Forecasting High-Frequency Financial Data Volatility Via Nonparametric Algorithms -Evidence from Taiwan Financial Market

Wo-Chiang Lee

Department of Finance and Banking
Aletheia University
E-mail: wclee@email.au.edu.tw
32 ,Chen Li Street ,Tamsui, Taipei County ,Taiwan

Abstract

This paper uses two computational intelligences algorithms-artificial neural networks(**ANN**) and genetic programming(**GP**) for forecasting four different horizon high frequency TAIEX financial data volatilities and compares the out sample forecasting performances with parametric volatility models, HISVOL, GARCH(1,1) ,EGRACH(1,1) and GJR-GARCH(1,1) models. Our results reveal that nonparametric computational intelligence algorithms are powerful for modeling high-frequency intraday financial data volatility.

Keywords: Integrated volatility, genetic programming, artificial neural networks

1. Motivation and Introduction

Volatility forecasting is an important task in financial markets, especially in asset allocation, risk management, security valuation, pricing derivatives and monetary policy making(Poon and Granger,2003).Even though the measurement and forecasting of volatility has attracted the interest of many researchers and practitioners, it remains a challenging econometric problem till now.

To this topic, there are two main questions must to be discussed. First of all, is volatility forecastable? If it does, then the second question is that which method will provide the best forecasting performance? For the first question, Poon and Granger(2003) surveyed 93 published and working papers that study forecasting performance and various volatility models and give us a clear answer, that is ,financial market volatility is clearly forecastable.

The second question is that how to estimate volatility well? Merton (1980) and Nelson (1992) suggested that volatility may be estimated arbitrarily well through the use of sufficiently finely sampled high-frequency returns over any fixed time interval.

However, integrated volatility (IV) has been served as a realized volatility model which calculated from the cumulative squared intraday returns of the underlying securities at high frequencies as defined by Anderson et al. (1998).

Many types of models are used to forecast realized volatility[1]. In addition ,we can witness a phenomenal growth in the application of computational intelligence methodologies to forecasting problems in economics and finance in recent years. They are powerful to handle nonlinear relationships. In this study, we address nonparametric computational intelligences models, namely artificial neural networks(hence **ANN**) and genetic programming(hence **GP**) for forecasting the realized high-frequency time series volatility and compare the test sample forecasting performance with parametric models ,namely HISVOL, GARCH(1,1) and asymmetric EGARCH(1,1),GJR-GARCH(1,1). The reason of taking parametric models as benchmark is that parametric model such GARCH-type are easy to estimate and readily interpretable. It is also the simplest model that describes volatility clustering. Furthermore, it is widely used by both academics and practitioners and be a good benchmark volatility model. As my best knowledge, only few studies had applied GP to forecast volatility. Chen,S-H. and C.-H. ,Yeh.(1997) first proposed a time-variant and non-parametric approached to estimate volatility based on recursive genetic programming(RGP).Besides, the model can simultaneous detecting the structure change. Neely ,C. J. and P. A. Weller(2002) used GP in the volatility forecasting of exchange market and illustrated the strengths and weaknesses. The paper also showed that GP did consistently outperform the GARCH(1,1) model on MAE and model error bias at all horizons.This encourage us using this methodology

[1] The reader is referred to Poon and Granger(2003) for a review of this topic.

for investigation the Taiwan Weighted Stock Price Index(TAIEX) high-frequency time series data.

The paper is arranged as follows. Section 2 presents the forecasting methodologies and evaluation criteria of our empirical study. Section 3 gives the data description and experimental setup. Section 4 is the results analysis and we give some concluding and remarks in section 5.

2. Forecasting Methodology and Evaluation Criteria

As already mentioned above, we utilizes two computational intelligences(CI) and HISVOL,GARCH(1,1) ,EGARCH(1,1),GJR-GARCH(1,1) models for volatility forecasting. However, what is computational intelligence? Chen,S.-H. and Paul P. Wang(eds.)(2002) gived a clear definition: "Computation intelligence is a new development paradigm of intelligence systems which has resulted from a synergy between fuzzy sets, artificial neural networks ,evolutionary computations, machine learning ...etc., broadening computer science,physics,engineering,mathematics,statistics,phychology,and the social sciences alike."

In our empirical study, **HISVOL** model is based on past standard deviations. This group starts on the basis that $\sigma_{t-\tau}$ which is known or can be estimated at time t-1.Eg. the random walk model, σ_{t-1} can be used as a forecast for σ_t .While GARCH-Type models are GARCH(1,1),EGARCH(1,1) and GJR-GARCH(1,1).

In addition to, we set the **GP** parameters as table 1. five simulations are conducted for TAIFEX, namely **GP-Run1 to GP-Run 5** respectively.

Table1 Tableau for Genetic Programming

Population size(N)	50
Number of tree created by complete growth	25
Number of tree created by partial growth	25
Function set	{+,-,*,/,sin,cos, $\sqrt{\ }$, log,power}
Terminal set[2]	{R_{t-1},R_{t-2},…..R_{t-10},R}
Criterion of fitness (F)	Sum of squared Errors
Number of generations(n)[3]	200

[2] For comparison reason,we follow Chen,S-H. and C.-H.Yeh.(1997)'s terminal set.Chen,S. –H. and T.-W.,Kuo(2002) also consider return series{R_{t-1},….R_{t-10}}.

[3] When the number of generations(n) is 200,the convergence met.

In our ANN empirical study, we take series {R_{t-1}, $R_{t-2,\ldots,}$ R_{t-10}} as input variables, the output is volatility which is represented as equation (1).

$$\sigma^2_{ANN} = f(R_{t-1,} R_{t-2,\ldots\ldots\ldots\ldots} R_{t-10}) \qquad (1)$$

The parameters in ANN are setting as follows:

1) hidden unit is 10,15,20 respectively.
2) transfer function is hyperbolic.
3) learning algorithms is backpropagation.

We compares the three volatility models with five performance measure indexes. namely, mean absolute error (**MAE**),mean square error(**MSE**),mean absolute percentage error(**MAPE**) and **Theil's U**,**VaR backtesting.**

The idea of backtesting is derived from Sadorsky,P.(2004) which compared the forecasting performance of several well known volatility forecasting models by the backtesting of VaR.

PF test (Kupiec ,1995) is one of the most popular tests applied in backtesting. Let x be the number of "failures[4]" in a sample of size n. If the VaR model is correct, x follows the binomial distribution with parameter (n,x). The null hypothesis is, that the forecasting model is correct, and LR test statistic is equation(2)

$$LR = -2\ln[(1-p)^{n-x} p^x] + 2\ln[(1-\frac{x}{n})^{n-x} (\frac{x}{n})^x] \qquad (2)$$

$$Ho : \hat{p} = p$$
$$H1 : \hat{p} \neq p$$

Where, p is the loss probability of the VaR model to be tested. Under the null hypothesis the PF test statistic presented in the equation (2) follows Chi-square distribution with one degree of freedom.

3. Data Description and Experimental Setup

Our empirical data got from Taiwan Economic Journal**(TEJ)** stock databank. Where, the 10 minute ,30minute,1 hour horizon dataset ranges from Dec. 1, 2003 until Dec. 31, 2003 intra day data for per one minute trading. The full sample is divided into two subperiods, the training period is over 9:01, Dec. 1 2003 to 10:48 Dec.24 ,2003,total 4210 observations. The test sample period is over 10:49 Dec.24,2003 to

[4] It is the number of cases in which loss exceeds the one forecasted by VaR model.

13:30 Dec. 31,2003 ,total 2000 observations. Whereas, the one day horizon dataset ranges from Jan. 2, 2002 until Dec. 31, 2004. The full sample is also divided into two subperiods, the training period is over 9:00, Jan. 2 2002 to 13:30 April 13 ,2004,total 500 observations. The test sample period is over 9:00 April 14,2004 to 13:30 Dec. 31,2004 ,total 200 observations.

As mentioned above, a better approach is to sum intraday returns to measure **realized daily volatilit**y (i.e., integrated volatility) more accurately. We measures integrated volatility using ten irregularly spaced intraday observations. If $S_{i,t}$ is the *i*th observation on date *t*, we define

$$R_{it} = 100 \ln \frac{S_{i+1,t}}{S_{i,t}} \tag{3}$$

Then the 10 ,30 minute ,1 hour and one-day horizon integrated volatility on data t is measured by equation (4)

$$\sigma^2{}_{it} = \sum_{i=1}^{T} R^2{}_{i,t} \tag{4}$$

Where ,i=1,2,…..T,T=9(one-day),10(10 min),30(30 min),60(1-hour) and $R_{i,t}$ represents one minute return for 10 minute,30 minute,1 hour horizon, and 30 minute return for one-day horizon respectively.

4. Experimental Results

From the summary statistics of the four different horizon return and volatility and the trend of this two time series. We find that all the returns and volatility reject the hypothesis of normal distribution. Besides, most of the series exhibit the positive skewness and high kurtosis. Volatility clustering are founded on this series.

In each GP-runs, we can find a volatility formula. Eg. figure 1 illustrates the 10 minute horizon GP-Run3 tree. Then ,we can obtain the best 10-minute horizon GP-Run3 volatility formula as equation (5)

$$\sigma^2_{GP-Run3} = \sqrt{R^2_{t-2}R^2_{t-5} + 2R^2_{t-2}R_{t-5} + R^2_{t-2}} \tag{5}$$

Overall the summary statistics for forecasting models. At the 10 minute horizon, HISVOL can outperform all the forecasting models, ANN and GP models are also better than GARCH-type model. EGARCH model has a high performance than GARCH and GJR-GARCH model. Whereas the 30

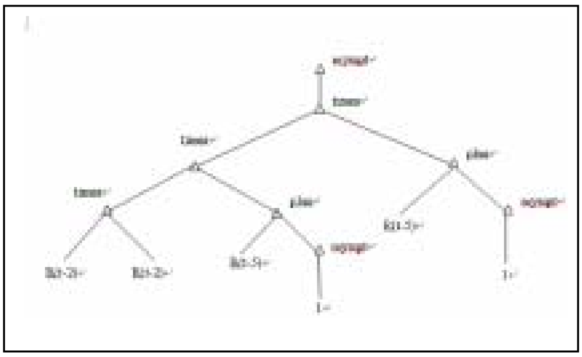

Fig. 1: GP-Run3 volatility tree-10 minute horizon

minute horizon, HISVOL still outperform all the forecasting models, ANN and GP models are also better than GARCH-type model. Yet, GP model has a better average performance than ANN model, EGARCH model has a better performance than GARCH and GJR-GARCH model.

It is worthy to attention that at a longer 1 hour horizon. HISVOL can not still outperform all the forecasting models. The ANN and GP models can beat other models. HISVOL is also better than GARCH-type models.

When the time horizon increase to one-day, ANN and GP models still have a high average performances. However, the HISVOL can not beat all the GARCH-type models.

Overall these forecast comparison, we finds the average performance, ANN model can outperform all the models, GP is ranked as second, then HISVOL ,EGARCH,GARCH and GJR-GARCH model.

We compares the VaR backtesting LR TEST results at the 1%,2.5%,5% significance level respectively and finds that a longer horizon has a high probability of accept the null hypothesis than a short horizon.

A kernel density plot can be considered a refinement of a histogram or frequency plot. In this paper, we evaluates the kernel estimate of the forecast errors density (forecast minus realized volatility) for HISVOL,GARCH(1,1),EGARCH(1,1),GJR-GARCH(1,1),ANN and GP models. Fig. 2 presents one of the four different horizon results. The dotted vertical line denotes zero. The evidences from all the different horizons show that 10 minute horizon, the GP error distribution are closer to zero than those of the forecasting models. However, in 30 minute ,one day horizon kernel density plot respectively. We find GP error distribution is still closer to zero than those of the forecasting models. Besides, the other models locate on the right side of zero. While the 1 hour horizon kernel density plot, the ANN error distribution

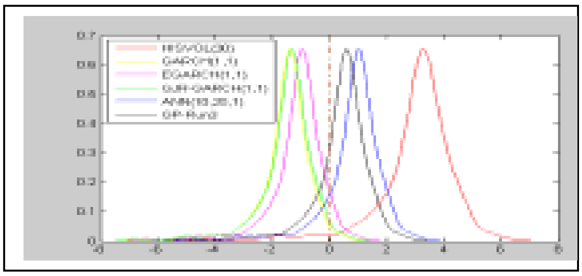

Fig. 2 :The kernel estimates of the errors densities of forecasting models-one day horizon

is still closer to zero than those of the forecasting model. Yet, the other models locate on the two side of zero. To sum up, we find GP volatility forecasting model is superior to other models when judged by kernel density plot in most of different time horizons. ANN model can outperform other models in 1 hour horizon.

5. Concluding Remarks

In this paper, we utilizes two computational intelligences nonparametric algorithms to undergo learning experiences and gather expertise to remember, predict, make decisions for high frequency time series data. We also compares the forecasting performance of several well known parametric volatility models.

Overall these forecast summary statistics show that average performance, ANN model can outperform all the models, GP is the second best model, then HISVOL ,EGARCH,GARCH and GJR-GARCH model.

As mentioned above, a few literatures has shown the results of GP and ANN model in volatility forecasting by inter day data. In particular, our empirical results also show that the GP and ANN do reasonably well in forecasting out sample volatility than other parametric models in most of performance indicators except a more shorter horizon. The reason is that these techniques can learn and represent non-linear relationships, adapt to new environments, and manage uncertainty and imprecision.

An interesting topic for further research might therefore be to consider different frequency financial intraday data. Additionally ,considering the genetic learning algorithms for fuzzy rule or neuro-fuzzy algorithms.

6. References

[1]. Andersen T. G. and T. Bollerslev, “Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts,” *International Economics Review*, 39, pp.885-905,1998.

[2]. Chen,S.-H. and Paul P. Wang (eds.), “Computational Intelligence in Economics and Finance,” Springer.2002.

[3]. Chen,S-H. and C.-H.Yeh., “Using Genetic Programming to Model Volatility in Financial Time Series: The Case of Nikkei 225 and S&P 500,” *Proceedings of the 4th JAFEE International Conference on Investments and Derivatives* (JIC'97), Aoyoma Gakuin University, Tokyo, Japan, July 29-31,pp.288-306,1997.

[4]. Chen,S.–H. and T.-W.,Kuo, “Are Efficient Markets Really Efficient? Can Financial Economic Tests Convince Machine Learning People?,” No.13 of Chen,S.-H. and Paul P. Wang (eds.) “*Computational Intelligence in Economics and Finance*,”Springer.pp.288-296.,2002.

[5]. Kupiec, P.,“Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal of Derivatives*, Winter, pp.73 - 84.,1995.

[6]. Merton, R.C., “On Estimating the Expected Return on the Market: An Exploratory Investigation ,”*Journal of Financial Economics*, 8, pp.323-361,1980.

[7]. Nelson, D. B., “Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model,” *Journal of Econometrics*, 52, pp.61-90,1992.

[8]. Neely ,C. J. and P. A. Weller, (2002), “Using a Genetic Program to Predict Exchange Rate Volatility,” *Genetic Algorithms and Genetic Programming in Computational Finance*,edited by S. H. Chen, Kluwer Academic Publishers ,pp.263-279,2002.

[9]. Poon, S. and C. Granger,“ Forecasting Volatility in Financial Markets: A Review, ” *Journal of Economic Literature*, 41, pp.478-539,2003.

[10]. Sadorsky,P.,“A Comparison of Some Alternative Volatility Forecasting Models For Risk Management,” *Proceedings of the Second International Conference on Financial Engineering and Application*, 2004.