Combining stock market volatility forecasts using an EWMA technique*

Dong Jing-rong
School of Economics and Management
Chongqing Normal University
Chongqing, P.R. China, 400047

Abstract-**Forecasting stock market volatility is an important and challenging task for both academic researchers and business practitioners. The recent trend to improve the prediction accuracy is to combine individual forecasts using a simple average or weighted average where the weight reflects the inverse of the prediction error. In the existing combining methods, however, the errors between actual and predicted values are equally reflected in the weights regardless of the time order in a forecasting horizon. In this paper, we present a new approach where the forecasting results of the Generalized Autoregressive Conditional Heteroscedastic (GARCH), the Exponential Generalized Autoregressive Conditional Heteroscedastic(EGARCH), and random walk models arecombined based on a weight that reflects the inverse of the Exponentially Weighted Moving Average (EWMA) of the Mean Absolute Percentage Error (MAPE) of each individual prediction model. The results of an empirical study indicate that the proposed method has a better accuracy than the Generalized Autoregressive Conditional Heteroscedastic (GARCH),Exponential Generalized Autoregressive Conditional Heteroscedastic (EGARCH) and random walk models, and also combining methods based on using the MAPE for the weight.**

Ⅰ. INTRODUCTION

While traditional financial economics research has tended to focus upon the mean of stock market returns, in recent times the emphasis has shifted to focus upon the volatility of these returns. Moreover, the international stock market crash of 1987 has increased the focus of regulators, practitioners and researchers upon volatility. These concerns have led researchers to examine the level and stationarity of volatility over time. Specifically, research has been directed toward examining the accuracy of volatility forecasts obtained from various econometric models.

There is a large literature on forecasting volatility. Many econometric models have been used. However, no single model is superior. Using US stock data, for example, Brooks (1998) finds the Generalized Autoregressive Conditional Heteroscedastic (GARCH) models outperform most competitors[1]. Brailsford and Faff (1996) (hereafter BF) find that the models are slightly superior to most simple models for forecasting Australian monthly stock index volatility[2]. Using data sets from Japanese and Sigaporean markets respectively, however, Tse (1991) and Tse and Tung (1992) find that the exponentially weighted moving average models provide more accurate forecasts than the GARCH model[3,4]. Dimson and Marsh (1990) find in the UK equity market more parsimonious models such as the smoothing and simple regression models perform better than less parsimonious models, although the GARCH models are not among the set of competing models considered[5].

The weakness of most previous studies is their dependence on a single model that is expected to capture all aspects of the volatility formation process. An alternative solution to overcome the limitation is to combine individual forecasts based on models of different specifications and/or information sets[6]. Armstrong (2001) reported that an equally weighted combination of forecasts reduced the average forecasting error by 12.5%[7]. Bates and Granger (1969) advocated the use of a weighted average when combining forecasts with the weight being calculated from the variance and covariance of the different forecasting errors[8]. Similarly Russell and Adam (1987), Schwaerzel and Rosen (1997)combined individual forecasts using weights that were obtained from the mean squared error, mean absolute error, or Mean Absolute Percentage Error (MAPE) of the individual models[9,10]. Menezes et al. (2000)presented a detailed review on combining models and covered the simple average method, regression-based methods, and the switching method[11]. Chan et al. (2004) used quality control techniques to decide when one needs to recalculate the combining weights[12]. Their approach is relatively new. However, they still obtained the weight in terms of the average error from the entire training data set. In the existing combining methods, the errors between the actual and predicted values are equally reflected in the weights regardless of their time order in a forecasting and thus this kind of weight is slow to react to dynamic changes.

In this paper, we propose an approach in which the results obtained using the GARCH, Exponential Generalized Autoregressive Conditional Heteroscedastic (EGARCH), and random walk approaches are combined using a weight that reflects the inverse of the Exponentially Weighted Moving Average (EWMA) of the Absolute Percentage Error (APE) of each of the prediction models. The results of an empirical study based on Shenzhen stock market data indicate that the proposed method has a better accuracy than the GARCH, EGARCH, and random walk models, and also combining methods based on using the MAPE for the weight.

The remainder of this paper is organized as follows.

*This research was supported by National Social Science Foundation of china under Grant04CY012 and the Excellent Young Teacher Foundation of National Education Committee under Grant2002350

Section 2 presents the individual prediction models along with combining algorithms. In Section 3, a comparison study based on Shenzhen stock market data is performed and the results are summarized. Conclusions are drawn in Section 4.

II. FORECASTING MODEL

In this section we describe four different models applicable to stock market volatility prediction: (i) GARCH; (ii) EGARCH; (iii) random walk; and (iv) combination methods including EWMA and MAPE approaches.

A. GARCH model

The generalized autoregressive conditional heteroscedastic (GARCH) model was developed by Bollerslev and is an extension of the autoregressive conditional heteroscedastic (ARCH) model introduced by Engle (1982) to allow for a more flexible lag structure[13,14]. The GARCH model involves the joint estimation of a conditional mean and a conditional variance equation. Usually, a GARCH (1,1) is found to be sufficient to adequately model volatility. The GARCH (1,1) model employed in this study can be described as follows:

$$\begin{cases} r_t = \mu + \varepsilon_t, \quad where \quad \varepsilon_t \left| \Omega_{t-1} \sim N(o, h_t) \right. \\ h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \qquad \omega > 0, \alpha > 0, \beta > 0, \end{cases} \tag{1}$$

Where r_t is the return on the stock index measured as the logarithm of relative price change and h_t is the conditional volatility. Ω_{t-1} is the information set available at time t.

B. EGARCH model

A limitation of the GARCH model described above is that the conditional variance responds to positive and negative innovations in the same manner. However, there is a body of evidence that suggest that the restriction is not empirically valid, in other words, it has been noted that often negative shocks to the conditional mean equation have a larger effect upon volatility than positive shock. One model which remove the assumption of symmetric responses of volatility to shocks of different sign are the exponential GARCH (EGARCH) model proposed by Nelson(1991)[15].The EGARCH(1,1) model used in this study can be described as follows:

$$\begin{cases} r_t = \mu + \varepsilon_t \qquad where \quad \varepsilon_t \left| \Omega_{t-1} \sim N(0, h_t) \right. \\ \ln(h_t) = \omega + \beta \ln(h_{t-1}) + \gamma \left(\dfrac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \alpha \left[\dfrac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\dfrac{2}{\pi}} \right] \end{cases} . \tag{2}$$

In the EGARCH model, the asymmetry arises from the direct inclusion of the term in ε_{t-1}, normalized by the standard deviation of the data. There are no nonnegativity restrictions on these parameters in the EGARCH model. This is a particularly useful property which significantly simplifies the estimation of parameters and avoids a number of possible difficulties in a negative estimation of GARCH model.

C. Random walk model

When time series data have an irregualr pattern, naive models such as the random walk model are difficult to beat. The simple random walk model implies that the trend remains roughly constant throughout the whole series without any persistent upward or downward drift. The random walk model can be described as

$$\hat{\sigma}_t(RW) = \sigma_{t-1}, \tag{3}$$

where σ_t is the daily volatility measure. Hence it assumes that the best forecast of today's volatility is yesterday's observed volatility.

D. EWMA combination

We introduce the EWMA approach in an effort to combine the individual results obtained from the GARCH, EGARCH and random walk approaches. In the EWMA combining method, the weight associated with each individual forecasts is determined in such a manner so as to more reflect recent performances rather than historic performances. In other words, our proposed weights reflect changes in the model performance as a function of the time *t*.

In general, the EWMA of time series data s_t can be defined as follows (Montgomery, 1996)[16].

$$Z_t = \lambda s_t + (1-\lambda) s_{t-1}, \tag{4}$$

where $0 < \lambda \le 1$ is a constant and $Z_0 = \bar{s}$.

It is well known that using the EWMA approach Z_t is reduced to the form of a weighted average of all the previous samples:

$$Z_t = \lambda \sum_{j=0}^{t-1} (1-\lambda)^j s_{t-j} + (1-\lambda)^t Z_0, \tag{5}$$

with s_t being defined as the APE of the individual prediction models at time t. That is $s_t = \left| \sigma_t - \hat{\sigma}_t^i \right| / \sigma_t$ where σ_t is the actual volatility at time t and $\hat{\sigma}_t^i$ is the predicted volatility of time t obtained using model i. Subsequently, Z_0 can be obtained as the MAPE so that the EWMA of the APE of prediction model *i* at time t can be expressed as

$$EWMA_{i,t} = \lambda \sum_{j=0}^{t-1} (1-\lambda)^j \frac{\left| \sigma_{t-j} - \hat{\sigma}_{t-j}^i \right|}{\sigma_{t-j}} + (1-\lambda)^t \times MAPE_i, \tag{6}$$

where $MAPE_i$ is the mean of the APE of model *i*.

Next, we propose an EWMA combining method that uses the inverse of the EWMA as the weight for the individual forecasting model. In this case the weight w_i can be written as

$$w_{i,t} = \frac{EWMA_{i,t}^{-1}}{\sum_{i=1}^{N} EWMA_{i,t}^{-1}}, \tag{7}$$

where N is the total number of forecasts to be combined.

Based on Equation (7), we can combine individually predicted results $\hat{\sigma}_{t-j}^i$ directly using weight $w_{i,t}$ as follows:

$$\hat{\sigma}_c = \sum_{i=1}^{N} w_{i,t} \hat{\sigma}_t^i . \tag{8}$$

In this way, a predicted value at time t from an individual model associated with a smaller EWMA gets a larger weight and thus recent performances are favored over historical performances.

In the existing MAPE combining method the weight $w_{i,t}$ is written as

$$w_{i,t} = \frac{MAPE_{i,t}^{-1}}{\sum_{i=1}^{N} MAPE_{i,t}^{-1}} \quad . \tag{9}$$

III. EMPIRICAL STUDY

A. Data description and research method

The data analysed in this paper are the daily Shenzhen Stock Exchange closing component stock price index for the period 2 January 2001 to 30 July 2006. The Shenzhen Stock Exchange publishes a daily composite index that is based on the weighted market capitalization of all listed companies. The data was obtained from Datastream. Daily returns are identified as the difference in the natural logarithm of the closing index value for two consecutive trading days.

Table 1 contains the number of return observations for the stock index and statistics testing the null hypothesis of independence and identically distributed normal variates. The descriptive statistics for the return series are mean, standard deviation, skewness, kurtosis and Ljung-Box statistics LB(12) and LB^2 (12) for the return series and the squared return series. Under the assumption of normality, skewness and kurtosis have asymptotic distributions N(0,6/T) and N(3,24/T), respectively, where T is the number of return observations. The return distribution is positively skewed,indicating that the distribution is non-symmetric. Furthermore, the relatively large value of kurtosis statistics suggests that the underlying data are leptokurtic, or fattailed and sharply peaked about the mean when compared with the normal distribution.

The Ljung-Box LB(12) statistics for the cumulative effect of up to the twelfth order autocorrelation in the return exceeds 21.026 (5% critical value from a chisquared distribution with 12 degrees of freedom). It indicates that there is some evidence for serial correlation in the stock return series that should be accounted for in the mean equation.

Even if there was a lack of serial correlation, evidence would imply only that the series was uncorrelated, and no conclusion could be drawn on independence. The Ljung-Box LB2(12) statistics for the squared return provides us with a test of intertemporal dependence in the variance. TheLB2(12) statistical value 119.34, which exceeds 26.217 (1% critical value from a chi-squared distribution with 12 degrees of freedom), rejects significantly the zero correlation null hypothesis. In other words, the distribution of the next squared return depends not only on the current squared return but also on several previous squared returns, which will result in volatility clustering. These results clearly reject the independence assumption for the time series of daily stock returns. In sum, there are dependence, non-normality, thicktails and volatility clustering in the time series data of Shenzhen daily stock returns.

The approach taken in this paper is one-step-ahead forecasts. One-step-ahead prediction is useful in evaluating the adaptability of a forecasting model. Since our main goal is to evaluate the volatility forecasting performance of four models, we wish to consider a reasonably large hold-out sample. Therefore, the sample data set of daily Shenzhen component index prices is divided into two parts. The first part is from 2 January 2001 to 31 December 2005. The second part is from 2 January 2006 to 30 July 2006. The second part of the data set serves as the test or comparison period in which the out of sample forecasts from the models are compared. The first part of the data set is reserved for estimating the initial parameters of the models. Furthermore, since it is not a priori assumed that one model necessarily dominates other models over the whole sample, we repeat our modelling and forecasting exercise for different subsamples. We thus fit the models to a sample of five years, generate a one-step-ahead forecast, delete the first observation from the sample and add the next one, and generate again a one-step-ahead-forecast. This process continues until we get a volatility forecasts for each day from 2 January 2006 to 30 July 2006.

TABLE I
SAMPLE STATISTICS FOR DAILY RETURNS (JANUARY, 2, 2001-JULY, 30, 2006)

NO.of obs	Mean (%)	Standard deviation (%)	Skewness coefficient	Kurtosis	LB(12) return	LB^2(12) squared return
1449	-0.0785	2.175	0.8993	11.4332	22.8945	119.34

B. Out-of-sample model forecast results

Two commonly used loss functions or error statistics: the root mean squared error (RMSE) and the mean absolute percentage error (MAPE) are employed to measure the performance of out-of-sample model forecast results. They are defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\sigma_t - \hat{\sigma}_t)^2} \quad ,$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\sigma_t - \hat{\sigma}_t\right| / \sigma_t \quad ,$$

where σ_t and $\hat{\sigma}_t$ denote the actual volatility and the forecasted volatility forecast in time t, respectively.

Table 2 reports the root mean squared forecast error(RMSE) and the mean absolute percentage error (MAPE) for each of the individual models and each of the combining methods for the out-of-sample period 2 January 2006 to 30 July 2006. In terms of RMSE and MAPE, the EWMA($\lambda = 0.5, \lambda = 0.9$) and MAPE Combining methods are better than the individual forecasting methods. Within the Combining methods, the EWMA combining method outperforms the MAPE combining method based on RMSE and MAPE. Within the individual forecasting models, the EGARCH model has almost the same accuracy as the GARCH model with both these models being more accurate than the random walk model using either measure of performance.

It is interesting to note that the EWMA ($\lambda = 0.1$) combining method have a worse performance than the GARCH and EGARCH models and also the MAPE

TABLE II
OUT-OF-SAMPLE FORECASTING PERFORMANCE OF COMPETING MODELS FOR THE VOLATILITY OF STOCK INDEX (JANUARY 2, 2006-JULY 30, 2006)

Error statistics / Model	RMSE	MAPE
GARCH(1,1)	0.00875	2.8543
EGARCH(1,1)	0.00864	2.84765
Random Walk	0.00972	2.9643
MAPE Combining	0.00823	2.9213
EWMA Combining($\lambda = 0.1$)	0.00934	2.9612
EWMA Combining ($\lambda = 0.5$)	0.00812	2.9111
EWMA Combining($\lambda = 0.9$)	0.00777	2.8893

combining method. This is probably due to the fact that the random walk model performed poorly in predicting Shenzhen Stock market Volatility. Thus, it would appear that the effect of combining the individual forecasts is smaller, when the difference in the accuracies among individual models is higher. On the other hand, EWMA ($\lambda = 0.9$) combining method has a better accuracy than any other method. The performance of the EWMA method is dependent on the λ value. When λ is large, the effect of combining the individual forecasts is clear. This is because as the value of λ increase, a larger weight is assigned to the EGARCH and GARCH forecasts than to the random walk forecast.

Ⅳ. CONCLUSION

Stock market volatility forecasting is a widely researched area in finance literature. The performance of forecasting models of varying complexity has been investigated according to a range of measures and generally mixed results have been recorded. On the one hand some argue that relatively simple forecasting techniques are superior, while others suggest that the relative complexity of ARCH-type models is worthwhile. The weakness of most previous studies is their dependence on a single approach that is expected to capture all aspects of the volatility formation process. In this paper we seek to extend previous studies by combining individual forecasts based on models of different specifications and/or information sets to produce improved volatility forecasts.

In an effort to improve the accuracy of forecasting stock market volatility, we have proposed a method to combine forecasts that uses the inverse of the EWMA of the APE as the weight for an individual forecast. A key advantage of our proposed method is that a forecast associated with a smaller EWMA is given a large weight. This means that recent performances are considered to a greater extent than more historical performances. The performance of the proposed method and four other forecasting models in predicting Shenzhen Stock market volatility have been investigated.

The results of our empirical study indicate the following. First, the GARCH and EGARCH models have a better accuracy than the random walk model. Secondly, the combining methods generally outperform the individual models. However, the EWMA ($\lambda = 0.1$) combining method have a worse performance than the GARCH and EGARCH models. This result indicates that the combining method with small λ may be ineffective when the difference in accuracy for the individual forecasts is large. Lastly, the proposed EWMA combining method ($\lambda = 0.5, \lambda = 0.9$) has a better accuracy than the MAPE combining method.

REFERENCES

[1]C. Brooks, "Predicting stock index volatility: Can market volume help?", Journal of Forecasting, 1998,17, 59-80.
[2]T.J. Brailsford, and R.W. Faff, "An evaluation of volatility forecasting technique", Journal of Banking and Finance, 1996 , 20, 419-38.
[3]Y. K. Tse, "Stock return volatility in the Tokyo stock market", Japan and the World Economy, 1991, 3, 285-98.
[4]Y. K. Tse and S. H. Tung, "Forecasting volatility in the Singapore stock market", Asia Pacific Journal of Management, 1992,9, 1-13.
[5]E. Dimson and P. Marsh, "Volatility forecasting without data-snooping", Journal of Banking and Finance, 1990, 14, 399-421.
[6]R. Clemen, "Combining forecasts: a review and annotated bibliog-raphy", International Journal of Foreca-sting, 1989, 5, 559-583.
[7]J. S. Armstrong, "Combining Forecasts in Principles of Forecasting: A Handbook for Researchers and Practitioners", 2001,pp.417-439.
[8]J. M. Bates and C. Granger, "The combination of forecasts. Operational Research Quarerly", 1969,20, 451-468.
[9]T. D. Russll and E. E. Adam, "An empirical evaluation of alternative forecasting combinations", Management Science, 1987,33(10), 1267-1376.
[10]R. Schwaerzel and B. Rosen, "Improving the accuracy of financial time series prediction using ensemble networks and high order statistics", in Proceedings of the International Conference on Neural Networks,Houston, TX, 1997,Volume 4, pp. 2045-2050.
[11]L. M. Menezes,D. W. Bunn and J. W. Taylor, "Review of guidelines for the use of combined forecasts", European Journal of Operational Research, 2000,120, 190-204.
[12]C. K. Chan, B. G. Kingsman and H. Wong, "Determinging when to update the weights in combined forecasts for product demand-an application of the C-USUM technique", European Journal of Operational Research, 2004,153, 757-768.
[13]T. Bollerslev, "Generalized autoregressive conditional heteroscedasticity", Journal of Econometrics, 1986, 31, 307-27.
[14]R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", Econometrica, 1982,50, 987-1007.
[15]D.B.Nelson, "Conditional heteroskedasticity in asset return:A new approach", Econometrica, 1991,59,347-370.
[16] D.C.Montgomery, Introduction to statistical quality control, Wiley, New York, NY, 1996.