International Journal of Statistics and Applied Mathematics 2019; 4(6): 45-47

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2019; 4(6): 45-47

www.mathsjournal.com
Received: 19-09-2019
Accepted: 21-10-2019

Dr. Mohammed Abou Elfettouh Ghazal
Professor, Department of Mathematical Statistics, Faculty of Science, Damietta University, Damietta, Egypt

Dr. Asaad Ahmed Gad Elrab
Department of Mathematics, Faculty of Science, Al-Azhar University, Cairo, Egypt

Wafa Hamed Abd Allah
Department of Mathematical Statistics, Faculty of Science, Mathematics Department at Damietta University, Egypt

Corresponding Author:
Wafa Hamed Abd Allah
Department of Mathematical Statistics, Faculty of Science, Mathematics Department at Damietta University, Egypt

A hybrid statistical empirical mode decomposition with neural network in time series forecasting

Dr. Mohammed Abou Elfettouh Ghazal, Dr. Asaad Ahmed Gad Elrab and Wafa Hamed Abd Allah

Abstract
Scope Extending of Empirical Mode Decomposition called Statistical Empirical Mode Decomposition (SEMD), recently proposed by Kim *et al*. (2012) invented a new data analysis technique for nonlinear and non-stationary time series. By breaks a time series into a small number of independent and concretely implicational intrinsic modes functions based on scale separation, SEMD explains the generation of time series data from a novel perspective. This study illustrates a statistical empirical mode decomposition based on neural network learning paradigm (SEMD-NN) for forecasting Egypt stock market. By the criteria of some statistic loss functions, SEMD-NN outperforms Holt-winters family model, empirical mode decomposition based on neural network (EMD-NN) and ensample empirical mode decomposition and neural network (EEMD-NN) in improving forecast accuracy.

Keywords: Hybrid statistical, empirical mode decomposition, neural network, time series forecasting

1. Introduction

Time series signal originally seeks for a appropriate model to fit the data; this is complicated by the fact that the data is classically non-stationary, with non-linear relationships between past and future values and behaviour happening simultaneously at diverse time scales. Nowadays Hilbert Huang Transformation which known as Empirical Mode decomposition (EMD) is a method considered to decompose a signal into its intrinsic modes [3], with each mode constrained to a limited frequency band and has seen extensive usage in the field of financial signal analysis. What makes EMD attractive in financial analysis is that it is an empirically based technique that is a posteriori and adaptive, allowing the data to speak for itself. No a priori assumptions are necessary, as is the case with conventional time-frequency techniques such as Fourier or wavelet analyses. The time-frequency components obtained from EMD can simplify this task by allowing one to examine the series for one intrinsic mode function (IMF) at a time and over time horizons that are most favorable for the respective IMFs. While EMD is conventionally uttalized to analyse the individual modes of a time series, usage of the method as a filter has also been identified [4, 5].

An advantage of EMD-filtering is that the data still retains its nonlinearity and non-stationary, which is not the case when using conventional filtering techniques. While EMD filtering can be used to separate the systematic behaviour of the time series from noise at the selected time scale, a technique is still required to model future behaviour based on the historic values of the filtered signal. Unfortunately as any statistics technique this method struggling with some issues. Consequently Kim and Oh in 2012 [8] suggests new techniques which considered as extensions of EMD called Statistical Empirical Mode Decomposition (SEMD) which work out better than EMD.

Artificial Neural Networks (ANN's) are a widely used machine learning technique that simulates the structure of a biological neural network in order to model arbitrary relationships between a set of inputs and a set of outputs [2]. The structure of the neural network consists of nodes distributed across input, hidden and output layers, connected by weighted connections and activation functions [6]. This structure gives neural networks the built-in property to identify nonlinear relationships between input and output variables, making it ideal for application to nonlinear domains such as financial prediction. This research proposes an ANN model applied to data filtered with a novel SEMD-filtering technique for multi-step prediction of stock market. The purpose of the prediction will be to maximize the returns of an investor by identifying the most exploitable forecast horizon and the optimal sampling period using empirical methods. The input data forming part of the training set will be filtered to improve the signal to noise ratio for the selected forecast horizon at the appropriate time scale. The SEMD-filtered ANN model will be tested on multiple exchange rates and will be compared with an EMD-ANN and EEMD applied to unfiltered data as well as to a random walk model in terms of accuracy of predictions and simulated returns on an investment.

In this paper, motivated by the work in Yu *et al.* [9], we provide statistical empirical mode decomposition based neural network (SEMD-NN) for forecasting volatilities of Egyptian stock market. Our aim here is to compare the propose SEMD-NN with other competitive approaches including Holt-winters family models, ANN for forecasting the volatility of Egypt stock market. The rest of this paper is organized as follows. We will introduce the theory of SEMD and numeral network in section 2. Section 3 is dedicated to the introduction of proposed method. 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 concludes our paper.

2. Scope Extending of Empirical Mode Decomposition (Statistical Empirical Mode Decomposition)

The proposed algorithm is designed for considering noisy signals that are used in the field of statistics. Thus, we call the proposed algorithm statistical EMD (SEMD). Formally, the SEMD algorithm can be stated as follows:

1. ((Modified sifting) Take a signal x to be decomposed, and extract the first mode $h_{1,\lambda}$ by using a smoothing technique.

(A-1) Identify the local maxima (minima) z of the signal $h^0_{1,\lambda}$ where $h^0_{1,\lambda}$ is the original signals x.

(A-2) Construct an upper envelop $\hat{u}_\lambda$ ((lower envelope $\hat{\ell}_\lambda$) by applying a smoothing technique with a smoothing parameter λ to the maxima (minima) z.

(A-3) Compute the local mean $m_\lambda = \frac{1}{2}(\hat{u}_\lambda + \hat{\ell}_\lambda)$ by the average of both the envelopes, and then obtain a candidate intrinsic mode

$$h^1_{1,\lambda} = h^0_{1,\lambda} - m_\lambda.$$

(A-4) Repeat steps (A-1)–(A-3) for the signal $h^1_{1,\lambda}$ until the signal $h^j_{1,\lambda}$ at the jth iteration satisfies the IMF conditions.

(A-5) Decompose the signal $x = h_{1,\lambda} + r_\lambda$, where $h_{1,\lambda}$ λ is defined as the limit of $h^j_{1,\lambda}$ and r_λ is the remaining signal.

B. (Conventional sifting) If the remaining signal $r_\lambda = x - h_{1,\lambda}$ has an intrinsic oscillation mode, then r_λ can be further decomposed by conventional sifting. See; [7].

2.1 Neural Network

The neural network model introduced in this paper is the multilayer feed-forward network which is the most basic and regularly used one in financial applications. We identify 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.

3. SEMD-NN Model

The original price signal is first decomposed into a 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 possibly will be precisely predicted [9]. The volatility of price series is defined as

$$\sigma_t = \sum_{j=0}^{h} \frac{c_j(t)}{y_t} \qquad (1),$$

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

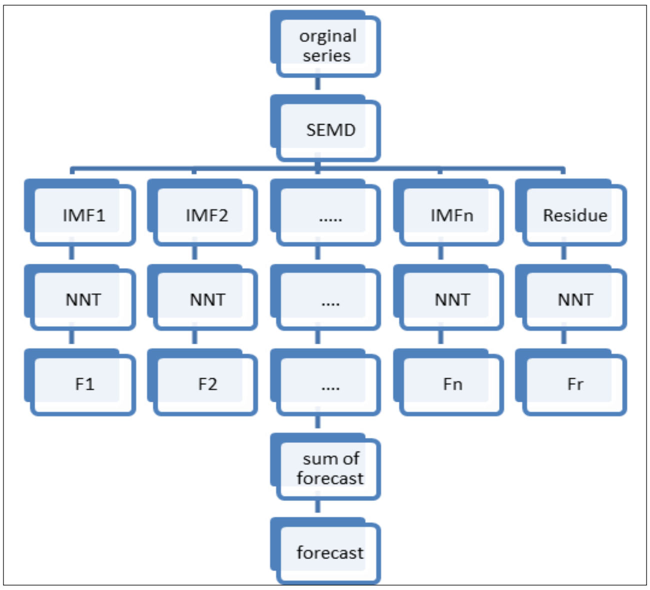

Fig 1: shows the follow chart of the methodology

3.1 Forecasting Scheme and Evaluation

In the estimation of volatility forecast of real date, the actual volatility is indirectly apparent and hence it has to be estimated. In this study, we make a decision to measure $\sigma_t$as the squared difference between the return and its mean value, that is

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

Where $\bar{y}_t$ is the mean of returns. We calculate the forecasting performance using two standard statistical criteria: mean absolute forecast error (MAE), mean squared error (RMSE). Where *N* represents the total number of the predicted values. σ_t^2 denotes the predicted conditional variance. σ_t^2 represents the actual volatility [1]. 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. The better is the performance of prediction. The data in this study is comprised of the daily stock closing price indices of the Egypt stock market from 24 May 2010 to 24 May 2018. These stock price indices *t p* are then transformed into daily returns *t y* by 100 times their log difference:

$$y_t = 100 \times (\ln(p_t) - \ln(p_{t-1})) \qquad (3),$$

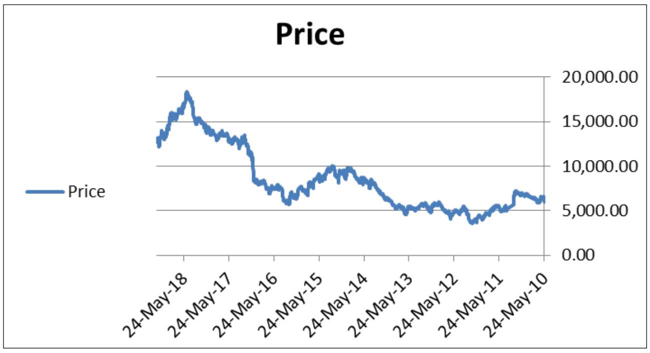

Fig 2: Price indices and returns in Egypt stock markets.

The sample data consists of 2085 daily returns. The daily series and the returns of Egypt stock market is depicted in Fig. 1. This figure exhibits the typical volatility clustering phenomenon with periods of unusually large volatility followed by periods of relative tranquility. We use the SEMD method to decompose the original price series of Egypt market.

4. Empirical Results

In this section the performances of SEMD based on neural network, EMD based neural network, and EEMD based neural network in the volatility of both Egypt stock market 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 500 values. The evaluation sample spanned from the 501 through the 2085th 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 1: shows the results of these models, shows The MAE and RMSE statistics favor that SEMD-NN performs best in forecasting the magnitude of volatility.

Markets	Models	MAE	RMSE
Egypt	EMD-NN	2.116	2.437
	EEMD-NN	2.244	5.501
	SEMD-NN	2.066	2,087

5. Conclusion

We present an (SEMD-NN) for forecasting volatilities of EGYPT stock market. By exploring data's intrinsic modes, SEMD-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 and EEMD-NN. SEMD-NN presents the most excellent forecasting performance.

5.1 Availability of data and materials

All data generated and analyzed during this study are included within this article.

5.2 Competing interest

The authors declare that they have no competing interests

6. Acknowledgements

This work of Abd Allah WH was supported by the Mathematical Statistics-Faculty of science- Damietta University of Egypt. This research was supported by Ghazal MA. Damietta University, Faculty of Science at the Mathematical Department.

7. References

1. Ding Z, Granger CWJ, Engle RF. A long memory property of stock market returns and a new model, Journal of Empirical Finance. 1993; 1:83-106.
2. Dunis C, Huang X. Forecasting and trading currency volatility: an application of recurrent neural regression and model combination, Journal of Forecasting. 2002; 21:317-354.
3. Engle R. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation, Econometrica. 1982; 50:957-1008.
4. Engle RF. Discussion: Stock market volatility and the crash of 87, Review of Financial Studies. 1990; 3:103-106.
5. Ferland R, Lalancette S. Dynamics of realized volatilities and correlations: an empirical study, Journal of Banking and Finance. 2006; 30(7):2109-2130.
6. Huang NE, Shen Z, Long SR. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis, Process of the Royal Society of London. A. 1998; 454:903-995.
7. Huang NE, Wu ML, Qu WD, Long SR, Shen SSP. Applications of Hilbert-Huang transform to non-stationary financial time series analysis, Applied Stochastic Models in Business and Industry. 2003; 19:245-268.
8. Kim D, Kim KO, HS Oh. Extending the scope of empirical mode decomposition by smoothing EURASIP Journal on Advances in Signal, 2012, Springer.
9. Yu L, Wang SY, Lai KK. Forecasting Crude Oil Price with an EMD-based Neural Network Ensemble Learning Paradigm, Energy Economics. 2008; 30:2623-2635.