AIX MARSEILLE SCHOOL OF ECONOMICS

IT TECHNIQUES FOR FINANCE

Stock Market Return Evaluation: An Approach Based on Vector Autoregressive

ZAETTA PAUL
M2 EBF 2017/2018

Abstract

This study is based on the forecast of stock market indices. We consider indices as assets, more specifically as exchange-traded funds (ETFs). It is an investment fund traded on stock exchanges much like equities. This type of fund tracks an index, such as a stock market index. ETFs are one of the most popular types of exchange-traded products. ETFs generally have higher daily liquidity and lower fees than mutual fund shares.

Nowadays, one of the most important problems against companies is the liquidity asset. We can mention a few rules such as those of Basel II and III, which partly aim to force companies to have liquid asset. ETFs can therefore be an alternative because of their liquidity.

In addition, the stock market indices provide a good representation of the current economic situation. Consequently, the behavior of them is crucial in investment decision-making.

The data on stock market indices consist of the S&P 500, CAC 40 and Nikkei 225 composite indices for United States, France and Japan respectively. The observations of the stock market indices used in this study come from the yahoo finance website. The measurements have been gathered for the period ranging from 16th December 2002 to 14th December 2017, consisting of approximately 3500 data points by series.

The aim of this study is to forecast some stock market indices using a Vector Autoregressive Moving Average (VARMA) model in order to have a quick idea if this model can be used as a tool in investment decisions.

Keywords: confidence interval forecast, multivariate model, point forecast, VARMA model, stock market index.

I. Descriptive statistics

The chart below allows us to have an overview of the performance of stock market indices, which are the S&P 500, CAC 40 and Nikkei 225 (see Figure 1). These prices are standardized in order to compare the evolution on the same scale over the whole period considered. We can notice some correlation between all theses stock market indices. In addition, we can obviously see that the trends of each series over the entire period are very similar. In this figure we observe three important trends. The first trend is before the 2007-2008 financial crisis (before the 1000^{th} observation). The second trend is the crisis period corresponding to the bearish market. And the last trend concerns the post-crisis period with a long bullish trend (after the 1500^{th} observation).

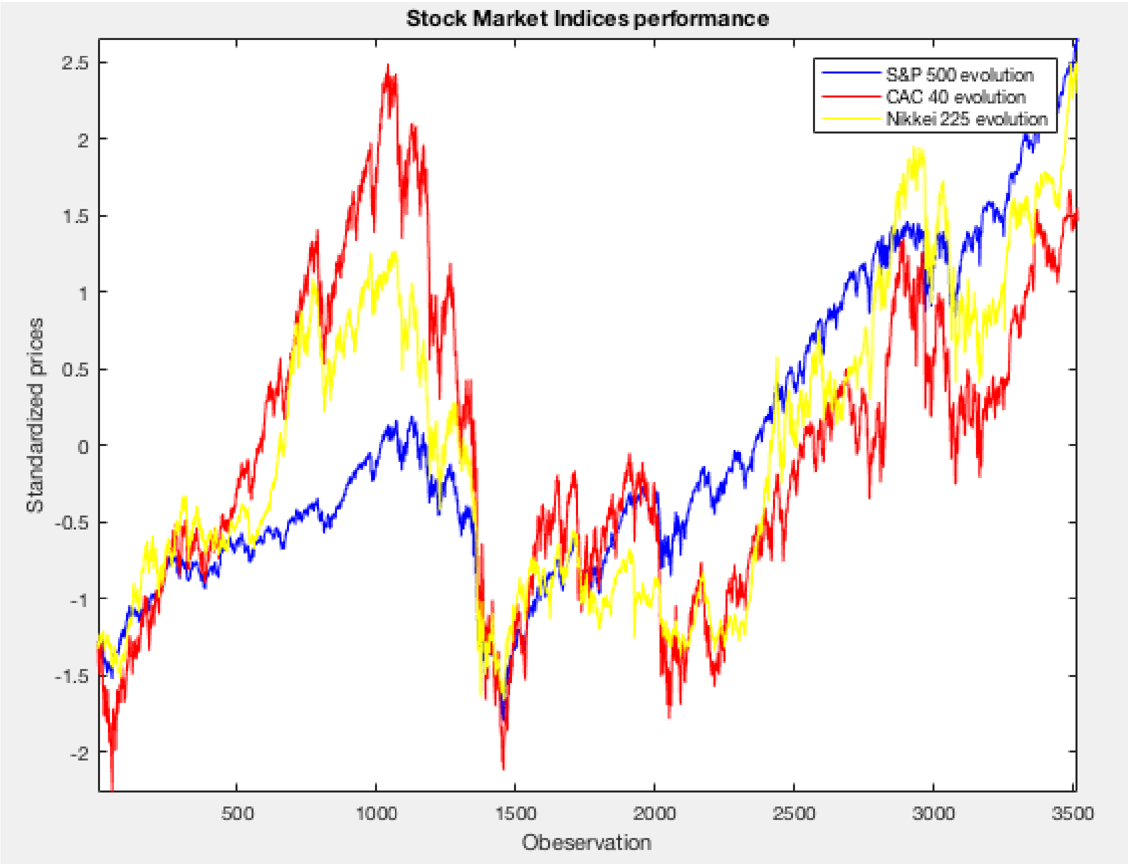

Figure 1. Episode of 15 years with stock market indices performance (with a temporal resolution of 1-day).

The market returns (indexed by i) at time t are computed as follows

$$r_{i,t} = log(P_{i,t}/P_{i,t-1}) * 100$$

where $P_{i,t}$ and $P_{i,t-1}$ are the closing prices at time t and t-1, respectively.

The figure below shows us the returns for each series over the entire period considered (see Figure 2). We can apply the Augmented Dickey-Fuller test (ADF) to these returns in order to explore the existence of unit roots in individual series. This test consists of two steps. The first step is to check whether or not autocorrelation is present in the sample of the time series. The second step is the same procedure as before but on the residuals. The results obtained form this test indicate us that all returns are stationary.

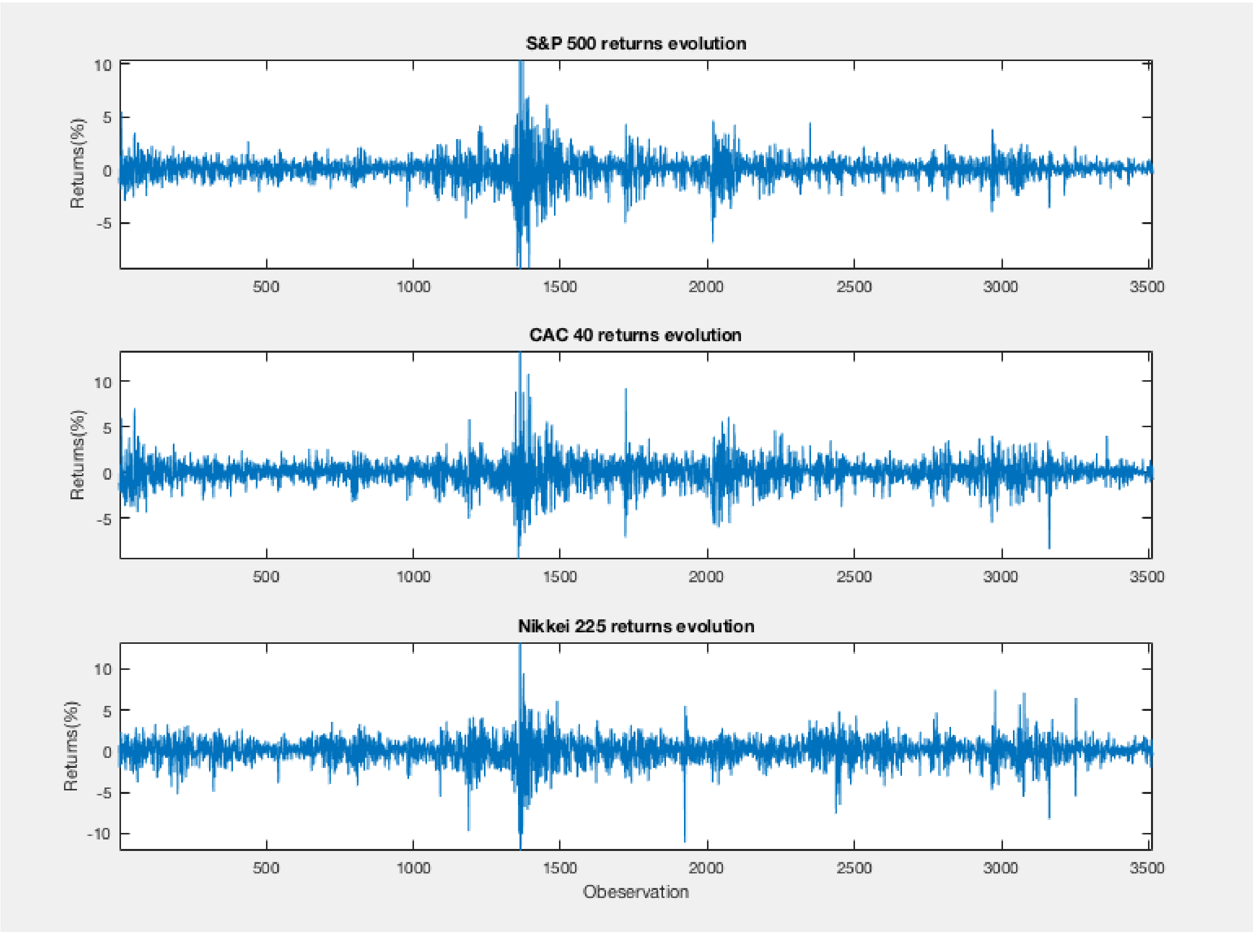

Figure 2. Episode of 15 years with stock market returns (with a temporal resolution of 1-day)

The next figure (see Figure 3) shows the different autocorrelation functions on the residuals from returns of each series. These bar charts and the Ljung-Box test confirm the choice to use the ADF test because they confirm that the residuals are not correlated, because otherwise the ADF test is inefficient.

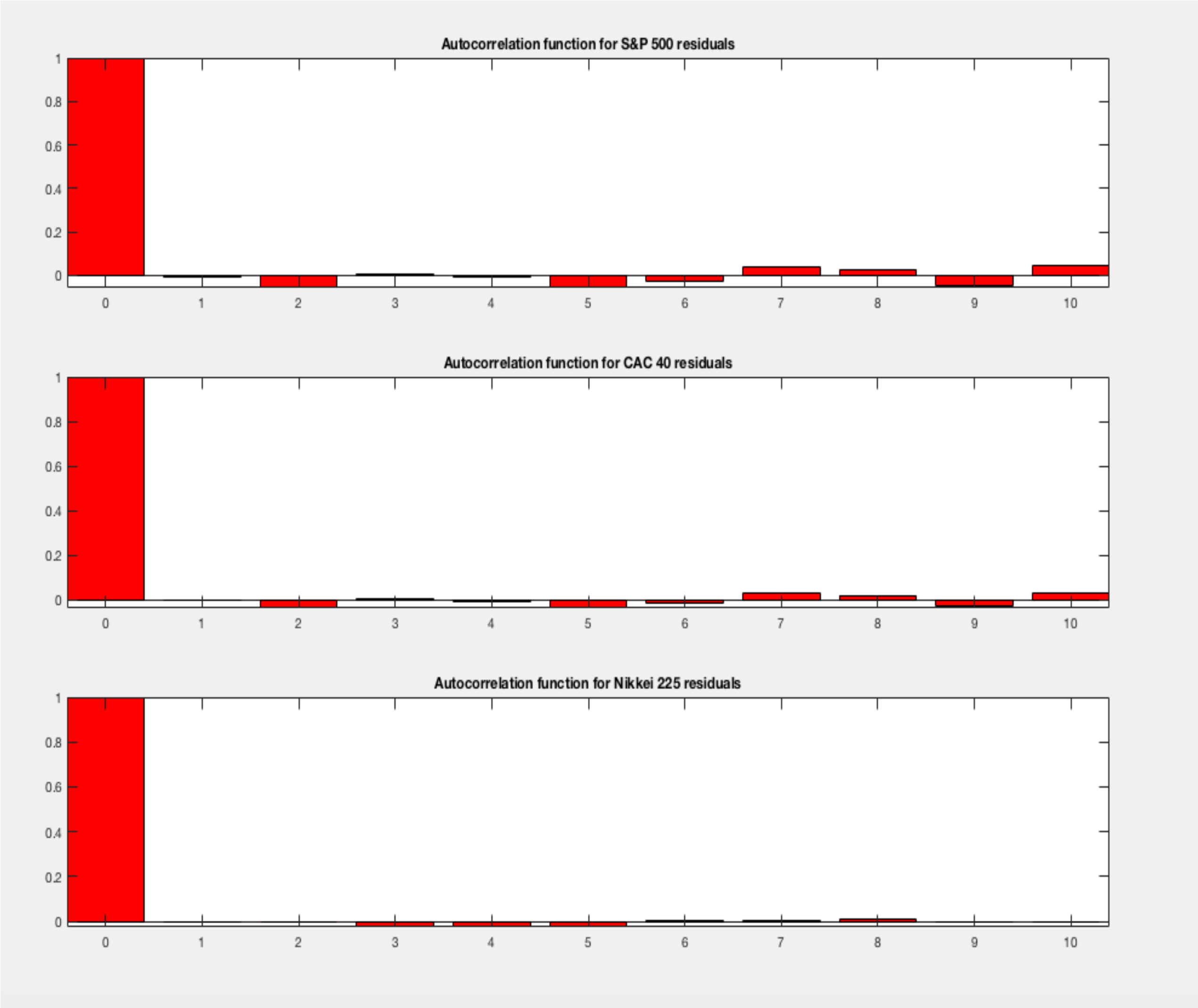

Figure 3. Autocorrelation functions for all stock market indices considered (with a number of lags equal to 10).

II. Model specification

2.1. The Model

To model the stock market return we used a Vector AutoRegressive Moving Averages (VARMA) model. This model is the natural multivariate extension of univariate ARMA processes. This type of model is useful in the sense that it reveals the cross-correlations between series and exceeding the isolated analysis of the data series.

The VARMA(1,1) can be written as follows

$$Y_t = C + \phi_1 Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Where, Y_t is a vector $(N \times 1)$ of N obeserved time series at time t, C the vector $(N \times 1)$ of constants, ϕ_1 is $(N \times N)$ autoregressive parameter matrix while θ_1 is moving average parameter matrix also of dimension $(N \times N)$ and ε_t a vector $(N \times 1)$ of independent innovations, distributed accorded to a multivariate Gaussian.

2.2. Estimation

The VARMA(1,1) model in vector form can be written as follows

$$Y = ZX + \varepsilon$$

It's equivalent to

$$\begin{pmatrix} Y_{1,2} & Y_{2,2} & Y_{3,2} \\ Y_{1,3} & Y_{2,3} & Y_{3,3} \\ \vdots & \vdots & \vdots \\ Y_{1,T} & Y_{2,T} & Y_{3,T} \end{pmatrix} = \begin{pmatrix} 1 & Y_{1,1} & Y_{2,1} & Y_{3,1} \\ 1 & Y_{1,2} & Y_{2,2} & Y_{3,2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & Y_{1,T-1} & Y_{2,T-1} & Y_{3,T-1} \end{pmatrix} \begin{pmatrix} c_1 & c_2 & c_3 \\ a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,2} & \varepsilon_{2,2} & \varepsilon_{3,2} \\ \varepsilon_{1,3} & \varepsilon_{2,3} & \varepsilon_{3,3} \\ \vdots & \vdots & \vdots \\ \varepsilon_{1,T} & \varepsilon_{2,T} & \varepsilon_{3,T} \end{pmatrix}$$

In the same way as an ARMA model, we proceed to two simple regressions (Ordinary Least Square). The first regression allows us to obtain the Autoregressive coefficients. Then, the lagged values of residuals (from the first OLS) are added and consistent Moving Average coefficients can be estimated.

III. Forecasting returns

In order to forecast returns, we used the Autoregressive coefficients estimated using the VARMA model. We can observe on the figure below some predictions of the model over three days (see Figure 4). The time 1 represents the last observation from the sample, so the prediction starts at time t = 2.

We can see from Figure 4 that the best expected return for the next period is the S&P 500 when we place ourselves on time 1. While the expected return on the CAC 40 is the worst for period 2, which is the 15th December 2017.

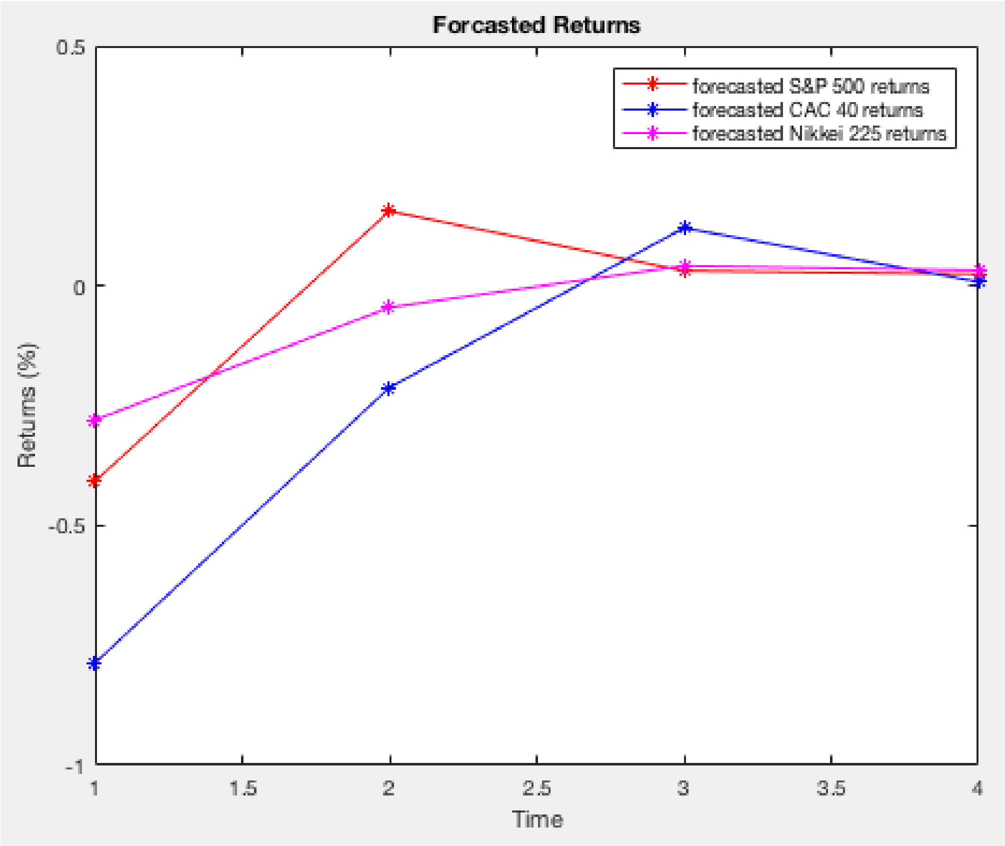

Figure 4. Point forecasts of the returns of the stock market indices considered for the next three days.

Assuming that the errors respect the i.i.d (Independent and Identically Distributed) assumption and the normality assumption, the confidence interval for the horizon h can be written as follows

$$y_{t+h} \in \left[\hat{y}_{t+h} + t^{\alpha/2} \left(\sum_{j=0}^{h-1} b_j^2 \right)^{1/2} \hat{\sigma}_\varepsilon \right]$$

where $\hat{y}_{t+h}$ is the point forecast, $\varepsilon_t \; i.i.d \; (0, \sigma_\varepsilon^2)$ and b_j is a factor, which penalizes forecasts far from the present ($b_0 = 1$).

The figure below shows the point forecasts for each stock market indices for the 15th December 2017 and the 90% confidence intervals associated (see Figure 5). In this case, α is equal to 0.10 and the horizon h is equal to one. The star point of period 2 represents the point forecast for each stock market index and the crosses represent the 90% confidence interval associated with the point forecast.

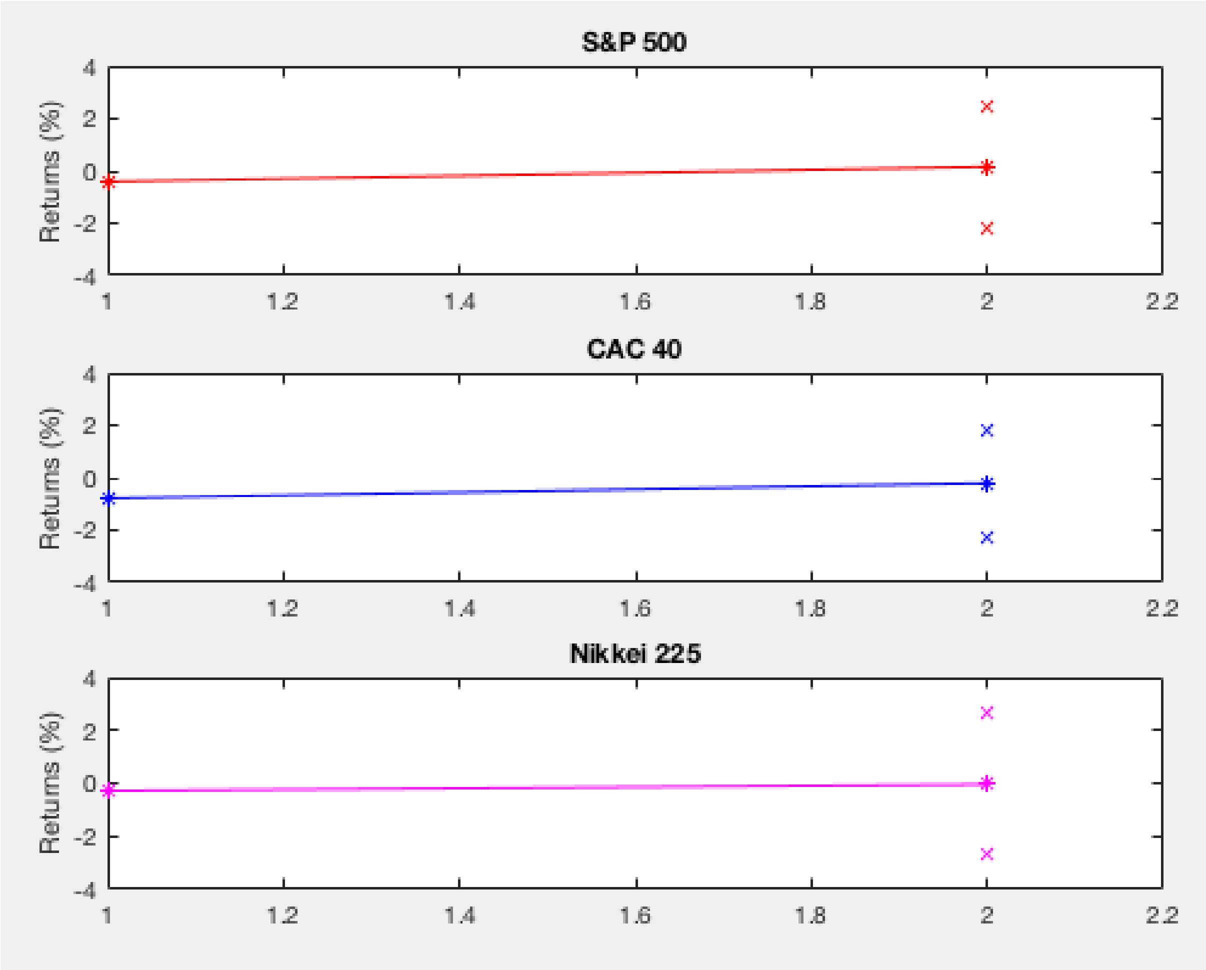

Figure 5. Forecasting returns (with a 90% confidence interval) of the stock market indices for the next day, which is the 15th December 2017 (period 2).

In order to illustrate the results and evaluate them very quickly, we applied the same model to a subsample. In this way, we can compare the forecast with reality. We obtained slightly different coefficients from those in Figure 4.

Therefore in the next figure, we plotted some forecasts, which illustrate the quality of the forecasts by comparing the point forecasts with the reality (see Figure 6). More precisely, we estimate the next three daily stock market indices from date 5 using the VARMA(1,1) model, which are represented by the red curves.

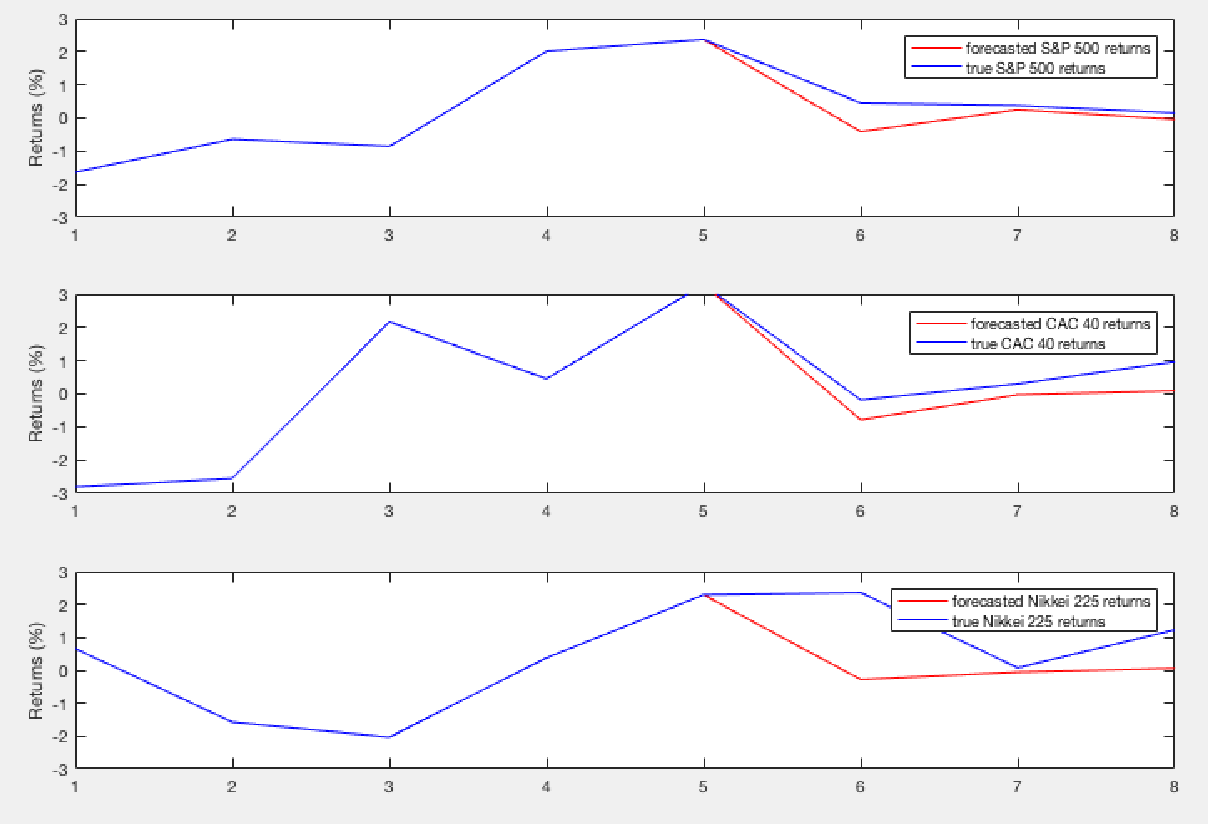

Figure 6. Comparison of stock market indices forecasts with a VARMA(1,1) model and reality.

IV. Conclusion and Discussion

The VARMA model is part of the class of mean reversion models. Consequently, this model is not really appropriate to forecast values with a horizon above one. But, forecasting with a horizon equal to one can give us information on the market direction for the next period, which can be combined with other decision-making tools. In our view, this model provides suitable forecast when the economy is stable, but when something unexpected happens, the model will not be able to anticipate it.

Furthermore, we observed at the beginning that prices seem to have some correlations. It may be interesting to analyze and forecast their correlations. This is an important aspect of Markowitz' s modern theory of portfolio (1959). He says that negative correlation helps to diversify the portfolio and achieve an efficient portfolio.

References

Khalfaoui, R., Boutahar, M., 2013. Portfolio Risk Evaluation: An Approach Based on Dynamic Conditional Correlations Models and Wavelet Multi-Resolution Analysis.

Mainassara, Yacouba Boubacar., Francq, Christian., 2010. Estimation des modèles VARMA structurels avec innovations linéaires non corrélées mais non indépendantes.

Markowitz, H., 1959. Efficient Diversification of Investements.