Economic Modelling 29 (2012) 2583–2590

Contents lists available at SciVerse ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

Empirical mode decomposition–based least squares support vector regression for foreign exchange rate forecasting

Chiun-Sin Lin [a], Sheng-Hsiung Chiu [b], Tzu-Yu Lin [b,*]

[a] Department of Business and Entrepreneurial Management, Kainan University, No.1 Kainan Road, Luzhu Shiang, Taoyuan 33857, Taiwan
[b] Department of Management Science, National Chiao Tung University, No. 1001, Ta-Hsueh Road, Hsinchu City 30010, Taiwan

ARTICLE INFO

Article history:
Accepted 11 July 2012

Keywords:
Empirical mode decomposition
Least-squares support vector regression
Foreign exchange rate forecasting
Intrinsic mode function

ABSTRACT

To address the nonlinear and non-stationary characteristics of financial time series such as foreign exchange rates, this study proposes a hybrid forecasting model using empirical mode decomposition (EMD) and least squares support vector regression (LSSVR) for foreign exchange rate forecasting. EMD is used to decompose the dynamics of foreign exchange rate into several intrinsic mode function (IMF) components and one residual component. LSSVR is constructed to forecast these IMFs and residual value individually, and then all these forecasted values are aggregated to produce the final forecasted value for foreign exchange rates. Empirical results show that the proposed EMD-LSSVR model outperforms the EMD-ARIMA (autoregressive integrated moving average) as well as the LSSVR and ARIMA models without time series decomposition.

1. Introduction

Financial time series forecasting has come to play an important role in the world economy as a result of its ability to forecast economic benefits and influence countries' economic development; it has attracted increasing attention from academic researchers and business practitioners for its theoretical possibilities and practical applications (Hadavandi et al., 2010; Lu et al., 2009). Since the breakdown of financial market boundaries in order to enhance the efficiency of capital funding, for example, the Bretton Woods system of monetary management was officially ended in the 1973; currencies traded internationally has become crucial economic indices for international trade, financial markets, the alignment of economic policy by governments, and corporate financial decision-making. However, it is widely known that financial time series forecasting has shortcomings, including its inherent nonlinearity and non-stationarity (Huang et al., 2010; Lu et al., 2009). Therefore, financial time series forecasting is one of the most challenging tasks in the financial markets.

For modeling financial time series, autoregressive integrated moving average (ARIMA) models have been popular and are widely chosen for academic research observing the behavior of foreign exchanges and stock markets, because of their statistical properties and forecasting performance (Khashei et al., 2009). However, some problems arise when forecasting financial time series with ARIMA models, as follows. First is the characteristic linear limitation of ARIMA models, in contrast to real-word financial time series, which are often nonlinear (Khashei et al., 2009; Zhang, 2001; Zhang et al., 1998) and are rarely pure linear combinations. Second is the robustness limitation of ARIMA models—the ARIMA model selection procedure depends greatly on the competence and experience of the researchers to yield desired results. Unfortunately, choice among competing models can be arbitrated by similar estimated correlation patterns and may frequently reach inappropriate forecasting results.

With the recent development of machine-learning algorithms, several methods have been utilized that work more effectively than the traditional linear model in time series forecasting problems. For example, the support vector machine (SVM) is a novel machine-learning approach. SVM's generalization capability in obtaining a unique solution (Lu et al., 2009) and structural risk minimization principle (SRM) in achieving high performance (Duan and Stanley, 2011) have drawn attention to SVM's research applications. Support vector regression (SVR) is the regression model of SVM (Vapnik, 2000). It has been applied to investigate the forecasting ability of financial time series. Lu et al. (2009) used SVR to construct a stock price forecasting model, and Huang et al. (2010) and Ni and Yin (2009) both implemented SVR in exchange rate forecasting models. However, the training phrase of SVR is a time-consuming process when there is a lot of data to deal with. Therefore, least squares support vector regression (LSSVR), proposed by Suykens and Vandewalle (1999), has been applied in much literature as an alternative (He et al., 2010; Khemchandani et al., 2009); it is a simplified version of traditional SVR that alters inequality constraints into equal conditions and employs a squared loss function, which is a differential setting relative to traditional SVR (Wang et al., 2011), to achieve higher calculation speed and efficiency while retaining the advantage of the structural risk minimization principle.

* Corresponding author. Tel.: +886 921865025.
E-mail address: blackmallows@gmail.com (T.-Y. Lin).

0264-9993/$ – see front matter
http://dx.doi.org/10.1016/j.econmod.2012.07.018

When we model financial time series using LSSVR or ARIMA, we must remember that these financial time series are inherently nonlinear and non-stationary. If we ignore this problem, it will result in worse forecasting. The property of financial time series and the divide and conquer principle (Yu et al., 2008) are important in constructing a financial time series forecasting model. Therefore, hybrid models are widely used to solve the limitations in financial time series forecasting. Empirical mode decomposition (EMD) is suitable for financial time series in terms of finding fluctuation tendency, which simplifies the forecasting task into several simple forecasting subtasks. EMD as a time–frequency resolution approach offers a new way by which the -stationary and nonlinear behavior of time series can be decomposed into a series of valuable independent time resolutions (Tang et al., 2012). It also can reveal the hidden patterns and trends of time series, which can effectively assist in designing forecasting models for various applications (An et al., 2012; Guo et al., 2012). Guo et al. (2012), for example, decomposed wind-speed series using EMD and then forecasted them using a feed-forward neural network, whereas Chen et al. (2012) proposed an EMD approach combined with an artificial neural network for tourism-demand forecasting.

In this paper, EMD and LSSVR are used to present a financial time series forecasting model for foreign exchange rate, in which consideration of the decomposed financial time series structure will increase the accuracy and practicability of the proposed model in terms of overcoming the nonlinearity and non-stationarity limitations to the linear statistical model. The proposed approach is compared with the combination of EMD and ARIMA as well as with the existing LSSVR and ARIMA models, and it is shown that the proposed model can yield more accurate results. Three financial time series are used as illustrative examples, as follows: USD/NTD exchange rate, JPY/NTD exchange rate, and RMB/NTD exchange rate.

2. Methodology

2.1. Empirical mode decomposition

Empirical mode decomposition (EMD) is a nonlinear signal-transformation method developed by Huang et al. (1998, 1999). It is used to decompose a nonlinear and non-stationary time series into a sum of intrinsic mode function (IMF) components with individual intrinsic time scale properties. According to Huang et al. (1998), each IMF must satisfy the following two conditions. First, the number of extreme values and zero-crossings either are equal or differ at the most by one; and second, the mean value of the envelope constructed by the local maxima and minima is zero at any point. The detail-decomposition process of EMD is presented by Huang et al. (1998). Suppose that a data time series can be decomposed according to the following procedure.

(1) Identify all the local maxima and minima of $x(t)$.
(2) Obtain the upper envelope $x_u(t)$ and the lower envelope $x_l(t)$ of the $x(t)$.
(3) Use the upper envelope $x_u(t)$ and the lower envelope $x_l(t)$ to compute the first mean time series $m_1(t)$, that is, $m_1(t)=[x_u(t)+x_l(t)]/2$.
(4) Evaluate the difference between the original time series $x(t)$ and the mean time series and get the first IMF $h_1(t)$, that is, $h_1(t)=x(t)-m_1(t)$. Moreover, we see whether $h_1(t)$ satisfies the two conditions of an IMF property. If they are not satisfied, we repeat steps 1–3 of the decomposition procedure to eventually find the first IMF.
(5) After we obtain the first IMF, a repeat of the above steps is necessary to find the second IMF, until we reach the final time series $r(t)$ as a residue component that becomes a monotonic function, which is suggested for stopping the decomposition procedure (Huang et al., 1999).

The original time series $x(t)$ can be reconstructed by summing up all the IMF components and one residue component as Eq. (1), as follows.

$$x(t)=\sum_{i=1}^{n}h_i(t)+r(t). \tag{1}$$

2.2. Least squares support vector regression

The support vector machine (SVM) developed by Vapnik (1995, 2000) is based on the SRM principle. It aims to minimize the upper bound of the generalization error, instead of the empirical error as in other neural-network methods such as back-propagation networks (BPN). SVM explores not only the problem of classification but also the regression application of forecasting. Vapnik et al. (1997) proposed support vector regression (SVR) as an SVM regression estimation model, introducing the concept of the ε-loss function.

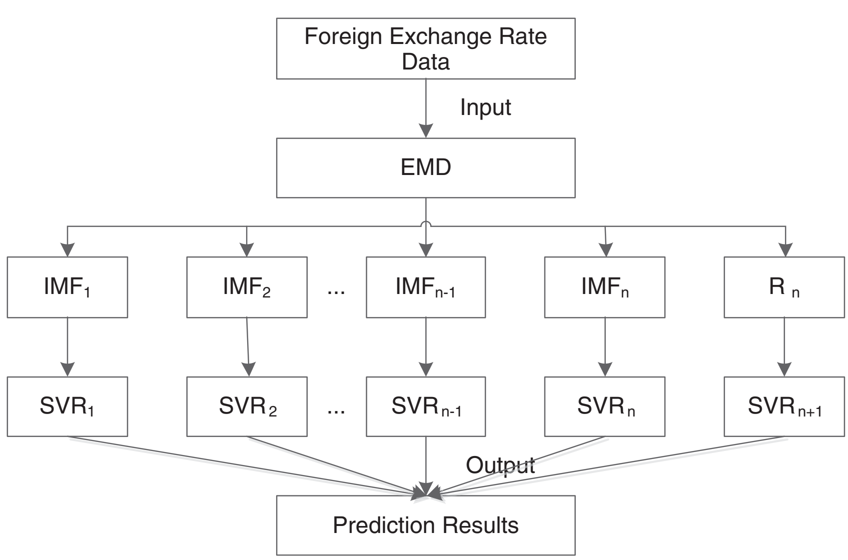

Fig. 1. The proposed EMD-LSSVR forecasting model for foreign exchange rate.