Stat Methods Appl (2011) 20:201–220
DOI 10.1007/s10260-010-0152-x

A wavelet-based approach for modelling exchange rates

Boubaker Heni · Boutahar Mohamed

Received: 13 January 2010 / Accepted: 30 September 2010 / Published online: 20 October 2010

Abstract This paper proposes a new approach, based on the recent developments of the wavelet theory, to model the dynamic of the exchange rate. First, we consider the maximum overlap discrete wavelet transform (MODWT) to decompose the level exchange rates into several scales. Second, we focus on modelling the conditional mean of the detrended series as well as their volatilities. In particular, we consider the generalized fractional, one-factor, Gegenbauer process (GARMA) to model the conditional mean and the fractionally integrated generalized autoregressive conditional heteroskedasticity process (FIGARCH) to model the conditional variance. Moreover, we estimate the GARMA-FIGARCH model using the wavelet-based maximum likelihood estimator (Whitcher in Technometrics 46:225–238, 2004). To illustrate the usefulness of our methodology, we carry out an empirical application using the daily Tunisian exchange rates relative to the American Dollar, the Euro and the Japanese Yen. The empirical results show the relevance of the selected modelling approach which contributes to a better forecasting performance of the exchange rate series.

Keywords Exchange rates · Forecasting · GARMA-FIGARCH · Wavelets

1 Introduction

Long-range dependence of the exchange rates series has been largely provided in numerous empirical studies. Several authors find evidence of long memory in exchange rates and their volatility (Cheung 1993; Baillie and Bollerslev 1994). Unfortunately,

B. Heni · B. Mohamed (✉)
Luminy Faculty of Science, GREQAM and University of Méditerranée, Marseille, France
e-mail: boutahar@univmed.fr

B. Heni
e-mail: heni.boubaker@etumel.univmed.fr

this specification does not allow to take into account persistent seasonal or cyclical behaviour. Therefore, Gray et al. (1989) proposed the GARMA (Gegenbauer Autoregressive Moving Average) process allowing for a damped sinusoidal pattern in the autocorrelation function that decays toward zero at a hyperbolic rate. The spectrum of a GARMA model is bounded at all frequencies except at the frequency that determines the cycle in the autocorrelation function. This frequency is called Gegenbauer frequency or G-frequency. Woodward et al. (1998) suggest a k-factor extension of the single frequency Gegenbauer process that allows for long memory behaviour to be associated with k singularities in the spectral density function located in the frequency range $[0, \pi]$. The statistical properties of the k-factor GARMA processes are discussed in details in Woodward et al. (1998) and Ferrara and Guégan (2001).

An important assumption for the application of the ARFIMA and GARMA models is the stationarity of the series. Previous studies consider the series of exchange rate in first difference. This induces that the exchange rates series are non stationary in level but stationary in first difference. To examine the stationarity of the series, the authors consider the traditional unit root tests (ADF, PP, KPSS). However, these tests are criticized for being biaised in the presence of long memory behaviour or structural breaks (Perron 1997, 2006).

To resolve this problem, we adopt a new analysis based on the wavelet multiscaling theory. This approach is able to separate long memory dependence from short memory dependence as well as from change-points or structural breaks (Percival and Walden 2000). More generally, wavelets allow processes to be non stationary with time varying stochastic properties.

The main advantage of the wavelet analysis is to decompose each time series into several scales and preserve thus the informational content present in the series. This concept is known as multiresolution analysis (Mallat 1989). Compared to standard Fourier analysis, wavelets provide a localized analysis in the time domain as well as in the frequency domain. More interestingly, wavelets represent functions that have discontinuities or sharp peaks; decompose and reconstruct finite non stationary signals. (See Percival and Walden 2000, for an exhaustive overview of wavelet techniques in time series analysis).

The objective of this paper is double. First, we adopt a multiresolution analysis of the level exchange rates using the maximum overlap discrete wavelet transform (MODWT). Second, we model the detrended series. For that, we consider a generalized long memory processes to model the conditional mean. A particular interest is accorded to conditional variance, which has approved to be time varying, and a FIGARCH model is considered to represent a finite persistence in the volatility.

This model is able to represent the seasonal persistence besides the double long memory behaviour. Since the estimation methods have been largely discussed in the literature, we find different procedures for estimation. We consider an extension of the wavelet-based approximate maximum likelihood estimator developed by Whitcher (2004).

This paper is organized as follows. The next section describes the wavelet approach. Section 3 presents the GARMA-FIGARCH model. Section 4 discusses the empirical results. Section 5 presents a comparative study of the predictive performance

of the selected model with ARFIMA-FIGARCH forecast performances and Sect. 6 concludes the paper.

2 Basic concepts of wavelet theory

The wavelet transform is a time-scale representation which describes the time evolution of a given signal on a scale-by-scale basis. It is analogous to standard Fourier transform, yet the complex exponentials are replaced by wavelet functions that—via dilation and translation operations—allow a flexible time-frequency resolution[1] and enable to describe local characteristics of a given function in a parsimonious way.

It should be stressed that a wavelet function proceeds as a special filter possessing specific properties. In fact, the wavelet transform can be defined in terms of a high-pass wavelet filter $\{h_l, l = 0, \ldots, L-1\}$ and its associated low-pass scaling filter satisfying the "quadrature mirror" relationship given by $g_l = (-1)^{l+1} h_l, l = 0, \ldots, L-1$.

Daubechies (1992) defined a useful class of wavelet filters: the Daubechies compactly supported wavelet filters and distinguishes between two choices: the extremal phase filters $D(L)$ and the least asymmetric filters $LA(L)$(see Daubechies (1992) for further review on the filtering terminology).

In this paper, we apply a new class of wavelet filters known as Minimum Bandwidth Discrete-Time (MBDT) wavelet filters developed by Morris and Peravali (1999). Compared to length L-Daubechies wavelet filters, the MBDT filters (which we denote by $MB(L)$) offer superior frequency localization properties given a filter of the same length.

In the next section, we provide a brief review of the main concepts of wavelet theory, namely, Discrete Wavelet Transforms (DWT), Multiresolution Analysis (MRA) and Discrete Wavelet Packet Transforms (DWPT).

2.1 Discrete wavelet transforms

The Discrete Wavelet Transforms (DWT) is any wavelet transform for which the wavelets are discretely sampled. The wavelet transform decomposes a series (signal) into a set of basis functions. These basis functions are called wavelets.

Let X be an N dimensional vector whose elements are real-valued time series $\{X_t, t = 0, \ldots, N-1\}$ of dyadic length $N = 2^J$. The length N vector of discrete wavelet coefficients is obtained via: $\mathbf{w} = WX$ where W is an $N \times N$ orthonormal matrix defining the DWT. This matrix is composed of the wavelet and scaling filter coefficients arranged on a row-by-row basis. The vector $\mathbf{w}$ of the DWT coefficients and the matrix W can be partitioned in subvectors and submatrices, respectively

$$\mathbf{w} = (\mathbf{w}_1, \ldots, \mathbf{w}_J, \mathbf{v}_J)' \quad \text{and} \quad W = (W_1, \ldots, W_J, V_J)',$$

[1] The term "frequency" is employed here in a broad sense because a wavelet function -not being an infinite sinusoid- is defined over a finite domain and can cover a whole frequency range. Unlike the Fourier transform, the wavelets are localiszd both in time and in scale.

where $\mathbf{w}_j = W_j X$, $j = 1, \ldots, J$ is an $N/2^j$ dimensional vector of wavelet coefficients associated with changes on scale $\tau_j = 2^{j-1}$ and $\mathbf{v}_J = V_J X$ is an $N/2^J$ dimensional vector of scaling coefficients associated with averages on scale $\tau_J = 2^J$. The scaling coefficients at the coarsest level J describe global features of the data while the wavelet coefficients corresponding to frequency bands $(1/2^{j+1}, 1/2^j]$ depict local features at finer scales $1, \ldots, J$. In practice, the DWT is implemented via a pyramid algorithm (Mallat 1989).

Let $\{h_l\}_{l=0}^{L-1}$ and $\{g_l\}_{l=0}^{L-1}$ denote Daubechies wavelet and scaling filters, respectively. The pyramid algorithm is based upon the two following iterations

$$w_{j,t} = \sum_{l=0}^{L-1} h_l v_{j-1,(2t+1-l) \mod N_{j-1}},$$

and

$$v_{j,t} = \sum_{l=0}^{L-1} g_l v_{j-1,(2t+1-l) \bmod N_{j-1}}, \quad t = 0, \ldots, N_j - 1,$$

where $v_{0,t} = X_t$ and $N_j = N/2^j$. If we write $\{w_{j,t}, t = 0, \ldots, N_j - 1\}$ as $\mathbf{w}_j$ and $\{v_{j,t}, t = 0, \ldots, N_j - 1\}$ as $\mathbf{v}_j$, then if $N = 2^J$, the pyramid algorithm is complete after J iterations yielding $\mathbf{w}_1, \ldots, \mathbf{w}_J, \mathbf{v}_J$. This defines the full DWT.

It should be noticed that if N is a multiple of 2^{J_0}, we say that we carry out a partial DWT. Moreover, using the partial DWT, we can formulate an additive decomposition of a time series $\{X_t, t = 0, \ldots, N - 1\}$ where N is an integer multiple of 2^{J_0}. In fact, we have

$$X = W'\mathbf{w} = \sum_{j=1}^{J_0} W_j' \mathbf{w}_j + V_{J_0}' \mathbf{v}_{J_0} = \sum_{j=1}^{J_0} d_j + s_{J_0}.$$

The latter equation defines a Multiresolution Analysis (MRA) of X. Whereas the jth level wavelet detail $d_j = W_j' \mathbf{w}_j$ describes changes at a scale $\tau_j = 2^{j-1}$, s_{J_0} is referred to the J_0th level smooth and is associated with averages at a scale $\tau_{J_0} = 2^{J_0}$.

It is worthy to note that s_{J_0} represents the smooth baseline trend component of the time series, while the wavelet details d_j, $j = 1, \ldots, J_0$ capture the higher frequency oscillations and represent finer-scale deviations from the smooth trend as j decreases.

A modified version of the DWT is the Maximal Overlap Discrete Wavelet Transform (MODWT) (Percival and Walden 2000). The MODWT algorithm carries out the same filtering steps as the standard DWT but does not subsample (decimate by 2) and therefore the number of scaling and wavelet coefficients at every level of the transform is the same as the number of sample observations. The concepts of partial MODWT "partial MODWT" and MODWT-based multiresolution analysis are defined in a similar way to those of the DWT.

Percival and Walden (2000) list several properties that distinguish the MODWT from the DWT. For the present needs, it is enough to mention that the MODWT can

handle any sample size and that the details and the smooth component of the MODWT are associated with zero phase filters. This means features in the original time series may be suitably aligned with those of the MODWT-based MRA.

2.2 Discrete wavelet packet transforms

The discrete wavelet packet transform (DWPT) (Percival and Walden 2000) is a generalization of the DWT which splits the whole frequency band $[0, 1/2]$ into individual and regularly spaced intervals.

For a given time series X of dyadic length $N = 2^J$, the jth level DWPT is an orthonormal transform yielding an N dimensional vector of wavelet packet coefficients $\left(W_{j,2^{j-1}}, W_{j,2^{j-2}}, \ldots, W_{j,0}\right)'$ where each $W_{j,n}, n = 0, \ldots, 2^j - 1$, has $N_j = N/2^j$ dimension and is associated with the frequency interval $\kappa_{j,n} = [n/2^{j+1}, (n+1)/2^{j+1}]$. Let $\{h_l\}_{l=0}^{L-1}$ and $\{g_l\}_{l=0}^{L-1}$ the Daubechies wavelet and scaling filters, respectively, starting with the recursion $X = W_{0,0}$, the tth elements of $W_{j,n}$ are computed through the following filtering steps

$$W_{j,n,t} = \sum_{l=0}^{L-1} u_{n,l} W_{j-1,[n/2],(2t+1-l) \mod N_{j-1}}, t = 0, \ldots, N_j - 1, \tag{1}$$

where

$$u_{n,l} = \begin{cases} g_l & \text{if } n \bmod 4 = 0 \text{ or } 3 \\ h_l & \text{if } n \bmod 4 = 1 \text{ or } 2 \end{cases}.$$

Here $[.]$ denotes the integer part operator.

It is remarkable to note that the collection of doublets (j, n) (also called nodes) is known as a wavelet packet tree and will be denoted by $\mathcal{T} = \{(j, n) : j = 0, \ldots, J; n = 0, \ldots, 2^j - 1\}$. An orthonormal basis $\mathcal{B} \subset \mathcal{T}$ is obtained when a collection of DWPT coefficient vectors yield a disjoint and nonoverlapping complete covering of the frequency range $[0, 1/2]$ called a 'disjoint dyadic decomposition'. Hence, in matrix notation, a vector of DWPT coefficients is obtained via $\mathbf{w}_\mathcal{B} = W_\mathcal{B} X$ where $W_\mathcal{B}$ is an orthonormal $N \times N$ matrix defining the DWPT through the basis $\mathcal{B}$.

Analogously to the MODWT, the downsampling step in the DWPT can also be removed using a variant of this transform known as the MODWPT which relies on rescaled versions of the filter $u_{n,l}$ introduced above.

2.3 Wavelet-based detrending methodology

In the following, we assume that the observed time series $\{X_t, t = 0, \ldots, N-1\}$ can be modelled additively by

$$X_t = T_t + u_t, \ t = 0, \ldots, N-1, \tag{2}$$

where T_t is an unknown deterministic trend component and u_t is a zero mean stochastic process.

By virtue of the wavelet multiscaling approach, we exploit the fact that the MRA permits to separate the trend component from the stochastic component. The idea behind is that the trend can be associated with a smooth slowly varying dynamic on large scales (low frequencies) while the small scale (high frequency) part of the signal may still be purely stochastic. An important implication of the level J_0-partial MRA structure is that, the deterministic trend component will be captured mainly by the scaling coefficients at the coarse scale J_0 whereas the detail levels, which represent fine-scale deviations from the smooth trend, can then be regarded as zero mean stochastic process.

Consequently, the detrending procedure is done by subtracting the wavelet smooth. This wavelet-based detrending approach corresponds to the application of a low-pass filter affecting only the frequency range $[0, 1/2^{J_0+1}]$. Using a multiresolution analysis in this way, the periodicities in the frequency interval $]1/2^{J_0+1}, 1/2]$ are well preserved in both amplitudes and phases when compared to a traditional differencing (of integer order) of the time series.

3 GARMA model: presentation and estimation methods

3.1 The k-factor GARMA model

The k-frequency GARMA model generalizes the ARFIMA model, by allowing for periodic or quasi-periodic movement in the data. The multiple frequency GARMA model is defined as follows

$$\phi(L)\prod_{i=1}^{k}(1-2\nu_i L+L^2)^{d_i}(u_t-\mu)=\theta(L)\varepsilon_t,$$

where the parameters ν_i provide information concerning the periodic movement in the data, $\phi(L)$ and $\theta(L)$ are polynomials in the lag operator L such that all roots of $\phi(z)$ and $\theta(z)$ lie outside the unit circle and ε_t is a white noise disturbance sequence, k is a finite integer, $|\nu_i| < 1, i = 1, \ldots, k$, d_i is a fractional differencing parameter, μ is the mean of the process, $\lambda_i = \cos^{-1}(\nu_i), i = 1, \ldots, k$, denote the Gegenbauer frequencies (G-frequencies). This model was initially proposed by Gray et al. (1989). The k-frequency GARMA model is stationary when $|\nu_i| < 1$ and $d_i < 1/2$ or when $|\nu_i| = 1$ and $d_i < 1/4$ (see Gray et al. 1989), the model exhibits long memory when $d_i > 0$.

For a single frequency GARMA model, when $\nu = 1$ the model reduces to an ARFIMA(p, d, q) model, and when $\nu = 1$ and $d = 1/2$, the process is an ARIMA model. Finally, when $d = 0$ we get a stationary ARMA model. Cheung (1996b) determines the spectral density function and shows that for $d > 0$, the spectral density function has a pole at $\lambda = \cos^{-1}(\nu)$, which ranges from 0 to π. It is important to note that when $|\nu| < 1$, the spectral density function is bounded at the origin for

GARMA processes, and thus does not suffer from many of the problems associated with ARFIMA models (see Cheung 1996a).

The GARMA model may prove to be an important long memory model in that it relaxes several aspects of the ARFIMA model. First, it is clear that the GARMA model allows for more diversity in the covariance structure of a variable witnessed both through the autocorrelation function and the spectral density function. Further, we can extend the k-factor GARMA model discussed above by inserting a fractional filter in the conditional variance equation. For this purpose, we propose GARMA-FIGARCH model which combines both the seasonality and long memory dependence and is able to capture the long memory behaviour in the mean and in the conditional variance. This choice will give much more flexibility compared to ARFIMA-FIGARCH model.

We consider the following k-factor GARMA process with FIGARCH-type innovations to take into account the presence of a time-varying conditional variance.

In the conditional mean equation, we fit a k-factor GARMA process formally given by

$$\phi(L)\prod_{i=1}^{k}(1-2v_iL+L^2)^{d_i}(u_t-\mu)=\theta(L)\varepsilon_t,$$
$$\varepsilon_t \mid I_{t-1} \sim N(0,h_t).$$

The variance equation is modelled by a FIGARCH process

$$h_t=\omega+\left\{1-[1-\beta(L)]^{-1}\varpi(L)(1-L)^{\delta}\right\}\varepsilon_t^2,$$

where I_{t-1} is the information set available at time $t-1$. Here, the conditional variance is modelled using a FIGARCH-type process. The innovations of the k-factor GARMA process are assumed to be normally distributed. $\phi(L)=1+\phi_1L+\cdots+\phi_pL^p$, $\theta(L)=1+\theta_1L+\cdots+\theta_qL^q$, $\beta(L)=1+\beta_1L+\cdots+\beta_PL^P$ and $\varpi(L)=1+\varpi_1L+\cdots+\varpi_QL^Q$ are polynomials in the lag operator L whose roots are distinct and lie outside the unit circle.

In this paper, we place particular interest on the case where $k=1$. The method is proposed to estimate the long memory parameter d, the homoscedastic short memory parameters $(\phi_1,\ldots,\phi_p,\theta_1,\ldots,\theta_q)'$ and the heteroscedastic short memory parameters $(\beta_1,\ldots,\beta_P,\varpi_1,\ldots,\varpi_Q)'$ to use the orthonormal wavelet packet decompositions as approximate whitening filters for seasonal long memory processes in order to develop estimation techniques for the GARMA-FIGARCH parameters. The next paragraph deals with the wavelet-based estimation procedure in more details.

3.2 Wavelet-based estimation procedure

In this section, we present the wavelet-based approximate maximum likelihood estimator introduced by Whitcher (2004). The estimation procedure focuses primarily on seasonal persistent processes i.e. one factor GARMA$(0,d,0)$ models. We then consider generalizing this procedure to the case of one factor GARMA-FIGARCH

adaptation which provides a useful way for analyzing the relationship between the conditional mean and conditional variance of a process exhibiting the long memory property.

Let Y be a realization of a zero mean stationary seasonal persistent (SP) process with unknown parameters $\Theta = (d, \lambda_G, \sigma_\varepsilon^2)$ where d is the seasonal fractional differencing parameter, λ_G is the Gegenbauer frequency and $\sigma_\varepsilon^2 > 0$ denotes the variance of the innovation process. The spectral density function (SDF) of this process is

$$S_Y(\lambda) = \sigma_\varepsilon^2 \left[2\left|\cos(2\pi\lambda) - \cos(2\pi\lambda_G)\right|\right]^{-2d}, \; -1/2 < \lambda < 1/2. \tag{3}$$

Given the likelihood function for Y under the assumption of stationarity and gaussianity, $\mathcal{L}(\Theta \mid Y) = (2\pi)^{-N/2} |\Sigma_Y|^{-1/2} \exp\left(-\frac{1}{2} Y' \Sigma_Y^{-1} Y\right)$, Whitcher (2004) synthesizes wavelet analysis with seasonal long-memory processes to construct an approximate Whittle-type ML estimator. He exploits the fact that, under an intuitive choice of the DWPT's orthonormal basis $\mathcal{B}$, the DWPT generates wavelet coefficients that are approximately uncorrelated thus eliminating the effect of the singularity in the spectrum on each scale of the MRA. Thanks to this decorrelation property, we can approximately diagonalize the variance/covariance matrix Σ_Y of the SP process Y in the following manner: $\Sigma_Y \approx \widehat{\Sigma}_Y = W_{\mathcal{B}}' \Omega_N W_{\mathcal{B}}$ where $W_{\mathcal{B}}$ is an orthonormal $N \times N$ matrix defining the DWPT through the basis $\mathcal{B}$ and Ω_N is a diagonal matrix containing the band-pass variances of Y.

Namely, the band-pass variance for an SP process with spectrum given by Eq. (3) in the frequency interval $\kappa_{j,n} = [n/2^{j+1}, (n+1)/2^{j+1}]$ is

$$\begin{aligned} \omega_{j,n}^2 &= 2 \int_{n/2^{j+1}}^{(n+1)/2^{j+1}} S_Y(\lambda) d\lambda \\ &= 2 \int_{n/2^{j+1}}^{(n+1)/2^{j+1}} \sigma_\varepsilon^2 \left[4\left|\cos(2\pi\lambda) - \cos(2\pi\lambda_G)\right|^2\right]^{-d} d\lambda. \end{aligned} \tag{4}$$

The approximate Log-Likelihood function is given by

$$\begin{aligned} \widehat{\mathcal{L}}\left(d, \lambda_G, \sigma_\varepsilon^2 \mid Y\right) &= \log\left(\left|\widehat{\Sigma}_Y\right|\right) + Y' \widehat{\Sigma}_Y^{-1} Y \\ &= N \log(\sigma_\varepsilon^2) + \sum_{(j,n)\in\mathcal{B}} \left[N_j \log\left(\varpi_{j,n}^2\right) + \frac{W_{j,n}' W_{j,n}}{\sigma_\varepsilon^2 \varpi_{j,n}^2} \right], \end{aligned}$$

where $\varpi_{j,n}^2 = \omega_{j,n}^2 / \sigma_\varepsilon^2$ denote the rescaled band-pass variances whereas the DWPT coefficients $W_{j,n}$ are defined by Eq. (1).

Through properties of diagonal and orthonormal matrices, Whitcher (2004) derives the reduced approximate Log-Likelihood function which is given by the following equation

$$\widehat{\mathcal{L}}(d, \lambda_G, \mid Y) = N \log\left(\widehat{\sigma}_{\varepsilon}^2(d, \lambda_G)\right) + \sum_{(j,n)\in\mathcal{B}} \frac{W'_{j,n} W_{j,n}}{\varpi_{j,n}^2},$$

where the innovation's variance is previously estimated

$$\widehat{\sigma}_{\varepsilon}^2(d, \lambda_G) = \frac{1}{N} \sum_{(j,n)\in\mathcal{B}} \frac{W'_{j,n} W_{j,n}}{\varpi_{j,n}^2}.$$

Minimizing numerically $\widehat{\mathcal{L}}(d, \lambda_G, \mid Y)$ in this case over the parameter space $(-1/2, 1/2) \times (0, 1/2)$ leads to approximate MLEs for the fractional differencing parameter d and the Gegenbauer frequency λ_G, respectively.

A way of determining the 'best basis' $\mathcal{B}$, from all the possible orthonormal partitions, is done by carrying out statistical white noise tests (such as portmanteau test) at each node (j, n) to yield approximately flat SDF for the DWPT coefficient vectors thus ensuring approximate decorrelation property of the transform (see Whitcher (2004) for more details).

In this paper, we suggest extending Whitcher's method to handle one factor GARMA-FIGARCH parameterization by performing two steps. In the first step, we incorporate short memory components in order to capture the small scale (high frequency) features of the time series. This extension entertains p autoregressive and q moving average parameters thus producing a GARMA -type specification whose power spectrum is given formally by

$$S_{p,d,q}(\lambda) = \sigma_{\varepsilon}^2 \frac{\left|\theta\left(e^{-2i\pi\lambda}\right)\right|^2}{\left|\phi\left(e^{-2i\pi\lambda}\right)\right|^2 \left[2\left|\cos(2\pi\lambda) - \cos(2\pi\lambda_G)\right|\right]^{2d}}.$$

Constructing the approximate log-likelihood function only differs by substituting the new spectral density function $S_{p,d,q}$ for S_Y in Eq. (4) i.e. in the computation of the band pass variances. In the second step, the FIGARCH parameters are estimated on the squared residuals.

4 Empirical results

4.1 Data

The data set consists of the nominal daily exchange rates of the Tunisian exchange rates relative to US dollar, the Euro and the Japanese Yen from the 1st January 1999 to the 31st March 2008. The data are issued from the Datastream and are transformed in logarithm form. Table 1 reports the key descriptive statistics of the exchange rates series.

Both skewness and excess kurtosis statistics indicate that the series tend to have a fatter-tail distribution than a normal distribution. Further evidence on the nature of departure from normality may be obtained from the Jarque-Bera test statistics which

Table 1 Summary statistics of the exchange rates

	LUSD	LEuro	LYEN
Mean	0.271	0.380	-4.471
Median	0.269	0.383	-4.469
Maximum	0.417	0.597	-4.278
Minimum	0.086	0.213	-4.660
Std Dev	0.067	0.124	0.062
Skewness	0.085	0.131	-0.149
Excess-Kurtosis	2.481	1.506	4.448
Jarque-Bera	28.804	222.352	211.239
Probability	0.000	0.000	0.000
Observations	2,320	2,320	2,320

exhibit that the null hypothesis of normality should be rejected at the l% significant level. This result is not surprising and it is often found in the empirical literature on exchange rates. It implies several extreme values relative to the standard normal distribution.

4.2 Stationarity

In what follows, we check the stationarity of the series. As showed by Fig. 1, the exchange rates appear to have a nonstationary behaviour, in the sense that they do not converge towards their long term means and testify a great instability.

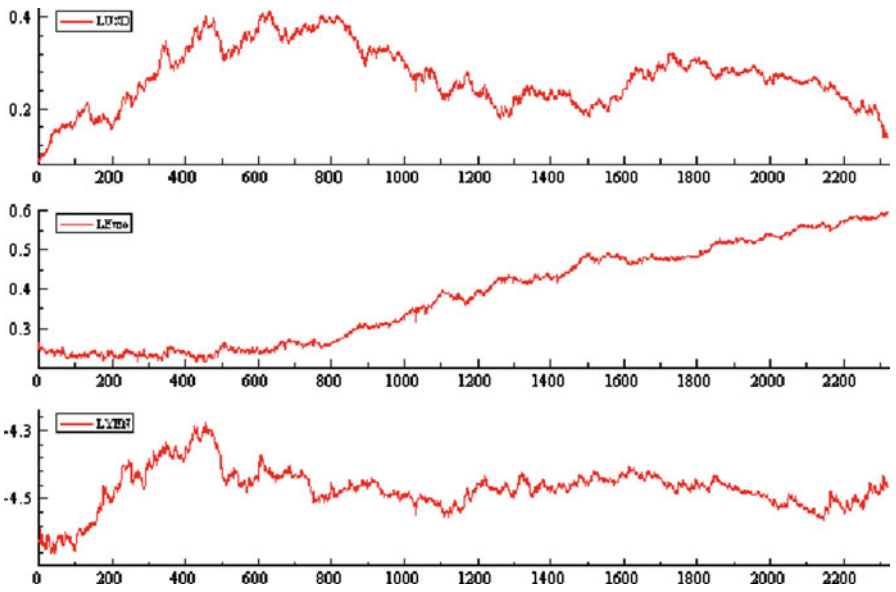

Fig. 1 Evolutions of the logarithm of exchange rates in level

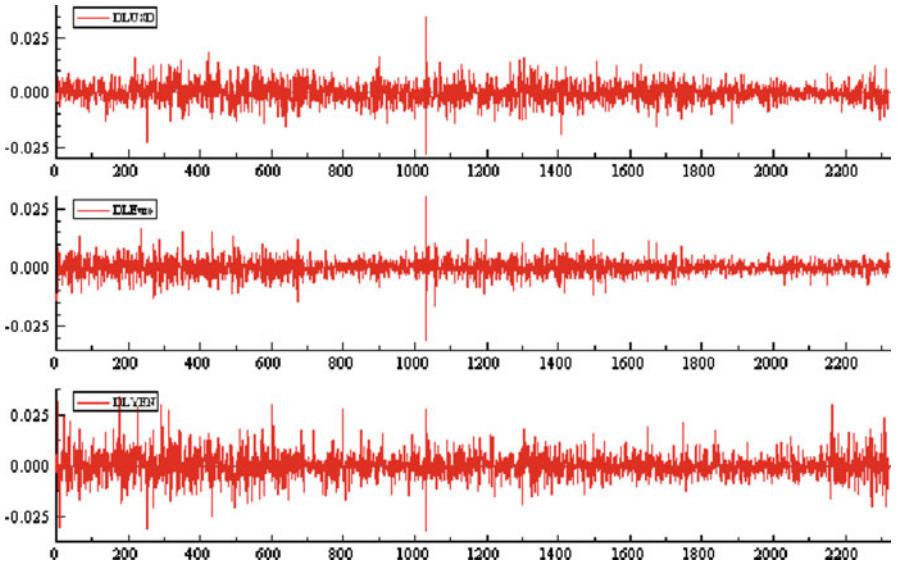

Fig. 2 Evolutions of the logarithms of exchange rates in first difference

In order to deal with the apparent non stationarity in the mean, we consider the first differences of natural logarithms of daily exchange rates (in percentage) (Fig. 2) $r_{t+1} = 100 \log(s_{t+1}/s_t)$ for $t = 0, \ldots, N-1$, where s_t represents the levels of the real exchange rates and N is the total number of observations. The differenced series seem to fluctuate randomly around zero, while the variance varies over time and shows clear patterns of volatility clustering phenomenon. To assess the stationarity of the exchange rates series in level we apply three types of unit root tests: the Dickey-Fuller test (DF or ADF) (Dickey and Fuller 1979, 1981), the Phillips-Perron test (PP) (Phillips and Perron 1988) and the KPSS test (Kwiatkowski et al. 1992). The obtained results are reported in Table 2.

For the three series of the logarithms of exchange rates, and while referring to ADF and PP test statistics, we accept the null hypothesis of unit root at 1% significance level whatever the retained model is (with or without constant). This result is confirmed by the KPSS test according to which we reject the null hypothesis of stationarity for all series at 1% significance level. Thus, the analysed series are governed by an $I(1)$ process. Now, we examine the stationarity of the series in first differences. The obtained results are reported in Table 3.

According to the ADF and PP test statistics, we reject the null hypothesis of a unit root at 1% significance level for the three series. However, while referring to the KPSS test statistics, the null hypothesis of stationarity is rejected for the DLUSD and DLEuro at 5% and 10% significance levels, respectively. The retained processes for the these series are neither $I(0)$ nor $I(1)$. For the DLYEN, we can state that it is stationary $I(0)$, since results issued from the KPSS test confirm those of the ADF and PP tests in accepting the null hypothesis of stationarity at 10% significance level. We conclude that these tests have low power to discriminate between an $I(1)$ process and a fractionally integrated $I(d)$ process with $d < 1$.

Table 2 Results of unit root test in level

	Lag	τ	τ_μ	$Z(t_\alpha)$	$Z(t_{\alpha\bullet})$	$\widehat{\eta}_\mu$
LUSD	1	0.124	−2.024	0.118	−2.011	12.376***
	4	0.083	−2.110	0.110	−2.021	4.973***
	8	0.051	−2.124	0.096	−2.142	2.776***
	15	0.063	−2.157	0.082	−2.211	1.576***
	24	−0.006	−2.241	0.074	−2.231	1.062***
LEuro	1	0.638	2.992	0.514	2.827	60.789***
	4	0.779	3.416	0.641	3.104	44.237***
	8	0.781	3.343	0.702	3.242	24.611***
	15	0.796	3.470	0.734	3.404	13.878***
	24	0.470	3.069	0.758	3.375	8.919***
LYEN	1	0.753	−2.378	0.746	−2.372	10.876***
	4	0.690	−2.107	0.658	−2.354	5.645***
	8	0.658	−2.092	0.623	−2.349	3.812***
	15	0.619	−2.283	0.602	−2.347	2.241***
	24	0.603	−2.269	0.587	−2.351	1.687***

Notes: τ (resp. $Z(t_\alpha)$) and τ_μ (resp. $Z(t_{\alpha\bullet})$) are the ADF (resp. PP) test statistics for the models without and with constant. $\widehat{\eta}_\mu$ is the statistics of KPSS test where the residuals are issued from the regression with a constant. ***, ** and * denote significance at 1, 5 and 10% level, respectively

Table 3 Results of unit root test in first differences

	Lag	τ	τ_μ	$Z(t_\alpha)$	$Z(t_{\alpha\bullet})$	$\widehat{\eta}_\mu$
DLUSD	1	−33.303***	−33.311***	−47.896***	−47.899***	0.489**
	4	−20.437***	−20.4461***	−47.892***	−47.894***	0.463**
	8	−14.973***	−14.989***	−47.900***	−47.903***	0.441*
	15	−11.841***	−11.861***	−47.926***	−47.924***	0.425*
	24	−8.422***	−8.437***	−47.948***	−47.945***	0.417*
DLEuro	1	−27.662***	−27.792***	−54.393***	−54.495***	0.523**
	4	−36.726***	−36.924***	−54.350***	−54.525***	0.497**
	8	−22.270***	−22.551***	−54.745***	−54.995***	0.470**
	15	−16.261***	−16.631***	−54.932***	−55.305***	0.389*
	24	−11.096***	−11.574***	−55.126***	−55.763***	0.357*
DLYEN	1	−18.087***	−18.641***	−48.784***	−48.635***	0.334
	4	−15.441***	−15.451***	−48.529***	−48.212***	0.295
	8	−13.861***	−13.991***	−48.349***	−48.251***	0.269
	15	−10.884***	−10.911***	−48.220***	−48.016***	0.258
	24	−8.465***	−8.516***	−48.018***	−48.026***	0.248

Notes: τ (resp. $Z(t_\alpha)$) and τ_μ (resp. $Z(t_{\alpha\bullet})$) are the ADF (resp. PP) test statistics for the models without and with constant. $\widehat{\eta}_\mu$ is the statistics of KPSS test where the residuals are issued from the regression with a constant. ***, ** and * denote significance at 1, 5 and 10% level, respectively

Table 4 Estimates of the orders of integration in first differences

	DLUSD	DLEuro	DLYEN
$\widehat{d}_w$	0.003	0.004	0.002
t-Student	0.467	0.346	0.561
Prob	0.644	0.722	0.575
$\widehat{d}_{MODWT}$	0.004	0.006	0.002
t-Student	0.675	0.478	0.479
Prob	0.499	0.633	0.632
$\widehat{d}_R$	0.009	0.007	−0.012
t-Student	0.438	0.195	−0.586
Prob	0.661	0.846	0.558

Therefore, we propose to estimate the fractional order of integration for each series in first difference using Robinson's semiparametric method, Whittle-type pseudo-maximum likelihood estimator and a wavelet-based maximum likelihood estimator in order to check the results given by the standard unit root tests.

Empirical results reported in Table 4 show that the estimated values of integration orders are not significant. These results show evidence that the exchange rates series are stationary in first differences with short-term dependence structure. It is well known that the conventional unit root are biased in the presence of long memory and the power of these tests decreases when the deterministic trend is included. For that, we adopt a new approach based on wavelet analysis that is ideally suited to distinguish short from long memory. The principle of this approach is to decompose a non stationary process into its components, each of which is associated with a particular frequency band.

4.3 Detrending exchange rate series

We propose to decompose the three exchange rates series, for that we carry out a partial MODWT at level $J = 8$ using the $LA(8)$ wavelet filter. The MODWT-based MRAs are displayed in Figs. 3–5. The choice of the $LA(8)$ wavelet is motivated by the fact that, empirically and compared to other wavelet filters, the $LA(8)$ filter is more appropriate to depict periodicities in the series of studies (LUSD, LEuro and LYEN).

Figures 3–5 illustrate the plot of the wavelet smooth s_8 and the wavelet details d_j, $j = 1, \ldots, 8$. The figure investigation reveals the existence of a strong periodicity at a scale $j = 7$ associated with the frequency interval $[1/128, 1/256]$, corresponding to 128–256 day oscillation. The wavelet smooth s_8 associated with the frequency range $[0, 1/512]$ appears to be capturing the trend component for the three exchange rate series: LUSD, LEuro and LYEN. Thus, ignoring s_8 and summing up over of the eight zero mean wavelet details (by definition of the MODWT and because of the properties of the wavelet filter $LA(8)$ previously discussed), we may define a wavelet-based technique for detrending time series. This approach corresponds to the application of a low-pass filter affecting only frequencies in the range $[0, 1/512]$.

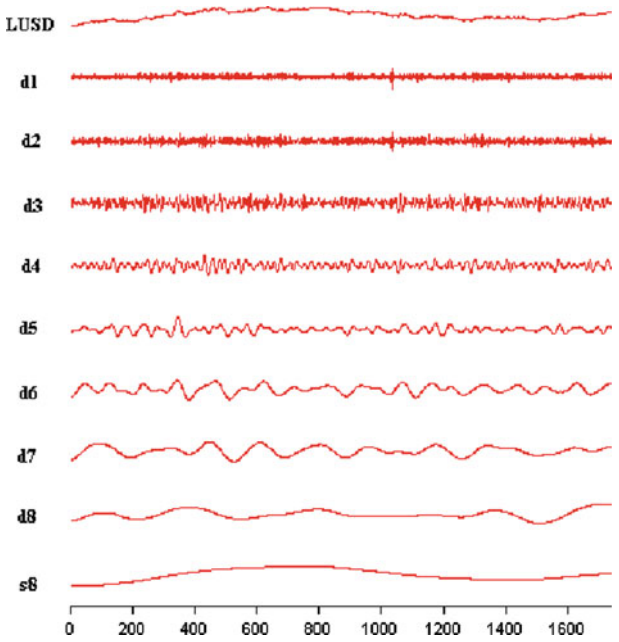

Fig. 3 Partial MODWT, J = 8, of the LUSD series using the LA(8) wavelet filter and reflection boundary conditions

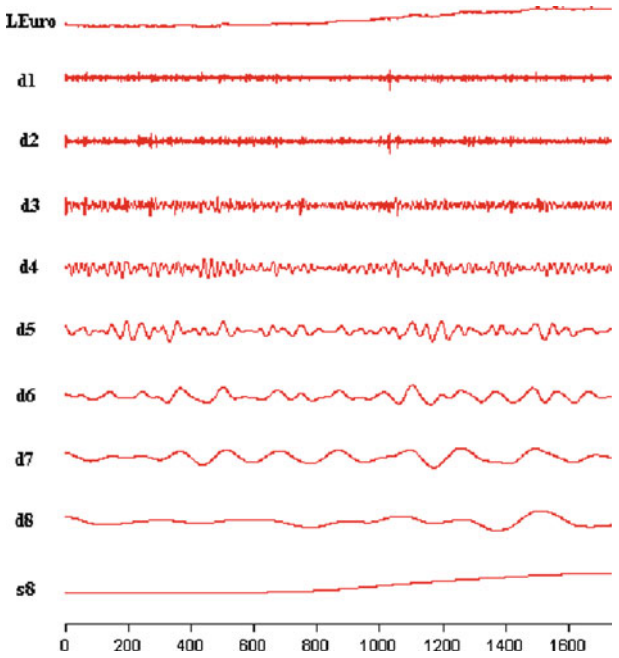

Fig. 4 Partial MODWT, J = 8, of the LEuro series using the LA(8) wavelet filter and reflection boundary conditions

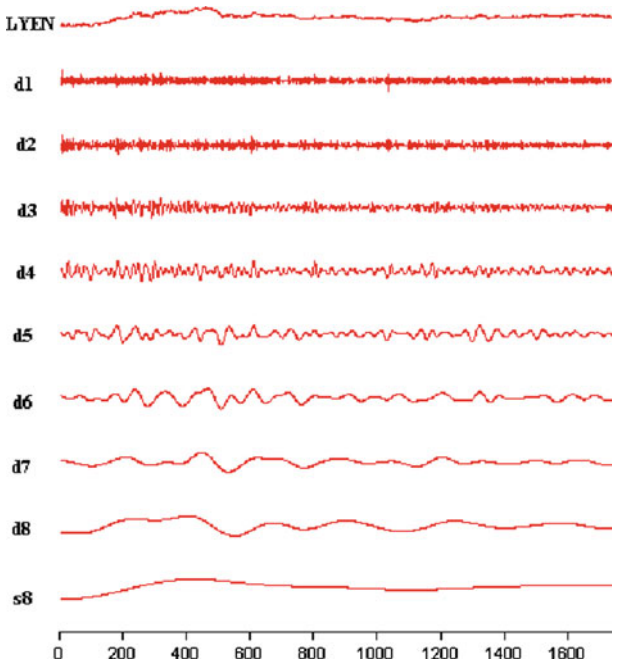

Fig. 5 Partial MODWT, J=8, of the LYEN series using the LA(8) wavelet filter and reflection boundary conditions

4.4 Seasonal modelling of the detrended exchange rates

In this section, we model the detrended exchange rates series using a generalized long memory process to take into account a potential periodic persistence associated with the seasonal components observed in the exchange rates series (Figs. 3–5).

The analysis of Fig. 6 allows us to draw several conclusions. First, the empirical autocorrelations of the detrended exchange rate series show clearly long-term dependence over the series of studies, along with a strong cyclic behaviour of relatively long period. This persistence detected is confirmed by the analysis of Table 5 which illustrates statistics of Box-Pierce test corrected of ARCH effects using Diebold's procedure (Diebold 1987) for 1,000 lags ($Q(1{,}000)$). We can thus assert the existence of a seasonal dependence in the exchange rate series. This behaviour could not be detected while using a standard long memory model to estimate the order of integration for the differenced series. Moreover, the smoothed periodograms of the series show clearly a strong explosion (peak) at a very low frequency (very close to zero). As we already mentioned, this type of behavior is typical to a generalized long memory process.

Thus, the seasonal long memory behaviour observed from the ACFs and the smoothed periodograms suggest fitting a GARMA process with time-dependent heteroskedasticity to the logarithms of the filtered exchange rate series.

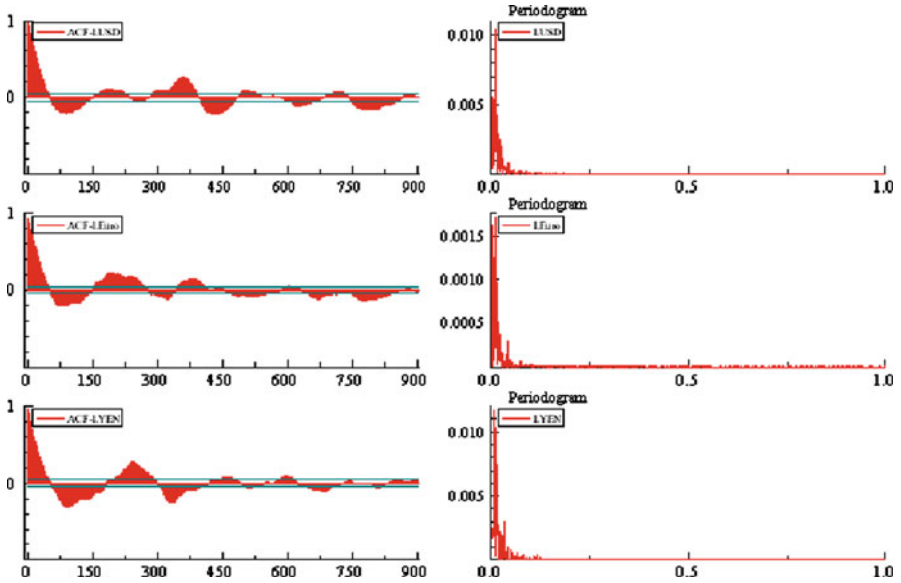

Fig. 6 Correlograms and Periodograms of the detrended exchange rates

Table 5 ARCH-corrected Box-Pierce test of the detrended exchangerates series

	LUSD	LEuro	LYEN
$Q(1000)$	2,556.440 (0.000)***	2,200.272 (0.000)***	2,405.992 (0.000)***

Note: Marginal levels of significance are indicated between square brackets. *** denotes significance at 1% level

4.5 Estimation results

In this section, we propose to estimate the parameters of the one factor GARMA process with FIGARCH innovations defined in Sect. 3.1. The advantage of such a modelling process is to take into account the persistence in the volatility shocks that we can still observe in the second order moments even after detrending the exchange rate series. An interesting feature of the FIGARCH model is that it nests both the GARCH model for $\delta = 0$, and the IGARCH specification for $\delta = 1$. If $0 < \delta < 1$, there is a long term dependence in the conditional variance indicated by an hyperbolic decay of the autocorrelation and autocovariance functions. This model allows dealing such various behaviours noted previously for exchange rates (see Baillie et al. 1996; Ling and Li 1997).

As indicated by Whitcher (2001), we use wavelet filters of length $L > 8$. In fact, the longer the wavelet filter in time domain is, the better the frequency resolution around the G-frequency will be. According to Whitcher (2001), wavelet filters of length $L > 8$ are needed to adequately filter seasonal persistent processes. Under this condition, we apply several filters $(D(L), LA(L), MB(L)\ldots)$ to estimate the long and short memory parameters.

Table 6 Estimates of the 1 factor GARMA-FIGARCH model

	LUSD	LEuro	LYEN
(p, d, q) (P, δ, Q)	$(1, d, 0)$ $(1, \delta, 1)$	$(1, d, 0)$ $(1, \delta, 1)$	$(1, d, 0)$ $(1, \delta, 1)$
Wavelet filter	D 16	MB 16	D 16
μ	0.001 (0.014)	0.000 (0.011)	0.000 (0.020)
d	0.476 (4.124)***	0.442 (3.874)***	0.452 (4.012)***
θ_1	–	–	–
ϕ_1	0.547 (4.763)***	0.721 (5.025)***	0.487 (4.273)***
λ_G	0.0072 (3.514)***	0.0068 (3.604)***	0.0069 (3.427)***
ω	0.004 (0.857)	0.000 (0.953)	0.000 (0.739)
δ	0.374 (3.344)***	0.385 (3.854)***	0.398 (4.338)***
β_1	0.587 (4.867)***	0.581 (4.521)***	0.598 (3.584)***
ϖ	0.430 (8.999)***	0.492 (8.657)***	0.426 (9.624)***
Skewness	−0.018	0.027	0.010
Excess of Kurtosis	0.425	0.524	0.761
$Q(20)$	11.897	13.714	13.963
$Q^2(20)$	7.457	7.763	8.478
$BDS(5)$	−0.002**	−0.257**	0.009*
Log-likelihood	4,981.451	5,803.697	4,725.024

Note: The student's t-statistics are indicated between brackets. ***, ** and * denote significance at 1, 5 and 10% level, respectively

We consider the wavelet-based approximate maximum likelihood technique to estimate the AR and MA polynomials for the one factor GARMA process. In order to select the lag, we rely on Schwarz and Hannan-Quinn criteria. Assuming a FIGARCH type specification in the conditional variance implies that the residuals are not independent and identically distributed (i.i.d). This problem can be overcome while considering a Normal distribution by standardizing the residuals $z_t = \widehat{\varepsilon}_t h_t^{-1/2}$. Measures of skewness and excess of kurtosis, those of Box-Pierce $Q(20)$ and $Q^2(20)$ and finally the BDS test statistics proposed by Brock et al. (1987) for $m = 5$ and $l = 1$ are calculated on the standardized residuals of the estimated models. Table 6 summarizes the estimation results.

The obtained results show significant values for both the seasonal long memory parameter and the G-frequency. The value of G-frequency is similar for the three countries and is very close to zero. This indicates the existence of a seasonality of long period and reinforces the choice of a seasonal persistent specification in the mean equation of the filtered exchange rate series.

On the other hand, the estimation results show that the fractional order of integration in the scedastic function of the selected GARMA-FIGARCH specification is highly significant at 1% significance level. This means that the volatility of the detrended exchange rate series is fractionally integrated. The value of the fractional parameter lays between 0.374 and 0.398 shows a higher degree of persistence of volatility for the detrended exchange rate series.

5 Predictive performance of the 1 factor GARMA-FIGARCH process: a comparative study

At the end of this article, we evaluate the out of sample forecasts for the exchange rates. For that, we use data over the period January 1, 1999–December 29, 2006 for estimation purpose and consider those from January 2, 2007 to March 31, 2008 for measuring forecast performance. Given that the volatility prediction can vary across forecast horizons, we consider one, five, ten and fifteen step ahead forecasts ($s = 1, 5, 10, 15$).[2] For the trend prediction, the problem is simplified by the fact that this component, captured essentially by the approximation level s_8 is very smooth.

We compare the forecast performance of GARMA-FIGARCH model which modelise the detrended exchange rate, to the classical ARFIMA-FIGARCH model, applied to exchange rates series in first difference, using three evaluation criteria (*MSE, MAPE* and *LL*)[3] for all series.

The empirical results show that the GARMA-FIGARCH model presents a higher degree of persistence in the volatility than that revealed by ARFIMA-FIGARCH model. Notice, indeed, that the long memory parameters of the FIGARCH model estimated in the forecasts for the three exchange rate series considered in first difference varie between 0.2 and 0.25. Furthermore, we observe from Table 7 that the GARMA-FIGARCH specification outperforms the ARFIMA-FIGARCH models since the GARMA-FIGARCH prediction errors are the smallest for all evaluation criteria and all forecast horizons. Given that the ARFIMA-FIGARCH model has produced the largest prediction errors, we can conclude that differencing the exchange rates can reduce the predictive performance significantly.

6 Conclusion

The principal objective of this paper is to propose a new approach to model the dynamics of exchange rates series without using a differencing operator in order to preserve the informational contents of the series. First, we use the wavelet-based detrending procedure. Then, we consider a more general class of long memory processes which take into account the persistent seasonal component besides the long range dependence observed in the conditional mean and the volatility. More precisely, we estimated a GARMA-FIGARCH model using the wavelet based approximate maximum likelihood estimator developed by Whitcher (2004).

The obtained results are encouraging and show that the exchange rate series displays evidence of generalized long memory. In particular, the seasonal long memory parameter, the G-frequency and the fractional differencing parameters are similar for all the series of studies. The predictive performance of the selected model provides evidence of the power of the selected model compared to ARFIMA-FIGARCH: The

[2] Diebold (1998) underlines that the long term forecast horizon is about ten to twenty days.

[3] MSE: the Mean Square Error, MAPE: the Mean Absolute Prediction Error expressed as a percentage and LL: the Logarithmic Loss function.

Table 7 Out of sample forecasts for ARFIMA-FIGARCH and GARMA-FIGARCH models

	Criterion	$s = 1$	$s = 5$	$s = 10$	$s = 15$
LUSD					
ARFIMA-FIGARCH	MSE	0.512	0.213	0.201	0.191
	MAPE	7.156	3.605	3.121	2.229
	LL	6.795	4.863	2.057	1.756
GARMA-FIGARCH	MSE	0.269	0.132	0.087	0.068
	MAPE	4.850	2.854	3.135	2.210
	LL	5.112	2.055	1.858	1.507
LEuro					
ARFIMA-FIGARCH	MSE	0.322	0.219	0.178	0.158
	MAPE	5.146	4.128	3.131	2.123
	LL	6.638	4.165	2.467	2.326
GARMA-FIGARCH	MSE	0.223	0.116	0.076	0.055
	MAPE	4.142	2.128	2.117	2.103
	LL	4.282	3.233	2.079	1.264
LYEN					
ARFIMA-FIGARCH	MSE	0.743	0.542	0.387	0.252
	MAPE	6.453	5.793	4.063	3.498
	LL	7.742	5.972	4.362	2.645
GARMA-FIGARCH	MSE	0.428	0.374	0.169	0.103
	MAPE	4.632	3.675	3.250	2.904
	LL	5.433	4.539	3.227	1.376

proposed approach can be useful for a better understanding of the structure of exchange rates series.

References

Baillie RT, Bollerslev T (1994) Cointegration, fractional cointegration and exchange rate dynamics. J Finance 49:737–745

Baillie RT, Bollerslev T, Mikkelsen HO (1996) Fractionally integrated generalized autoregressive conditional heteroskedasticity. J Econom 74:3–30

Brock WA, Dechert WD, Scheinkman JA (1987) A test for independence based on the correlation dimension. Working Paper, University of Wisconsin at Madison

Cheung YW (1993) Long memory in foreign-exchanges rates. J Bus Econ Stat 11:93–101

Chung C-F (1996a) Estimating a generalized long memory process. J Econom 73:237–259

Chung C-F (1996b) A generalized fractionally integrated autoregressive moving-average process. J Time Ser Anal 17:111–140

Daubechies I (1992) Ten lectures on wavelets. Volume 61 of CBMS-NSF regional conference series in applied mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia

Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 74:427–431

Dickey DA, Fuller WA (1981) The likelihood ratio statistics for autoregressive time series with a unit root. Econometrica 49:1057–1072

Diebold FX (1987) Testing for serial correlation in the presence of ARCH. Proceedings of the American Statistical Association, Business and Economic Statistics Section, pp 323–328
Diebold F (1998) Elements of forecasting. South-Western College Publishing, Cincinnati, OH
Ferrara L, Guégan D (2001) Forecasting with k-Gegenbauer processes: theory and applications. J Forecast 20:581–601
Gray HL, Zhang N-F, Woodward WA (1989) On generalized fractional processes. J Time Ser Anal 10: 233–257
Kwiatkowski D, Phillips P, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationnary against the alternative of unit root. J Econom 54:159–178
Ling S, Li WK (1997) Fractional autoregressive integrated moving-average time series conditional heteroskedasticity. J Am Stat Assoc 92:1184–1194
Mallat S (1989) A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans Pattern Anal Mach Intell 11:674–693
Morris JM, Peravali R (1999) Minimum-bandwidth discrete-time wavelets. Signal Process 76:181–193
Percival DB, Walden AT (2000) Wavelet methods for time series analysis. Cambridge University Press, Cambridge
Perron P (1997) Further evidence from breaking trend functions in macroeconomic variables. J Econom 80:355–385
Perron P (2006) Dealing with structural breaks. In: Patterson K, Mills TC (eds) Palgrave handbook of econometrics, vol 1, econometric theory. Palgrave Macmillan, Basingstoke pp 278–352
Phillips PCB, Perron P (1988) Testing for a unit root in time series regression. Biometrika 75:335–346
Whitcher B (2001) Simulating gaussian stationary processes with unbounded spectra. J Comput Graph Stat 10:112–134
Whitcher B (2004) Wavelet-based estimation for seasonal long-memory processes. Technometrics 46: 225–238
Woodward WA, Cheng QC, Gray HL (1998) A k-factor GARMA long-memory model. J Time Ser Anal 19:485–504