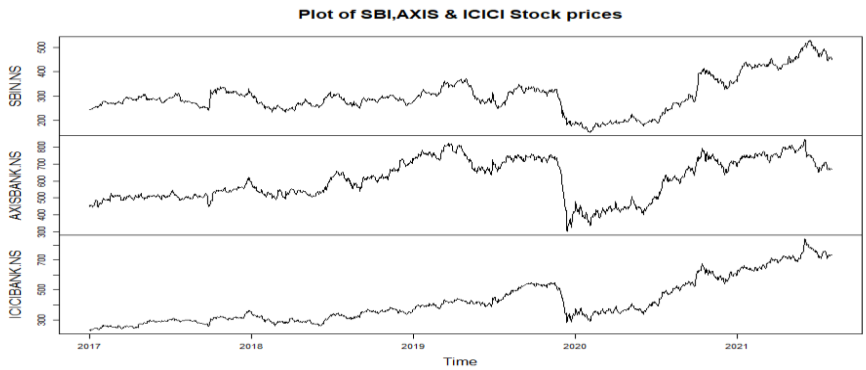

Figure 1: The time profile of stock prices of SBI, AXIS and ICICI

SBIN.NS	AXISBANK.NS	ICICIBANK.NS
-1.7711	-2.3859	-2.0978
0.6752	0.415	0.5369

Table 1: Represents ADF test results

Fitting GARCH and VAR Model:

Consider, Y_1 as the stock price of SBI bank, Y_2 as the stock price of AXIS bank, Y_3 as the stock price of ICICI bank. Generally, order determination of GARCH model is done by observing PACF plot of squared return series. But Hansen & Lunde suggested that GARCH (1,1) model is enough in practical to explain the volatility of the financial time series. Thus we fitted GARCH(1,1) model for the share prices under study.

Further, we fitted VAR model of order 5, which is selected using the information criteria. The system of equations is given below. Observe that $\phi_{11,4} Y_{1,t-4}$ notation in the system of equation represents the influence of SBI on itself at 4^{th} lag. $\phi_{13,4} Y_{3,t-4}$ represents the influence of ICICI on SBI at 4^{th} lag and so on.

Equation 1: SBIN.NS = SBIN.NS.l1 + AXISBANK.NS.l1 + ICICIBANK.NS.l1 + SBIN.NS.l2 + AXISBANK.NS.l2 + ICICIBANK.NS.l2 + SBIN.NS.l3 + AXISBANK.NS.l3 + ICICIBANK.NS.l3 + SBIN.NS.l4 + AXISBANK.NS.l4 + ICICIBANK.NS.l4 + SBIN.NS.l5 + AXISBANK.NS.l5 + ICICIBANK.NS.l5

Equation 2: AXISBANK.NS = SBIN.NS.l1 + AXISBANK.NS.l1 + ICICIBANK.NS.l1 + SBIN.NS.l2 + AXISBANK.NS.l2 + ICICIBANK.NS.l2 + SBIN.NS.l3 + AXISBANK.NS.l3 + ICICIBANK.NS.l3 + SBIN.NS.l4 + AXISBANK.NS.l4 + ICICIBANK.NS.l4 + SBIN.NS.l5 + AXISBANK.NS.l5 + ICICIBANK.NS.l5

Equation 3: ICICIBANK.NS = SBIN.NS.l1 + AXISBANK.NS.l1 + ICICIBANK.NS.l1 + SBIN.NS.l2 + AXISBANK.NS.l2 + ICICIBANK.NS.l2 + SBIN.NS.l3 + AXISBANK.NS.l3 + ICICIBANK.NS.l3 +

SBIN.NS.l4 + AXISBANK.NS.l4 + ICICIBANK.NS.l4 + SBIN.NS.l5 + AXISBANK.NS.l5 + ICICIBANK.NS.l5

	Estimate	Std. Error	t value	Pr(>|t|)
SBIN.NS.l1	-0.0194	0.03744	-0.518	0.60452
AXISBANK.NS.l1	-0.02	0.02096	-0.955	0.33981
ICICIBANK.NS.l1	0.05441	0.03106	1.752	0.08003 .
SBIN.NS.l2	-0.0334	0.03742	-0.893	0.37177
AXISBANK.NS.l2	0.00607	0.02102	0.289	0.77266
ICICIBANK.NS.l2	0.03774	0.03117	1.211	0.22612
SBIN.NS.l3	0.00425	0.03752	0.113	0.90976
AXISBANK.NS.l3	-0.0271	0.02094	-1.293	0.19633
ICICIBANK.NS.l3	0.00483	0.03096	0.156	0.87613
SBIN.NS.l4	0.02639	0.03759	0.702	0.48279
AXISBANK.NS.l4	-0.0558	0.02089	-2.67	0.00769 **
ICICIBANK.NS.l4	0.08388	0.03113	2.695	0.00714 **
SBIN.NS.l5	0.07732	0.03756	2.059	0.03973 *
AXISBANK.NS.l5	-0.0173	0.02084	-0.832	0.4058
ICICIBANK.NS.l5	0.0559	0.03102	1.802	0.07182 .

Table 2: Representing the summary of VAR(5) model with respect to SBI

From Table 2, we can say that parameters ICICIBANK.NS.11 (i.e., $\phi_{13,1}$ $Y_{3,t-1}$), AXISBANK.NS.14 (i.e., $\phi_{12,4}Y_{2,t-4}$),ICICIBANK.NS.14 (i.e., $\phi_{13,4}$ $Y_{3,t-4}$),SBIN.NS.15 (i.e., $\phi_{11,5}Y_{1,t-5}$) & ICICIBANK.NS.15 (i.e., $\phi_{13,5}Y_{3,t-5}$) are enough to explain the equation 1.

	Estimate	Std. Error	t value	Pr(>|t|)
SBIN.NS.l1	0.082395	0.071837	1.147	0.2516
AXISBANK.NS.l1	-0.08038	0.040209	-1.999	0.0458 *
ICICIBANK.NS.l1	0.085613	0.059583	1.437	0.151
SBIN.NS.l2	0.093131	0.071797	1.297	0.1948
AXISBANK.NS.l2	0.016932	0.040322	0.42	0.6746
ICICIBANK.NS.l2	-0.05685	0.059793	-0.951	0.3419
SBIN.NS.l3	-0.06392	0.071992	-0.888	0.3748
AXISBANK.NS.l3	-0.02654	0.04017	-0.661	0.5089
ICICIBANK.NS.l3	0.041656	0.059393	0.701	0.4832
SBIN.NS.l4	0.065793	0.072122	0.912	0.3618
AXISBANK.NS.l4	-0.0652	0.04008	-1.627	0.104
ICICIBANK.NS.l4	0.103031	0.059719	1.725	0.0847 .
SBIN.NS.l5	0.173986	0.072057	2.415	0.0159 *
AXISBANK.NS.l5	-0.0017	0.039975	-0.043	0.966
ICICIBANK.NS.l5	0.024677	0.059517	0.415	0.6785

Table 3: Representing the summary of VAR(5) model with respect to AXIS

From Table 3, we can say that parameters AXISBANK.NS.l1 (i.e., $\phi_{22,1}$ $Y_{2,t-1}$), ICICIBANK.NS.l4 (i.e., $\phi_{23,4}$ $Y_{3,t-4}$), SBIN.NS.l5 (i.e., $\phi_{21,5}$ $Y_{1,t-5}$), are enough to explain the equation 2.

	Estimate	Std. Error	t value	Pr(>|t|)
SBIN.NS.l1	-0.0124	0.05167	-0.24	0.8102
AXISBANK.NS.l1	0.00415	0.02892	0.143	0.886
ICICIBANK.NS.l1	-0.0253	0.04285	-0.59	0.5554
SBIN.NS.l2	-0.0402	0.05164	-0.779	0.4364
AXISBANK.NS.l2	0.00374	0.029	0.129	0.8975
ICICIBANK.NS.l2	0.00091	0.043	0.021	0.9831
SBIN.NS.l3	0.06578	0.05178	1.27	0.2042
AXISBANK.NS.l3	-0.0051	0.02889	-0.175	0.8608
ICICIBANK.NS.l3	-0.1013	0.04272	-2.372	0.0179 *
SBIN.NS.l4	0.0334	0.05187	0.644	0.5198
AXISBANK.NS.l4	-0.0165	0.02883	-0.572	0.5674
ICICIBANK.NS.l4	0.04789	0.04295	1.115	0.2651
SBIN.NS.l5	0.08873	0.05182	1.712	0.0871 .
AXISBANK.NS.l5	0.02735	0.02875	0.951	0.3416
ICICIBANK.NS.l5	0.03782	0.04281	0.884	0.3771

Table 4: Representing the summary of VAR(5) model with respect to ICICI.

From Table 4, we can say that parameters ICICIBANK.NS.l3 (i.e., $\phi_{33,3}\, Y_{3,t-3}$), SBIN.NS.l5 (i.e., $\phi_{31,5}\, Y_{1,t-5}$), are enough to explain the equation 3.

DATE	forcast.of.SBI.using.VAR	forcast.using.GARCH	actuall.value.of.SBI.
03-01-2022	452.0042	647.1	470.8
04-01-2022	452.5932	842.5	483.5
05-01-2022	452.859	1037.9	492.4
06-01-2022	452.4504	1233.2	491.7
07-01-2022	452.3181	1428.5	491.25
10-01-2022	452.3167	1623.8	503.65
11-01-2022	452.3593	1819	505.95
12-01-2022	452.3745	2014.2	510.25
13-01-2022	452.3217	2209.4	511.35

Table 5: Representing the forecast values of SBI using VAR and GARCH methods

DATE	forcast.of.axis.using.VAR	forcast.using. GARCH	actuall.value.of.AXIS
03-01-2022	677.4594	768.46	696.35
04-01-2022	678.2476	868.44	709.15
05-01-2022	678.8828	968.79	726.9
06-01-2022	677.675	1069.51	730.3
07-01-2022	677.2965	1170.6	730.6
10-01-2022	677.3208	1272.05	742.8
11-01-2022	677.4274	1373.87	743.25
12-01-2022	677.4568	1476.05	746.85
13-01-2022	677.3592	1578.59	740.7

Table 6: Representing the forecast values of AXIS using VAR and GARCH methods.

DATE	forcast.of.icici.using.VAR	forcast.using.GARCH	actuall.value.of.ICICI
03-01-2022	735.4802	1172.4	764.7
04-01-2022	735.7643	1608.9	772.85
05-01-2022	735.7785	2045.3	788.05
06-01-2022	734.9803	2481.5	785.05
07-01-2022	734.7448	2917.5	793.25
10-01-2022	734.7847	3353.3	810.75
11-01-2022	734.9313	3788.9	810.65
12-01-2022	734.9543	4224.3	823.75
13-01-2022	734.8393	4659.6	824.7

Table 7: Representing the forecast values of ICICI using VAR and GARCH methods.

From the Table 5,6 and 7 it is evident that the forecast values of all the share prices under study using VAR model is closer to the actual values than that of using the GARCH (1,1).

	MAPE	
SBI	VAR	8.66588
	GARCH	185.869
AXSIS	VAR	7.076909
	GARCH	59.958
ICICI	VAR	7.710201
	GARCH	262.6368

Table 8: Representing accuracy of VAR model and GARCH model based on MAPE

By observing above table 8 since MAPE value of different banks with respect to VAR model is less than GARCH model. Thus our study conclude that VAR forecast is more accurate to adopt in the real life.

IV. Conclusion

From the descriptive analysis, we can say that stock prices of SBI and ICICI are positively skewed whereas stock price of AXIS bank is negatively skewed.
A stock with negative skewness is one that generates frequent small gains and few extreme or significant losses in the time period considered. On the other hand, a stock with positive skewness is one that generates frequent small losses and few extreme gains.

From the time profile of SBI, AXIS & ICICI we observe that the data is not stationary and contains trend component. The fluctuations in the three data sets are similar over the time period. Thus, the variations in the stock price over the time period is not much affecting on the accuracy of the fitted model. Since there is cointegration between the share prices under study, we can conclude that there is influence of stock prices of SBI, AXIS, ICICI banks on each other.

Based on the comparison of actual & forecasted values and accuracy measure we can say that VAR(5) model out performs the GARCH (1,1). If we use only the univariate class models to analyse and forecast the data which is having inter-connection, it will mislead the investors. Thus, it is very important to handle the relationships between the different stock prices under study than ignoring it and carrying out a univariate time series analysis for prediction.

REFERENCE

[1]. Castán-Lascorz, M.A., Jiménez-Herrera, P. Troncoso,P.A. and Asencio-Cortés,G. (2022). A new hybrid method for predicting univariate and multivariate time series based on pattern forecasting. Information Sciences: an International Journal, 586. 611–627. https://doi.org/10.1016/j.ins.2021.12.001

[2]. Diamantis, K., Spiliotis, E., Petropoulos, F. and Vassilis, A. (2021). On the selection of forecasting accuracy measures. Journal of the Operational Research Society. 73. 1-18. DOI:10.1080/01605682.2021.1892464

[3]. Fildes, R., and Goodwin, P. (2007). Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting. Interfaces, 37(6), 570–576. http://www.jstor.org/stable/20141547

[4]. Hansena, P. R. and Lunde, A. (2005). A Forecast Comparison of Volatility Models: Does Anything Beat A Garch(1,1)?. Journal Of Applied Econometrics, 20, 873–889. DOI: 10.1002/jae.800

[5]. Hossain, A., Kamruzzaman, M. and Ali, M. (2015). Vector Autoregressive (VAR) Modeling and Projection of DSE. Chinese Business Review, 14, 273-289. DOI:10.17265/1537-1506/2015.06.001

[6]. Iberedem, I. and Okoro, B. (2016). Forecasting Stocks with Multivariate Time Series Models. International Journal of Mathematics and Statistics Invention, 4, 32-44

[7]. Kolassa, S. and Martin, R. (2011). Percentage Errors Can Ruin Your Day (and Rolling the Dice Shows How). International Institute of Forecasters, 23, 21-27

[8]. Runkle, D. E. (2002). Vector Autoregressions and Reality, Journal of Business & Economic Statistics, 20(1), 128-133. DOI: 10.1198/073500102753410435

[9]. Stefani, D., Jacopo, A. Caelen, O., Hattab, D., Borgne, L. and Bontempi, G. (2019). A Multivariate and Multi-step Ahead Machine Learning Approach to Traditional and Cryptocurrencies Volatility Forecasting: Methods and Protocols. DOI:10.1007/978-3-030-13463-1_1

[10]. Stock, J.H. and Watson, W. M. (2001). Vector Autoregressions. Journal of Economic Perspectives, 15 (4), 101-115. DOI: 10.1257/jep.15.4.101