



ELSEVIER

Physica A 289 (2001) 543–556

PHYSICA A

www.elsevier.com/locate/physa

# Differentiating intraday seasonalities through wavelet multi-scaling

Ramazan Gençay<sup>a,\*</sup>, Faruk Selçuk<sup>b</sup>, Brandon Whitcher<sup>c</sup>

<sup>a</sup>*Department of Economics, University of Windsor, Windsor, Ontario, Canada N9B 3P4*

<sup>b</sup>*Department of Economics, Bilkent University, Bilkent 06533, Ankara, Turkey*

<sup>c</sup>*Eurandom, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

Received 14 June 2000

---

## Abstract

It is well documented that strong intraday seasonalities may induce distortions in the estimation of volatility models. These seasonalities are also the dominant source for the underlying misspecifications of the various volatility models. Therefore, an obvious route is to filter out the underlying intraday seasonalities from the data. In this paper, we propose a simple method for intraday seasonality extraction that is free of model selection parameters which may affect other intraday seasonality filtering methods. Our methodology is based on a wavelet multi-scaling approach which decomposes the data into its low- and high-frequency components through the application of a non-decimated discrete wavelet transform. It is simple to calculate, does not depend on a particular model selection criterion or model-specific parameter choices. The proposed filtering method is translation invariant, has the ability to decompose an arbitrary length series without boundary adjustments, is associated with a zero-phase filter and is circular. Being circular helps to preserve the entire sample unlike other two-sided filters where data loss occurs from the beginning and the end of the studied sample. © 2001 Elsevier Science B.V. All rights reserved.

*PACS:* C52; C53

*Keywords:* Intraday seasonalities; Multi-scaling; High-frequency foreign exchange process; Wavelets

---

## 1. Introduction

It is well documented that strong intraday seasonalities may induce distortions in the estimation of volatility models. These periodicities are also the dominant source for

---

\* Corresponding author. Fax: +1-519-9737096.

*E-mail address:* gencay@uwindsor.ca (R. Gençay).

the underlying misspecifications of the various volatility models. Therefore, an obvious route is to filter out the underlying intraday seasonalities from the high-frequency data. In the literature, it has been demonstrated that practical estimation and extraction of the intraday periodic component of the return volatility are both feasible and indispensable for meaningful intraday studies. Earlier studies of modeling intraday seasonalities are provided by Müller et al. [1], Dacorogna et al. [2] and Andersen and Bollerslev [3]. In Ref. [2], a time-invariant polynomial approximation to the daily activity in the distinct geographical regions of the foreign exchange market is adopted.<sup>1</sup> This type of de-seasonalization is appropriate for foreign exchange markets but may not be directly applicable to the stock markets data.

In Ref. [3], intraday periodicities in volatility are modeled with flexible Fourier form (FFF) as a nonlinear regression model. This approach is not market specific so that it is easily applicable to any high-frequency data such as stock or foreign exchange series. The results in Ref. [3] indicate that FFF is successful in extracting most of the intraday seasonalities, but short-term intraday periodicities remain left in the filtered returns. The estimation of the FFF regression involves selecting the interaction terms, truncation lag for the Fourier expansion and dummy variables to minimize distortions. The model selections are based on choosing models which best match the basic shapes of the periodic pattern with a minimal number of parameters. In particular, the position of the dummy variables which are included to minimize the *distortions* are based on the researcher's view of the data and are therefore model specific.

In this paper, we propose a simple method for extracting intraday seasonality which is simple to calculate and can easily be implemented as it does not depend on a particular model selection criterion or parameter choices. The proposed method is based on a wavelet multi-scaling approach which decomposes the data into its low- and high-frequency components through the application of a non-decimated discrete wavelet transform. There are two important findings which result from the methodology of this paper.

First, we can construct a model-free estimate of the foreign exchange rate volatility which is entirely disentangled from its intraday seasonalities. One way of eliminating intraday seasonalities is to work with daily and weekly aggregate data. A recent study, in this direction, is made by Andersen et al. [5] where daily volatility estimates are constructed from high-frequency data. The drawback of the Andersen et al. [5] paper is that the theoretical underpinnings are based on diffusion-theoretic motivations which are highly parametric. The validity of this approach is also based on the asymptotic approximations of a volatility measure where the number of data points per period may well be below asymptotic requirements. Andersen et al. [5] eliminate various features from the data such as weekends, several fixed holidays, moving holidays and days with 15 longest zero returns. In our approach, there is no data elimination except the weekends. In another study, Fisher et al. [6] use five different seasonal adjustment filters before analyzing the multifractality of intra-daily and daily DM–USD series. As

---

<sup>1</sup> The details of this method is presented in Ref. [4].

compared to other studies, the results of this paper indicate that our volatility estimator is free from any short-lived intraday seasonalities captured by the autocorrelations of the long-term volatility. Furthermore, our approach does not involve subjective data elimination and it is robust to misspecifications as it is fully nonparametric. Our second contribution is that we can obtain filtered and standardized returns free of intraday and inherent seasonalities. These filtered and standardized returns do not suffer from the short-lived intraday seasonality contaminations.

We briefly introduce the wavelet methodology in Section 2, including the discrete wavelet transform, a non-decimated wavelet transform – the maximal overlap discrete wavelet transform – and multiresolution analysis. In Section 3, an example from simulated data is presented. Section 4 reports the empirical results of wavelet-based multiresolution analysis of high-frequency exchange rate returns and volatility. We conclude in Section 5.

## 2. Wavelet methodology

The wavelet transform is a powerful mathematical tool that has received more and more attention in the statistical and financial communities. The power of wavelets is their ability to analyze (decompose) features which vary over both time and scale. In the past, this was achieved through the short-time Fourier transform (STFT) where the Fourier transform was applied to a portion of the signal through a sliding window (typically Gaussian). This partitions the time–frequency plane into a square grid whose dimension depends upon the window used. The wavelet transform differs from the STFT by using an entirely different set of basis functions (not sinusoids) which adaptively partition the time–frequency plane to better capture the range of low- to high-frequency events. Detailed introductions to the theory of wavelets and wavelet transforms may be found in, for example, [7–9]. An extensive wavelet methodology from the economics and finance perspective is also available in [10].

A wavelet is defined to be a function  $\psi(t)$ , whose collection of  $(j, k)$ th-order translations followed by dilations

$$\psi_{j,k}(t) \equiv 2^{-j/2} \psi(2^{-j}t - k), \quad j, k \in \mathbf{Z} = \{0, \pm 1, \pm 2, \dots\},$$

form an orthonormal basis for  $L^2(\mathbf{R})$  – the space of all square-integrable functions. This means that each basis function depends on two parameters, the scale  $j$  and locations  $k$ , whereas the Fourier basis functions only depend on a single parameter – frequency. Any continuous function may be used as a wavelet if it satisfies a weak admissibility condition.<sup>2</sup> Fig. 1 shows examples of common wavelet basis functions.

<sup>2</sup> The function rapidly decreases to zero as  $t \rightarrow \pm\infty$  and oscillates ( $\int \psi(t) dt = 0$ ). Additional conditions may be imposed on  $\psi(t)$ , such as more vanishing moments  $\int t^m \psi(t) dt = 0$ , where  $m = 0, 1, \dots, M - 1$ , or more continuous derivatives  $\psi(t) \in C^m$ . A popular class of basis functions is the Daubechies family of compactly supported wavelets [8, Section 6.2].

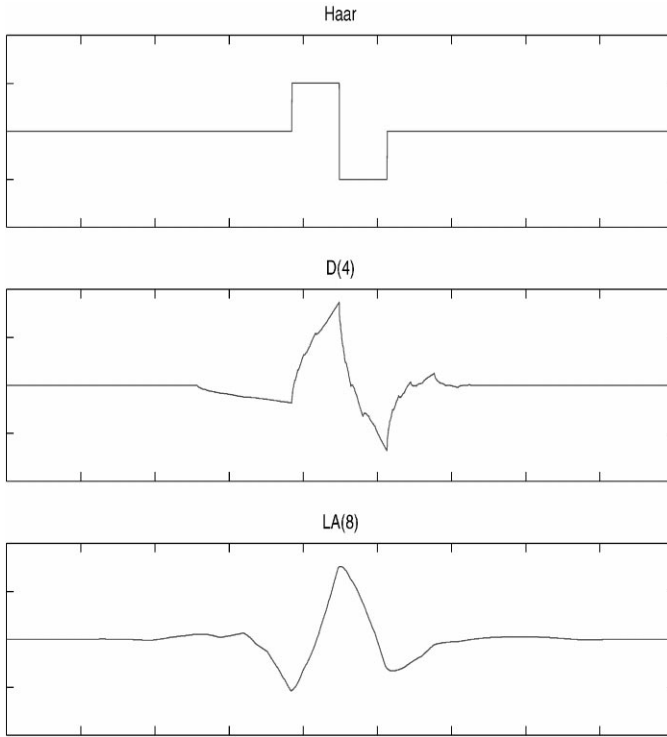


Fig. 1. Daubechies compactly supported wavelet functions: Haar wavelet based on two non-zero coefficients (also corresponds to the extremal phase wavelet based on two non-zero coefficients), the extremal phase wavelet based on four non-zero coefficients and the least asymmetric wavelet based on eight non-zero coefficients.

Let  $x(t)$  be a finite-energy signal<sup>3</sup> so that  $x(t)$  belongs to  $L^2(\mathbf{R})$ . We can analyze any function  $x(t)$  by projecting it against the wavelet basis functions. If we define

$$\langle x, y \rangle \equiv \int_{-\infty}^{\infty} x(t)y(t) dt$$

to be the inner product between the two finite-energy signals  $x(t)$  and  $y(t)$ , then the wavelet coefficient with scale parameter  $j$  and translation parameter  $k$  is equal to

$$w_{j,k} \equiv \langle x, \psi_{j,k} \rangle = \int_{-\infty}^{\infty} x(t)\psi_{j,k}(t) dt = 2^{-j/2} \int_{-\infty}^{\infty} x(t)\psi(2^{-j}t - k) dt .$$

The wavelet coefficient  $w_{j,k}$  is both localized in time and scale, unlike a Fourier coefficient which is local in frequency and global in time. If we assume the wavelet function is approximately an ideal high-pass filter, then the frequency-domain support of  $w_{j,k}$  at time  $k$  is approximately  $[-2^{-j}, -2^{-j-1}) \cup (2^{-j-1}, 2^{-j}]$ . The time-domain support of  $w_{j,k}$  is given by  $[k2^{-j}, (k+1)2^{-j}]$ . Thus, the wavelet function “zooms” in and out

<sup>3</sup>  $\int x^2(t) dt < \infty$ .

according to scale, becoming more narrow when analyzing high-frequency features and wider when analyzing low-frequency features.

A signal expansion via a orthogonal wavelet basis can be interpreted as an aggregation of details across all scales, thus providing a reconstruction formula

$$x(t) = \sum_{j=-\infty}^{\infty} d_j(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} w_{j,k} \psi_{j,k}(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \langle x, \psi_{j,k} \rangle \psi_{j,k}(t).$$

### 2.1. The discrete wavelet transform

Let  $\{h_l\} \equiv \{h_0, \dots, h_{L-1}\}$  denote the wavelet filter coefficients of a Daubechies compactly supported wavelet family, examples are given in Fig. 1, and let  $\{g_l\} \equiv \{g_0, \dots, g_{L-1}\}$  be the corresponding scaling filter coefficients, defined via the quadrature mirror relationship  $g_m = (-1)^{m+1} h_{L-1-m}$ . The wavelet filter  $\{h_l\}$  is associated with unit scale and we assume it satisfies  $\sum_{l=0}^{L-1} h_l^2 = 1$  and  $\sum_{l=0}^{L-1} h_l h_{l+2k} = 0$ .

The transfer function of a wavelet filter  $H_k = \sum_{m=0}^{N-1} h_m e^{-i2\pi mk/N}$ , for  $k=0, \dots, N-1$ , describes its band-pass nature. The wavelet filter  $\{h_l\}$  approximates an ideal high-pass filter, the accuracy of the approximation increasing with the filter length  $L$ , so that  $H_k$  has support on frequencies  $[-\frac{1}{2}, -\frac{1}{4}] \cup (\frac{1}{4}, \frac{1}{2}]$ . Let  $G_k$  denote the DFT of  $\{g_l\}$ . The scaling filter approximates an ideal low-pass filter implying  $G_k$  has support on frequencies  $[-\frac{1}{4}, \frac{1}{4}]$ .

Now define the wavelet filter  $\{h_{j,l}\}$  for scale  $2^{j-1}$  as the inverse DFT of

$$H_{j,k} = H_{2^{j-1}k \bmod N} \prod_{l=0}^{j-2} G_{2^l k \bmod N}, \quad k = 0, \dots, N-1.$$

The transfer function  $H_{j,k}$  can be interpreted as coming from successive low-pass filtering (averaging) operations on increasing scales and a final high-pass filtering (differencing) operation. The resulting wavelet filter associated with scale  $2^{j-1}$  has length  $\min\{N, L_j\}$ , where  $L_j \equiv (2^j - 1)(L - 1) + 1$ . Also, define the scaling filter  $\{g_{J,l}\}$  for scale  $2^{J-1}$  as the inverse DFT of

$$G_{J,k} = \prod_{l=0}^{J-1} G_{2^l k \bmod N}, \quad k = 0, \dots, N-1.$$

The transfer function  $G_{j,k}$  can be interpreted as coming from a sequence of low-pass filtering (averaging) operations at increasing scales resulting in a low-frequency band-pass filter.

Let  $\mathcal{W}$  be an  $N \times N$  matrix defining a  $J$ th-order partial orthonormal DWT based upon a Daubechies wavelet filter of even length  $L \leq N$ . The rows of  $\mathcal{W}$  consist of circularly shifted (by multiples of 2) versions of the zero-padded wavelet filters for scale  $2^{j-1}$ , defined via

$$\mathbf{h}_j \equiv [h_{j,0} \ 0 \ \dots \ 0 \ h_{j,L_j-1} \ h_{j,L_j-2} \ \dots \ h_{j,2} \ h_{j,1}]^T, \tag{1}$$

where the non-zero wavelet filter coefficients are in reverse order. Constructing a matrix from all possible circular shifts, at a particular scale  $2^{j-1}$ , of Eq. (1) yields the

sub-matrix  $\mathcal{W}_j$ . This allows us to think of the orthonormal matrix  $\mathcal{W}$  being comprised of several sub-matrices, each one stacked on top of the other; i.e.,

$$\mathcal{W} = \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{W}_2 \\ \vdots \\ \mathcal{W}_J \\ \mathcal{V}_J \end{bmatrix}.$$

For example, when  $L = 4$  and  $N > 4$  we get

$$\mathcal{W}_1 = \begin{bmatrix} h_{1,1} & h_{1,0} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & h_{1,3} & h_{1,2} \\ h_{1,3} & h_{1,2} & h_{1,1} & h_{1,0} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & h_{1,3} & h_{1,1} & h_{1,1} & h_{1,0} & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & h_{1,3} & h_{1,2} & h_{1,1} & h_{1,0} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & h_{1,3} & h_{1,2} & h_{1,1} & h_{1,0} \end{bmatrix}, \tag{2}$$

where  $\mathcal{W}_1$  is a  $N/2 \times N$  matrix whose rows are  $\mathbf{h}_1$  circularly shifted by  $2m - 1$  for  $m = 1, \dots, N/2$ . The remaining sub-matrices  $\mathcal{W}_2, \dots, \mathcal{W}_J$  are defined similarly to Eq. (2), being shifted by  $2^j m - 1$  for  $m = 1, \dots, N/2^j$ , and  $\mathcal{V}_J$  is identical in dimension to  $\mathcal{W}_J$  but contains circularly shifted versions of  $\mathbf{g}_J$ , instead of  $\mathbf{h}_J$ , by  $2^j m - 1$  for  $m = 1, \dots, N/2^j$ .

The scale-dependent shifts are equivalent to downsampling (or decimation of) the filtered output and insure that the transform is orthogonal. This is because the filters were designed to be orthogonal to their even shifts. In practice, the rows of the matrix  $\mathcal{W}$  are not explicitly constructed, but instead the DWT is implemented via the pyramid algorithm of Mallat [11]. The pyramid algorithm applies wavelet coefficients to the input series and subsamples the output one scale at a time.

When applied to a vector of observations  $\mathbf{X}$ , the DWT yields  $N$  wavelet coefficients  $\mathbf{W} = \mathcal{W}\mathbf{X}$ , which can be organized into  $J + 1$  vectors  $\mathbf{W} = [\mathbf{W}_1 \dots \mathbf{W}_J \mathbf{V}_J]^T$ , similar to  $\mathcal{W}$  above, where  $\mathbf{W}_j$  is a length  $N/2^j$  vector of wavelet coefficients associated with changes on a scale of length  $2^{j-1}$  and  $\mathbf{V}_J$  is a length  $N/2^J$  vector of scaling coefficients associated with averages on a scale of length  $2^J$ .

Like the DFT, orthonormality of the matrix  $\mathcal{W}$  implies that the DWT is an energy preserving transform so that  $\|\mathbf{W}\|^2 = \|\mathbf{X}\|^2$ . Given the structure of the wavelet coefficients, the energy in  $\mathbf{X}$  is decomposed on a scale by scale basis via

$$\|\mathbf{X}\|^2 = \sum_{j=1}^J \|\mathbf{W}_j\|^2 + \|\mathbf{V}_J\|^2, \tag{3}$$

where  $\|\mathbf{W}_j\|^2$  is the energy of  $\mathbf{X}$  due to changes at scale  $2^{j-1}$  and  $\|\mathbf{V}_J\|^2$  is the energy due to changes at scales  $2^{J-1}$  and higher.

2.2. The maximal overlap discrete wavelet transform

The DWT is a very useful operation, but does not possess all the attributes which may be desirable for certain applications. Problems are induced by the downsampling, or decimation, involved in computing the transform. In response to this, a non-decimated wavelet transform has been developed – the maximal overlap DWT (MODWT) (see, for example, Ref. [12]). The MODWT goes by several names in the literature, such as the ‘stationary DWT’ by Nason and Silverman [13] and the ‘translation-invariant DWT’ by Coifman and Donoho [14].

The MODWT gives up orthogonality in order to gain other features the DWT does not possess; such as the translation invariance and the ability to decompose an arbitrary length series without boundary adjustments. It does this by *not* decimating the filtered output at each scale. A consequence of this is that the wavelet and scaling coefficients must be rescaled in order to retain the energy-preserving property of the DWT. A thorough discussion of the MODWT will appear in [15, Chapter 5].

The notation follows from the DWT, with the  $J$ th order partial MODWT being defined by  $\tilde{\mathbf{W}} = \tilde{\mathcal{W}} \mathbf{X}$ , where  $\tilde{\mathbf{W}}$  is composed of  $J + 1$  length  $N$  vectors,  $\tilde{\mathbf{W}}_1, \dots, \tilde{\mathbf{W}}_J$  and  $\tilde{\mathbf{V}}_J$ , which can be arranged in the following manner:

$$\tilde{\mathbf{W}} \equiv [\tilde{\mathbf{W}}_1 \ \tilde{\mathbf{W}}_2 \ \dots \ \tilde{\mathbf{W}}_J \ \tilde{\mathbf{V}}_J]^T .$$

The vector of wavelet coefficients  $\tilde{\mathbf{W}}_j$  is associated with changes of length  $2^{J-1}$  and  $\tilde{\mathbf{V}}_J$  is associated with averages of lengths  $2^{J-1}$  and higher – just like the DWT.

Similar to the matrix  $\mathcal{W}$  for the DWT, the matrix  $\tilde{\mathcal{W}}$  is also made up of  $J + 1$  sub-matrices, each of them  $N \times N$ , and may be expressed via

$$\tilde{\mathcal{W}} = \begin{bmatrix} \tilde{\mathcal{W}}_1 \\ \tilde{\mathcal{W}}_2 \\ \vdots \\ \tilde{\mathcal{W}}_J \\ \tilde{\mathcal{V}}_J \end{bmatrix} .$$

In this case, when  $L = 4$  and  $N > 4$ , we have

$$\tilde{\mathcal{W}}_1 = \begin{bmatrix} \tilde{h}_{1,0} & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \tilde{h}_{1,3} & \tilde{h}_{1,2} & \tilde{h}_{1,1} \\ \tilde{h}_{1,1} & \tilde{h}_{1,0} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \tilde{h}_{1,3} & \tilde{h}_{1,2} \\ \tilde{h}_{1,2} & \tilde{h}_{1,1} & \tilde{h}_{1,0} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{h}_{1,3} \\ \tilde{h}_{1,3} & \tilde{h}_{1,2} & \tilde{h}_{1,1} & \tilde{h}_{1,0} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \tilde{h}_{1,3} & \tilde{h}_{1,2} & \tilde{h}_{1,1} & \tilde{h}_{1,0} & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tilde{h}_{1,3} & \tilde{h}_{1,2} & \tilde{h}_{1,1} & \tilde{h}_{1,0} & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & \tilde{h}_{1,3} & \tilde{h}_{1,2} & \tilde{h}_{1,1} & \tilde{h}_{1,0} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \tilde{h}_{1,3} & \tilde{h}_{1,2} & \tilde{h}_{1,1} & \tilde{h}_{1,0} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \tilde{h}_{1,3} & \tilde{h}_{1,2} & \tilde{h}_{1,1} & \tilde{h}_{1,0} \end{bmatrix} , \tag{4}$$

where  $\widetilde{\mathcal{W}}_1$  is a  $N \times N$  matrix, and the rows of the matrix  $\tilde{\mathbf{h}}_1 = \mathbf{h}_1/2^{1/2}$  are simply the rescaled wavelet filter coefficients circularly shifted by  $m - 1$  for  $m = 1, \dots, N$ . In general, let  $\tilde{\mathbf{h}}_j \equiv \mathbf{h}_j/2^{j/2}$  and  $\tilde{\mathbf{g}}_j \equiv \mathbf{g}_j/2^{j/2}$  be, respectively, the rescaled wavelet and scaling filter coefficients required to construct  $\widetilde{\mathcal{W}}$ . The remaining sub-matrices  $\widetilde{\mathcal{W}}_2, \dots, \widetilde{\mathcal{W}}_J$  are constructed similarly to Eq. (4) and  $\widetilde{\mathcal{V}}_J$  has the same structure as  $\widetilde{\mathcal{W}}_J$  only using circularly shifted scaling coefficients instead of wavelet coefficients. Circular shifting for all scales is identical to that of Eq. (4). In practice, a pyramid scheme is utilized similar to that of the DWT (see Ref. [12]). For time series of dyadic length<sup>4</sup> the MODWT may be sub-sampled and rescaled to obtain an orthonormal DWT.

Percival and Mofjeld [12] showed that the MODWT is an energy-preserving transform in the sense that

$$\|\mathbf{X}\|^2 = \sum_{j=1}^J \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_J\|^2.$$

This allows for a scale-based analysis of variance of a time series similar to spectral analysis via the DFT. In a wavelet analysis of variance, the individual wavelet coefficients are associated with a band of frequencies and specific time scale whereas Fourier coefficients are associated with a specific frequency only. Percival [16] showed that the MODWT-based estimator of wavelet variance to be asymptotically more efficient estimator over the DWT-based estimator.

### 2.3. Multiresolution analysis

Using the DWT, we may formulate an additive decomposition of a series of observations. The notation and terminology used here closely follows Percival and Walden [17]. Let  $\mathcal{D}_j \equiv \mathcal{W}_j^T \mathbf{W}_j$  for  $j=1, \dots, J$ , define the  $j$ th level *wavelet detail* associated with changes in  $\mathbf{X}$  at scale  $2^{j-1}$ . The wavelet coefficients  $\mathbf{W}_j$  represent the portion of the wavelet analysis (decomposition) for scale  $2^{j-1}$ , while  $\mathcal{D}_j$  is the portion of the wavelet synthesis for the same scale. The final wavelet detail is defined to be  $\mathcal{D}_{J+1} \equiv \mathcal{V}_J^T \mathbf{V}_J$ , and is equal to the sample mean of the observations.

A multiresolution analysis (MRA) may now be defined via  $\mathbf{X} = \sum_{j=1}^{J+1} \mathcal{D}_j$ . Thus, the wavelet details form an additive decomposition of the original series. Let  $\mathcal{S}_j \equiv \sum_{k=j+1}^{J+1} \mathcal{D}_k$  define the  $j$ th level *wavelet smooth* for  $0 \leq j \leq J$ , where  $\mathcal{S}_{J+1}$  is defined to be a vector of zeros. Whereas the wavelet detail  $\mathcal{D}_j$  is associated with variations at a particular scale,  $\mathcal{S}_j$  is a cumulative sum of these variations and will be smoother and smoother as  $j$  increases. In fact,  $\mathbf{X} - \mathcal{S}_j = \sum_{k=1}^j \mathcal{D}_k$  so that only lower scale details (high-frequency features) will be apparent. These lower scale details may be represented through the  $j$ th level *wavelet rough*  $\mathcal{R}_j \equiv \sum_{k=1}^j \mathcal{D}_k$  for  $1 \leq j \leq J + 1$ , where  $\mathcal{R}_0$  is defined to be a vector of zeros. Hence, the vector of observations may be represented via a wavelet smooth and rough, i.e.,

$$\mathbf{X} = \mathcal{S}_j + \mathcal{R}_j \quad \text{for all } j.$$

<sup>4</sup> A time series has dyadic length if  $N = 2^m$  for some positive integer  $m$ .

An analogous MRA may be performed utilizing the MODWT via  $\mathbf{X} = \sum_{j=1}^{J+1} \tilde{\mathcal{D}}_j$ , where  $\tilde{\mathcal{D}}_j \equiv \tilde{\mathcal{W}}_j^T \tilde{\mathbf{W}}_j$  for  $j = 1, \dots, J$ . We may also define the MODWT-based wavelet smooths and roughs to be  $\tilde{\mathcal{P}}_J \equiv \sum_{k=j+1}^{J+1} \tilde{\mathcal{D}}_k$  and  $\tilde{\mathcal{R}}_j \equiv \sum_{k=1}^j \tilde{\mathcal{D}}_k$ , respectively. A key feature to an MRA using the MODWT is that the wavelet details and smooth are associated with zero-phase filters; see, for example, Percival and Walden [17, Section 5.4]. Thus, interesting features in the wavelet details and smooth may be aligned perfectly with events in the original time series. This attribute is not available through the DWT since it downsamples the output of the filtering operations.

### 3. An example

The presence of seasonalities (periodicities) in a long memory process may obscure the underlying low-frequency dynamics. Specifically, the periodic component pulls the calculated autocorrelations down, giving the impression that there is no persistence other than particular periodicities. Consider the following AR (1) process with a periodic component:

$$y_t = \alpha + \beta y_{t-1} + \sum_{i=1}^4 3.0 S_{it} + \varepsilon_t, \quad t = 1, \dots, T, \tag{5}$$

where  $S_{it} = \sin((2\pi/P_i)t) + \eta v_{it}$ ,  $\alpha = 0.0$ ,  $y_0 = 1.0$ ,  $\beta = 0.95$  and  $T = 1000$ . Periodic components are  $P_1 = 3$ ,  $P_2 = 4$ ,  $P_3 = 5$ , and  $P_4 = 6$  so that the process has 3, 4, 5, and 6 period stochastic seasonality. The random variables  $\varepsilon_t$  and  $v_{it}$  are identically and independently distributed disturbance terms with zero mean. The signal-to-noise ratio,  $\eta$ , in each seasonal component is set to 0.30.

Fig. 2 presents the autocorrelograms of the simulated AR (1) process with and without the periodic components. The autocorrelogram of the AR (1) process without seasonality (excluding  $\sum 3.0 S_{it}$  from the simulated process) starts from a value of 0.95 and decays hyperbolically as expected. However, the autocorrelogram of the AR (1) process with the seasonality starts from 0.40 and indicates the existence of a periodic component. The underlying long memory persistence of the AR (1) process in the absence of the seasonality component is entirely obscured by these periodic components.

A well-designed seasonal adjustment procedure, therefore, should *clean* the data from its seasonal components and leave the underlying inherent non-seasonal structure intact. In the example above (see Fig. 2), the solid line is the autocorrelogram of the non-seasonal AR (1) dynamics and the dotted lines are the autocorrelogram of the de-seasonalized series with the method proposed in this paper. The simulated AR (1) process in Eq. (5) is decomposed into a wavelet smooth and 2 wavelet details.<sup>5</sup> The wavelet detail  $\tilde{\mathcal{D}}_1$  (associated with changes on the unit scale) captures frequencies

<sup>5</sup> The Daubechies least asymmetric wavelet family of length  $L = 8$  was utilized in all wavelet transformations. The results are not very sensitive to the choice of the wavelet family as long as the underlying process is stationary or an integer difference of the process is stationary.

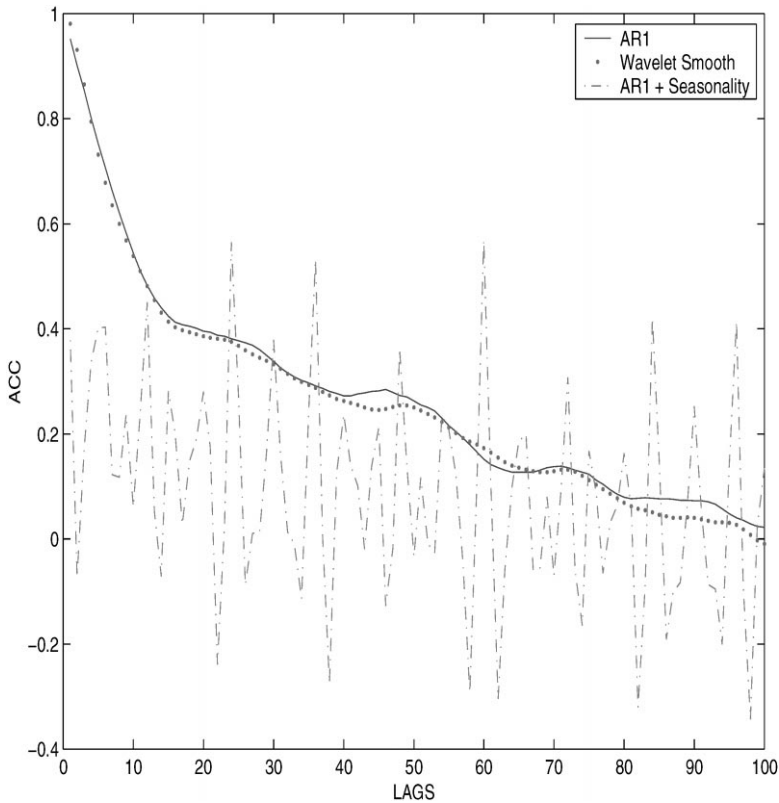


Fig. 2. Sample autocorrelograms for the simulated AR1 process (—), AR1 plus seasonality process (- - - - -), and MODWT smooth of the AR1 plus seasonality process (· · · · ·).

$\frac{1}{4} \leq f \leq \frac{1}{2}$ , i.e., any oscillation with a period length of 2–4. Similarly,  $\tilde{\mathcal{D}}_2$  contains frequencies  $\frac{1}{8} \leq f \leq \frac{1}{4}$ , any oscillation with a period length of 4–8. Therefore, it is expected that wavelet smooth only contains long-memory dynamics and is free of seasonalities. As Fig. 2 demonstrates, our methodology successfully uncovers the long-memory dynamics without imposing any spurious persistence into the filtered series.

#### 4. Empirical findings

The studied data sets are the 5-min Deutschemark – US Dollar (DEM–USD) and Japanese Yen – US Dollar (JPY–USD) price series for the period from October 1, 1992 to September 29, 1993. The data set<sup>6</sup> is provided by Olsen and Associates in Zurich, Switzerland. Bid and ask prices at each 5 min interval are obtained by linear interpolation over time as in [1,2]. Prices are computed as the average of the logarithm

<sup>6</sup> This data set is also known as HFDF-I data set.

of the bid and ask prices:

$$P_t = \frac{1}{2}[\log P(\text{bid})_t + \log P(\text{ask})_t] \quad \text{and} \quad t = 1, \dots, 74,800. \quad (6)$$

Olsen and Associates applied data cleaning filters to the price series (as received from Reuters) in order to correct for data errors and to remove suspected outliers. We also removed the weekend quotes from Friday 21:05 GMT to Sunday 21:00 GMT. Apart from this, we did not apply any further filtering to the data set nor did we exclude any data points.<sup>7</sup> Continuously, compounded 5-min returns are calculated as the log difference of the prices and presented as

$$r_t = (\log P_t - \log P_{t-1})100. \quad (7)$$

In Figs. 3(a) and (b) autocorrelograms of the 5-min absolute return series are presented. These figures show that the intra-daily absolute returns exhibit strong intraday

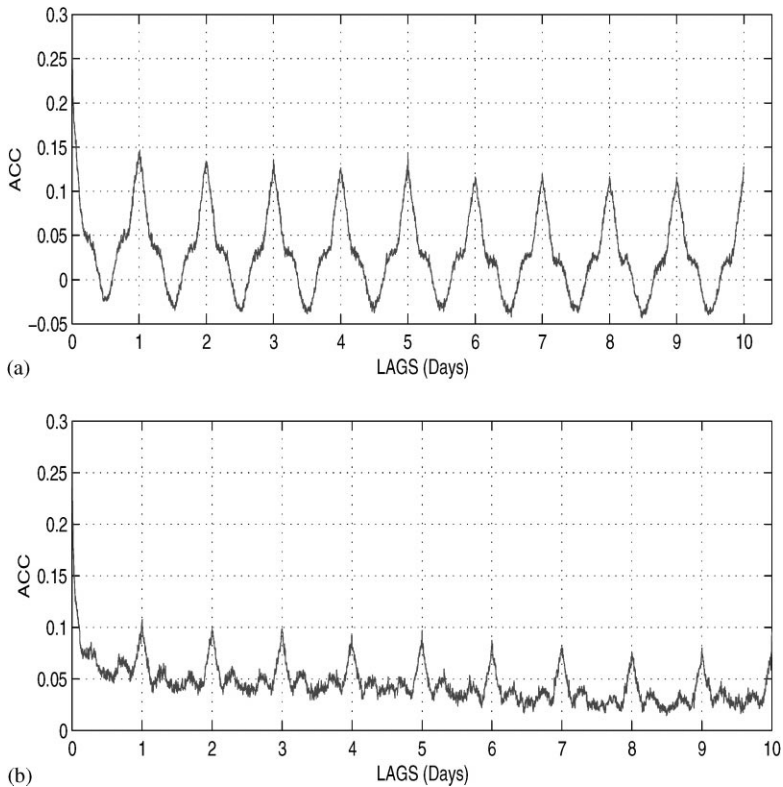


Fig. 3. Sample autocorrelogram for the 5-min absolute returns of (a) Deutschemark–US Dollar spot exchange rate and (b) Japanese Yen–US Dollar spot exchange rate from October 1, 1992 to September 29, 1993.

<sup>7</sup> Andersen et al. [5] utilized a longer (10 years) sample of DEM–USD and JPY–USD series. They removed weekends and several (mostly North American) holidays from the sample. They have also excluded the days containing “15 longest zero and constant runs”. Andersen and Bollerslev [3,17] analyzed the same data set. They also removed the weekend quotes from their sample.

seasonalities. This phenomenon is well known and reported extensively in the literature (see, for example, Refs. [2,3]).

Our model of intraday returns is similar to that in Andersen and Bollerslev [3]

$$r_t = v_t s_t \varepsilon_t, \quad (8)$$

where  $r_t$  is the raw returns,  $v_t$  is the long-term volatility,  $s_t$  is the seasonal volatility and  $\varepsilon_t$  is the identically and independently distributed innovations. Squaring both sides of Eq. (8), taking in natural logarithm and dividing both sides by two leads to

$$\log|r_t| = \log|v_t| + \log|s_t| + \log|\varepsilon_t|. \quad (9)$$

Eq. (9) provides an additive separation of the long memory, and the seasonal decomposition of volatility in the process. We use  $\log|r_t|$  to obtain the MODWT decomposition of the DEM–USD and the JPY–USD series. An eight level MODWT is utilized to decompose the  $\log|r_t|$  at the 5-min frequency. The Daubechies least asymmetric family of wavelets ( $LA(8)$ ) was utilized in maximal overlap discrete wavelet transformation. The highest level detail (level 8 detail) captures frequencies  $\frac{1}{512} \leq f \leq \frac{1}{256}$ ; i.e., any oscillation with a period length of 256–512. Since there are 288 5-min returns per day, details from 1 to 8 will contain all intra-day periodicities. The filtered returns are defined as

$$r_t(f) = \frac{r_t}{|f_t|}, \quad (10)$$

where  $\log(f_t) = \log f_{t,1} + \log f_{t,2} + \log f_{t,3} + \dots + \log f_{t,8}$ .  $\log f_t$  corresponds to intra-day seasonal volatility ( $s_t$ ) and high-frequency components of the innovations ( $\varepsilon_t$ ) obtained from MODWT details. For example,  $\log f_{t,1}$  is the first detail in MODWT and it contains 10–20 min periodicities and the highest frequency part of the innovations. Similarly,  $\log f_{t,2}$  is the detail 2 and it contains 20–40 min periodicities and the second highest frequency part of the innovations. The highest detail  $\log f_{t,8}$  contains 1280 min (approximately 21 h) to 2560 min (approximately 43 h) periodicities. The filtered absolute returns, therefore, are free from any intra-day periodicities and innovations.

For a long-memory process (see Ref. [18]), the autocovariance function at lag  $k$  satisfies  $\gamma(k) \sim \lambda k^{-\alpha}$ , where  $\lambda$  is the scaling parameter and  $\alpha \in [0, 1]$ . A leading example is the fractionally integrated process for which  $\alpha = 1 - 2d$  and  $d$  is the order of fractional integration. In Ref. [3], the fractional order of integration is estimated as  $d = 0.36$  for the same DEM–USD series utilized in this paper. Andersen et al. [5] calculate six  $d$  estimates from various volatility measures for the DEM–USD and JPY–USD series. These six  $d$  estimates vary from 0.346 to 0.448. In our calculation below, we therefore set  $d = 0.4$  to represent the average of these six estimates. In Fig. 4, we present the autocorrelograms of the filtered 5-min absolute returns along with the estimated autocorrelogram of a long memory process with  $d = 0.4$ .

The autocorrelograms for the filtered absolute returns exhibit hyperbolic decay. The rate of this decay mimics the hyperbolic decay observed in a fractionally integrated process with the fractional integrating parameter,  $d = 0.4$ . This decay rate is similar across both DEM–USD and JPY–USD series.

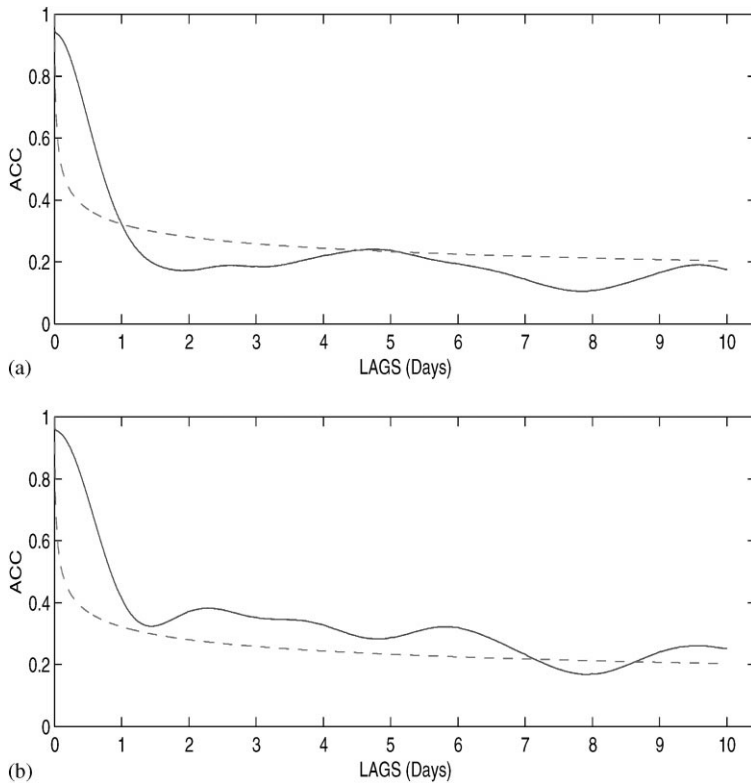


Fig. 4. Sample autocorrelogram for the MODWT filtered 5-min absolute returns of (a) Deutschemark–US Dollar spot exchange rate and (b) Japanese Yen–US Dollar spot exchange rate from October 1, 1992 to September 29, 1993. The dotted line is the autocorrelogram for the estimated hyperbolic decay rate for  $d = 0.40$ , i.e.,  $k^{2.00.40-1} = k^{-0.20}$  where  $k$  is the number of lags.

## 5. Conclusions

In this paper, we have proposed a simple method for intraday seasonality extraction which is free of model selection parameters which affect other intraday seasonality filtering methods. Our methodology is based on a wavelet multi-scaling approach which decomposes the data into its low- and high-frequency scales through the application of a non-decimated discrete wavelet transformation. It is simple to calculate and can easily be implemented as it does not depend on a particular model selection criterion and model specific parameter choices. The proposed filtering method has the translation invariance property, has the ability to decompose an arbitrary length series without boundary adjustments, possesses the zero-phase property and it is circular. The circularity property helps to preserve the entire sample unlike other two-sided filters where data loss occurs from the beginning and the end of the studied sample.

## Acknowledgements

Ramazan Gençay gratefully acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada.

## References

- [1] U.A. Müller, M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwarz, C. Morgeneegg, Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis, *J. Banking Finance* 14 (1990) 1189–1208.
- [2] M.M. Dacorogna, U.A. Müller, R.J. Nagler, R.B. Olsen, O.V. Pictet, A geographical model for the daily and weekly seasonal volatility in the foreign exchange markets, *J. Int. Money Finance* 12 (1993) 413–438.
- [3] T.G. Andersen, T. Bollerslev, Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns, *J. Finance* 52 (1997) 975–1005.
- [4] M.M. Dacorogna, R. Gençay, U.A. Müller, R.B. Olsen, O.V. Pictet, *An Introduction to High Frequency Finance*, Academic Press, New York, 2001, forthcoming.
- [5] T.G. Andersen, T. Bollerslev, F.X. Diebold, P. Labys, The distribution of exchange rate volatility, 2000, *JASA*, forthcoming.
- [6] A. Fisher, L. Calvet, B. Mandelbrot, Multifractality of Deutschemark/US Dollar exchange rates, manuscript, Cowles Foundation, Yale University, 1997.
- [7] C.K. Chui, *An Introduction to Wavelets, Wavelet Analysis and its Applications*, Vol. 1, Academic Press, San Diego, 1992.
- [8] I. Daubechies, *Ten Lectures on Wavelets*, CBMS–NSF Regional Conference Series in Applied Mathematics, Vol. 61, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [9] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, 1998.
- [10] R. Gençay, F. Selçuk, B. Whitcher, *Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press, New York, 2001, forthcoming.
- [11] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1989) 674–693.
- [12] D.B. Percival, H.O. Mofjeld, Analysis of subtidal coastal sea level fluctuations using wavelets, *J. Amer. Stat. Assoc.* 92 (1997) 868–880.
- [13] G.P. Nason, B.W. Silverman, The stationary wavelet transform and some statistical applications, in: A. Antoniadis, G. Oppenheim (Eds.), *Wavelets and Statistics*, Lecture Notes in Statistics, Vol. 103, Springer, New York, 1995, pp. 281–300.
- [14] R.R. Coifman, D.L. Donoho, Translation-invariant de-noising, in: A. Antoniadis, G. Oppenheim (Eds.), *Wavelets and Statistics*, Lecture Notes in Statistics, Vol. 103, Springer, New York, 1995, pp. 125–150.
- [15] D.B. Percival, A.T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge, 2000.
- [16] D.B. Percival, On estimation of the wavelet variance, *Biometrika* 82 (1995) 619–631.
- [17] T.G. Andersen, T. Bollerslev, DM-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer-run dependencies, *J. Finance* 53 (1998) 219–265.
- [18] J.R.M. Hosking, Asymptotic distribution of the sample, mean, autocovariances and autocorrelations of long memory time series, *J. Econometrics* 73 (1996) 261–284.