Predicting Daily Exchange Rate with Singular Spectrum Analysis

Hossein Hassani [a], Abdol S. Soofi [b*], Anatoly A. Zhigljavsky [a]

[a] *Group of Statistics, Cardiff School of Mathematics, Cardiff University, CF24 4AG, UK.*

[b] *Department of Economics, University of Wisconsin-Platteville, Platteville, WI 53818-3099, USA.*

March 6, 2007

Abstract

This paper uses univariate and multivariate singular spectrum analysis for predicting the value and the direction of changes in the daily pound / dollar exchange rate. To perform the forecast, we also use the daily dollar exchange rates with respect to Euro and Japanese yen. We use the random walk model as a benchmark model to evaluate performances of the singular spectrum analysis as a prediction method. The empirical results show that the forecast based on the multivariate singular spectrum analysis compares favorably to the forecast of the random walk model, both for predicting the value and the direction of changes in the daily pound / dollar exchange rate.

Keywords: Singular spectrum analysis, forecasting exchange rate, random walk model, Diebold-Mariano test statistic.

*Corresponding author. Tel.:(608) 342-1834; Fax: (608) 342-1122.
E-mail addresses: hassanih@cf.ac.uk (H. Hassani), Soofi@uwplatt.edu (A.S. Soofi), ZhigljavskyAA@cardiff.ac.uk (A.A. Zhigljavsky)

1 Introduction

Explaining the behaviors and accurate prediction of the exchange rates has proved to be rather challenging tasks for economists over the last three decades. Publication of Meese and Rogoff (1983) which showed a simple random walk model can outperform both linear stochastic time series and structural econometric models in predicting the exchange rates has generated the voluminous literature of exchange rate economics. The main focus of these empirical works were to develop methods that could outperform a random walk model in predicting the exchange rates, although attempts were made to develop theories that could explain the volatility of floating currencies also.

Accurate prediction of financial asset prices, however, is seriously doubted by those financial economists who believe in efficiency of financial markets. Efficient Market Hypothesis (EMH) in its weak form implies that the returns of financial asset prices are white noise processes consisting of independent, identically distributed random variables[1]. Furthermore, the white noise nature of the returns implies that the series at the level follows a random walk and is unpredictable.

In spite of the popularity of EMH, mostly in the academic circles, a vast literature dealing with predictions of the financial asset prices exits. Reviewing the empirical exchange rate economics literature one could discern two strands of research in the field that closely follow fundamentalist and chartist (technical analyst) schism that prevails in prediction of equity prices in the stock markets. In the context of exchange rate economics, the fundamentalists believe that exchange rates are determined by the money supply, the price level, national income, interest rates, productivity, and other relevant economic variables. The chartists, on the other hand, argue that explaining volatility and accurate predictions of the exchange rates by economic fundamentals is at best futile. They reason that, in spite of daily variations of the exchange rates, the fundamental economic variables seldom, if at all, change in the very short run, making the fundamentals unlikely explanatory variables, at least, in the short-run. Accordingly, the chartists attempt to use historical prices of currencies to unravel the underlying dynamics of the exchange rates, and by modeling the dynamics predict future evolutions of the data generating processes of these currencies (Frankel and Froot, 1990).

The most prominent models used in predicting the exchange rates in the fundamentalist tradition include the purchasing power parity theory (Frenkel, 1981; Corbae and Ouliaris, 1988; Soofi, 1998), sticky-price monetary model (Frankel, 1979), the Balassa-Samuleson productivity differential model (DeGrgorio and Wolf, 1994), the behavioral equilibrium exchange rate model (Clark and McDonald, 1999), and the interest rate parity model (Chinn and Meredith, 2004).

Time series analyses of exchange rates, both linear and nonlinear, attempt to predict the exchange rates by using the historical data of interest and without considering the fundamental economic variables that economic theory purports to cause the exchange rate behaviors.

The earlier empirical works in the latter strand of exchange rate economics often used linear stochastic models such as ARIMA process, however, recent development in nonlinear dynamical systems theory, methods of time-delay embedding, and phase space reconstruction has opened up the possibility of testing for presence of nonlinear, deterministic structure in the dynamics of the exchange rates. For example, Soofi and Cao (2002a), Soofi and Galka (2003), and Cao and Soofi (1998) and references therein are attempts in prediction and understanding the underlying dynamics of the exchange rates using methods and algorithms from dynamical systems theories that are rarely used in the main stream financial economics.

Cheung *et al.*(2005) provides a comprehensive comparative analysis of these competing struc-

[1]White noise follows a power law in the form of $f^{-\beta}$, where f is the frequency and β is the spectral exponent with $\beta = 0$.

tural econometric models of exchange rates against a random walk as a benchmark model using quarterly data. The study finds evidence that the structural models outperform the random walk model. We find random walk modeling of quarterly data questionable, for the dynamics of underlying data generating process of the exchange rates could go through radical changes over three months. It is one thing to argue that today's price is a reasonable predictor of tomorrow's price; however, claiming that today's price is a good predictor of the asset's price 90-day hence should be taken with a grain of salt.

In this paper we aim to predict daily pound/dollar, yen/dollar, and Euro/dollar exchange rates using singular spectrum analysis (SSA). Furthermore, we compare the prediction results with those of a random walk model and use Diebold-Mariano test statistics to rule out the comparative results are chance occurrences. Finally, to gain a better understanding of prediction accuracy of the methods, we examine the cumulative distribution of the absolute errors of the competing forecasting methods used in this study.

The next section of the paper discusses SSA method. The time series data and empirical results are presented in Sections 3 and 4. Section 5 discusses the implications of the results of this study for efficeint market hypothesis. Finally, Section 6 makes concluding remarks.

2 Singular Spectral Analysis

The main purpose of SSA is to decompose the original series into a sum of series, so that each component in this sum can be identified as either a trend, periodic or quasi-periodic (perhaps, amplitude-modulated), or noise. This is followed by a reconstruction of the original series.

2.1 Informal description

The main idea of the Basic SSA is as follows.

Consider the real-valued nonzero time series $Y_T = (y_1, \ldots, y_T)$ of sufficient length T. Let $K = T - L + 1$, where L ($L \leq T/2$) is some integer called the window length. Define the matrix

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_T \end{pmatrix} \quad (1)$$

and call it the trajectory matrix. Obviously $x_{ij} = y_{i+j-1}$ so that the matrix $\mathbf{X}$ has equal elements on the diagonals $i + j =$const.

We then consider $\mathbf{X}$ as a multivariate data with L characteristics and $K = T - L + 1$ observations. The columns X_j of $\mathbf{X}$, considered as vectors, lie in an L-dimensional[2] space $\mathbb{R}^L$. Define the matrix $\mathbf{XX}^T$. Singular value decomposition (SVD) of $\mathbf{XX}^T$ provides us with the collections of L eigen–values $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_L \geq 0$ and the corresponding eigen–vectors $P_1, P_2, \ldots, P_L$, where P_i is the normalised eigen–vector corresponding to the eigenvalue λ_i ($i = 1, \ldots, L$). Note that one can apply SVD to a variety of matrices; for example, in addition to $\mathbf{XX}^T$, it is customary to use either the covariance or correlation matrix computed from $\mathbf{X}$, treated as a multivariate data matrix, see Golyandina et al (2001).

[2]Selecting the optimal embedding dimension, L is an important step in accurate phase space reconstruction of a singular time series, discussion of which is beyond the scope of this work. See Cao(1997) on optimal determination of embedding dimension L.

A group of l (with $1 \leq l < L$) eigen–vectors determine an l-dimensional hyperplane in the L-dimensional space $\mathbb{R}^L$ of vectors X_j. The distance between vectors X_j ($j = 1, \dots, K$) and this l-dimensional hyperplane can be rather small (it is controlled by the choice of the eigenvalues) meaning that the projection of $\mathbf{X}$ into this hyperplane approximates well the original matrix $\mathbf{X}$. If we choose the first l eigen–vectors $P_1, \dots, P_l$, then the squared L_2-distance between this projection and $\mathbf{X}$ is equal to $\sum_{j=l+1}^{L} \lambda_j$. According to the Basic SSA algorithm, the L-dimensional data is projected onto this l-dimensional subspace and the subsequent averaging over the diagonals allows us to obtain an approximation to the original series.

Let us now formally describe this algorithm next.

2.2 Formal description of the Basic SSA

Let us have a time series $Y_T = (y_1, \dots, y_T)$. Fix L ($L \leq T/2$), the window length, and let $K = T - L + 1$.

2.2.1 Algorithm 1.

(*Basic SSA*)

Step 1. (*Computing the trajectory matrix*): transfers a one-dimensional time series $Y_T = (y_1, \dots, y_T)$ into the multi-dimensional series $X_1, \dots, X_K$ with vectors $X_i = (y_i, \dots, y_{i+L-1})' \in \mathbf{R}^L$, where $K = T - L + 1$. Vectors X_i are called *L-lagged vectors* (or, simply, *lagged vectors*). The single parameter of the embedding is the *window length* L, an integer such that $2 \leq L \leq T$. The result of this step is the trajectory matrix $\mathbf{X} = [X_1, \dots, X_K] = (x_{ij})_{i,j=1}^{L,K}$.

Step 2. (*Constructing a matrix for applying SVD*): compute the matrix $\mathbf{X}\mathbf{X}^T$.

Step 3. (*SVD of the matrix* $\mathbf{X}\mathbf{X}^T$): compute the eigenvalues and eigen–vectors of the matrix $\mathbf{X}\mathbf{X}^T$ and represent it in the form $\mathbf{X}\mathbf{X}^T = P\Lambda P^T$. Here $\Lambda = diag(\lambda_1, \dots, \lambda_L)$ is the diagonal matrix of eigenvalues of $\mathbf{X}\mathbf{X}^T$ ordered so that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_L \geq 0$ and $P = (P_1, P_2, \dots, P_L)$ is the corresponding orthogonal matrix of eigen–vectors of $\mathbf{X}\mathbf{X}^T$.

Step 4. (*Selection of eigen–vectors*): select a group of l ($1 \leq l \leq L$) eigen–vectors $P_{i_1}, P_{i_2}, \dots, P_{i_l}$.

The grouping step corresponds to splitting the elementary matrices $\mathbf{X}_i$ into several groups and summing the matrices within each group. Let $I = \{i_1, \dots, i_p\}$ be a group of indices $i_1, \dots, i_p$. Then the matrix $\mathbf{X}_I$ corresponding to the group I is defined as $\mathbf{X}_I = \mathbf{X}_{i_1} + \cdots + \mathbf{X}_{i_p}$.

Step 5. (*Reconstruction of the one-dimensional series*): compute the matrix $\tilde{\mathbf{X}} = ||\tilde{x}_{i,j}|| = \sum_{k=1}^{l} P_{i_k} P_{i_k}^T \mathbf{X}$ as an approximation to $\mathbf{X}$. Transition to the one–dimensional series can now be achieved by averaging over the diagonals of the matrix $\tilde{\mathbf{X}}$.

2.3 Multivariate Singular Spectrum Analysis: MSSA

Multivariate (or multichannel) SSA is an extension of the standard SSA to the case of multivariate time series (see e.g. Broomhead and King (1986)). It can be described as follows. Assume that we have an M-variate time series $y_j = \left(y_j^{(1)}, \dots, y_j^{(M)}\right)$, where $j = 1, \dots, T$ and let L be window length. Using (1), we can define the trajectory matrices $\mathbf{X}^{(i)}$ ($i = 1, \dots, M$) of the one-dimensional time series $\{y_j^{(i)}\}$ ($i = 1, \dots, M$). The trajectory matrix $\mathbf{X}$ can then be defined as

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(M)} \end{pmatrix}. \tag{2}$$

The other stages of the Basic Multivariate SSA procedure are identical to the Basic SSA as described in Algorithm 1 with an obvious modification that the diagonal averaging should be applied to each of the M components separately.

There are numerous examples of successful application of the multivariate SSA (see, for example, Plaut and Vautard, 1994; Danilov and Zhigljavsky, 1997), but the theory of multivariate SSA is yet to be developed.

2.4 SSA Forecasting Algorithm

Forecasting by SSA can be applied to the time series that approximately satisfy linear recurrent formulae (LRF). That is why we start with the LRF. An important property of the SSA decomposition is the fact that, if the original time series Y_T satisfies a linear recurrent formula (LRF):

$$y_{i+d} = \sum_{k=1}^{d} a_k y_{i+d-k}, \qquad 1 \le i \le T-d \tag{3}$$

of some dimension d with some coefficients $a_1, \ldots, a_d$, then for any T and L there are at most d nonzero singular value in the SVD of the trajectory matrix $\mathbf{X}$; therefore, even if the window length L and $K = T-L+1$ are larger than d, we only need at most d matrices $\mathbf{X}_i$ to reconstruct the series.

The fact that the series y_t satisfies an LRF (3) is equivalent to its representability as a sum of products of exponentials, polynomials and harmonics, that is as

$$y_t = \sum_{k=1}^{q} \alpha_k(t)\, e^{\mu_k t} \sin\left(2\pi\omega_k t + \varphi_k\right) \tag{4}$$

Here $\alpha_k(t)$ are polynomials, μ_k, ω_k and φ_k are arbitrary parameters. The number of linearly independent terms q in (4) is less than or equal to d. The class of series that can be approximated by the series satisfying LRFs of the form (3) (or, equivalently, by the time series of the form (4) with small number of terms) is very broad. We may also be interested in some periodic (perhaps, amplitude-modulated) components of the original series and in the trend, which is a residual of the time series when the noise and all oscillatory components of the series are removed.

SSA forecasting algorithm is based on a fact which, roughly specking, states the following; if the number of terms r in the SVD of the trajectory matrix $\mathbf{X}$ is smaller than the window length L, then the series satisfies some LRF of some dimension $d \le r$.

Let us formally describe the forecasting algorithm under consideration (for more information see Golyandina et al.(2001)):

Algorithm input:
(a) Time series $Y_T = (y_1, \ldots, y_T)$.
(b) Window length L, $1 < L < T$.
(c) Linear Space $\mathfrak{L}_r \subset \mathbf{R}^L$ of dimension $r < L$. It is assumed that $e_L \notin \mathfrak{L}_r$, where $e_L = (0, 0, \ldots, 1) \in \mathbf{R}^L$.

(d) Number M of points to forecast for.

Notations and comments:
(a) $\mathbf{X} = [X_1, \ldots, X_K]$ is the trajectory matrix of the time series Y_T.
(b) $P_1, \ldots, P_r$ is an orthonormal basis in $\mathfrak{L}_r$.
(c) $\widehat{\mathbf{X}} = [\widehat{X}_1 : \ldots : \widehat{X}_K] = \sum_{i=1}^{r} P_i P_i^T \mathbf{X}$. The vector $\widehat{X}_i$ is the orthogonal projection of X_i onto the space $\mathfrak{L}_r$.
(d) $\widetilde{\mathbf{X}} = \mathcal{H}\mathbf{X} = [\widetilde{X}_1 : \ldots : \widetilde{X}_K]$ is the result of the Hankellization of the matrix $\widehat{\mathbf{X}}$.
(e) For any vector $Y \in \mathbf{R}^L$ we denote by $Y_\vartriangle \in \mathbf{R}^{L-1}$ the vector consisting of the last $L-1$ components of the vector Y, while $Y^\triangledown \in \mathbf{R}^{L-1}$ is the vector of the first $L-1$ components of the vector Y.
(f) We set $v^2 = \pi_1^2 + \ldots + \pi_r^2$, where π_i is the last component of the vector P_i $(i = 1, \ldots, r)$.
(g) Suppose that $e_L \notin \mathfrak{L}_r$. (In the other words, we assume that $\mathfrak{L}_r$ is not a vertical space). Then $v^2 < 1$. It can be proved that the last component y_L of any vector $Y = (y_1, \ldots, y_L)^T \in \mathfrak{L}_r$ is a linear combination of the first components $(y_1, \ldots, y_{L-1})$ (see Golyandina et al.(2001), Chap. 5):

$$y_L = a_1 y_{L-1} + \ldots + a_{L-1} y_1.$$

Vector $A = (a_{L-1}, \ldots, a_1)$ can be expressed as

$$A = \frac{1}{1-v^2} \sum_{i=1}^{r} \pi_i P_i^\triangledown$$

and dose not depend on the choice of a basis $P_1, \ldots, P_r$ in the linear space $\mathfrak{L}_r$. In the above notations, define the time series $Y_{T+M} = (y_1, \ldots, y_{T+M})$ by the formula

$$y_i = \begin{cases} \widetilde{y}_i & \text{for } i = 1, \ldots, T \\ \sum_{j=1}^{L-1} a_j y_{i-j} & \text{for } i = T+1, \ldots, T+M \end{cases} \tag{5}$$

The numbers $y_{T+1}, \ldots, y_{T+M}$ from the M terms of the SSA recurrent forecast. Let us define the linear operator $\mathcal{P}^{(r)} : \mathfrak{L}_r \mapsto \mathbf{R}^L$ by the formula

$$\mathcal{P}^{(r)} Y = \begin{pmatrix} Y_\vartriangle \\ A^T Y_\vartriangle \end{pmatrix}, \quad Y \in \mathfrak{L}_r$$

If setting

$$Z_i = \begin{cases} \widetilde{X}_i & \text{for } i = 1, \ldots, K \\ \mathcal{P}^{(r)} Z_{i-1} & \text{for } i = K+1, \ldots, K+M \end{cases} \tag{6}$$

the matrix $\mathbf{Z} = [Z_1, \ldots, Z_{K+M}]$ is the trajectory matrix of the series Y_{T+M}. Therefore, (6) can be regard as the vector form of (5).

2.5 Bootstrap averaged series

Experimentally, we find that bootstrapping the original series would reduce noise and improve forecasting accuracy. Accordingly, let us consider a method of constructing bootstrap averaged series for the signal $Y_T^{(1)}$ of the original series. Under suitable choice of window length L and the corresponding eigentriples, we have the representation $Y_T = \widetilde{Y}_T^{(1)} + \widetilde{Y}_T^{(2)}$, where $\widetilde{Y}_T^{(1)}$ (the reconstructed series) approximates $Y_T^{(1)}$ and $\widetilde{Y}_T^{(2)}$ is the residual series. Suppose now that we have a (stochastic) model for the residual $\widetilde{Y}_T^{(2)}$ (for instance, we can postulate some model for $Y_T^{(2)}$, and since $\widetilde{Y}_T^{(1)} \approx Y_T^{(1)}$, we apply the same model for $\widetilde{Y}_T^{(2)}$ with the estimated parameters).

Then, simulating N independent copies $Y_{T,i}^{(2)}$ of the series $\widetilde{Y}_T^{(2)}$, we obtain N series $Y_{T,i} = \widetilde{Y}_T^{(1)} + \widetilde{Y}_{T,i}^{(2)}$ and produce N reconstructing results $\widetilde{y}_{T,i}^{(1)}$.

When the sample $\widetilde{y}_{T,i}^{(1)}$ $(1 \leq i \leq T)$ of the reconstruction results is obtained, we can calculate its bootstrap averaged series by averaging the bootstrap results. The simplest model for $Y_T^{(2)}$ is the Gaussian white noise model. The corresponding hypotheses can be checked with the help of the standard test for randomness and normality.

3 Time series data

3.1 Data

We shall use three series of daily exchange rates: pound/dollar (UK), Euro/dollar (EU) and yen/dollar (Japan). We scale each data series according to $y_t \to y_t / \parallel Y_T \parallel \quad t = 1, \ldots, T$, where $\parallel Y_T \parallel^2 = \sum_{t=1}^{T} y_t^2$.

It should be noted that the results of the SSA analysis depend on the scale of the data. To make sure that all series we are dealing with have the same scale (weight) we adopt the normalization method introduced above.

Fig. 1 shows these three (rescaled) series over the period 3-Jan-2000 to 8-Dec-2006, in these prediction exercises. Each of these series contain 1810 points.

Its very clear that the UK and EU series are highly correlated (indeed, the linear correlation coefficient between UK and EU series is about 0.975). It also looks like the yen/dollar exchange rate series behaves differently than pound/dollar and Euro/dollar (EU) series (the correlation coefficient between UK and Japan is -0.55) [3]. It must be mentioned that this correlation only shows the relationship between the main trends of the series.

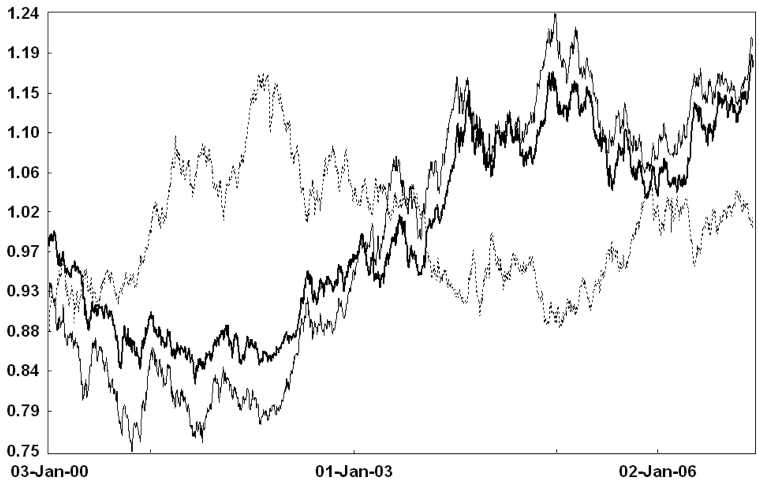


Figure 1: The exchange rate series UK (thick line), EU (thin line) and Japan (dashed line) exchange rate series over the period 2000 to 2006.

[3]We note that the correlations between yen/dolar and pound/dollar as well as yen/dollar and Eu/dollar are 0.577 and 0.487, respectively.

3.2 Trend Analysis

The main discrepancy between SSA and classical time series analysis lies in the notion of trend. For the SSA technique, trend is any slowly varying component of the series which does not contain cyclical / seasonal components. As we do not have obvious periodic components in the series, we only need to extract the trend of these data sets and for trend extraction small window length should suffice (for more information about selection of the SSA parameter see Golyandina et. all (2001), chap. 1 and 2).

Fig 2 shows the extracted trend of the original series of UK (thick line), EU (thin line) and Japan (dashed line) which are obtained from the first eigentriple and the window length $L = 30$. Note that we can build a more complicated approximation of the trend if we use some other eigentriples and smaller window length. However, the precision we would gain will be very small but the model of the trend will become much more complicated. The linear correlation coefficient between the trends of UK and EU series is 0.978 and between UK and Japan is -0.558. We see that the correlation coefficients have slightly increased (in the absolute values). This is due to smoothing. The change is very small but important for forecasting. We found that if we use bootstrap averaged series (which can be considered as smoothed versions of the series) rather than the original series, then the forecasting becomes more precise. This finding is in agreement with some results reported in the literature which indicate that reducing noise level may help us to get more accurate forecasts, especially in financial data and nonlinear series (for example, see Soofi and Cao (2002)).

To forecast UK exchange rate series, we shall use rescaled and then bootstrapped EU and Japan exchange rate series. Note that we use the original UK series in conjunction with rescaled and bootstrapped EU and Japan series.

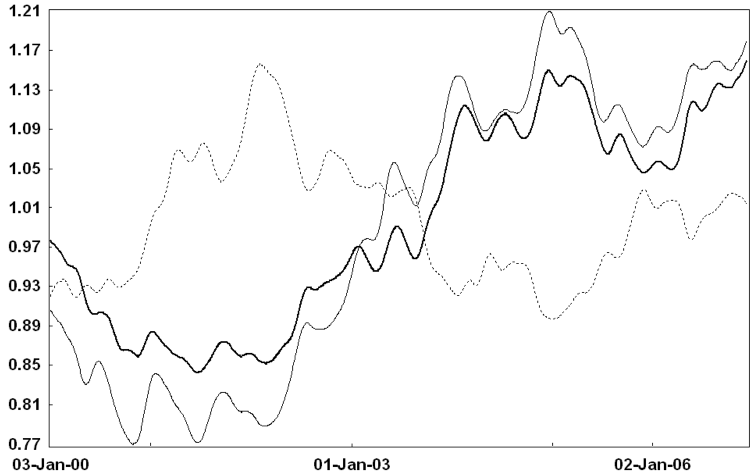


Figure 2: Trends of UK, EU and Japan exchange rate series which are obtained from the first eigentriple.

4 Results

In Table 1 we show the results of comparison of RW forecasts with forecasts made by SSA, MSSA, and MSSA (AI). In the first column for each forecasting procedure, we provide the values of the normalized root-mean-square error (RMSE):

$$\mathrm{RMSE} = \frac{[\sum_{i=1}^{N}(\hat{y}_{T+i} - y_{T+i})^2]^{1/2}}{[\sum_{i=1}^{N}(y_{T+i} - y_{T+i-1})^2]^{1/2}}. \tag{7}$$

Here N is the number of forecasted points and $\hat{y}_{T+i}$ is the one-step ahead forecasted value of y_{T+i}. The denominator in (7) is the root-mean-square error of the RW forecast.

If either RMSE $= 0$ or RMSE is very small, then the predictions are prefect or very accurate. If RMSE < 1, then the selected forecasting procedure outperforms the random walk. Alternatively, the values of RMSE are larger than 1, then the performance of the selected forecasting procedure is worse than the random walk forecast.

The second characteristic computed for each forecasting procedure is the so-called modified Diebold-Marino statistics (DM) defined by Harvey et al. (1977). As we forecast one-step ahead, the DM statistic is just $\bar{d}\sqrt{N(N-1)}\left(\sum_{i=1}^{N}(d_i - \bar{d})^2\right)^{-\frac{1}{2}}$ where $\bar{d} = N^{-1}\sum_{i=1}^{N} d_i$, $d_i = (\hat{y}_{T+i} - y_{T+i})^2 - (y_{T+i} - y_{T+i-1})^2$, and N represents out-of-sample observations. Large (in absolute values) negative values of DM statistics indicate superiority of the chosen method over the RW.

The third characteristic computed for each method is the direction of change criterion (DC). It shows the proportion of forecasts that correctly predict the direction of the series movement. Let Z_t $(t = T+1, \ldots, T+N)$ takes a value 1 if the forecast series correctly predicts the direction of change and 0 otherwise. The Moivre-Laplace central limit theorem implies that for large samples the test statistic $2(\bar{Z} - 0.5)N^{1/2}$ is approximately distributed as standard normal. When $\bar{Z}$ is significantly larger than 0.5, the forecast is said to have the ability to predict the direction of change. Alternatively, if $\bar{Z}$ is significantly smaller than 0.5, the forecast tends to give the wrong direction of change.

The rows of Table 1 correspond to different time periods where we aggregate the forecasts in different time intervals. The first row characterizes the first 10 forecasts, the second row - the first 20, etc. The last row summarizes the results for 60 forecasted data points.

We have selected 60 data points. The behaviour of the series in the chosen period looks very typical. As shown in Fig. 3 we have many changes of direction in the series, periods of slow and fast movements of the normalized rates.

We observed that the forecast is typically good when there is no sudden change of behaviour of the series at the forecast point. Alternatively, if there is such a change, the forecast is often misleading.

Table 1 shows that the results based on Basic SSA are slightly better than the results obtained using RW but the difference is not large. On the other hand, the difference between MSSA predictions and RW are significant with respect to all chosen criteria. This confirms the findings we have made from observing Figures 3 and 4. Of course, Table 1 confirms that the MSSA(AI) forecasts are the most accurate.

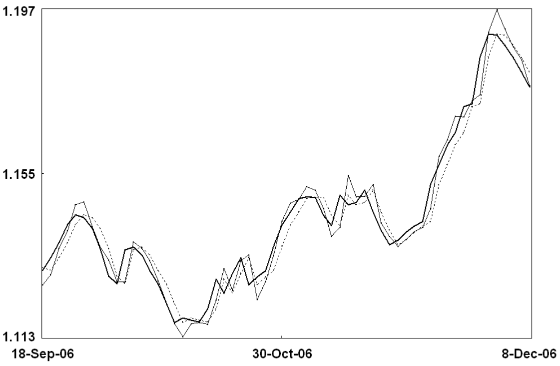
To consider the precision of the technique, we forecast all observations of the series UK from 18-Sep-2006 to 8-Dec-2006. We only perform one-step ahead forecasting based on the most up-to-date information available at the time of the forecast.

We select window length 3 for both Basic SSA and MSSA to forecast the UK series. The length of the series for all three series is the same and we also use the bootstrap averaged series instead of the original series for EU and Japan series.

To acquire a better understanding of forecasting accuracy of the methods, we examine the empirical cumulative distribution function for the absolute errors of the respective methods next.

N	SSA			MSSA			MSSA (AI)		
	RMSE	DM	DC	RMSE	DM	DC	RMSE	DM	DC
10	0.87	-0.46	0.90***	0.84	-0.94	0.80*	0.70	-1.56	0.90***
20	0.83	-1.00	0.90***	0.73	-1.91*	0.85***	0.63	-2.37***	0.85***
30	0.95	-0.36	0.80***	0.81	-1.55	0.83***	0.62	-2.77***	0.80***
40	1.02	0.16	0.75***	0.84	-1.64*	0.77***	0.61	-3.66***	0.77***
50	1.05	0.40	0.72***	0.81	-2.08**	0.76***	0.60	-4.08***	0.76***
60	1.04	0.38	0.73***	0.79	-2.45***	0.78***	0.62	-4.28***	0.80***

Table 1: Summary of the results for forecasting of UK exchange rate series with SSA, MSSA, RW. *, **, and *** indicate the significant results on the 10%, 5% and 1% levels, respectively.



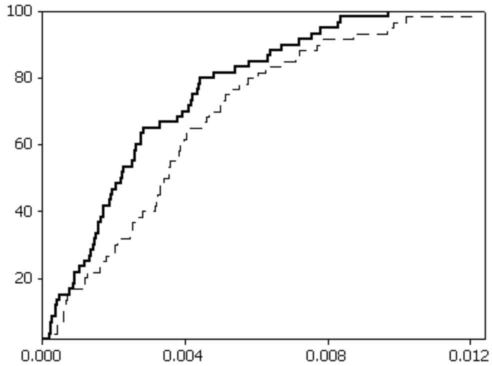


Figure 3: Left: Original series (thin line), forecasting results of MSSA (thick line) and Random Walk (dashed line); Right: Empirical cumulative distribution functions of the absolute errors for MSSA (thick line) and random walk (dashed line).

Fig. 3 (left) shows the original UK series (thick line) together with forecasted values obtained by MSSA (thin line) and the random walk model (dashed line) for the last 60 data points. In Fig. 3 (right) we display the empirical cumulative distribution function (CDF) for the absolute errors of the MSSA and RW forecasts. This plot shows that the empirical distribution of the RW errors stochastically dominates the distribution of the MSSA errors (that is, the RW errors are stochastically larger than the MSSA errors). The histograms of errors of the two forecasts (absolute values) is shown in Fig. 4. Note that the Kolmogorov-Smirnov test (the p-value is 0.90), indicates that the distribution of errors for the MSSA forecast does not contradict the hypothesis of normality.

Fig. 5 shows the scatterplot of the forecasting errors for the MSSA and RW. The white area in this plot shows the part of the plane where the absolute values of the MSSA forecast are smaller; the grey area shows the region where the absolute values of the MSSA forecast are larger.

We have performed calculations similar to the ones for MSSA for two other SSA forecasts. The first one is the Basic SSA forecast (one-dimensional, no EU and Japan series involved). We shall refer to this version simply as SSA. The second one is the MSSA forecasts (EU and Japan series are involved) when we use the information about EU and Japan time series in forecasting the UK series. We shall refer to this version of SSA as MSSA (AI), where AI stands for 'additional information'.

We do not provide the figures corresponding to Fig. 3–5 for the Basic SSA and MSSA (AI)

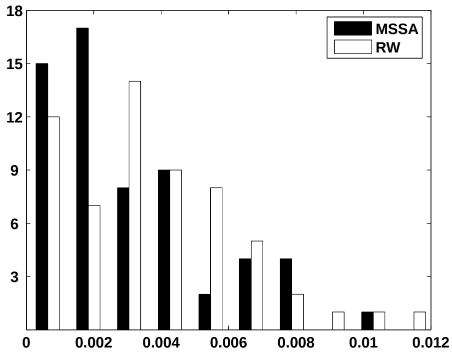


Figure 4: Histograms of the forecasting errors (absolute values) for MSSA and RW.

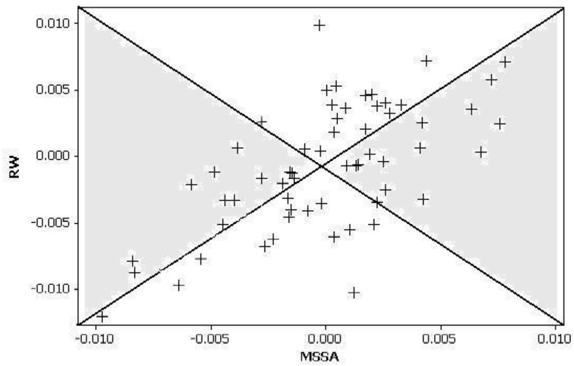


Figure 5: Scatterplot of the errors for MSSA versus RW forecasts.

forecasts. In SSA forecast, the forecasting errors are not obviously smaller (in probabilistic sense) than the errors of the RW forecast (see Table 1 for some characteristics of the these errors). The errors for the MSSA (AI) forecast are much smaller than for all other forecasts (see Table 1). This is not surprising though as the additional data used for forecast is highly correlated with the values we are forecasting.

5 Discussions

The empirical results of the present study are instructive in examining the efficient market controversy. Accordingly, we first present formal discussions of the martingale games, random walk processes, their relationship with the EMH, and then we elaborate on the implications of our findings for the EMH.

A stochastic process x_t follows a martingale if

$$E_t(x_{t+1}|\Omega_t) = x_t \tag{8}$$

where Ω_t is the information set at time t that includes x_t also. Equation (8) implies that if x_t follows a martingale the best forecast of x_{t+1} is x_t, given the information set Ω_t.

Alternatively, one could present a martingale as a "fair game"– meaning a game that is neither in your favor nor in your opponent's favor– as

$$E_t[(x_{t+1} - x_t)|\Omega_t] = 0 \tag{9}$$

The implication of the fair game model (9) is that the returns of the asset price x_t are unpredictable, given the information set Ω_t. Accordingly, the information set Ω_t is fully reflected in the asset price, and this is known as the EMH[4].

Note that one may restrict the information set Ω_t only to the asset's past price history, making alternative representation of (8) and (9) as

$$E(x_{t+1}|x_t, x_{t-1}, ...) = x_t \tag{10}$$

or

$$E(x_{t+1} - x_t|x_t, x_{t-1}, \ldots ...) = 0 \tag{11}$$

In the latter representation, again, the EMH suggests that the information contained in the price series of an asset is "instantly, fully, and perpetually" reflected in the asset's current price. Since the price series and the information contained in it are available to all market participants, no one can benefit by attempting to take advantage of the information contained in the price history of an asset by trading in the markets. This reasoning implies that the price movements in the most efficient market are completely random.

A random walk model without drift is represented as follows:

$$x_{t+1} = x_t + \eta_t \tag{12}$$

where η_t is *i.d.d.* , a white noise process, with zero mean. A random walk model is a martingale, but a more restrictive one, in the sense that it requires both independence of conditional expectation of price changes from the available information (as does the martingale) as well as independence of higher conditional moments (variance, skewness, and kurtosis) of the probability distribution of price changes.

What are the implications of our empirical findings for the EMH? Based on the results of SSA predictions which were based only on the past price history, we conclude that the currency markets are efficient and follow a random walk process. However, the results based on MSSA which are obtained by including other information, i.e. EU/dollar exchange rate, clearly point to inadequacy of the random walk in modeling exchange rate for predictions. Moreover, the superior results obtained from the direction of change method, also provide additional support for the view that currency markets may not be efficient in the sense discussed above.

6 Summary and conclusions

This paper used univariate and multivariate singular spectrum analysis in prediction of value and direction of changes (series moving up or down) in the daily pound/dollar exchange rates. We use a random walk as a benchmark model to compare performances of the SSA, MSSA, and direction of change criterion (DC) in these prediction exercises. We also used Diebold-Mariano statistics to validate the findings.

The empirical results and the test statistics show that MSSA and DC have outperformed random walk models for pound/dollar and EU/dollar exchange rates[5]. However, these methods could not outperform a random walk model using yen/dollar exchange rate.

[4]We are using EMH in a generic sense, to avoid further discussion of the types of efficient market hypothesis which is not germane to the issue here. We refer the interested reader to Campbell, Lo, and Mackinlay, 1997.

[5]We do not report the results for EU/dollar rate.

Explaining the difference in performances of these methods in predicting UK and EU rates on one hand and yen rate on the other, hinges upon co-movements of the series. As was pointed out earlier, UK and EU move in proximity of each other and have high correlation. However, the correlation between pound/dollar and yen/dollar rates is relatively small and negative. The high correlation between two series is a good indicator of accurate predictability of one series using the two series together in prediction exercises.

Another explanation might be that because of stochastic trends in all series[6] UK/dollar and EU/dollar might be cointegerated but UK/yen and EU/yen are not cointegrated. Of course, a linear combination of two cointegrated series is a stationary process, even though each series is non-stationary.

Given that the traditional structural econometric models of exchange rates have a poor record in prediction of the exchange rates in comparison to random walk models, we believe SSA and MSSA methods are highly promising. As is shown in this paper, the SSA method, at least in its multivariate representation, has decisively outperformed random walk models for two exchange rate series. Further methodological development in this field as well as extensive application of these methods in financial and economic data could prove to be indispensable for accurate prediction exercises.

[6] It is well-known that exchange rates under study here are integrated $I(1)$ series

References

Broomhead, D.S. and King, G.P. (1986). Extracting Qualitative Dynamics from Experimental Data. *Physica D* **20**, 217-236.

Campbell, J. Y., A. W. Lo, and A. C. Mackinlay (1997). *The Econometrics of Financial Markets*, Princeton University Press: Princeton.

Cao, L. (1997). Practical methods of determining minimum embedding dimension of a scalr time series, *Physica D*, 43-50.

Cheung, Y., M. D. Chinn, and A. G. Pasual (2005). Empirical exchange rate models of nineties: Are any fit to survive?, *Journal of International Money and Finance*, **24** 1150-1175.

Cao, L. and A. Soofi (1999). Nonlinear deterministic prediction of daily dollar exchange rates. *International Journal of Forecasting*, **15**, 421–430.

Chinn, M. D. (1997). Paper pushers or paper money? Empirical assessment of fiscal and monetary models of exchange rate determination. *Journal of Policy Modeling* **19**, 51-78

Chinn, M.D., Meredith, G. (2004). Monetary policy and long horizon uncovered interest parity, *IMF Staff Papers 51*, 409-430.

Clark, P. MacDonald, R. (1999). Exchange rates and economic fundamentals: a methodological comparison of BEERs and FEERs. In Stein, J., MacDonald, R. (Eds.), *Equilibrium Exchange Rates*. Kluwer, Boston, MA, pp:285-322.

Corbae, D. and S. Ouliaris (1988). Cointegration and tests of purchasing power parity, *Review of Economics and Statistics*, **70**, 508-511.

Danilov, D. and Zhigljavsky, A. (Eds.). (1997). *Principal Components of Time Series: the 'Caterpillar' method*, University of St. Petersburg Press. (In Russian.)

DeGregorio, J., Wolf, H. (1994). Terms of trade, productivity, and the real exchange rate. *NBER Working paper No. 4807.*

Frankel, J. and K. A. Froot(1990) Chartists, Fundamentalists, and Trading in the Foreign Exchange market, *American Economic Review*, **80**, 181-185.

Frankel, J.A., (1979). On the mark: a theory of floating exchange rates based on real interest differentials, *American Economic review* **69**, 610-622.

Frenkel, J.A. (1981). Flexible exchange rates, prices, and the role of 'news': Lessons from the 1970's, *Journal of Political Economy* **89**, 665-705.

Harvey, D. I., Leybourne, S. J. and Newbold, P. (1997). Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting*, **13**, 281-291.

Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and related techniques*. Chapman & Hall/CRC.

Meese, R. and Rogoff, K. (1983). Empirical exchange rate models of the seventies: do they fit out-of-sample? *Journal of International Economics* **14**, 3-24.

Plaut, G. and Vautard, R. (1994). Spells of low-frequency oscillations and weather regimes in the northern hemisphere. *J. Atmos. Sci.* **51**(2), 210-236.

Soofi, A. (1998) A fractional co-integration test of purchasing power parity: The case of selected members of OPEC. *Applied Financial Economics* ,**8**, 559-566.

Soofi, A. and L.Cao (2002a). (eds.) *Modelling and Forecasting Financial Data: Techniques of Nonlinear Dynamics*, Kluwer Academic Publishers: Boston.

Soofi, A., and L. Cao (2002b). Prediction and Volatility of Black Market Currencies: Evidence from Renminbi and Rial Exchange Rates. *International Journal of Theoretical and Applied Finance*, **5**, 659–666.

Soofi, A. and A. Galka (2003) Measuring the Complexity of Currency Markets by Fractal Dimension Analysis, *International Journal of Theoretical and Applied Finance*, Vol. 6, No. 6.