Modelling stock index volatility

Răduță Mihaela-Camelia*

Abstract

In this paper I compared seven volatility models in terms of their ability to describe the conditional variance. The models are compared out-of-sample using daily return data for five stock indices for the time period between 2003 and 2014. The results concluded that there is no ideal model for volatility forecasting, but asymmetric models and generalized error distribution tend to generate better forecasts than exponentially weighted moving average models.

Keywords: *stock indices, volatility models, rolling window, forecast evaluation.*

1. Introduction

Volatility is one of the most important concepts in the financial world. It can be used for detivatives valuation, for portfolio optimization and to calculate the market risk using Value-at-Risk models (Poon și Granger, 2003).

A problem often encountered which led to many conflicting empirical studies is the excess volatility, and selecting the most appropriate models to study volatility has become quite controversial. On the stock market it is necessary to understand this phenomenon especially for investors because the high volatility implies greater uncertainty that would result in important gains or losses (Islam et al., 2005).

This paper aims to model and forecast volatility of the most relevant stock indices from North America, Europe, Asia, South America and Australia. This can be achieved through volatility models. Currently, there are a large number of such models, but the most used are conditionally heteroscedastic models.

Studies that have examined this issue do not agree on a common conclusion, but they tend to favor the forecasts obtained using asymmetric models as they capture the leverage effect, which is encountered especially for shares.

Generally, my results are consistent with the findings of other research studies. More specifically, the use of asymmetric models and non-normal distributions generate better forecasts, while the models for which is not necessary to estimate the parameters give the worst results.

This paper is divided into four chapters. The first part is dedicated to theoretical and empirical studies in the literature. It summarizes the main features of financial data, the proxy variables used to substitute the real volatility and results obtained in the past literature for various indices and time periods.

* camelia.raduta91@yahoo.com

The second part focuses on the database and testing the features of financial returns for the selected indices.

In the third part is described in detail the research methodology and the volatility models, the types of proxy variables and the criteria applicable to the forecasts generated by using those models.

The last part includes the presentation of the results, respectively the estimates obtained using volatility models and performance valuation. The valuation was conducted taking into account various criteria, distributions of returns and volatility proxy.

2. Literature review

2.1. Stylized facts of financial returns

Financial time series have certain features that can not be explained by the linear structural models (Brooks, 2008), such as:

- *Leptokurtosis*: financial assets returns distribution exhibit fat tails; kurtosis estimates vary between 4 and 50 (Engle and Patton, 2001).

- *Volatility clustering*: periods of extreme returns are followed by periods with extreme returns. Therefore, the future expected volatility is influenced by today's shocks. Fama (1965) and Mandelbrot (1963) were among the first researchers that demonstrated this property in their work.

- *Leverage effect*: higher sensitivity of volatility due to a sharp decline in price than to a growth by the same amount; this feature is generally present on shares and stock indices (Engle and Patton, 2001).

Engle and Patton (2001) studied the ability of volatility models to capture the stylized facts of asset returns and to forecast conditional volatility.

To test the above properties, the authors used daily closing prices of the Dow Jones Industrial Average index for a 12 years period. According to the results, the return distribution is negatively skewed and kurtosis is very high, which means that the assumption of normality of returns distribution is rejected.

2.2. Modelling and forecasting volatility

A valuable article in literature is the one published by Hansen and Lunde (2005) in which they conducted a comprehensive analysis of the predictive power of volatility models. In this regard, they compared 330 GARCH models and their extensions using daily dollar-deutsche mark exchange rate data and IBM return data. The models were evaluated out-of-sample using six different loss functions.

Three important conclusions can be drawn from their study. For exchange rate data series was shown that GARCH (1,1) is not inferior to other more advanced models, but in terms of IBM returns the model is surpassed by other superior models, namely those that capture the leverage effect, the best performance being recorded by A-PARCH (2,2). The authors also investigated the impact that various types of distributions errors could have on forecasting performance. The effects are again divided between the two data series. If for exchange rate data, Student t-distribution has led to better results on average than the normal distribution, for data returns the things have reversed.

Because volatility can not be directly observed, a common method is to use proxies who could replace the true conditional variance. The literature uses the following methods:

- Standard deviation or daily returns variance;
- Squared returns: the disadvantage of this method is that squared returns is a noisy variable and therefore, the forecast quality is quite poor;
- Intra-daily range: it is easy to estimate and more efficient than squares returns (Louzis et al., 2013);
- Realized volatility, used by Hansen and Lunde in their research. The method is applicable to highly liquid capital markets and for small markets such data is affected by microeconomic problems such as unsynchronized trading (Silvey, 2007).
- Implied volatility which derives from options prices: all information regarding an option are observable in the market (excluding volatility). By solving the Black-Scholes equation, the implied volatility can be calculated (which is the estimate that the market assigns to the volatility of the underlying asset. Implied volatility proved to be a good indicator in predicting variance (Poon and Granger, 2003).

An elaborate analysis of volatility forecasting methods is conducted by Poon and Granger (2005). They compared 93 studies comprising various financial asset classes to determine which model predicts volatility with the best accuracy. All articles forecasted the volatility through the out-of sample technique and the models used were based on time series (historical volatility, heteroscedastic models and stochastic models) and implied volatility derived from options prices.

The results argue that implied volatility offers a better forecast than the volatility which is estimated through time series models because option price includes all current and future expectations of volatility.

Despite the complexity and flexibility of stochastic models, they recorded the worst result, and historical and heteroscedastic models generate similar predictions. In addition, by providing more information, high frequency observations generate better forecasts, especially for short time horizons.

Using historical volatility (including realized volatility) can be useful if there are no options in the market for a particular asset. The advantage of this method is that the true volatility can be computed accurately and forecasts can be improved when taking into account high frequency observations.

The studies from the literature have examined the phenomenon of volatility for different periods of time, geographic areas, financial assets, etc. Given the diversity of these papers, results are not uniform. However, some model specifications are clearly superior to others. Many articles suggest the use of asymmetric models to forecast volatility because they take into account the leverage effect, thus generating better forecasts. Also, the use of non-normal distribution (t-Student or generalized) is preferred to normal distribution as it explains better the thick tails feature of return distributions, as it was demonstrated in most articles.

3. The data used for the analysis

The data consists of daily closing prices adjusted for dividends of the following stock indices: Standard and Poors 500 (S & P 500), Financial Times Stock Exchange 100 (FTSE 100), Nikkei Stock Average (Nikkei 225), Bolsa de Valores do Estado de São Paulo (Ibovespa) and All Ordinaries (AORD) for the period January 2, 2003 - December 31, 2014 provided by *Yahoo! Finance* website. Unadjusted prices were converted into of returns time series, as follows:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

where Rt is the logarithmic return of the stock index in t period, Pt is the index value in t period and P_{t-1} is the index value in $t-1$ period.

Since the volatility is not a variable constant, but it rather changes over time, I used the rolling window method, recalibrating the models every year. The period between 2003-2014 was divided into two periods: the estimation period, which consists of approximately 1,250 observations (five years of daily observations) and the forecast period formed which consists of 1700-1800 observations (seven years of daily returns). I estimated the parameters for the first sample data (2003-2007) and generated forecasts for the next 250 observations. The models were then adjusted by rolling the window forward to capture the effect of parameters changes and this adjustment continued until the end of the whole considered period.

4. Methodology

4.1. Volatility models

Because volatility is not directly observable, it can be estimated through nonlinear models, the most known being the conditional volatility heteroscedastic models.

This paper aims at modelling stock indices using seven different volatility models: EWMA, GARCH-N (normal distribution of errors), GARCH-GED (generalized distribution of errors), IGARCH-N (normal distribution of errors), IGARCH-GED (generalized distribution of errors), EGARCH-N (normal distribution of errors), EGARCH-GED (generalized distribution of errors). According to previous theoretical and empirical research, financial returns data series do not follow the normal distribution, the forecasts generated by the volatility models being better when using another distribution, so I introduced in my study two types of distribution returns to verify if the same rule applies for this data.

EWMA model (Exponentially Weighted Moving Average) uses historical observations to illustrate the dynamic features of volatility, recent information having a greater impact on volatility forecasting. Therefore, the largest weight is associated with recent observations, while the older observations have a weight that decreases exponentially over time.

$$\sigma_{t+1}^2 = \lambda\sigma_t^2 + (1-\lambda)R_t^2$$

The ARCH model (AutoRegressive Conditionally Heteroscedastic) can be used when the homoscedasticity hypothesis is not respected. For financial series, the variance is not constant and can be modeled through ARCH models.

Because ARCH models have many limitations it appeared the **GARCH** models (they are less prone to violate the non-negativity restrictions), the most common of which is GARCH (1,1):

$$\sigma_{t+1}^2 = \omega + \alpha R_t^2 + \beta \sigma_t^2, \qquad \text{with } \alpha + \beta < 1$$

Thus, future volatility can be interpreted as a weighted average of the squared returns and variance from the current period.

If the sum of the two coefficients from GARCH model is equal to 1, GARCH (1.1) becomes:

$$\sigma_{t+1}^2 = \omega + (1-\beta) R_t^2 + \beta \sigma_t^2$$

This model is called the Integrated GARCH (**IGARCH**) and was first developed by Engle and Bollerslev.

One of the stylized facts implies that negative returns can influence the variance in a bigger proportion than positive returns. Explicit, a negative stock return lowers the company's equity, which means that the company becomes more risky, and assuming debt levels remain constant. This way, the GARCH models can be changed in order to capture the leverage effect. One example is **EGARCH** model, which is described by the following equation:

$$\ln(\sigma_{t+1}^2) = \alpha_0 + \beta \ln \sigma_t^2 + \alpha_i \left|\frac{R_t}{\sigma_t}\right| + \gamma \frac{R_t}{\sigma_t}$$

4.2. Forecast evaluation

4.2.1. Evaluation criteria

The use of volatility models depends on their ability to accurately forecast future volatility. Therefore, I conducted a series of out-of-sample forecasts to determine which models perform better.

Forecast evaluation can be a difficult process because the forecast is compared with a volatility proxy and not with its true value. The first method of assessing the accuracy of forecasts generated by the volatility models is the simple regression, in which the dependent variable is the proxy and the independent variable is the forecasted variance:

$$\hat{\sigma}^2 = b_0 + b_1 h + \varepsilon$$

The ranking is done according to the value of R squared.

The other two evaluation criteria refer to the use of loss functions. According to Patton, only two out of nine functions are robust, respectively Mean Squared Errors and Quasi Likelihood.

$$\text{MSE: } L(\hat{\sigma}^2, h) = (\hat{\sigma}^2 - h)^2$$

$$\text{QLIKE: } L(\hat{\sigma}^2, h) = \ln h + \frac{\hat{\sigma}^2}{h}$$

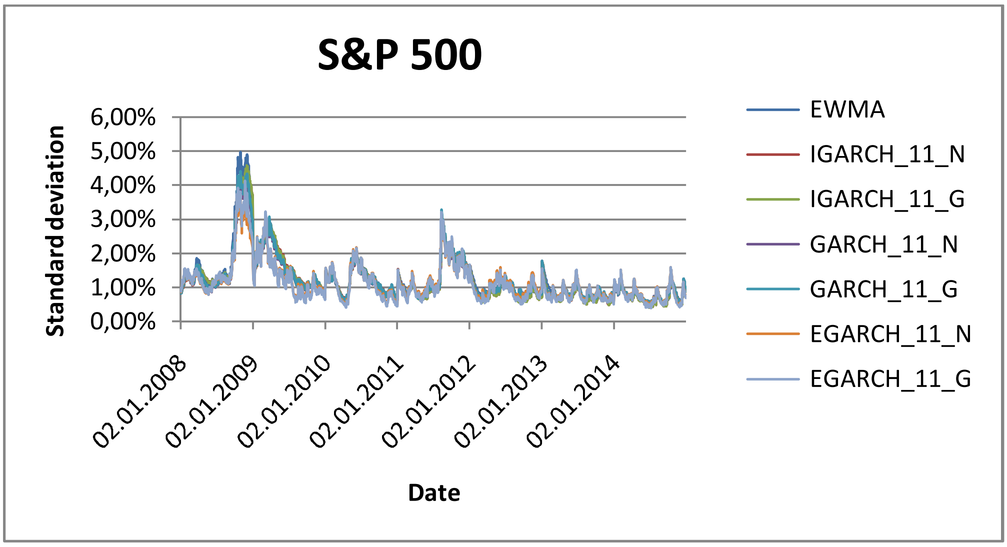
4.2.2. Volatility proxies

The simplest proxy is squared returns, which can be obtained using daily closing prices. Another variable that can replace real volatility is based on the logarithmic difference between the maximum and minimum price recorded during the day, often called as range.

5. Results

In the below graph is shown the forecasts for S&P 500 index. As you can see, there are no great differences between the forecasts. For all stock indices, the impact of financial crisis is evident on the performance of the capital market.

Graph 5.1. *S&P 500 forecasted volatility during 2008-2014*



Source: Own calculations

Generally, the asymmetric model performed best for all capital markets, except Brazil. This can be seen in table 5.1, where I aggregated the results for all volatility models, for all evaluation criteria and for all indices. On average, the best result is generated by EGARCH model with generalized distribution errors (EGARCH_11_G), and the second by EGARCH with normal distribution. Places 3 and 4 are occupied by GARCH with both types of distributions, followed by IGARCH, and the last one is EWMA. My results are similar to those obtained by Hansen and Lunde (2005) as the best models are those that allows the leverage effect and to those of Awartani and Corradi (2005) who claimed that the poorest performance is attributed to the RiskMetrics model.

Table 5.1. *The average ranking of volatility models*

Modele	S&P 500	FTSE 100	Nikkei 225	IBOVESPA	AORD	Total
EWMA	4,2	6,3	6,0	4,5	5,5	**5,3**
IGARCH_11_N	6,2	4,5	4,5	4,2	5,5	**5,0**
IGARCH_11_G	5,5	5,2	5,2	2,7	6,0	**4,9**
GARCH_11_N	4,2	3,7	2,8	2,3	4,5	**3,5**
GARCH_11_G	4,5	4,7	3,7	1,3	3,5	**3,5**
EGARCH_11_N	2,5	2,2	2,2	7,0	1,3	**3,0**
EGARCH_11_G	1,0	1,5	3,7	6,0	1,7	**2,8**

Source: Own calculations

Regarding the results for the two types of proxy used, they rank asymmetric models on the top, while the models for which is not necessary parameters estimation obtained the lowest scores.

A different way of interpreting the results is the ranking of models based on the frequency with which they occupied a particular place (table 5.2). For example, EGARCH with generalized distribution errors was ranked first 14 times out of 30 possible options[†], while EWMA never finished on first place. One can see the superiority of EGARCH and GARCH models with non-normal distribution compared with the normal distribution. In table 5.2 I also presented the results for the bottom of the ranking. Again, EWMA is the model with the most appearances on seventh place, but this time there is no longer a clear distinction between distributions errors.

Table 5.2. *The ranking for the first and last place regarding forecast evaluation*

Modele	1	7
EWMA	0	9
IGARCH_11_N	2	7
IGARCH_11_G	1	3
GARCH_11_N	2	1
GARCH_11_G	5	2
EGARCH_11_N	6	6
EGARCH_11_G	14	2

Source: Own calculations

Another widely discussed aspect in the literature is the use of non-normal distributions for predicting volatility. Theoretically, they should model better the returns compared to the normal distribution. My results confirm the previous findings. Thus, for EGARCH and IGARCH models, generalized distribution provides better forecasts than normal distribution. For GARCH model it can not be made a clear distinction between the two distributions. However, on average, the non-normal distribution is better than normal distribution. Also, in table 5.3 it can be observed that the model with the best results in forecasting volatility is EGARCH with generalized distribution.

[†] 5 indices * 3 criteria * 2 proxies

Table 5.3. The ranking according to the type of distribution errors

Modele	Distribuția normală	Distribuția generalizată
IGARCH	5,0	4,9
GARCH	3,5	3,5
EGARCH	3,0	2,8

Source: Own calculations

Also, at the indices level it can be analyzed which type of distribution is more efficient (table 5.4). The results are again divided, respectively for S&P 500, IBOVESPA and AORD indices the forecasts using generalized distribution are superior to those using the normal distribution, but for FTSE 100 and Nikkei 225 normal distribution proved to be useful for predicting the true volatility.

Table 5.4. The ranking of returns distribution among the stock indices

Indici	Distribuția normală	Distribuția generalizată
S&P 500	4,3	3,7
FTSE 100	3,4	3,8
Nikkei 225	3,2	4,2
IBOVESPA	4,5	3,3
AORD	3,8	3,7

Source: Own calculations

In conclusion, it is difficult to select a single model for volatility forecasting for all the data series, but the results converge in designating EGARCH model with generalized distribution as the one that generate forecasts with high accuracy because it incorporates all three stylized facts of the return series.

6. Conclusions

In this paper I compared the performance of forecasts generated by different volatility models for five indices relevant to the capital markets across the globe, namely: S&P 500, FTSE 100, Nikkei 225, IBOVESPA and AORD. Before the model estimation, it was necessary to test the statistical properties of returns. The results showed that the data series do not follow a normal distribution, are negatively skewed and show fat tails. Volatility clustering is validated, meaning that volatile returns tend to cluster in time. It also confirms the leverage effect, that volatility is more sensitive to a sharp decline in price than to a growth by the same amount.

All volatility models are statistically significant, the parameter values indicating that shocks in the conditional variance are highly persistent. In general, the model that best characterizes the returns behavior proved to be EGARCH with generalized distribution errors because it reflects the properties previously tested.

On average, for S&P 500, FTSE 100 and AORD indices, EGARCH models with normal distribution and generalized distribution performed the best, which is in line with the results of Awartani and Corradi (2005). For Nikkei 225, EGARCH and GARCH with normal distribution had the best results, followed by the same models, but with generalized distribution. For IBOVESPA things have radically changed, GARCH models having the highest forecast accuracy and asymmetric models having the lowest forecast accuracy.

The results for the two proxy variables support previous statements, so the weakest models are EWMA and IGARCH, and the best are asymmetrical models. In addition, EGARCH with generalized distribution errors came most often in first place, while EWMA had the most appearances in last place.

Finally, the type of distribution errors is particularly important in predicting volatility. According to the results obtained in this paper and in the literature, the use of generalized distribution errors significantly improves forecasts quality.

As future research, I recommend analyzing a possible correlation between the volatility from the various capital markets and its possible transmission between those markets.

References

1. Alberg, D., Shalit, H., Yosef, R. (2008). *Estimating stock market volatility using asymmetric GARCH models*. Applied Financial Economics, 18, 1201–1208.
2. Andersen, T. G., Bollerslev, T. (1998). *Answering the skeptics: Yes, standard volatility models do provide accurate forecasts*. International Economic Review, 39, 885-905.
3. Awartani, B. M. A., Corradi V. (2005). *Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries*. International Journal of Forecasting, 21, 167-183.
4. Bentes, S. R. (2015). *A comparative analysis of the predictive power of implied volatility indices and GARCH forecasted volatility*. Physica A, 424, 105–112.
5. Brooks, C. (2008*). Introductory Econometrics for Finance*, Cambridge University Press, pp. 379-427.
6. Chiang, T. C., Doong, S.-C. (2001). *Empirical Analysis of Stock Returns and Volatility: Evidence from Seven Asian Stock Markets Based on TAR-GARCH Model*. Review of Quantitative Finance and Accounting, 17, 301-318.
7. Christoffersen, P. F. (2003). *Elements of financial risk management*. Academic Press, 31-39.
8. Curto, J. D., Pinto, J. C., Tavares, G. N. (2009). *Modeling stock markets' volatility using GARCH models with Normal, Student's t and stable Paretian distributions*. Stat Papers, 50, 311–321.
9. Ederington, L. H., Guan, W. (2005). *Forecasting volatility*. Journal of Futures Markets, 25, 465-490.
10. Engle, R. F., Patton, A. J. (2001). *What good is a volatility model?*. Quantitative Finance, 1, 237-245.
11. Fama, E. F. (1965). *The Behaviour of Stock-Market Prices*. Journal of Business, 38, 34-105.
12. Hansen, P. R., Lunde, A. (2005). *A forecast comparison of volatility models: Does anything beat a GARCH (1,1)?*. Journal of Applied Econometrics, 20, 873-889.
13. Islam, S. M. N., Watanapalachaikul, S. (2005). *Empirical Finance: Modelling and Analysis of Emerging Financial and Stock Markets*. Physica-Verlag Heidelberg, 123-146.
14. Liu, W., Morley, B. (2009). *Volatility forecasting in the Hang Seng Index using the GARCH approach*. Asia-Pacific Financial Markets, 16, 51-63.
15. Louzis, D. P., Xanthopoulos-Sisinis, S., Refenes, A. P. (2013). *The role of high frequency intra-daily data, daily range and implied volatility in multi-period Value-at-Risk forecasting*. Journal of Forecasting, 32, 561-576.
16. Mandelbrot, B. (1963).*The Variation of Certain Speculative Prices*. Journal of Business, 36, 394-419.

17. Niguez, T.-M. (2008). *Volatility and VaR forecasting in the Madrid Stock Exchange*. Spanish Economic Review, 10, 169–196.
18. Patton, A. J. (2011). *Volatility forecast comparison using imperfect volatility proxies*. Journal of Econometrics, 160, 246–256.
19. Poon, S. H. (2005). *A Practical Guide to Forecasting Financial Market Volatility*. John Wiley and Sons, pp. 10-14.
20. Poon, S. H., Granger, C. (2003). *Forecasting Volatility in Financial Markets: A Review*. Journal of Economic Literature, XLI, pp. 478–539.
21. Poon, S. H., Granger, C. (2005). *Practical Issues in Forecasting Volatility*. Financial Analysts Journal, 61, 45-56.
22. Quen, T. Y., Hoong, T. S. (1992). *Forecasting Volatility in the Singapore Stock Market*. Asia Pacific Journal of Management, 9, 1-13.
23. Silvey, T. A. (2007). *An investigation of the relative performance of GARCH models versus simple rules in forecasting volatility*. In Forecasting Volatility in the Financial Markets (Third Edition), Quantitative Finance (pp. 101-129). Oxford: Elsevier.
24. So, M. K. P., Xu, R. (2013). *Forecasting Intraday Volatility and Value-at-Risk with High-Frequency Data*. Asia-Pacific Financial Markets, 20, 83–111.
25. Zivot, E., Wang, J. (2006). *Modeling financial time series with S-PLUS*. Springer-Verlag, pp. 313.