2011 International Conference on Economics and Finance Research
IPEDR vol.4 (2011) © (2011) IACSIT Press, Singapore

Forecasting Volatility of Chinese Composite Index Based on Empirical Mode Decomposition and Neural Network

Jingfeng Xu
China Institute for Actuarial Science
Central University of Finance and Economics
Beijing, China
Email: jingfeng.caida@gmail.com

Jian Liu
School of Banking and Finance
University of International Business and Economics
Beijing, China
Email: liujian8210@gmail.com

***Abstract*—Empirical Mode Decomposition (EMD), recently proposed by Huang et al. [12], appears to be a novel data analysis method for nonlinear and non-stationary time series. By decomposing a time series into a small number of independent and concretely implicational intrinsic modes based on scale separation, EMD explains the generation of time series data from a novel perspective.**

This paper presents an empirical mode decomposition based on neural network learning paradigm (EMD-NN) for forecasting volatilities of Shanghai A shares (Shanghai) and Shenzhen A shares (Shenzhen). By the criteria of some statistic loss functions, EMD-NN outperforms GARCH family models (GARCH, EGARCH, GJR), moving average and neural network in improving predictive accuracy.

Keywords-Empirical Mode Decomposition; GARCH; EGARCH; GJR; Moving average; Neural network.

I. INTRODUCTION

Modeling volatility plays a central role in the booming financial markets since it is important for pricing derivatives, calculating measure of risk and hedging. A large number of time series based volatility models have been developed since the introduction of ARCH model of Engle [5]. So far in the literature, the predominant model of the past is the GARCH model by Bollerslev [2], who generalizes the seminal idea on ARCH by Engle [5]. The popularity of the GARCH model is due to its ability to capture volatility clustering and leptokurtosis. To improve the forecasting ability, many nonlinear extensions of the GARCH model have been proposed. For instance, Nelson [13] introduced the Exponential GARCH (EGARCH) model in order to model asymmetric variance effects. Another asymmetric model is the GJR model introduced by Glosten et al. [8], which has been designed to capture the leverage effect between the asset return and the volatility. Other models such as APARCH, AGARCH, TGARCH and QGARCH models have also been developed (by Ding, Granger and Engle [3], Engle [6], Zakoian [17] and Sentana [14]) for the flexibility of the model.

It is difficult to identify above models which is superior in volatility forecasting because asset returns often do not contain sufficient information and the conditional variance is unobserved. To obtain more accurate predictions, semi-parametric approaches such as Neural Network (NN) have been successfully shown for modeling and forecasting volatility (Taylor [15]; Dunis and Huang [4]; Ferland and Lalancette [7]; Bildirici and Ersin [1]). The main reason for using the neural network is its ability to best approximate any nonlinear functions arbitrarily without a priori assumptions on data distribution (Haykin [9]). In addition, the neural network can capture the stylized characteristics of financial returns such as leptokurtosis, volatility clustering, and leverage effects.

In recent years, a new method, Empirical Mode Decomposition (EMD) introduced by Huang et al. [10], has now been applied to financial data. The core of EMD is to decompose data into a small number of independent and nearly periodic intrinsic modes based on local characteristic scale, which is defined as the distance between two successive local extrema in EMD. Each derived intrinsic mode is dominated by scales in a narrow range. This decomposition method is adaptive, and, therefore, highly efficient. Since the decomposition is based on the local characteristic time scale of the data, it is applicable to non-linear and non-stationary processes. Huang et al. [11], [12] used the EMD approach as a filtering technique to measure the volatility of the market. Yu et al. [16] presented an EMD based on ANN learning paradigm in crude oil price. They decomposed the crude oil price series into intrinsic mode functions (IMFs), and used a neural network model to predict the tendencies of each IMF so that formulating an ensemble output for the original crude oil price series.

In this paper, motivated by the work in Yu et al. [16], we present an empirical mode decomposition based neural network learning paradigm (EMD-NN) for forecasting volatilities of Shanghai A shares (Shanghai) and Shenzhen A shares (Shenzhen). Our goal here is to compare the proposed EMD-NN with other competitive approaches including GARCH family models (GARCH, EGARCH, GJR), moving average, ANN for forecasting the volatility of China stock market.

The rest of this paper is organized as follows. We will introduce the theory of EMD in section 2. Section 3 is devoted to the introduction of other volatility models. In section 4 we consider the statistical criteria employed in time-series analysis to assess the adequacy of models, and present the data collected for this experiment. Section 5 discusses the volatility forecasting performance of all models for the real data. Section 6 concludes our paper.

II. EMPIRICAL MODE DECOMPOSITION

Empirical mode decomposition (EMD) is a method developed by Huang which is adaptive and appears to be suitable for non-linear and non-stationary signal analysis. It is based on the simple assumption that any signal consists of different simple intrinsic modes of oscillations. EMD can extract these intrinsic modes from the original time series, based on the local characteristic scale of data itself, and represent each intrinsic mode as an intrinsic mode function (IMF). Each IMF must satisfy two basic conditions (Huang et al. [10]):

1. In the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one. Note, either local minima or local maxima are extrema.

2. At any point, the mean value of the envelope, one defined by the local maxima (upper envelope) and the other by the local minima (lower envelope) is zero.

Based on this definition, IMFs can be extracted from the data series according to the following sifting procedure:

1) Identify all the local extrema, including local maxima and local minima, of time series $x(t)$.

2) Generate its upper and lower envelopes, $e_{\min}(t)$ and $e_{\max}(t)$, with cubic spline interpolation.

3) Calculate the point-by-point mean $m(t)$ from upper and lower envelopes:

$$m(t) = \frac{e_{\min}(t) + e_{\max}(t)}{2}.$$

4) Extract the mean from the time series and define the difference of $x(t)$ and $m(t)$ as $d(t)$:

$$d(t) = x(t) - m(t).$$

5) Check properties of $d(t)$: a. If it is an IMF, denote $d(t)$ as the i th IMF and replace $x(t)$ with the residual $r(t) = x(t) - d(t)$. The i th IMF is often denoted as $c_i(t)$ and the i is called its index; b. If it is not, replace $x(t)$ with $d(t)$;

6) Repeat above steps until the residual satisfies some stopping criterion.

The sifting process can be stopped when it satisfies any of the following predetermined criteria: 1) either the component $d(t)$ or the residue $r(t)$ becomes so small that it is less than the predefined value of the substantial consequence; 2) the residue $r(t)$ becomes a monotonic function from which no more IMF can be extracted.

The maximum number of IMFs from a data series of length N is $\log_2^N$. The resulting set of IMFs $c_i(t)$, and the final residue $r(t)$, are linearly related to the original data series, $x(t)$, as

$$x(t) = \sum_{i=1}^{n} c_i(t) + r(t),$$

Where n is the number of IMFs.

The definition of the EMD method implies that the extracted IMF components are orthogonal, i.e. linearly independent. However, as noted by Huang et al. [10] orthogonality depends on the decomposition method, and is a requirement of linear decomposition systems, where as the EMD is a non-linear method. Still Huang et al. [10] provided expressions for calculating local orthogonality indexes and defined corresponding IMFs orthogonal when the index is less than or equal to 0.1.

The volatility in Huang et al. [11], [12] is defined as the ratio of the absolute value of the IMF component(s) to the signal at any time:

$$\sigma_t = \frac{x_h(t)}{x(t)},$$

where $x_h(t) = \sum_{j=1}^{h} c_j(t)$. The resulting volatility is a function of time and h shows over what time scale it is applicable. The unit of this variability parameter is the fraction of the market value. It is a simple and direct measure of the market volatility.

The advantages of EMD can be briefly summarized as follows: First, it can reduce any data, from non-stationary and nonlinear processes, into simple independent intrinsic mode functions; Second, since the decomposition is based on the local characteristic time scale of the data and only extrema are used in the sifting process, it is local, self-adaptive, concretely implicational and highly efficient.

III. MODEL FRAMEWORK

The variable we analyze is the daily financial return, y_t, converted from the corresponding price, p_t, using the continuously compound transformation

$$y_t = \log(p_t) - \log(p_{t-1}).$$

In the modeling of the conditional volatility, it is informative to consider the conditional mean and variance of y_t given conditional on the information set, $\mathcal{F}_{t-1}$, that is,

$$\mu_t = E(y_t \mid \mathcal{F}_{t-1}), \quad \sigma_t^2 = Var(y_t \mid \mathcal{F}_{t-1}).$$

A. Moving average model

Normally, we expect that new data carry more information. To reflect this possibility, a moving average method is employed. The choice of the moving average estimation period is arbitrary. In this paper, 5 days is set. The moving average model can be expressed as

$$\sigma_t^2=\frac{1}{5}\sum_{i=1}^{5}\sigma_{t-i}^2 .$$

B. GARCH model

The GARCH (p, q) model has generally been found to be the most appropriate of the standard ARCH family of models. We begin with a brief review of the GARCH (p, q) model developed by Bollerslev [2] for the simulation of returns.

$$y_t=\sigma_t\epsilon_t,\quad \sigma_t^2=\alpha_0+\sum_{i=1}^{p}\alpha_i y_{t-i}^2+\sum_{j=1}^{q}\beta_j\sigma_{t-j}^2,$$

where $\{\epsilon_t\}$ is a sequence of independently and identically distributed random variables with mean zero and unit variance, $\alpha_0>0,\alpha_i\ge 0,\beta_j\ge 0$, $\sum_{i=1}^{max(p,q)}(\alpha_i+\beta_j)<1$ guarantees the non-negativity of variance.

C. EGARCH model

As with ARCH and GARCH models, parameters of the conditional variance are subject to non-negativity constraints. An alternative way of describing the asymmetry in variance is the use of the EGARCH (exponential GARCH) model proposed by Nelson [13]. The EGARCH (p, q) model is defined as

$$\log(\sigma_t^2)=\omega+\sum_{i=1}^{q}\alpha_i g(z_{t-i})+\sum_{j=1}^{p}\gamma_j\log(\sigma_{t-j}^2),$$

where

$$g(z_t)=\theta z_t+\gamma(|z_t|+E|z_t|),\, z_t=\epsilon_t/\sigma_t,$$

θ and γ are real constants.

D. GJR model

Another volatility model commonly used to handle leverage effects is the GJR model. See Glosten, Jagannathan, and Runkle [8]. A GJR (p, q) model assumes the form

$$\sigma_t^2=\alpha_0+\sum_{i=1}^{p}(\alpha_i+\lambda_i N_{t-i})y_{t-i}^2+\sum_{j=1}^{q}\beta_j\sigma_{t-j}^2$$

where N_{t-i} is an indicator for negative y_{t-i} , that is, if $y_{t-i}<0$, $N_{t-i}=1$; Otherwise, $N_{t-i}=0$. α_i , λ_i and β_j are non-negative parameters satisfying conditions similar to those of GARCH models.

E. Neural Network

The neural network model used in this study is the multilayer feed-forward network which is the most basic and commonly used one in economic and financial applications. We specify that the feed-forward network has three hidden layers of sigmoid neurons followed by an output layer of linear neurons, each using a tan-sigmoid differentiable transfer function to generate the output. We choose the fast training Levenberg-Marquardt algorithm as the training algorithm. The value of the learning rate parameter is set to be 0.05.

F. EMD-NN Model

The original price series are first decomposed into a finite, and often small, number of intrinsic mode functions (IMFs). Then a three-layer feed-forward neural network (FNN) model is used to model each of the extracted IMFs, so that the tendencies of these IMFs could be accurately predicted. The volatility of price series is defined as

$$\sigma_t=\frac{\sum_{j=0}^{h}c_j(t)}{y_t},$$

where $c_j(t)$ is defined as in above Section.

IV. Forecasting Scheme and Evaluation

In the evaluation of volatility forecast of real date, the actual volatility is not directly observable and hence it has to be estimated. In this paper, we decide to measure σ_t as the squared difference between the return and its mean value, that is

$$\sigma_t^2=(y_t-\bar{y}_t)^2,$$

where $\bar{y}_t$ is the mean of returns.

We evaluate the forecasting performance using three standard statistical criteria: mean absolute forecast error (MAE), mean squared error (RMSE) and the hit rate (HR), expressed as follows:

$$MAE=\frac{1}{n}\sum_{t=1}^{N}|\sigma_t^2-\hat{\sigma}_t^2|,$$

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

$$HR(\%)=\frac{100}{n}\sum_{t=1}^{N}q_t,$$

$$q_t=\begin{cases}1,(\sigma_{t+1}^2-\sigma_t^2)(\hat{\sigma}_{t+1}^2-\hat{\sigma}_t^2)\ge 0\\ 0,otherwise\end{cases}$$

where N represents the total number of the predicted values. $\hat{\sigma}_t^2$ denotes the predicted conditional variance. σ_t^2 represents the actual volatility. Both MAE and RMSE measure the average magnitude of forecasting error without considering their direction, but RMSE is more useful when large errors are particularly undesirable. The smaller the values of them, the closer are the predicted value to the actual ones. HR measures how often the model gives the correct direction of change of volatility. The larger is the value of HR, the better is the performance of prediction.

The data in this study is comprised of the daily stock closing price indices of the Shanghai A shares (Shanghai) and Shenzhen A shares (Shenzhen) from January 2003 to

October 2006. These stock price indices p_t are then transformed into daily returns y_t by 100 times their log difference:

$$y_t = 100 \times (\ln(p_t) - \ln(p_{t-1})).$$

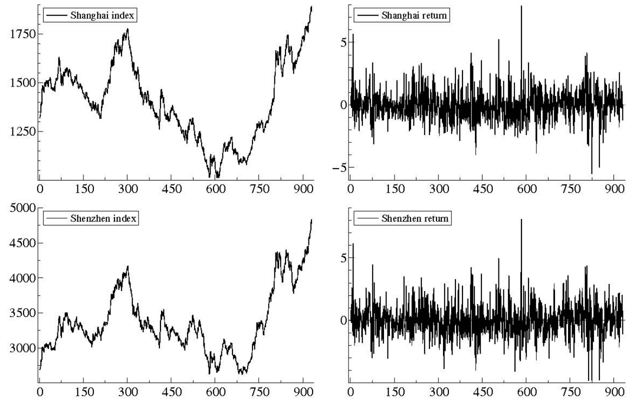

Figure 1. Price indices and returns in Shanghai and Shenzhen stock markets.

	Shanghai		Shenzhen	
	Statistics	p-value	Statistics	p-value
Mean	0.0372		0.0619	
Std.devn	1.2789		1.3648	
Skewness	0.4598		0.4005	
Ex.kutosis	2.6290		2.2241	
Normality	300.26	6.31e-66	216.32	1.06e-47
$Q(5)$	6.0197	0.3043	5.2274	0.3887
$Q^2(5)$	21.1046	0.0007	25.0629	0.0001
ARCH(5)	4.0355	0.0001**	4.7116	0.0003**

TABLE I. NORMALITY IS THE BERA-JARQUE NORMALITY TEST. $Q(5)$ IS THE LJUNG-BOX Q TEST, WITH LAG EQUAL TO 5. $Q^2(5)$ IS THE Q TEST FOR THE SQUARED RETURN SERIES. ** INDICATE SIGNIFICANCE AT THE 1% LEVELS.

The sample data consists of 930 daily returns. The daily series and the returns of Shanghai and Shenzhen indices are depicted in Fig. 1. This figure exhibits the typical volatility clustering phenomenon with periods of unusually large volatility followed by periods of relative tranquility. Table 1 reports the summary of descriptive statistics for Shanghai and Shenzhen returns. The Jarque and Bera statistics shows that we strongly reject the normality hypothesis for both series. For returns of both markets, the $Q(5)$ statistic of raw returns shows that we cannot reject that there is no significant autocorrelation, but the $Q^2(5)$ value reveals that there is significant autocorrelation in the squares returns. Engle's LM tests [5] for ARCH effect show significant evidence in support of GARCH effects. The daily returns of both stock markets have the same characteristics: the tail is thick, the distribution is not normal and the return is time-varying autocorrelation.

We use the EMD method to decompose the original price series of Shanghai market. The IMFs, shown in Fig. 2, are separated into high frequency parts and low frequency parts. Each IMF has some distinct characteristics. The residue (IMF8) is slowly varying around the long term mean. Therefore, it is treated as the long term trend during the evolution of stock price. Each sharp up or down of the low frequency component corresponds to a significant event, which should be representative of the effect of these events. The high frequency component, with the characteristics of small amplitudes, contains the effects of markets' short term fluctuations.

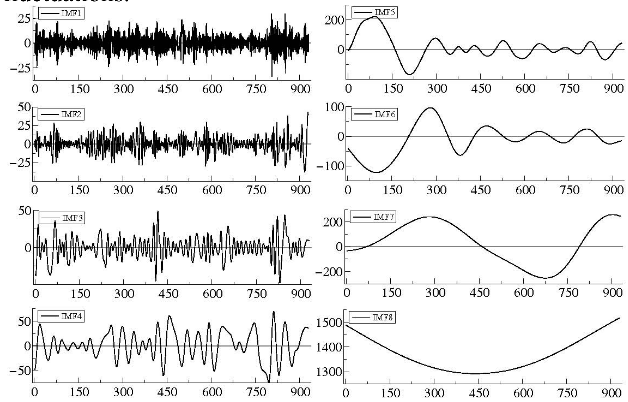

Figure 2. The EMD decomposition of price indices in Shanghai stock market.

Markets	Models	MAE	RMSE	HR
Shanghai	GARCH(1,1)	1.064	1.416	70.00
	EGARCH(1,1)	1.019	1.396	70.00
	GJR(1,1)	1.046	1.408	70.00
	MA	1.130	1.513	73.33
	NN	1.025	1.401	73.33
	EMD-NN	0.972	1.291	76.67
Shenzhen	GARCH(1,1)	1.242	1.489	73.33
	EGARCH(1,1)	1.254	1.505	73.33
	GJR(1,1)	1.226	1.479	73.33
	MA	1.118	1.588	76.67
	NN	1.215	1.463	73.33
	EMD-NN	1.126	1.384	76.67

TABLE II. VALUES OF VOLATILITY FORECASTING PERFORMANCES OF GARCH, EGARCH, GJR, MOVING AVERAGE (MA), NEURAL NETWORK (NN) AND EMD BASED ON NEURAL NETWORK (EMD-NN) IN SHANGHAI AND SHENZHEN STOCK MARKETS.

V. EMPIRICAL RESULTS

In this section the performances of moving average, GARCH (1, 1), EGARCH (1, 1), GJR (1, 1), neural network and EMD based on neural network in the volatility of both Chinese stock markets will be compared. The basic methodology is to estimate various models' parameters using a training sample and then form out-of-sample forecast. The training data are drawn from the former 900 values. The evaluation sample spanned from the 901 through the 930th data values is used to forecast the volatility. Here we adopt the recursive one-period-head forecasting scheme, which is employed with an updating sample window; the training data and forecasting data is carried out recursively by updating the sample with one observation each time, rerunning above approaches and recalculating the model parameters and corresponding forecasts.

Table 2 shows the results of these models, where GARCH family models are under the Gaussian distribution. The MAE, RMSE and HR statistics favor that EMD-NN performs best in forecasting the magnitude of volatility error and the correct direction of change of volatility among all models, except that MAE of the moving average is lowest in Shenzhen market. GJR model performs worst in both stock markets. The neural network outperforms GARCH family models in some statistics. We can't conclude the moving average behave well since it has unstable values of forecasting in both market.

VI. CONCLUSION

Due to difference of socialist and capitalist economic systems, the dynamics of the Chinese stock market exhibit some unique characteristics compared to other mature stock markets. They can't be captured by GARCH family models.

We present an empirical mode decomposition based neural network learning paradigm (EMD-NN) for forecasting volatilities of China stock market. By exploring data's intrinsic modes, EMD-NN not only helps discover the characteristics of the data but also helps understand the underlying rules of reality. Compared with other models, EMD-NN presents the most excellent forecasting performance.

REFERENCES

[1] M. Bildirici, O.O. Ersin, "Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange," Expert Systems with Applications, 2009, pp. 7355-7362.

[2] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," Journal of Econometrics, 31, 1986, pp. 307-327.

[3] Z. Ding, C.W.J. Granger, R.F. Engle, "A long memory property of stock market returns and a new model," Journal of Empirical Finance, 1, 1993, pp. 83-106.

[4] C. Dunis, X. Huang, "Forecasting and trading currency volatility: an application of recurrent neural regression and model combination," Journal of Forecasting, 21, 2002, pp. 317-354.

[5] R. Engle, "Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation," Econometrica, 50, 1982, pp. 957-1008.

[6] R.F. Engle, "Discussion: Stock market volatility and the crash of 87," Review of Financial Studies, 3, 1990, pp. 103-106.

[7] R. Ferland, S. Lalancette, "Dynamics of realized volatilities and correlations: an empirical study," Journal of Banking and Finance, 30 (7), 2006, pp. 2109-2130.

[8] L. Glosten, R. Jagannathan, D. Runkle, "On the relationship between the expected value and the volatility of the nominal excess return on stocks," Journal of Finance 46, 1993, pp. 1779-1801.

[9] S. Haykin, Neural Networks: A Comprehensive Foundations (2nd edn), Prentice Hall, Englewood Chiffs, NJ, 1999.

[10] N.E. Huang, Z. Shen, S.R. Long, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis," Process of the Royal Society of London. A, 454, 1998, pp. 903-995.

[11] N.E. Huang, M.L. Wu, S.R. Long, S.S.P. Shen, W.D. Qu, P. Gloersen, K.L. Fan, "Aconfidence limit for the empirical mode decomposition and the Hilbert spectral analysis," Proceedings of the Royal Society of London. A, 459, 2003, pp. 2317-2345.

[12] N.E. Huang, M.L. Wu, W.D. Qu, S.R. Long, S.S.P. Shen, "Applications of Hilbert-Huang transform to nonstationary financial time series analysis," Applied Stochastic Models in Business and Industry, 19, 2003, pp. 245-268.

[13] D. Nelson, "Conditional heteroskedasticity in asset returns: a new approach," Econometrica, 59, 1991, pp. 347-370.

[14] E. Sentana, "Quadratic ARCH models," Review of Economic Studies, 62(4), 1995, pp. 639-661.

[15] J. Taylor, "A quantile regression neural network approach to estimating the conditional density of multiperiod returns," Journal of Forecasting, 19, 2000, pp. 299-311.

[16] L. Yu, S.Y. Wang, K.K. Lai, "Forecasting Crude Oil Price with an EMD-based Neural Network Ensemble Learning Paradigm," Energy Economics, 30, 2008, pp. 2623-2635.

[17] J.M. Zakoian, "Threshold Heteroscedastic Models," Journal of Economic Dynamics and control, 18, 1994, pp. 931-955.