Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Physica A 387 (2008) 1585–1592
PHYSICA A
www.elsevier.com/locate/physa

Forecasting volatility of SSEC in Chinese stock market using multifractal analysis

Yu Wei*, Peng Wang

School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China

Received 8 November 2007
Available online 17 November 2007

Abstract

In this paper, taking about 7 years' high-frequency data of the Shanghai Stock Exchange Composite Index (SSEC) as an example, we propose a daily volatility measure based on the multifractal spectrum of the high-frequency price variability within a trading day. An ARFIMA model is used to depict the dynamics of this multifractal volatility (MFV) measures. The one-day ahead volatility forecasting performances of the MFV model and some other existing volatility models, such as the realized volatility model, stochastic volatility model and GARCH, are evaluated by the superior prediction ability (SPA) test. The empirical results show that under several loss functions, the MFV model obtains the best forecasting accuracy.

Keywords: Econophysics; Multifractal; Realized volatility; Stochastic volatility model; GARCH; Superior predictive ability

1. Introduction

Modeling and forecasting volatility in financial markets is a key issue in many important fields, such as derivative products pricing, portfolio allocation and risk measurement. The seminal paper of Engle [1] has paved the way for the development of a large number of so-called historical volatility models in which a time varying volatility process is extracted from financial return data. Many of these models can be regarded as variants of the generalized autoregressive conditional heteroskedasticity (GARCH) models [2]. A rival class for ARCH is associated with the stochastic volatility (SV) models [3].

Both GARCH and SV models are regularly used for the analysis of daily, weekly and monthly returns. However the recent widespread availability of intraday high-frequency prices of financial assets and the work done on them have shed new light on the concept of volatility: as a matter of fact, data sampled at regular intradaily intervals can be summarized into a measure called realized volatility (RV) which, under some assumptions, is a consistent estimator of the quadratic variation of the underlying diffusion process [4]. In principle, the volatility measures derived from high-frequency data should prove to be more accurate, hence allowing for forecast efficiency gains. Nevertheless, recently Ref. [5] shows that realized volatility is prone to all sorts of microstructure problems.

* Corresponding author.
E-mail address: ywei@home.swjtu.edu.cn (Y. Wei).

doi:10.1016/j.physa.2007.11.015

Since the suggestion of Mandelbrot [6] that multifractal is a powerful tool for depicting volatility complexities in financial markets, much research has been done in this field. However most of these studies focus on empirical tests of multifractality in different financial data sets. So we wonder whether multifractal analysis can contribute to the measurement and forecasting accuracy of volatility in financial markets. Taking high-frequency data of SSEC index in Chinese stock market as an example, first we propose a so-called multifractal volatility (MFV) measure based on the multifractal spectrum of high-frequency price movements within one trading day. Second similar to realized volatility, we also propose an ARFIMA process to model the dynamics of MFV and use a rolling-window method to forecast the volatility of SSEC one day ahead. Finally, we use a formal test for superior prediction ability (SPA) proposed by Ref. [7] to evaluate the forecasting performance of the MFV model and compare it to other popular volatility models, such as RV, SV and GARCH models. The empirical results show that under several loss functions, i.e., mean square error adjusted for heteroskedasticity (HMSE) and mean absolute error adjusted for heteroskedasticity (HMAE), the MFV model obtains the best forecasting accuracy.

This paper is organized as follows. In the next section, we introduce the sample data and discuss how daily and intraday returns are constructed. In Section 3, we discuss how realized volatility is derived from intraday returns and the ARFIMA model for RV. In Section 4, we introduce the calculation of the multifractal volatility measure from the multifractal spectrum of high-frequency price movements within one trading day. In Section 5, the historical volatility models are briefly described. The out-of-sample forecasting methodology and SPA test are discussed in Section 6, and in Section 7, the estimation and forecasting results are presented. Section 8 summarizes the conclusions.

2. The data, daily and intraday returns

The data for our empirical study consists of high-frequency (every 5 min) price quotes of the Shanghai Stock Exchange Composite Index (SSEC), the most important stock index in the Chinese stock market, during the period from 19 January 1999 to 30 December 2005, which contains totally $N = 1670$ trading days. The Shanghai Stock Exchange is open from 9:30 a.m. to 11:30 a.m. and then from 1:00 p.m. to 3:00 p.m., so there are 4 trading hours in a trading day, and there are 48 quotes (per 5 min) of index in a day (excluding the open price). The 5-min price quotes are denoted as $I_{t,d}$, $t = 1, 2, \ldots, N$ and $d = 0, 1, 2, \ldots, 48$. $I_{t,0}$ denotes the open price on day t and $I_{t,48}$ the close price quote. Here the daily return R_t is defined as

$$R_t = 100(\ln I_{t,48} - \ln I_{t-1,48}), \tag{1}$$

and the intraday high-frequency return $R_{t,d}$ is defined as

$$R_{t,d} = 100(\ln I_{t,d} - \ln I_{t,d-1}), \quad d = 1, 2, \ldots, 48. \tag{2}$$

3. Realized volatility and the ARFIMA model

It is generally accepted that squared daily returns provide a poor approximation of actual daily volatility. Ref. [4] first points out that more accurate estimates can be obtained by summing squared intraday returns. If we would apply their method directly in this paper, we would define realized volatility as

$$RV_t' = \sum_{d=1}^{48} R_{t,d}^2. \tag{3}$$

However, this definition ignores the information obtained in the overnight returns. In order to account for this problem, Ref. [8] suggests scaling the realized volatility in this way:

$$RV_t = \gamma RV_t', \tag{4}$$

where the so-called scale parameter γ is defined as

$$\gamma = \frac{N^{-1} \sum_{t=1}^{N} R_t^2}{N^{-1} \sum_{t=1}^{N} RV_t'}. \tag{5}$$

In empirical work on realized volatility [9], it is pointed out that the natural logarithms of realized volatility series, denoted as lnRV, can be modeled by a Gaussian dynamic process. Ref. [9] adopts the autoregressive fractionally integrated moving average (ARFIMA) to model this process.

The ARFIMA$(1, d, 1)$ model with mean μ for realized volatility can be given by

$$(1-\phi L)(1-L)^d(RV_t-\mu)=(1+\theta L)\varepsilon_t, \tag{6}$$

where L is the lag operator, coefficients d, ϕ and θ are fixed and unknown and ε_t is Gaussian white noise with mean zero and variance σ_ε^2. Since different ARFIMA model specifications produce rather similar results [10], we also consider the ARFIMA$(1, d, 1)$ model in the following empirical study.

4. Multifractal volatility measure and its model

Some recent works find that the multifractal spectrum of the high-frequency price movements within a trading day contains valuable volatility information of the assets [11–13]. So, in this section, we discuss how to construct a so-called multifractal volatility (MFV) measure from the daily multifractal spectrum and how to model the multifractal volatility series.

According to the method previously proposed [11–13], we also use the box-counting method to calculate the multifractal spectrum of SSEC. To be clear, here we denote the high-frequency price quotes of SSEC, excluding the open price, at time t as $I(t)$ and t running from 1 to $M = 48 \times 1670 = 80, 160$.

The index variation with time can be divided into many normalized boxes (time intervals) of size δ ($\delta < 1$), for example, in the case the box size can be 1/48, 1/24, 1/16, 1/12, 1/8, 1/6, 1/4, 1/3, 1/2 and 1.

Suppose we need m boxes to cover the 48, 5-min quotes in a trading day, $I(t)$, $t = 1, 2, \dots, 48$, and there are n quotes in each boxes. P_i is the average probability in the box i, then

$$P_i(\delta)=\frac{\sum_{j=1}^{n} I(i_j)}{\sum_{t=1}^{48} I(t)}, \quad i=1,2,\dots,m, \tag{7}$$

where I_{i_j} is the jth quote in the ith box in the trading day.

Then we can describe it to be multifractal as

$$P_i(\delta)\sim\delta^{\alpha}, \tag{8}$$

$$N_\alpha(\delta)\sim\delta^{-f(\alpha)}, \tag{9}$$

where α is the singularity or *Hölder* exponent of the subset of probabilities, $N_\alpha(\delta)$ the number of boxes of size δ with the same probability, and $f(\alpha)$ the fractal dimension of the α subset.

What is important for us is the statistical information about index fluctuations contained in multifractal spectra $f(\alpha)$, and we may use the partition function $S_q(\delta)$ to calculate it. The partition function $S_q(\delta)$ is defined and expressed as a power law of δ with an exponent $\tau(q)$, where q is the moment order ($-\infty < q < \infty$):

$$S_q(\delta)=\sum_{i=1}^{m} P_i^q(\delta)\sim\delta^{\tau(q)}. \tag{10}$$

In our calculation, $|q|_{\max}$ is 120, and $\tau(q)$ can be obtained from the slope of the linear part of $\ln S_q(\delta)$–$\ln\delta$ curve. Then $f(\alpha)$ can be obtained by performing Legendre transformation as follows:

$$\alpha=\frac{\mathrm{d}\tau(q)}{\mathrm{d}q}, \tag{11}$$

$$f(\alpha)=\alpha q-\tau(q). \tag{12}$$

Therefore the width of the multifractal spectrum can be expressed as $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$, where $\alpha_{\max}$ and $\alpha_{\min}$ are the maximum and minimum of α. Since the probability of each box is $P_i(\delta) \sim \delta^\alpha$ and $\delta < 1$, $\alpha_{\min}$ and $\alpha_{\max}$ denote

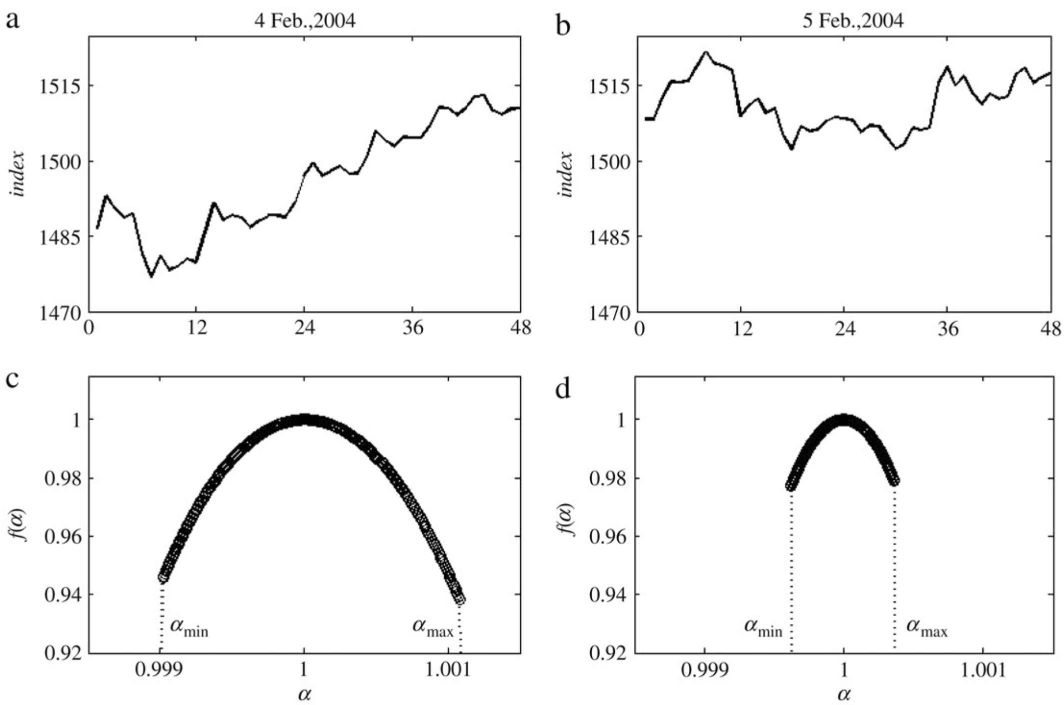

Fig. 1. The time dependence of the indexes of 4 Feb., 2004 (a); and 5 Feb., 2004 (b); the multifractal spectra of 4 Feb., 2004 (c); and 5 Feb., 2004 (d).

the values of the maximum and minimum probabilities. In other words, $\alpha_{\min}$ indicates the highest "price level" in that trading day and $\alpha_{\max}$ indicates the lowest "price level" in that trading day. Thus the larger the value of $\Delta\alpha$, the more violent the price fluctuation in the trading day, and $\Delta\alpha$ can be treated as a measure of daily volatility.

Fig. 1(a) and (b) show the high-frequency price movements of SSEC in two sequent trading days. Fig. 1(c) and (d) are the multifractal spectra of the two days, and the maximum *Hölder* exponent and the minimum one of each multifractal spectrum are $\alpha_{\max}$ and $\alpha_{\min}$. Visibly, the price varies more violently on 4 February, 2004 than on 5 February, 2004, so $\Delta\alpha$ of the first day is larger than that of the second day.

To make $\Delta\alpha$ comparable to realized volatility and other historical volatility measures, similar to the means of scaling realized volatility in Section 3, we formally define multifractal volatility measure (MFV) for day t as

$$MFV_t = \beta \Delta\alpha_t, \tag{13}$$

where the scale parameter β is defined as

$$\beta = \frac{N^{-1} \sum_{t=1}^{N} R_t^2}{N^{-1} \sum_{t=1}^{N} \Delta\alpha_t}. \tag{14}$$

As discussed in Section 3, realized volatility is computed by using every high-frequency data within the 48 intradaily quotes only about twice as illustrated by Eqs. (2)–(5). However MFV is constructed by $\Delta\alpha$, where every α is computed by using the box-counting method. To get the value of $\Delta\alpha$, every high-frequency data is used many times as illustrated by Eqs. (7)–(14). At this point, MFV may make fuller use of the statistical information in every high-frequency data than RV. Therefore we expect that MFV is a more accurate measurement of daily volatility than RV. Furthermore, the difference between MFV and other existing volatility measures is that RV, SV and GARCH model daily volatility directly by using daily or intradaily price (return) data, however MFV is constructed by indirectly using a kind of probability of multifractal measurement.

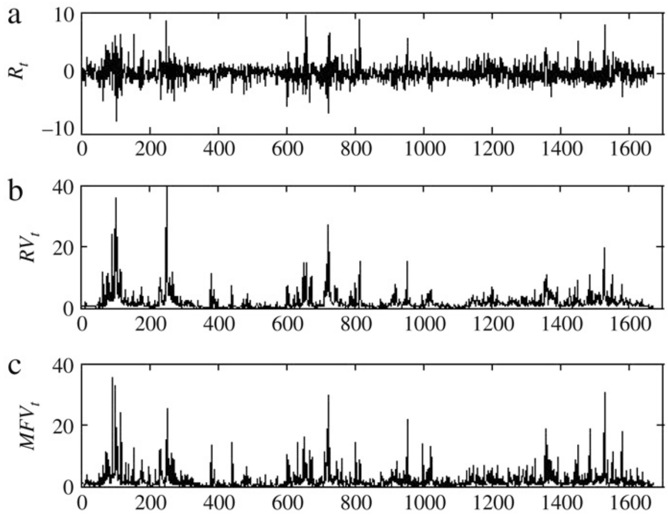

Fig. 2. The time dependence of daily returns of SSEC between 19 January 1999 and 30 December 2005 (a), the realized volatility (b) and multifractal volatility measures (c).

Fig. 2(a) gives the time dependence of daily returns of SSEC between 19 January 1999 and 30 December 2005, totally 1670 trading days. Fig. 2(b) and (c) give realized volatility and multifractal volatility measures respectively.

Fig. 2 shows that the multifractal volatility measure is quite similar to realized volatility measures, especially when there are large volatilities of daily returns. We further consider that many time series models, such as AR, MA and ARMA, can be treated as special forms of the ARFIMA model. So similar to realized volatility model in Section 3, we also propose an ARFIMA(1, d, 1) model with mean μ for the natural logarithms of multifractal volatility measures, denoted as lnMFV.

5. Historical volatility models

5.1. GARCH model

The most popular historical volatility model is generalized autoregressive conditional heteroskedasticity (GARCH) model. In our empirical investigation, the simple but effective GARCH(1, 1) model for daily returns is given by

$$\begin{aligned} R_t &= \sigma_t \varepsilon_t, \quad \varepsilon_t \sim NID(0, 1), \\ \sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \end{aligned} \tag{15}$$

with parameter restrictions $\omega > 0, \alpha \geq 0, \beta \geq 0$ and $\alpha + \beta \leq 1$.

5.2. Stochastic volatility model

The stochastic volatility model is an alternative model to GARCH and for daily returns it is given by

$$\begin{aligned} R_t &= \sigma_t \varepsilon_t, \quad \varepsilon_t \sim NID(0, 1), \\ \sigma_t^2 &= \sigma^{*2} \exp(h_t), \\ h_t &= \phi h_{t-1} + \sigma_\eta \eta_{t-1}, \quad \eta_t \sim NID(0, 1), \quad h_1 \sim NID(0, \sigma_\eta^2/\{1 - \phi^2\}), \end{aligned} \tag{16}$$

where σ^{*2} may be treated as a scale parameter, the persistence parameter ϕ is restricted to be positive and smaller than one to ensure stationarity. ε_t and η_t are assumed mutually uncorrelated.

6. Forecasting methodology and SPA test

For the empirical results given in the next section, we evaluate the forecasting performances of the four different volatility models described in Sections 3–5. The forecasting study is handled as follows. Each volatility model is estimated 420 times based on 420 estimation samples that include 1250 observations (the total sample includes 1670 observations). The first estimation sample includes the observations from $t = 1$ to 1250. Based on this estimation, a one-day ahead volatility forecast is made for time $t = 1251$. The second estimation sample starting at $t = 2$ and ending at $t = 1251$, is used to forecast the volatility at time $t = 1252$. Similarly, this process can be repeated 420 times and produces 420 one-day ahead volatility forecasts $\hat{\sigma}_m^2$, where $m = 1251, 1252, \ldots, 1670$.

Since the true volatility is not observable, the realized volatility measure in Eq. (4), denoted as RV_t is generally accepted to be an accurate estimator, and can therefore be used as a standard to evaluate the volatility forecasting performances for various volatility models [10]. Furthermore various forecasting criteria can be considered to assess the predictive accuracy of a volatility model. In this paper, we use 4 different accuracy statistics or loss functions, L_i, $i = 1, 2, 3$ and 4, as forecasting criteria:

$$L_1 : \text{MSE} = M^{-1} \sum_{m=H+1}^{H+M} (RV_m - \hat{\sigma}_m^2)^2, \tag{17}$$

$$L_2 : \text{MAE} = M^{-1} \sum_{m=H+1}^{H+M} |RV_m - \hat{\sigma}_m^2|, \tag{18}$$

$$L_3 : \text{HMSE} = M^{-1} \sum_{m=H+1}^{H+M} (1 - \hat{\sigma}_m^2 / RV_m)^2, \tag{19}$$

$$L_4 : \text{HMAE} = M^{-1} \sum_{m=H+1}^{H+M} |1 - \hat{\sigma}_m^2 / RV_m|, \tag{20}$$

where $H = 1250$, $M = 420$ and so $H + M = 1670$. MSE and MAE are the most popular loss functions in this comparison. HMSE and HMAE are MSE and MAE adjusted for heteroskedasticity.

The usual forecast comparison is based on different loss functions computed just in one single sample. However, the fact that a particular loss criterion is smallest for a particular model does not provide any information about its forecast superiority in other samples of the data set and in future samples of the data. Much works has focused on a testing framework for determining whether a particular model is better than another model [14]. The results in Ref. [14] and the important refinements in Ref. [15] constitute a framework that constructs a formal test for the superior prediction ability (SPA) of a benchmark or base model relative to a set of rival models.

The SPA test can be used for comparing the performances of two or more forecasting models. The forecasts are evaluated using a pre-specified loss function, and the "best" forecast model is the model that produces the smallest expected loss. The key method of computing SPA test statistic and its p-value requires bootstrap samples obtained by, for example, the stationary bootstrap procedure discussed in Ref. [16]. The technical details of SPA can be found in Refs. [7,10,15]. In summary, the p-value of a SPA test indicates the relative performance for a base model ($\mathbf{M}_0$) to an alternative model ($\mathbf{M}_k$). The higher the p-value, the better the forecasting performance of the base model $\mathbf{M}_0$. In particular, a p-value of about 1 can be treated as strong evidence that the base model is superior to the alternative ones.

7. Empirical results

7.1. *Some forecasting results*

The one-day ahead volatility forecasts for different models are constructed using the rolling forecasting methodology discussed in Section 6. To be clear, Fig. 3 shows only one piece of the forecasting results ($t = 1350, 1351, \ldots, 1550$, about 200 days) in which the realized volatility measures as the forecasting benchmark are displayed as dots. Fig. 3(a) shows the forecasting results of ARFIMA-lnRV and ARFIMA-lnMFV models and Fig. 3(b) the forecasting results of SV and GARCH models.

Visibly, GARCH and SV make higher volatility forecasts than ARFIMA-lnRV and ARFIMA-lnMFV as a whole. However to get reliable and robust evaluation, we need the results from SPA tests.

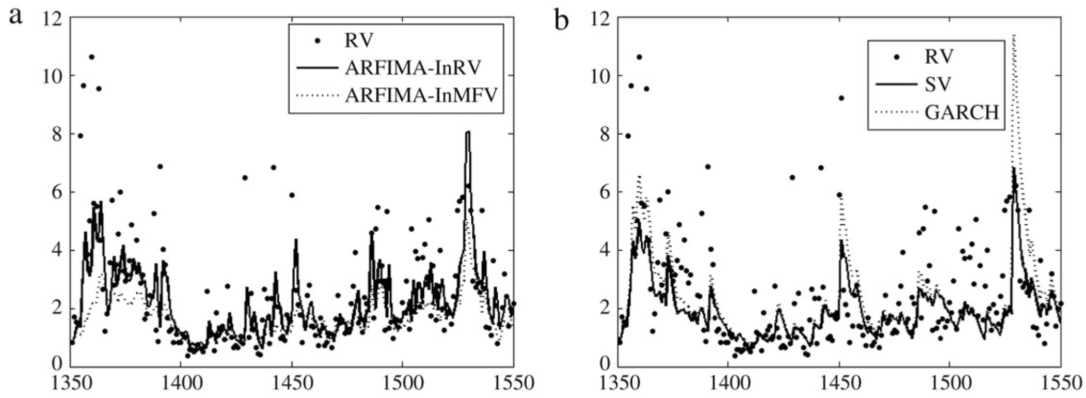

Fig. 3. One-day ahead volatility forecasts from ARFIMA-lnRV and ARFIMA-lnMFV (a); SV and GARCH (b) and realized volatility measures (as dots).

Table 1
Superior predictive ability (SPA) test for various volatility models

Loss function	Base model $\mathbf{M}_0$	Alternative models $\mathbf{M}_k$			
		ARFIMA-lnMFV	ARFIMA-lnRV	SV	GARCH
MSE	ARFIMA-lnMFV	–	0.007	0.076	0.082
	ARFIMA-lnRV	0.994	–	0.935	0.867
	SV	0.924	0.065	–	0.170
	GARCH	0.918	0.134	0.830	–
MAE	ARFIMA-lnMFV	–	0.035	0.415	0.740
	ARFIMA-lnRV	0.965	–	0.971	0.993
	SV	0.586	0.030	–	0.907
	GARCH	0.261	0.007	0.093	–
HMSE	ARFIMA-lnMFV	–	1.000	1.000	1.000
	ARFIMA-lnRV	0.001	–	0.940	1.000
	SV	0.000	0.060	–	1.000
	GARCH	0.000	0.001	0.000	–
HMAE	ARFIMA-lnMFV	–	0.993	1.000	1.000
	ARFIMA-lnRV	0.007	–	0.980	1.000
	SV	0.000	0.020	–	1.000
	GARCH	0.000	0.000	0.000	–

7.2. SPA tests

In this empirical test, we use 2000 times of bootstraps to obtain a SPA p-value between two competitive models. Table 1 presents the SPA results for various volatility models. The first column in Table 1 lists the four different loss functions and the second column gives the names of base model $\mathbf{M}_0$. The numbers in the table are the SPA p-values. As noted above, the higher the p-value, the better the forecasting performance of the base model ($\mathbf{M}_0$) than that of the alternative ones ($\mathbf{M}_k$).

Table 1 shows that the ARFIMA-lnRV model performs quite well in most of the cases, especially when it is compared with SV and GARCH models. When the loss functions of HMSE and HMAE are taken into account, ARFIMA-lnMFV outperforms all the other alternative models with almost all p-values of 1 (only a p-value of 0.993 with loss function of HMAE to ARFIMA-lnRV model). This result may indicate that multifractal volatility measures are more accurate and sensitive to the market variability and especially suitable for depicting the volatility heteroskedasticity. Table 1 also shows that the SV model does better volatility forecasts than does the GARCH model except for MSE as the loss function. Generally the volatility model base on high-frequency data, i.e., ARFIMA-lnRV and ARFIMA-lnMFV produce better volatility forecasts than those based on daily data, i.e., SV and GARCH model. This phenomenon may indicate the fact that there is more valuable information about market volatility in high-frequency data than in low-frequency data. Therefore volatility models constructed with high-frequency data should certainly be the direction for future research.

8. Conclusions

In this paper, taking about 7 years' high-frequency data of Shanghai Stock Exchange Composite Index (SSEC) as an example, we propose a daily volatility measure based on multifractal spectrum of the high-frequency price variability within a trading day. An ARFIMA$(1, d, 1)$ model is also constructed to depict the dynamics of the so-called multifractal volatility (MFV) measure. To testify the efficiency of the MFV measures, we compare the one-day ahead volatility forecasting performance of the MFV model with other three popular models, i.e., realized volatility model, SV and GARCH. The empirical results of the superior prediction ability (SPA) test show that the ARFIMA-lnMFV model outperforms all the other alternative models when the loss functions of HMSE and HMAE are taken into account. Furthermore, we find that volatility models based on high-frequency data, RV and MFV models, produce better volatility forecasts than those models based on daily data. These results suggest that the multifractal analysis of high-frequency data of financial assets may produce much valuable statistical information on volatilities and their dynamical characteristics. This information may help in the further research on derivative products pricing, portfolio allocation and financial risk management.

Acknowledgements

The author would like to acknowledge the referee for his valuable comments and Dr. P.R. Hansen for sharing the OX codes for SPA. The author also thanks the National Natural Science Foundation of China (70501025, 70771097) for providing financial support for this work.

References

[1] R.F. Engle, Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation, Econometrica 50 (1982) 987–1007.
[2] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31 (1986) 307–327.
[3] S.J. Taylor, Modeling Financial Time Series, John Wiley and Sons, Chichester, 1986.
[4] T.G. Andersen, T. Bollerslev, Answering the skeptics: yes, standard volatility models do provide accurate forecasts, International Economic Review 39 (1998) 885–905.
[5] R.F. Engle, G.M. Gallo, A multiple indicator model for volatility using intra-daily data, Journal of Econometrics 131 (2006) 3–27.
[6] B.B. Mandelbrot, A multifractal walk down Wall Street, Scientific American 298 (1999) 70–73.
[7] P.R. Hansen, A. Lunde, A forecast comparison of volatility models: does anything beat a GARCH(1, 1)? Journal of Applied Econometrics 20 (2005) 873–889.
[8] P.R. Hansen, A. Lunde, Consistent ranking of volatility models, Journal of Econometrics 131 (2006) 97–121.
[9] T.G. Andersen, T. Bollerslev, F.X. Diebold, H. Ebens, The distribution of realized stock return volatility, Journal of Financial Economics 61 (2001) 43–76.
[10] S.J. Koopman, B. Jungbacker, E. Hol, Forecasting daily variability of the S&P100 stock index using historical, realized and implied volatility measurements, Journal of Empirical Finance 12 (2005) 445–475.
[11] X. Sun, H. Chen, Z. Wu, Y. Yuan, Multifractal analysis of Hang Seng index in Hong Kong stock market, Physica A 291 (2001) 553–562.
[12] X. Sun, H. Chen, Y. Yuan, Z. Wu, Predictability of multifractal analysis of Hang Seng stock index in Hong Kong, Physica A 301 (2001) 473–482.
[13] Y. Wei, D. Huang, Multifractal analysis of SSEC in Chinese stock market: A different empirical result from Heng Seng index, Physica A 355 (2005) 497–508.
[14] H. White, A reality check for data snooping, Econometrica 68 (2000) 1097–1126.
[15] P.R. Hansen, A test for superior predictive ability, Journal of Business and Economic Statistics 23 (2005) 365–380.
[16] D.N. Politis, J.P. Romano, The stationary bootstrap, Journal of the American Statistical Association 89 (1994) 1303–1313.