Université de Sousse
Institut Supérieur de Gestion
&
Computational Mathematics Laboratory

Thèse

Pour obtenir le grade de

Docteur de l'Université de Sousse en Gestion

Wavelet and Its Applications in Financial and Economic Time Series

Présentée et soutenue publiquement le 25 Décembre 2012 par

Mohamed Essaied Hamrita

Devant le Jury

Monsieur le Professeur Slaheddine Hellara	Président du jury
Monsieur le Professeur Samir Ben Ammou	Directeur de thèse
Monsieur le Professeur Adel Alimi	Rapporteur
Monsieur le Professeur Ezzeddine Zahrouni	Examinateur
Madame le professeur Rossitsa Yalamova	Rapporteur

Université de Sousse

Institut Supèrieur de Gestion

Thèse de doctorat

en sciences de gestion

Wavelet and Its Applications in Financial and Economic Time Series

Élaborée par: Mohamed Essaied Hamrita

Sous la direction du professeur Samir Ben Ammou

Année universitaire 2011/2012

Remerciement

Je tiens tout d'abord à remercier chaleureusement mon directeur de thèse, Monsieur Samir Ben Ammou, pour la qualité exceptionnelle de son encadrement. Il m'a en effet guidé pendant toutes ces années, se rendant toujours disponible et relançant ma motivation lorsque celle-ci diminuait. Il a fait preuve à la fois d'énormément de pédagogie et de psychologie pour me faire avancer tout en me poussant peu à peu vers une plus grande autonomie. Et ce toujours dans la bonne humeur.

Je voudrais ensuite remercier Monsieur Anouar Ben Mabrouk, pour les innombrables discussions, scientifiques ou non, que nous avons eues, pour toutes les bonnes idées qu'il a eues et pour tous les conseils avisés qu'il m'a donnés, au sein du Computational Mathematics Laboratory depuis ma première année de thèse.

J'aimerais également exprimer toute ma reconnaissance à Monsieur Adel Alimi et Madame Rossitsa Yalamova qui ont accepté d'être mes rapporteurs. Je suis aussi très heureux que Monsieur Slaheddine Hellara et Monsieur Ezzeddine Zahrouni m'aient fait l'honneur de compléter mon jury.

Je voudrais de plus remercier toutes les autres personnes qui ont contribué au bon déroulement de ma thèse et à mon épanouissement. Pour finir, je remercie les personnes qui m'ont le plus soutenu au quotidien : ma femme, mes parents et mes amis.

A ma femme et ma petite Sarah

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction and motivation

The analysis of time series has often been difficult when data do not conform to well studied theoretical concepts. One of the most common statistical properties violated by time series is stationarity (a time series is considered weakly or second order stationary when it has a mean and autocovariance sequence that do not vary with time). In addition, tradition financial analysis was based on econometric time series modeling, almost exclusively in the time domain. Time series analysis has been developed in two complementary directions. One is the spectral analysis in frequency domain in communications engineering and another which is analysis of correlation in time domain in mathematical statistics and econometrics. Wavelet methods provide an efficient way to decompose time series from the time domain into scale (or frequency) domain and to localize the changes across time within different scales whiling holding the entropy conservation. The locality property and several other interesting properties relevant to scientific research, of the wavelets transformation enable us to analyze stochastic nonstationary processes. Fourier transform has been one of the most important techniques in spectrum analysis but lack the localization property. A refinement of the conventional Fourier transform is wavelet method.

The wavelet transforms is a power full mathematical tool that is receiving more and more attention by the statistical community. While most work is being done in the engineering and physical sciences, wavelet transforms have already proven useful in well established statistical fields such as nonparametric regression, classification, and time series analysis.

The basic idea of wavelets is to expand a function of interest in terms of basis functions which are all dilates and translates of a single wavelet function, called the mother wavelet.
The scale of dilation is a power of 2, and at each scale every integer translation of the function is considered. This means that a wavelet expansion gives a multiresolution analysis of a function of interest: the wavelet analysis gives an array of coefficients, where the coefficient at translate k of level j gives information about the original function localized in both scale (2^j) and time ($k/2^j$).

A topic of considerable recent interest is the studies of correlations in economic time series and, if so, how to best quantify these correlations. Correlation is one of the most important parameters that needs to be estimated in the context of Modern Portfolio Theory. The Capital Asset Pricing Model, CAPM, and the Arbitrage Pricing Theory (APT), (Campbell et al., 1997) use correlation as a measure of dependence between different financial instruments in order to arrive at an optimal portfolio selection. Surprisingly, correlation estimation has not received significant attention in the finance literature compared to other estimators such as volatility. In time series analysis, another central topic interest is the detection and location of nonstationary events in time series which may exhibit long memory structure.
First : change point detection is a well studied field in statistics. Detecting a change in variance has a much smaller amount of literature associated with it. Techniques include, but are not restricted to, Fourier methods, cumulative sum of squares methods, parametric time series methods and Bayesian methods. Re-

searchers have also investigated detecting and locating not single change of variance, but multiple changes. Techniques include a cumulative sum of squares method (Inclan and Tiao, 1994) and information criterion method (Chen and Gupta, 1997).

Second : the discrete wavelet transform (DWT) has been shown approximately decorrelate time series with long memory structure; see for example, (Tewfik and Kim, 1992) and (McCoy and Walden, 1996).

(Whitcher et al., 1999) takes advantage of this approximate *decorrelation* of the DWT and simplicity of cumulative sum of squares method to test for homogeneity of variance, on scale by scale basis, of long memory processes.

Another interest in time series analysis is the modeling the phenomenon of long-range dependence in observed data that show a strong persistence of their correlations by long-memory processes. Such data can typically be found in the applied sciences such as hydrology, geophysics, climatology and telecommunication (e.g.teletraffic data) but recently also in economics and in finance, e.g.for modeling (realized) volatility of exchange rate data or stocks.

In this study, we tried to explore those theories and techniques and applied wavelet analysis approach on financial and economic time series analysis. This study describe a new method to study the dynamics of stock prices and exam the reference dependent property financial markets.

The central goals of this study are: **first**, wavelet methods provide a new way to describe and exam stock prices; **second**, wavelets techniques help to detect the structure of economic and financial time series among different time horizons.

This thesis structure is as the following:

In the introductory chapter 1, we first review the basis of wavelet theory, and then provide a survey of wavelet applications in financial and economic time series and statistics.

Chapter 2 is motivated by two interesting questions:

Correlation analysis and detecting and locating change point. The wavelet covariance is shown to decompose the covariance between two stationary processes on scale by scale basis. The wavelet cross-covariance and cross-correlation are also defined in order to perform a more through scale by scale analysis of bivariate time series.

Fractional difference processes (FD) are introduced in chapter 3. In this chapter we explain the estimation methods of fractional differencing parameter in wavelet domain. We apply these estimators to the real stock data.

In chapter 4, we refer to the earlier work of analysis in the frequency domain. A different definition of causality is made, and its implications to the general idea of causality are discussed. The causality relationship between export, import and Growth Domestic Product (GDP) is analyzed using wavelet time-scale decomposition. An accurate analysis of co-integrated long-run equilibrium or causal relationships among these macroeconomic indicators is important. We consider the case of the Tunisian economy and explore and extract interesting relationships using wavelet technique. Using quarterly data over the time period 1961:1-2007:4, after filtering the series using wavelet technique, we have analyzed the time series properties of the exports, imports and economic growth variables in order to determine the appropriate functional form for testing the ELG hypothesis.

The study find that, the exports, imports and GDP are co-integrated. From the causality tests we have seen that there exist a bi-directional relationship between the Exports and GDP, no relevant causality between import and export growths at 10% level of significance and a bi-directional relationship between import and economic growths.

CHAPTER 2

Litterature in Wavelets

2.1 Introduction to Wavelets

2.1.1 Spectral analysis

Empirical finance is based on statistical analysis of stochastic financial time series. Emphasis on the study of descriptive statistical properties and autocorrelation structures and assume that the data generating process is second order stationary. Historically, time series analysis has been developed in two directions. One is focus on the study of spectral density function in frequency domain along the lines of communication engineering; another one is to study the statistical property of autocorrelation function in time domain in mathematical statistics. However, both of them are constrained under the stationarity paradigm and the relation of spectral density function and autocorrelation function can be bridged through the Fourier transformation. Normalized power density function is simply the Fourier transformation of autocorrelation function of a time series. Generally, spectral analysis orthogonally transforms a function (time series) into a series of components (other time series) with different amplitudes and frequencies, and the transformation process preserves the energy of original function and holds the additive property. Orthogonal Fourier transformation is ready to apply to deterministic function but cannot be directly apply to stochastic process. However, if we treat a time series as one realization of a stationary random process and with the help of Wiener's theory of Generalized Harmonic Analysis, we can define a power density function to characterize the whole random process.

Time series analysis for stationary random process has been well defined in both frequency and time domain. Unfortunately, real world data are generally reveal nonstationary properties, the traditional approach to deal with non-stationary processes has been to transform the non-stationary processes into stationary processes with some mathematical operations. Examples of the operations are differencing and detrending. But there will be serious problems if inappropriate method is used to eliminate trend, related literature can be found in (Enders, 1995).

2.1.2 Limitation of Spectral Analysis

Conventional spectral analysis has a limitation is that it is not suitable to analyze the abrupt changes, discontinuities, or other local irregularities in time series, strictly in other words, Fourier analysis, as an effective tool in spectral analysis, cannot be well applied in non stationary time series analysis. Stationarity property of a time series requires the relationships between consecutive observations in the series hold constant cross time. In time domain, this means there is constant autocorrelation functions cross time. In frequency domain, this means there is a series of certain frequency components, a particular frequency exists all the time in original time series, and the sum of them represents the original signal. Fourier transformation requires the smoothness of the time series. Although Fourier transformation does not require analyzed time series strictly periodical, it requires that the analyzed time series is integralable, so it can be used to study fairly smooth non-stationary time series. Fourier transformation itself is a smooth transformation, if the time series is not strictly stationary, in analysis process the transformation will force local irregular components into different frequency, which means that the energy of abrupt changes will be redistributed into different frequencies. In synthesis process the recomposed time series will smooth out the local volatility in original signal. This is particularly the case in financial time series because of their nonstationarity property. If the analyzed time series is highly non-stationary, synthesized approximation will reveal different patterns compare to true Data Generation Process (DGP) and have high Mean Square Error (MSE).

2.1.3 Wavelet properties

Wavelet transformation, basically, has seven properties. First is that wavelets can perfectly reconstruct the original signals. Second is the locality that localizes in both scale and time, which distinguish wavelet bases from classical orthonormal bases (e.g. Fourier). Also, by trading off the resolutions between scale and time, wavelets have the ability to cope partially with the limitations of Heisenberg's principle. The third is the decorrelation, or so called whitening, which means correlated signals in time domain become almost uncorrelated coefficients in the time-scale, domain. The fourth one is the property of disbalance energy, which means the original signal can be well described (approximately) by only a few energetic components (Lorentz curve in economics), which are measured by the large coefficients in wavelet domain. This property facilitates the thresholding technique. Another property is that wavelets filter data, similar to the third property. Depending on the support width of the wavelet functions, the wavelet transformation filters out different components of the signals. Another property is that wavelet can detect self-similar phenomena in different scales (frequencies) of the signals. The final one is that wavelets use an efficient pyramid algorithm, which enables the fast computation of the transformation.

2.1.4 Wavelet Analysis

Wavelet transformation provides an effective way to deal with some non-stationary properties of stochastic processes. The basic idea of wavelets method is to decompose a time series into a series of coefficients through a group of wave-like functions. The group of wave-like functions is dyadic translated from a single mother wavelet, which has certain required mathematic properties, e.g. usually they are orthogonal functions with unity energy. Each set of wavelet coefficients represent a piece of the time series in different time-scale and frequency-scale adaptively and the series of wavelet coefficients preserves the energy of original time series.

There are several advantages of wavelet methods in financial analysis. We can roughly summarize them into three main categories. First, we can directly study of certain nonstationary financial time series. Second, ready to study the local short term properties of financial behavior. Third, intuitively study and compare financial behavior in different timescale, e.g. intraday, daily, short term investing, and long term investment behavior. We will put detailed analysis into corresponding sections.

In general speaking, wavelet analysis is a refinement of Fourier analysis. Before the emerging of Fourier analysis, scientists study signal (time series) by looking at the signal components in time domain only. Although this way enables us to study the major repetitive components of a signal, complex signals with a multitude of components could not be accurately assessed. Application of the auto-correlation function on the time series allows us to indirectly obtain information about the frequencies present in the signal. However these techniques only provide a limited amount of additional information. The need to distinguish between components of a similar nature or hidden within a complex vibration signal led to the mathematical representation of these signals in terms of their orthogonal basis functions, a field of mathematics whose origins date back to Joseph Fourier's investigations into the properties of heat transfer. Difference between usual Fourier (sine wave) and a wavelet can be described by the localization property. Sine wave is localized in frequency domain only while wavelet is localized both in frequency and time domain.

Fourier Transform (FT) processes the raw signal (time series) by using a mathematic transformation, which transforms the signal from time domain into frequency spectrum. Processed signal tells us how many frequencies and how much (energy) of each frequency exists in the raw signal but it does not give us the time information (when a particular frequency appears across the time). If the signal is stationary, we don't need that information but unfortunately in real world most of our data sets, especially in financial and economic data, are non-stationary. One way to treat non stationary time series before wavelets is Windowed Fourier Transform (WFT). WFT can locate the time, which only transform part of a signal and that segment of signal is small enough that we can assume that portion of signal is stationary. By using a particular window function and shifting the window along the time dimension of the signal we can localize the frequency in the signal, and then we obtained a time-frequency representation of the signal. The transformation coefficients are the amplitudes of different frequencies at different time.
However, WFT has a resolution problem. The length of window is used for both high and low frequency components. This lack of adaptivity is not suitable for nonstationary signals, especially transient nonstationary signals. As we know, Heisenberg Uncertainty Principle states that we cannot know the exact time- frequency representation of a signal, what we can know are the bands of frequencies associated with the time intervals in the signal, and the resolutions in two domains are inversely related. Here we must impose a requirement on the width of the window function. Wavelet transformation is a solution to the problem. The

approach in WFT is to initially cut the signal into segments in time, and then to transform them into frequency representations and analyze them separately. The other approach in time- frequency analysis is first to decompose (filter) the signal into different frequency bands, and then cut these bands into segments in time and then analyze their energy distribution. The combination for the two approaches leads to the wavelet transforms.

Continuous Wavelet Transform (CWT) uses a particular wavelet, which has some required properties, as the window function (same logic in WFT). There are two main different between CWT and WFT. First one is that in CWT we use a wavelet to replace the cosine in WFT, which will give us much spikes information in the decomposed signal. Second one, the most significant characteristic, is the width of the window is changed for different frequencies. As in WFT, the windowed signal multiplied with window function then integral across time continuously. Continuous transformation contains highly redundant information because we decompose the time series from time domain into frequency and time domain adding one dimension. Discrete Wavelet Transform (DWT) discards the signal sample by two each time according to Nyquist's rule, which tells us that at lower frequencies (scale) the sampling rate can be decreased, and then orthogonal DWT decomposes the signal at different frequency bands (scale) with different resolution (shifting) and in each frequency band the existing frequency localized in a particular time according to that band's resolution. The orthogonal wavelet series approximation to a continuous time signal $f(t)$ is given by:

$$\begin{aligned} f(t) &= \sum_k S_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\psi_{J,k}(t) + \sum_k d_{J-1,k}\psi_{J-1,k}(t) + \dots \\ &+ \sum_k d_{1,k}\psi_{1,k}(t) \end{aligned}$$

J is the number of multi-resolution components, levels (scales), and k (shift) ranges from 1 to the number of coefficients in the specified component. $\phi_{j,k}$ and

$\psi_{j,k}$ are the approximating orthogonal wavelet functions given by

$$\phi_{j,k}(t) = 2^{-j/2}\phi\Big(\frac{t-2^j k}{2^j}\Big) \tag{2.1}$$

$$\psi_{j,k}(t) = 2^{-j/2}\psi\Big(\frac{t-2^j k}{2^j}\Big) \tag{2.2}$$

Wavelet coefficients are given by

$$S_{J,k} \approx \int \phi_{J,k} f(t)dt \tag{2.3}$$

$$d_{j,k} \approx \int \psi_{j,k}(t) f(t)dt \tag{2.4}$$

Given these coefficients, the wavelet series approximation of the original signal $f(t)$ is given by the sum of the smooth signal $S_{J,k}$, and the detail signals $D_{J,k}, D_{J-1,k}, \ldots, D_{1,k}$:

$$f(t) = S_{J,k} + D_{J,k} + D_{J-1,k} + \ldots + D_{1,k}$$

where

$$S_{J,k} = \sum_k s_{J,k}\phi_{J,k}(t),\ D_{J,k} = \sum_k d_{J,k}\psi_{J,k}(t), \ldots,\ \text{and } D_{1,k} = \sum_k d_{1,k}\psi_{J,k}(t) \tag{2.5}$$

DWT uses a fast pyramid algorithm to calculates the coefficients of the wavelet series approximation for discrete signal $f_1, \ldots, f_n$ of finite extent. The DWT maps the vector $\mathbf{f} = (f_1, \ldots, f_n)'$ to a vector of n wavelet coefficients $\mathbf{w} = (w_1, \ldots, w_n)'$. The vector $\mathbf{w}$ contains the coefficients $s_{J,k}, d_{J,k}, \ldots, d_{1,k},\ j = 1, 2, \ldots, J$ of the wavelet series approximation.

2.2 Wavelet Transforms

The wavelet analysis of a time series can be defined in terms of an orthonormal transform. Orthonormal transforms can be used to re-express a time series in a standardized form for further analysis. The original series and the transform are two representations of the same mathematical entity, and they contain the same information. However, the transform can be further manipulated, to reveal patterns of interest, and is in some ways more efficient representation of the original series. Fourier series is one example of an orthonormal transform and wavelet analysis is an extension of the discrete Fourier transform.

2.2.1 Multi-Resolution Analysis

In statistics, we are often faced with discretely-sampled signals and therefore we need to be able to perform wavelet decomposition of vectors, rather than continuous functions as above. The *Multi-Resolution Analysis* framework, first introduced by (Mallat, 1989) is commonly used to define discrete wavelet filters. The starting point is a scaling function ϕ and a Multi-Resolution Analysis of $\mathbb{L}^2(\mathbb{R})$, i.e. a sequence $\{V_j\}_{j\in\mathbb{Z}}$ of closed subspaces of $\mathbb{L}^2(\mathbb{R})$ such that

- $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis for V_0
- $\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset \ldots \subset \mathbb{L}^2(\mathbb{R})$
- $f(x) \in V_j \Longleftrightarrow f(2x) \in V_{j+1} \quad \forall\, j \in \mathbb{Z}$
- $\cap_j V_j = \{0\}$, $\overline{\cup_j V_j} = \mathbb{L}^2(\mathbb{R})$

The set $\{\sqrt{2}\phi(2t-k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis for V_1 since the map :

$$\begin{array}{rcl} V_0 & \longmapsto & V_1 \\ f(\cdot) & \longmapsto & \sqrt{2}f(2\cdot) \end{array}$$

is an isometry. The function ϕ is in V_1 so it must have an expansion

$$\phi(t) = \sqrt{2}\sum_k h_k\phi(2t-k) \quad ; \quad \{h_k\}_k \in l_2 \quad ; t \in \mathbb{R} \tag{2.6}$$

Once we have the scaling function ϕ, we use it to define the wavelet function (also called the *mother wavelet*) ψ. We define the latter in such a way that $\{\psi(x-k)\}_k$

is an orthonormal basis for the space W_0, being the orthogonal complement of V_0 in V_1

$$V_1 = V_0 \oplus W_0 \tag{2.7}$$

Defining $W_j = \text{span}\{\psi_{j,k} \ : \ k \in \mathbb{Z}\}$, we obtain that W_j is the orthogonal complement of V_j in V_{j+1}. We can write

$$V_{j+1} = V_J \oplus W_j = V_{j-1} \oplus W_{j-1} \oplus W_j = \ldots = V_0 \oplus \Big(\bigoplus_{i=0}^{j} W_i \Big) \tag{2.8}$$

There are precise procedures for finding ψ one ϕ is known (Daubechies, 1992). As $\{\sqrt{2}\phi(2\cdot -k)\}_k$ is an orthonormal basis for V_{-1}, its follows that both ϕ and ψ can be expressed as a linear combination of $\{\sqrt{2}\phi(2\cdot -k)\}_k$. In other words:

$$\phi(t) = \sqrt{2}\sum_k h_k \phi(2t-k) \tag{2.9}$$

$$\psi(t) = \sqrt{2}\sum_k g_k \psi(2t-k) \tag{2.10}$$

where

$$g_k = \sqrt{2}\int_{-\infty}^{+\infty} \phi(t)\phi(2t-k)dt \tag{2.11}$$

and

$$h_k = \sqrt{2}\int_{-\infty}^{+\infty} \psi(t)\phi(2t-k)dt \tag{2.12}$$

The two equations (2.9 and 2.10) are known as *dilation* and wavelet equation respectively, and they are used to move from the continuous wavelet transform (CWT) to the discrete wavelet transform (DWT).

In the two equations above, $\{h_k\}_k$ and $\{g_k\}_k$ behave as low and high pass filters respectively and they are such that:

$$g_k = (-1)^{k+1} h_{L-k-1} \ , \ k = 0, 1, \ldots, L-1$$

where L is the length of the filter.

A set of conditions must be satisfied by the coefficients before they can represent an orthonormal wavelet. A wavelet filter must sum to zero, have unit energy and be orthogonal to its even shift; in other words:

$$\sum_{k=0}^{L-1} h_k = 0 \tag{2.13}$$

$$\sum_{k=0}^{L-1} h_k^2 = 1 \tag{2.14}$$

$$\sum_{k=0}^{L-1} h_k g_{k+2l} = \sum_{-\infty}^{+\infty} h_k g_{k+2l} = 0 \tag{2.15}$$

for all $l \neq 0$, where the integer $2l$ is the index of the even displacement of the wavelet.

For the Haar wavelet,

$$\phi^{\text{Haar}} = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise.} \end{cases}$$

hence

$$g_k = \sqrt{2} \int_{-\infty}^{+\infty} \phi(t)\phi(2t-k)dt = \begin{cases} 1/\sqrt{2}, & k = 0, 1 \\ 0, & \text{otherwise.} \end{cases}$$

and $h_0 = g_1 = 1/\sqrt{2}$, $h_1 = -g_0 = -1/\sqrt{2}$.

The wavelet filter and the scaling filter can be described in terms of operators. Let $\mathbf{H}(.)$ and $\mathbf{G}(.)$ be the transfer function for $\{h_k\}_k$ and $\{g_k\}_k$ respectively. i.e.,

$$\mathbf{H}(f) = \sum_{-\infty}^{+\infty} h_k e^{-2i\pi fk} = \sum_{0}^{L-1} h_k e^{-2i\pi fk}$$

$$\mathbf{G}(f) = \sum_{-\infty}^{+\infty} g_k e^{-2i\pi fk} = \sum_{0}^{L-1} g_k e^{-2i\pi fk} = e^{-2i\pi f(L-1)} \mathbf{H}(\frac{1}{2} - f)$$

and let $\mathcal{H}(.)$, $\mathcal{G}(.)$ denote the squared gain functions:

$$\mathcal{H}(f) \equiv |\mathbf{H}(f)|^2 \ , \ \mathcal{G}(f) \equiv |\mathbf{G}(f)|^2$$

It is helpful to derive a rather simple condition that is equivalent to unit energy and orthogonality conditions and expressed in terms of the squared gain function:

$$\mathcal{H}(f) + \mathcal{H}(f + \frac{1}{2}) = 2 \quad \text{for all } f. \tag{2.16}$$

and

$$\mathcal{G}(f) + \mathcal{G}(f + \frac{1}{2}) = 2 \quad \text{or } \mathcal{H}(f) + \mathcal{G}(f) = 2 \quad \text{for all } f. \tag{2.17}$$

These relationship (2.16 and 2.17) are evident for the Haar and D(4) squared gain functions (Percival and Walden, 2000).

2.2.2 Discrete Wavelet Transform (DWT)

The Discrete Wavelet Transform (DWT) of a time series $\{X_t; t = 0, \ldots, N-1\}$ is an orthonormal transform. Let $\{W_n; n = 0, ..., N-1\}$ represent the DWT coefficients. Then we can write $\mathbf{W} = \mathcal{W}\mathbf{X}$ where $\mathbf{W}$ is a column vector of length $N = 2^J$ whose nth element is the nth DWT coefficient W_n and $\mathcal{W}$ is an $N \times N$ real-valued matrix defining the DWT and satisfying $\mathcal{W}^T\mathcal{W} = I_N$.

Orthonormality implies that $\mathbf{X} = \mathcal{W}^T\mathbf{W}$ and $\|\mathbf{W}^2\| = \|\mathbf{X}^2\|$. Hence DWT transform coefficients preserve the energy of original time series and the squared nth coefficient W_n^2 represents the contribution to the total energy.

The nth coefficient W_n is associated with a particular time scale or frequency band, and is located at a particular set of times. We can further group the $N \times N$

orthonormal DWT matrix $\mathcal{W}$ into $J+1$ sub matrices:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W_1} \\ \mathbf{W_2} \\ \vdots \\ \mathbf{W_J} \\ \mathbf{V_J} \end{bmatrix} = \mathcal{W}\mathbf{X} = \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{W}_2 \\ \vdots \\ \mathcal{W}_J \\ \mathcal{V}_J \end{bmatrix} \mathbf{X} = \begin{bmatrix} \mathbf{W}_1\mathbf{X} \\ \mathbf{W}_2\mathbf{X} \\ \vdots \\ \mathbf{W}_J\mathbf{X} \\ \mathbf{V}_J\mathbf{X} \end{bmatrix} \tag{2.18}$$

where $\mathcal{W}_j$ has dimension $N/2^j$; $\mathcal{V}_J$ is $1\times N$; $\mathbf{W}_J$ is a column vector of length $N/2^j$; and $\mathbf{V}_J$ contains the last element of $\mathbf{W}$.

The wavelet coefficients in the vector $\mathbf{W}_j$ are associated with differences of various orders in adjacent weighted averages over a scale of $\tau_j = 2^{j-1}$, while the scaling coefficient in $\mathbf{V}_J$ is equal to $\sqrt{N}$ times the sample mean $\overline{X}$ of X. As a special case, in the Haar DWT, each wavelet coefficient is the difference of adjacent averages over a scale τ_j.

The DWT energy preserving condition of orthonormality can be represented as

$$\|\mathbf{X}\|^2 = \|\mathbf{W}\|^2 = \sum_{j=1}^{J} \|\mathbf{W}_j\|^2 + \|\mathbf{V}_J\|^2 \tag{2.19}$$

$\|\mathbf{W}\|^2$ represents the contribution to the energy of $\{X_t\}$ due to the changes at scale τ_j.

Since $\|\mathbf{V}_j\|^2 = N\overline{X}^2$, we can decompose the sample variance as

$$\widehat{\sigma}^2_X = \frac{1}{N}\|X\|^2 - \overline{X}^2 = \frac{1}{N}\|W\|^2 - \overline{X}^2 = \frac{1}{N}\sum_{j=1}^{J}\|W_j\|^2 \tag{2.20}$$

$\|W_j\|^2/N$ represents the contribution to the sample variance of $\{X_t\}$ due to the changes at scale τ_j. If we define a DWT empirical power spectrum $\Big\{P_{\mathcal{W}}(\tau_j) :\ \tau_j = 1, 2, \ldots, N/2\Big\}$ for $\Big\{X_t\Big\}$ as

$$P_{\mathcal{W}}(\tau_j) = \frac{1}{N}\|W_j\|^2 \tag{2.21}$$

then we have

$$\sum_{j=1}^{J} P_{\mathcal{W}}(\tau_j) = \widehat{\sigma}_x^2$$

Now the wavelet synthesis of $\mathbf{X}$ is given by

$$\mathbf{X} = \mathcal{W}^T\mathbf{W} = \sum_{n=0}^{N-1} \mathcal{W}_n \mathbf{W}_n = \sum_{j=1}^{J} \mathcal{W}^T\mathbf{W} + \mathcal{V}^T\mathbf{V}$$

$\mathcal{W}_j$ and $\mathcal{V}_J$ are matrices by partitioning the rows of $\mathcal{W}$ commensurate with the partitioning of $\mathbf{W}$ into $\mathbf{W}_1, \ldots, \mathbf{W}_J$ and $\mathbf{V}_J$.

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

If we define $\mathcal{D}_j = \mathcal{W}_j^T \mathbf{W}_j$ for $j = 1, \ldots, J, \mathcal{D}_j$ is an N dimensional column vector whose elements are associated with changes in $\mathbf{X}$ at scale τ_j. Let $\mathcal{S}_j = \mathcal{V}_j^T \mathbf{V}_j$, which has all of its elements equal to the sample mean $\mathbf{X}$. We can now write

$$\mathbf{X} = \sum_{j=1}^{J} \mathcal{D}_j + \mathcal{S}_J \tag{2.22}$$

This defines a *Multi-Resolution Analysis* (MRA) of $\mathbf{X}$. The time series $\mathbf{X}$ is expressed as the sum of a vector of constants $\mathcal{S}_J$ and J other vectors $\mathcal{D}_j$, each of which contains a time series related to variation in $\mathbf{X}$ at a certain scale τ_j. We refer to $\mathcal{D}_j$ as the jth level wavelet details.

Let define

$$\mathbf{X} = \sum_{k=j+1}^{J} \mathcal{D}_k + \mathcal{S}_J \tag{2.23}$$

we refer $\mathcal{S}_J$ as the Jth level wavelet smooth for $\mathbf{X}$, since it is the difference between $\mathbf{X}$ and all the details from level 1 to level J

$$\mathbf{X} - \mathcal{S}_J = \sum_{k=1}^{J} \mathcal{D}_k \tag{2.24}$$

Because the orthonormality of $\mathcal{W}$,

$$\mathcal{D}_j^T \mathcal{D}_k = \mathbf{W}_j^T \mathcal{W}_j \mathcal{W}_k^T \mathbf{W}_k = \begin{cases} \mathbf{W}_j^T \mathbf{W}_j, & k = j \\ 0, & \text{otherwise.} \end{cases} \tag{2.25}$$

so the decomposition of the variance of $\mathbf{X}$ can be reexpressed on a scale by scale basis as

$$\widehat{\sigma}_X^2 = \frac{1}{N} \sum_{j=1}^{J} \|\mathcal{D}_j\|^2 \tag{2.26}$$

Above is a brief derivation of DWT decomposition (analyse) of the sample variance, and synthesis of X using matrix notation. In practice, the DWT is implemented through an algorithm known as pyramid algorithm using linear filtering operations. The pyramid algorithm is computational efficient. It can be conducted through several stages depending on what information we interest at. The details of the algorithm can be found in (Percival and Walden, 2000). The central ideas are that we use a basic wavelet filter (mother wavelet) and a scaling filter (father wavelet) extracting out the details (changes of moving averages) and smooth (averages) in the first level through mathematic convolution and down-sampling operations. The first level smooth is further processed by simultaneous stretches and shifts of the wavelet filter and scaling filter on a dyadic basis to extract changes and averages at the second level. The wavelet filter and scaling filter satisfy the orthonormality property, so that the whole process is an orthonormal transform. The algorithm preserves the energy of original time series X and thus its variance. The decomposition is an analysis of variance and the synthesis is the additional of all wavelet details and the wavelet smooth. There are several practical issues about using wavelets with finite time series like selection of wavelet form, boundary handling, and sample observations. We will discuss the corresponding details in the following sections.

Example: Oil Price Return

In this example, we apply a partial DWT ($J_p = 4$) to the daily Oil stock prices from December 17, 2003, to November 08, 2007.

A return series was computed via the first difference of the log-transformed prices that is, $r_t = \ln(p_{t+1}) - \ln(p_t)$. The returns series is plotted in the upper row of Figure 2.1.

There is an obvious increase in variance in the returns toward the latter third of the series. The length of the returns series is $N = 1024$, which is divisible by $2^4 = 16$, and therefore we may perform an order $J_p = 4$ partial DWT on it. The wavelet coefficient vectors $\mathbf{W}_1, ..., \mathbf{W}_4$ and scaling coefficient vector V_4 using the Haar wavelet are shown in the lower part of Figure 2.1. The first scale of wavelet coefficients $\mathbf{W}_1$ are filtering out the high-frequency fluctuations by essentially looking at adjacent differences in the data. There is a large group of rapidly fluctuating returns between observations 250 and 300. A small increase in the magnitude $\mathbf{W}_2$ is also observed between observations 250 and 300, but smaller than the unit scale coefficients. This vector of wavelet coefficients is associated with changes of scale). Since the Oil price return series exhibit low-frequency oscillations, the higher scale (low-frequency) vectors of wavelet coefficients $\mathbf{W}_3$ and $\mathbf{W}_4$ indicate large variations from zero. The same is true for the scaling coefficients, which are associated with averages of scale $2\tau_4$ or greater.

The same decomposition was performed using the LA(8) wavelet filter and provided in Figure 2.1. The interpretation for each of the vectors of coefficients is the same as in the case of the Haar wavelet filter. The wavelet coefficients will be different given the length of the filter is now eight versus two, and should isolate features in a specific frequency interval better since the LA(8) is a better

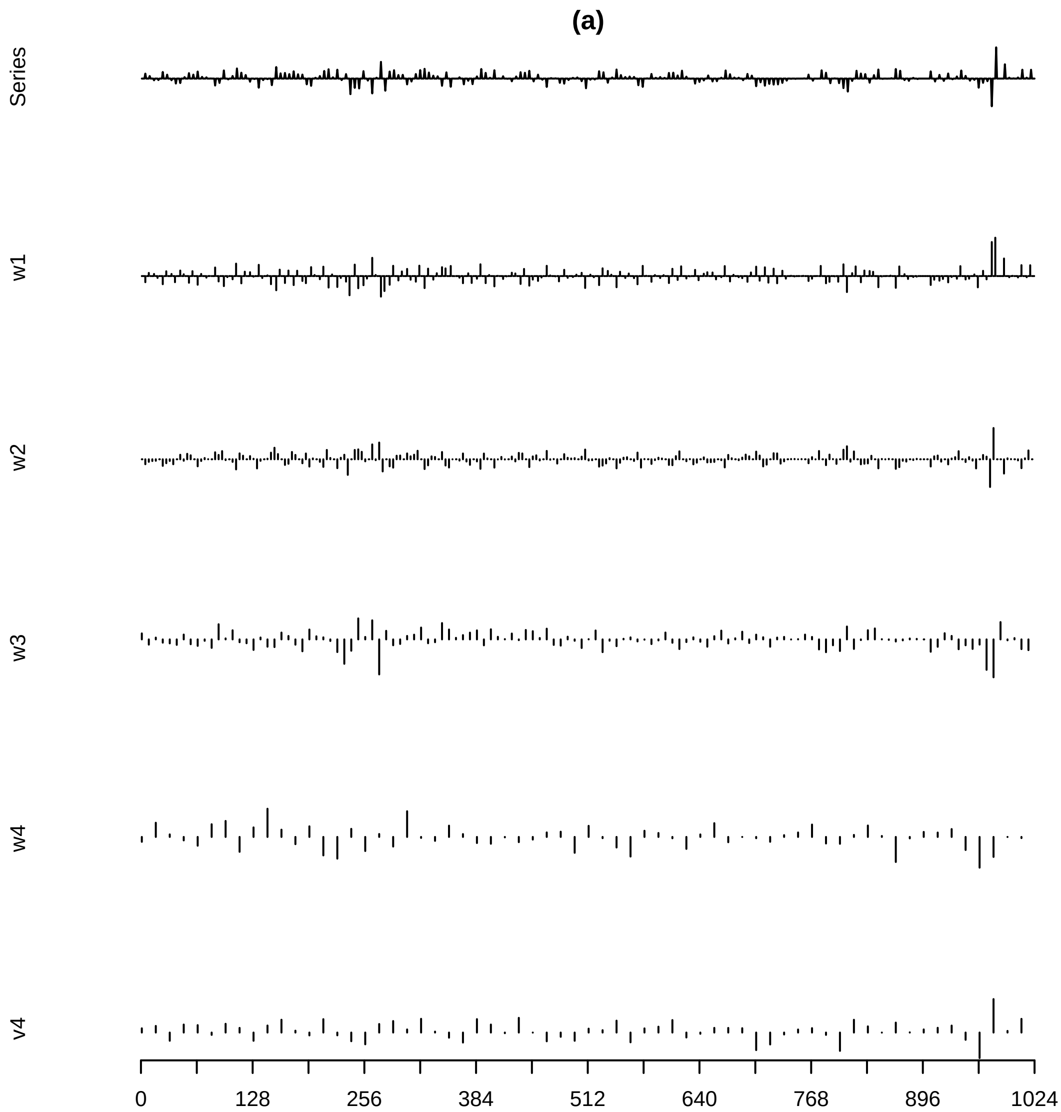

Figure 2.1: Wavelet decomposition of oil price return (Haar). The upper row is the original time series.

approximation to an ideal band-pass filter over the Haar wavelet. Note, the LA(8) wavelet coefficient vectors have been circularly shifted in order to better align features in the wavelet coefficients with the original time series. To be precise, the first vector of coefficients has been shifted to the left by two time units $\mathbf{W}_1^{-2}$ due to the specific phase properties of the LA(8) wavelet filter.

Let us now look at the volatility $v_t = |r_t|$ of Oil price series through a DWT Multi-Resolution Analysis. The volatility series is provided in the top row of

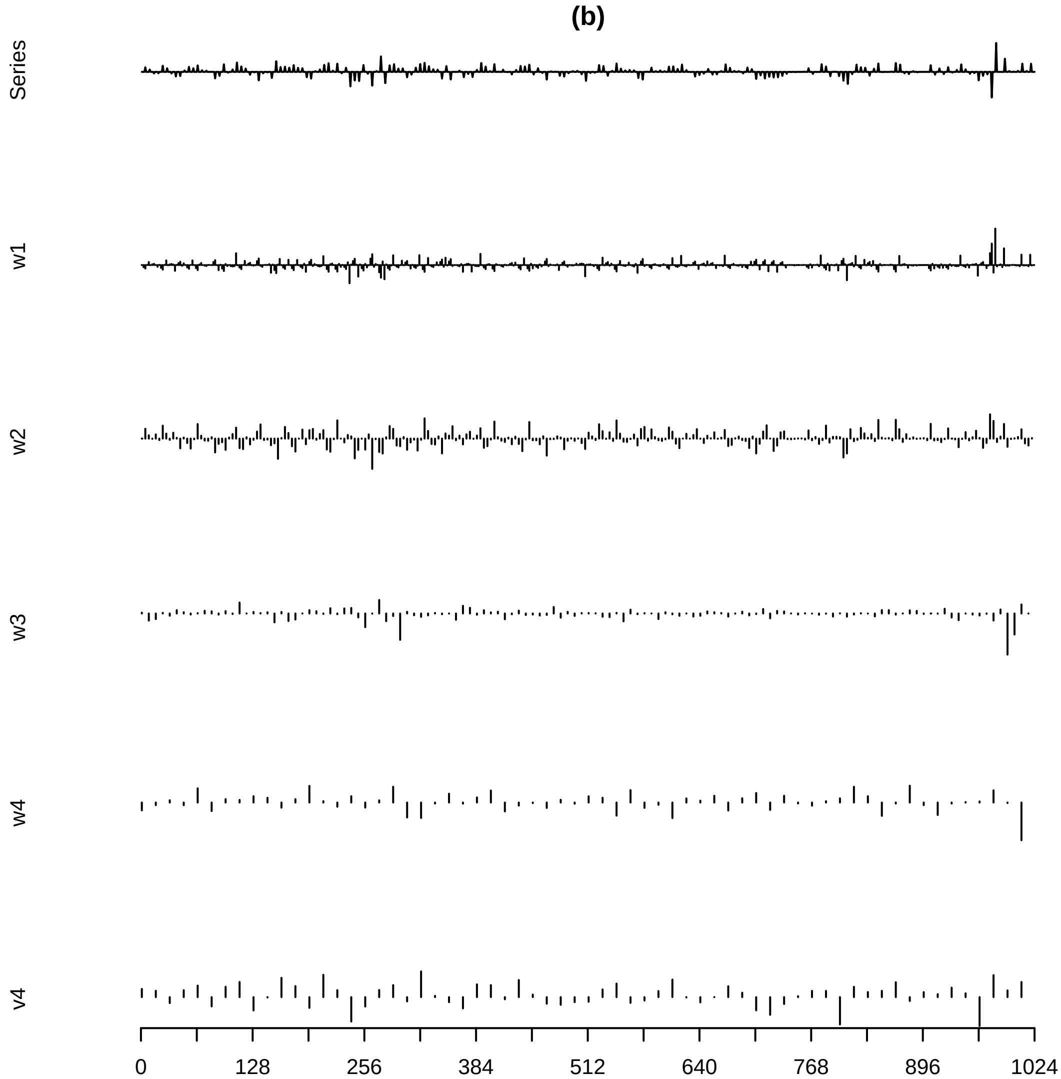

Figure 2.2: Wavelet decomposition of oil price return (LA(8)). The upper row is the original time series.

Figure 2.3 and exhibits the increase in volatility we observed in the return series. A DWT Multi-Resolution Analysis (MRA) using the Haar wavelet filter is given below the volatility series in Figure 2.3.

The five rows below the data display the first four wavelet details and wavelet smooth that form an additive decomposition. That is, at each time point t adding up the wavelet detail coefficients $d_{j,t}$, $j = 1, ..., 4$, and the wavelet smooth $s4, t$ will produce the coefficient v_t from the volatility series. All the information contained

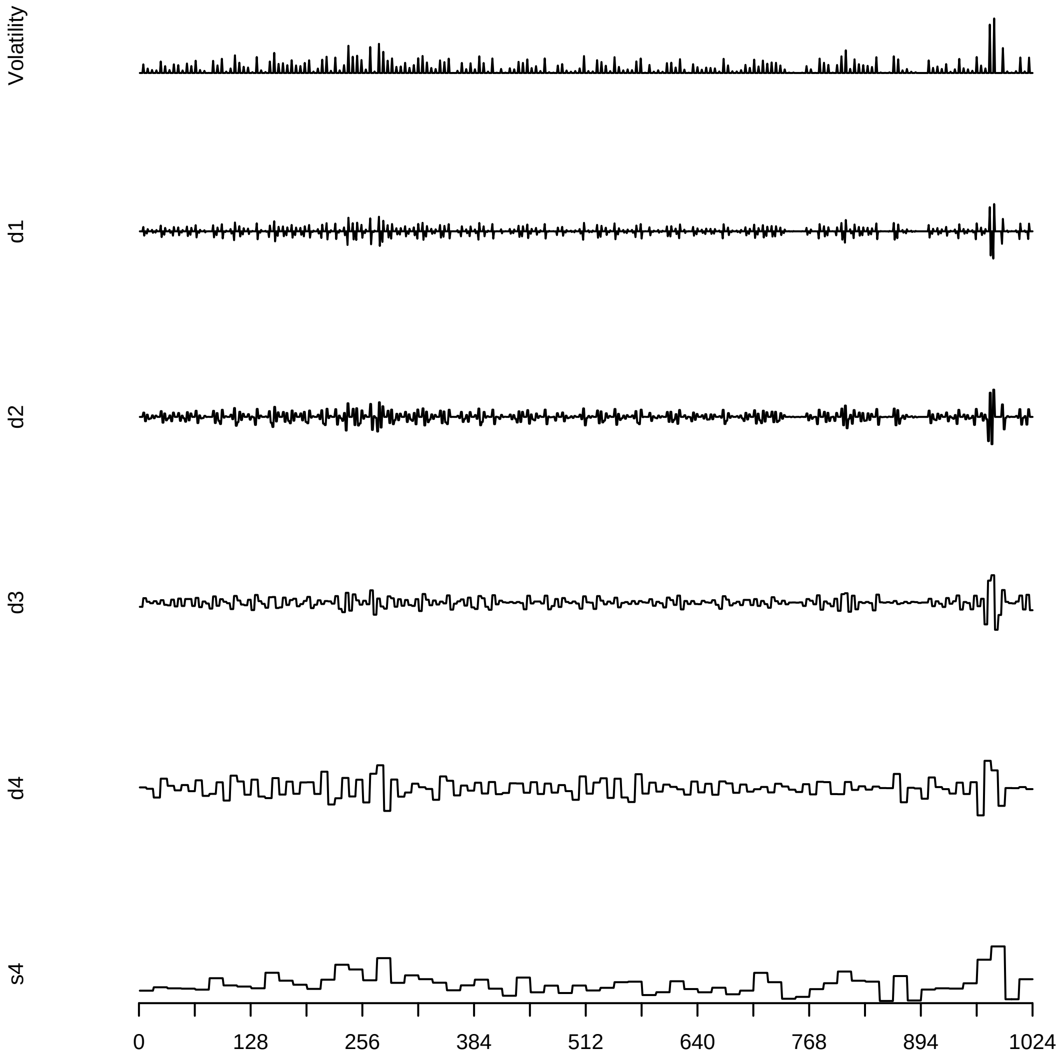

Figure 2.3: DWT Multi-Resolution Analysis of oil price volatility (Haar). The volatility series is plotted in the top row.

in the volatility series is perfectly captured in the MRA. No anomalies have been introduced by this procedure.

Figure 2.4 shows a similar MRA of the IBM volatility series, but instead of the Haar wavelet we use the LA(8) wavelet filter.

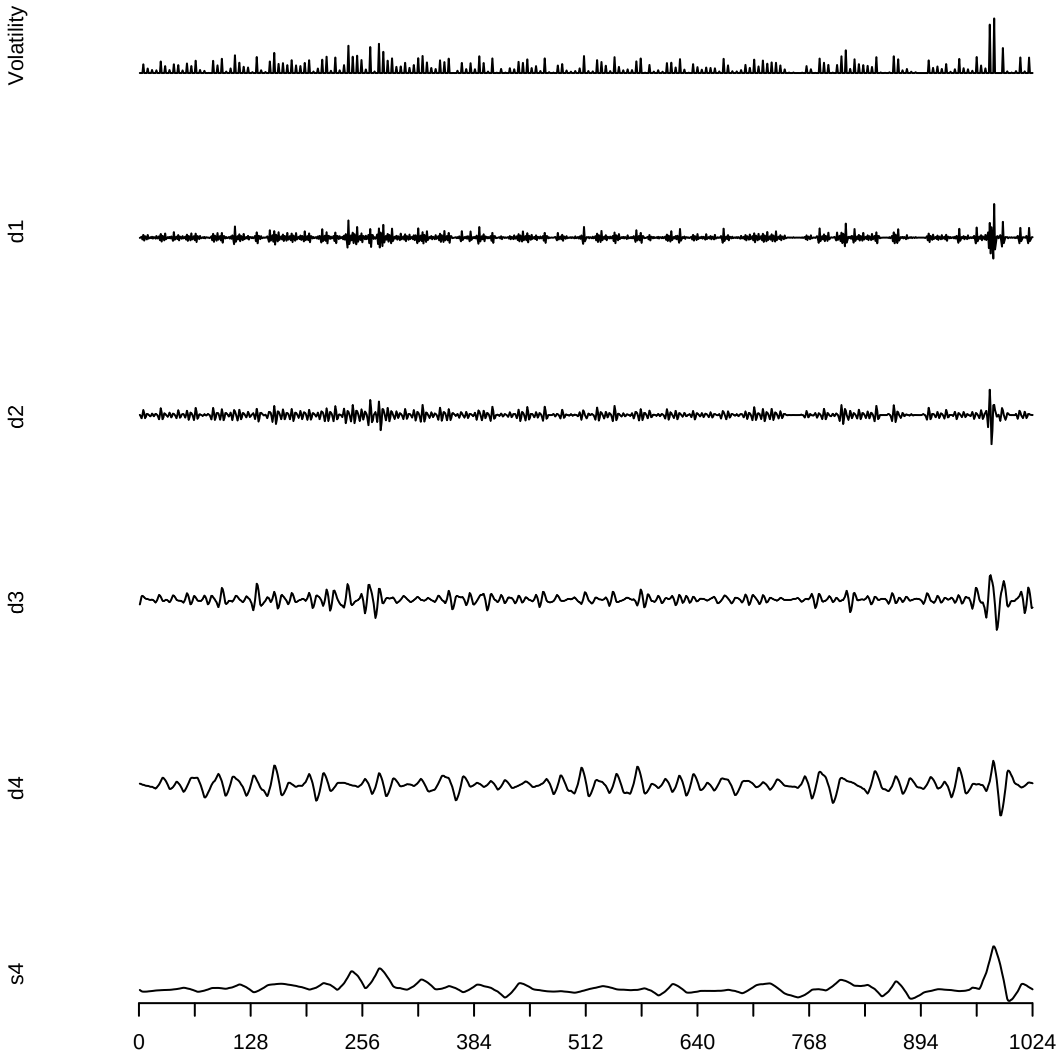

Figure 2.4: DWT Multi-Resolution Analysis of oil price volatility (LA(8))). The volatility series is plotted in the top row.

2.2.3 Maximum Overlap Discrete Wavelet Transform (MODWT)

However the orthonormal discrete wavelet transform (DWT), even if widely applied to time series analysis in many disciplines, has two main drawbacks:

- The dyadic length requirement (i.e. a sample size divisible by 2^J)[1],
- and the fact that the wavelet and scaling coefficients are not shift invariant due

[1] This condition is not strictly required if a partial DWT is performed.

to their sensitivity to circular shifts because of the decimation operation.

An alternative to DWT is represented by a non-orthogonal variant of DWT: the maximal overlap DWT (MODWT)[2].

In the orthonormal Discrete Wavelet Transform (DWT) the wavelet coefficients are related to non-overlapping differences of weighted averages from the original observations that are concentrated in space. More information on the variability of the signal could be obtained considering all possible differences at each scale, that is considering overlapping differences, and this is precisely what the maximal overlap algorithm does.[3] Thus, the maximal overlap DWT coefficients may be considered, the result of a simple modification in the pyramid algorithm used in computing DWT coefficients through not down-sampling the output at each scale and inserting zeros between coefficients in the wavelet and scaling filters.[4] In particular, the MODWT wavelet and scaling coefficients $\widetilde{w}_{j,t}$ and $\widetilde{v}_{j,t}$ are given by:

$$\widetilde{w}_{j,t} = \frac{1}{2^{j/2}} \sum_{l=0}^{L-1} \widetilde{h}_{j,l} X_{t-l}$$

$$\widetilde{v}_{j,t} = \frac{1}{2^{j/2}} \sum_{l=0}^{L-1} \widetilde{g}_{j,l} X_{t-l}$$

where the MODWT wavelet and scaling filters $\widetilde{h}_{j,l}$ and $\widetilde{g}_{j,l}$ are obtained by rescaling the DWT filters as follows:

$$\widetilde{h}_{j,l} = \frac{h_{j,l}}{2^{j/2}}$$

[2]The MODWT goes under several names in the wavelet literature, such as the "non-decimated DWT", "stationary DWT" (Nason and Silverman, 1995), "translation-invariant DWT" (Coifman and Donoho, 1995) and "time-invariant DWT".

[3]Indeed, the term maximal overlap refers to the fact that all possible shifted time intervals are computed. As a consequence, the orthogonality of the transform is lost, but the number of wavelet and scaling coefficients at every scale is the same as the number of observations.

[4]The DWT coefficients may be considered as a subset of the MODWT coefficients. Indeed, for a sample size power of two the MODWT may be rescaled and sub-sampled to obtain an orthonormal DWT.

$$\widetilde{g}_{j,l} = \frac{g_{j,l}}{2^{j/2}}$$

The MODWT wavelet coefficients $\widetilde{w}_{j,t}$ are associated with generalized changes of the data on a scale $\tau_j = 2^{j-1}$. With regard to the spectral interpretation of MODWT wavelet coefficients, as the MODWT wavelet filter $h_{j,l}$ at each scale j approximates an ideal high-pass with pass-band $f \in [1/2^{j+1}, 1/2^j[$[5], the τ_j scale wavelet coefficients are associated to periods $[2^j, 2^{j+1}[$.

The maximum overlap discrete wavelet transform (MODWT) is a modified version of DWT. The rational for this modification is because that the DWT is sensitive to the choice of a starting point for a time series. This sensitivity is due to the fact that in each stage of pyramid algorithm, the even shift wavelet and scaling filters and sub-sampling, we lost information on the time location. The MODWT eliminates this down-sampling by applying DWT pyramid algorithm twice picking up discarded outputs from DWT filter, circularly shift filtering, while preserving the ability to carry out the energy decomposition (analysis of variance) and the additive decomposition (multi-resolution analysis). Like DWT, MODWT has the following two basic properties:

The MODWT energy decomposition:

$$\|\mathbf{X}\|^2 = \|\widetilde{\mathbf{W}}\|^2 = \sum_{j=1}^{J} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_J\|^2 \tag{2.27}$$

and the MODWT additive decomposition:

$$\mathbf{X} = \sum_{j=1}^{J} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_J \tag{2.28}$$

Due to the time invariant property of MODWT, many import applications of wavelets are based on MODWT. The important idea of wavelet variance is defined

[5]On the other hand at scale τ_J the scaling filter $g_{J,l}$ approximates an ideal low-pass filter with pass-band $f \in [0, 1/2^{J+1}]$.

on MODWT coefficients.

Exemple: Oil Price series

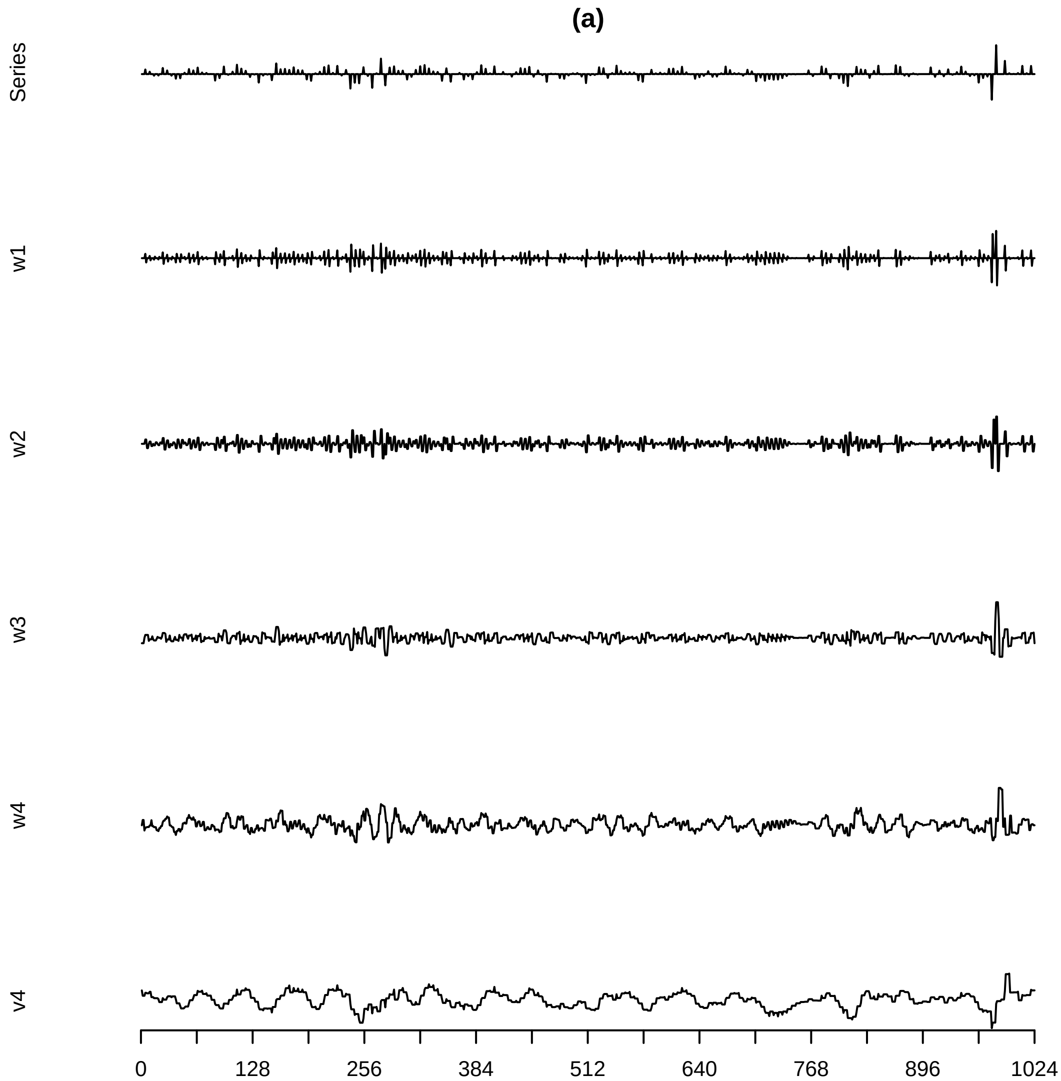

Figure 2.5: MODWT decomposition of oil price return series (Haar). The upper row is the original time series.

To provide an example wavelet analysis using the MODWT, we consider the oil price return series. This series has length $N = 1024$, but with the MODWT we are no longer limited to decomposing a sample size of dyadic length. The only limiting factor is the overall depth of the transform given by $J = log_2(N)$. We perform a J-level decomposition ($J = 4$) applying the maximum overlap discrete wavelet

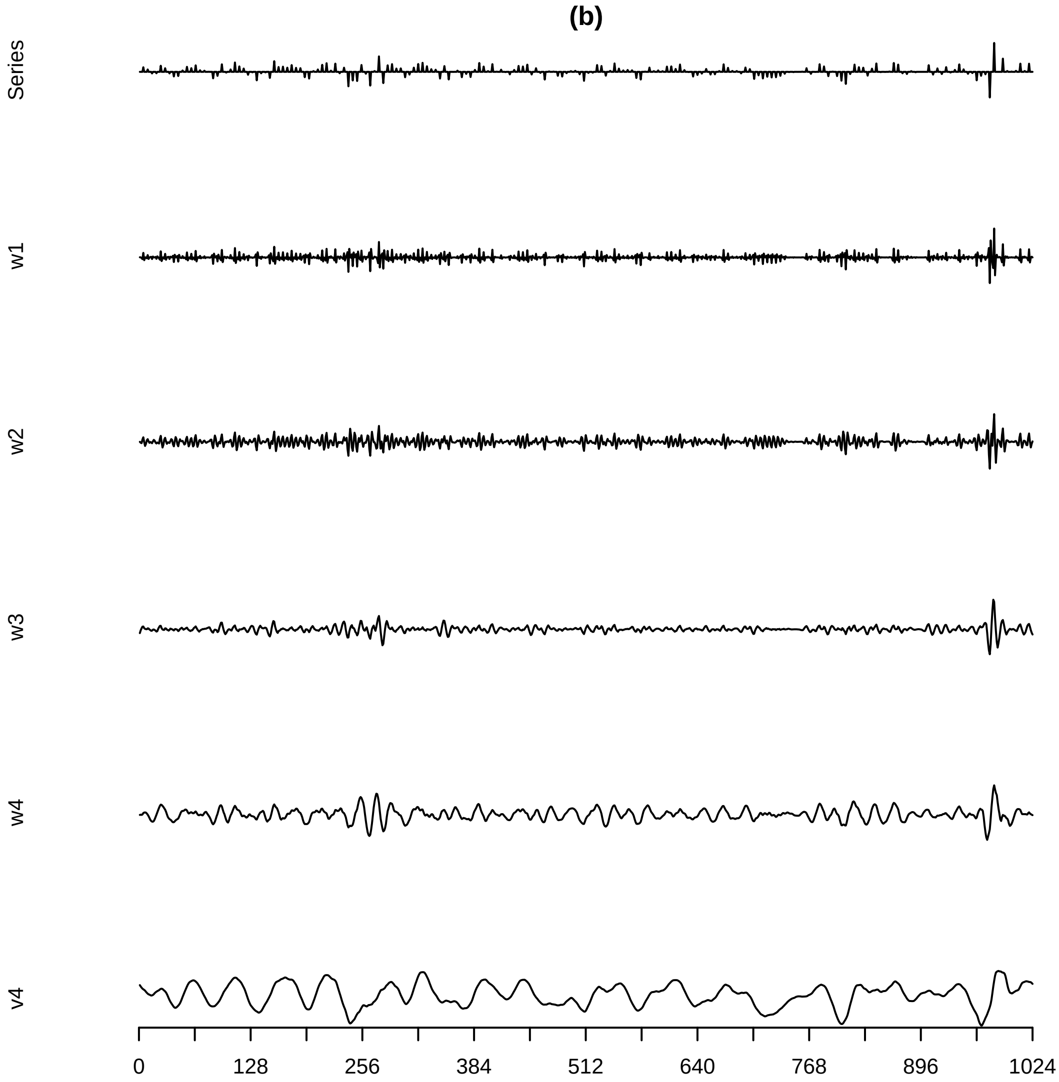

Figure 2.6: MODWT decomposition of oil price return series LA(8). The upper row is the original time series.

transform (MODWT) to aggregate daily oil price return series and using the Haar wavelet filter Figure 2.5 and the Daubechies least asymmetric (LA) wavelet filter of length $L = 8$, denoted as $LA(8)$, which is a fourth order filter based on four non-zero coefficients (Daubechies, 1992) as the order of the filter is equal to the number of vanishing moments (half the length of the filter). The application of the MODWT with a number of scales $J = 4$ produces five wavelet and scaling filter coefficients: v_4, w_4, w_3, w_2, w_1, where each wavelet scale is associated to a particular time period.

As the MODWT wavelet filter belongs to high-pass filter with pass-band given by the frequency interval $[1/2^{j+1}, 1/2^j[$ for scales $1 \leq j \leq J$, inverting the frequency range to produce a period of time we obtain, with monthly data, that wavelet coefficients associated to scale $\tau_j = 2^{j-1}$ are associated to periods $[2^j, 2^{j+1}[$. Thus, scale 1 represents frequencies corresponding to 2-4 day period dynamics, and scales 2, 3, and 4 correspond to 4-8, 8-16, and 16-32 days period dynamics, respectively. The first three time scales represent the short-run dynamics of a signal (corresponding to the very high- and high-frequency components), scale 4 roughly correspond to the standard business cycle time period (Stock and Watson 2004), while the trend is associated to the low-frequency components of a signal corresponding to the long-run elements. Figure 2.6 shows the time series of the oil price return series and the MODWT coefficient sequences $\widetilde{w}_{j,t}$ for levels $j = 1$ to $j = J = 4$ using $LA(8)$ wavelet filter.

2.3 Applications

2.3.1 Wavelet as filter

Many conventional concepts in time series analysis can be related to discrete wavelet transformation. Examples are moving average in time series estimation and prediction, windowed Fourier analysis to localize the structure changes, autocorrelation analysis to study the seasonality in time series, and Kalman filter in time series forecasting. Most of these techniques can be interpreted as a particular filter from signal processing engineering, though they focus their analysis on either in time domain or in frequency domain. A wavelet function is simply another special filter with certain properties.

Filter

A time series is a discrete sample realization of a random process indexed by time, a finite or infinite sequence. A linear filter linearly transforms a time series into another time series with a filter coefficient vector. Mathematically, output of a filter is the convolution of the input time series with filter coefficient vector. Depend on the relationship between present output with past, present, and future values of input and past values of output, types of filter can be classified into different categories. If the present output of a filter is a liner weighted average of past and present values of inputs as well as the past values of output, the filter is an infinite impulse response filter because of its response to an impulse input will be an infinite sequence. A difference equation is an example. If the present output is only related with the values of input, it is called a Finite Impulse Response filter (FIR). Causal filter is that the output of filter is determined by past and present values of the input only. Noncausal filter's output is a liner weighted average of past and present values of the input as well as the future values of the input. Example is a centered simple moving average. Filter operation is a convolution of input with filter coefficients, simply a weighted average, in time domain. Convolution transformation in time domain is correspondence to the filtering in frequency domain. In frequency domain, convolution transformation can be a low pass, high pass, or band pass filter. Convolution in time domain is equivalent to multiplication of the spectrum of the input by the frequency response of the filter in the Fourier domain. In other words convolution operation shapes the spectrum of the input time series according to the frequency response of the filter. The input frequency response of a filter is determined by its gain function and phase function in frequency domain. If the filter's gain function is large at the low frequencies band and small at high frequencies band, we call it a low-pass filter. Inversely, we call it a high-pass filter. There are also band-pass filters which only allow a certain part of frequencies band through the filters. Usually, filtering transformation induces phase shift, there will be a change in phase in output of the filter compare with the input time series.

Wavelet as filter

Wavelet transformation can be seen as a piece by piece information decomposition of a time series. Each piece of decomposed signal is expressed by one wavelet coefficient or one scaling coefficient. The coefficients are obtained through applying the wavelet filter and scaling filter to the original signal. Each wavelet coefficient is the convolution of a wavelet filter with the original signal. A wavelet filter is a special filter with certain properties, basic properties include compact support, unit energy, integration to zero, and orthogonal with its even shifts. The quadrature mirror filter of a wavelet filter is known as scaling filter. In DWT the wavelet filter is stretched and shifted in time domain and shrunk in the frequency domain according 2^j scale. In practice, the DWT transformation is implemented through a pyramid algorithm filtering procedure. Related detail of wavelet theory and transformation algorithm can be found in (Percival and Walden, 2000) and (Genacy et al., 2002). Here we only have a brief introduction on what wavelet decomposition do on the original time series, the implication of the results, and potential application in financial analysis. In time domain, wavelet filter is simply to difference the weighted average of adjacent observations in each half of the support; the result of scaling filter is the weighted average of adjacent observations in the whole support. Haar wavelet filter is simply the first order difference of the signal. Haar scaling filter is the moving average of the signal. In frequency domain, wavelet filter is a high pass filter and scaling filter is a low pass filter. The higher levels of wavelet filters are all band pass filters. Haar wavelet filter is a high pass filter but with poor cutoff frequency because Haar wavelet has only one vanishing moment. Haar wavelet is the only symmetry wavelet and it has good localization property in time domain but poor localization in frequency domain.

2.3.2 Wavelet Methods in Statistics

The properties and especially the fact that each discrete wavelet coefficient depends on just a portion of a time series leads to the possibility of effectively dealing with time series whose statistical characteristics evolve over time (Percival and Walden, 2000). And due to the property that wavelets are very good at compressing a wide range of signals into a small amount of large wavelet coefficients, a very large proportion of the coefficients can be set to zero without loss of important information. This property allows wavelets to deal with heterogeneous and intermittent behaviors (Silverman, 2000). Wavelet based methods have impacts on both theoretical and applied statistics and have been developed as a new descriptive statistical methodology. Statistical multi-scale modeling has become a new area of research. Wavelet methods in statistics are developing in areas such as nonparametric regression, nonparametric density estimation, time series modeling and forecasting. Bayesian approaches emerges with wavelet methods has a promising potentials in statistical modeling, especially for an adaptive processes. Following is a brief review of wavelet methods in statistical applications, which can be further used and developed in financial analysis.

Earliest application of wavelets in statistics is in the field of nonparametric statistical estimation. In the simplest curve estimation context, wavelet-based methods have many advantages, especially for models where the underlying trend is not homogeneous, but may include discontinuities or other local irregularities. The basic idea is to use wavelets to transform the data, shrink the transform coefficients and then invert the modified coefficients back to form an estimate of the original series. The use of linear wavelet methods in statistics can be found in (Antoniadis, 1991) and (Walter, 1992) for density estimation and (Antoniadis et al., 1999) for nonparametric regression. Non-linear nonparametric regression and smoothing are pioneered by (Donoho, 1992, 1993), (Donoho and Johnstone, 1994) and (Donoho and Johnstone, 1995) based on the wavelet shrinkage principle. Shrinkage estimation was founded by (Stein, 1995) when he found that the sample mean is not an admissible estimator of the multivariate normal mean. After then a variety of shrinkage rules were proposed.

Shrinkage rules have been used in earlier signal processing techniques such as Fourier series analysis. The problem is that the shrinking of coefficients affects the whole reconstructed signals. Because of the smoothing property of wavelets, shrinking rules can be applied locally on the coefficients in wavelet transformations. Applying shrinking technique in wavelet domain was developed by (Mallat, 1989) and (Donoho and Johnstone, 1995). The simplest wavelet non-linear shrinkage technique is thresholding (Donoho, 1995). Statistical optimality of wavelet shrinkage was explored, and Bayesian approaches are incorporated in the shrinkage thresholding rules selections. Some recent work can be found in (Donoho and Johnstone, 1998), (Donoho and Johnstone, 1999), and (Donoho and Yu, 1998).

Density estimation has a long history and wavelet -based density estimation is among the earliest contributions of wavelets in statistics. (Doukhan, 1988) and (Doukhan, 1990) first introduced linear wavelet density estimators and explored their mean-square errors. (Donoho and Johnstone, 1995) and (Gao, 1993) explored the non- linear density estimations using wavelets.

Wavelets applications have been well developed in another important area, time series analysis. There is vast literature on this topic and wavelets are so popular and are developing so quickly that a inclusive complete review is impossible at this point. Except the de-noising and function and density estimations we discussed before, wavelet methods have been explored in time-scale stochastic time series modeling and forecasting in general. Conventional spectral analysis for stationary processes using Fourier transformation has been extended to analyze non-stationary using time-scale analysis. Following is a brief review in several directions in which, I believe, wavelets have potential applications in financial economics analysis.

Time series wavelet de-noising and estimations use the shrinkage technique and spectral density estimation using thresholding following the same logic that has been discussed before for no time series samples.

Non-decimated wavelet transform (NDWT), so called stationary wavelet transformation, was introduced by (Dutilleux, 1989). Examples of NDWT can be found in (Johnstone and Silverman, 1997). Related references are (Shensa, 1992), (Coifman and Donoho, 1995), (Mallat and Hwang, 1992), (Nason and Silverman, 1995), and (Zheng et al., 1999).

(Johnstone and Silverman, 1997) discuss the correlated and nonstationary noise issue. In (Johnstone and Silverman, 1998), they discuss how to deal with irregularly spaced data in the wavelet domain. Another important topic is wavelet variance, discussed in (Percival and Walden, 2000) and (Genacy et al., 2002). Wavelet variance provides a scale-based analysis of variance complementary to traditional frequency-based spectral analysis. Wavelet variance (a similar concept is Allan variance, which is used to measure the frequency stability in time domain) is a global scale measure of behavior. The study of wavelet variance can be used to analyze and synthesize both stationary long memory processes and related non-stationary processes, e.g., fractional Brownian motion (fBm). Locally stationary wavelet process (LSW) and evolutionary wavelet spectrum (EWS) issues are discussed in (Nason et al., 1998, Nason and von Sachs, 1999). Many other pioneers' works we are unable to cover here. However, it is clear that there is ample room for substantial improvement of the powerful wavelet methods. Examples are exploring wavelets on irregular point sets by (Daubechies et al., 1999), etc.

2.3.3 About Wavelets in Economic and Financial Analysis

Wavelets theory has developed into a methodology for financial analysis, and wavelet analysis has been applied to a lot of situations, and most of them have had

favorable results. Wavelet analysis has remarkable impact on many fields, mainly on mathematics, signal processing, image analysis, data compression, geophysics, numerical analysis, and statistics.
Despite the special properties, however, wavelets applications in financial analysis have not been extensively used. Literature in this field is rare. Wavelets applications in finance can be grouped into three categories. The first category is that wavelets methods are used to study the non-stationary property of financial time series; associated topics include structure change, local stationary and long-memory process. The second category is that wavelet methods provide alternatives for forecasting. Wavelets applications in statistics, discussed above, closely relates to the above two categories, which are technically oriented. The third category is more oriented by the financial theories; they concern how wavelets decompositions can be used to improve the hypothesis testing on exiting theories and also can provide insights of financial phenomena and enhance the development of theories. Next is a brief literature review of wavelets in finance. Mark Jensen's paper (Jensen, 1997) is an introduction to application of wavelets to finance. In this paper he discussed the advantages of using wavelets to deal with high frequency irregular spaced financial data.

In a series of his papers (Jensen, 1999, 1998, 2000) on using wavelets of analysis long-memory process, he developed several methods to estimate the fractional differencing parameter in autoregressive, fractionally integrated, moving average model (ARFIMA).
Based on his wavelet OLS estimator, (Tkacz, 2001) studied the order of integration of interest rates for U. S. and Canada, and find that most rates are mean-reverting in the very long run, with the fractional order of integration increasing with the term to maturity.

(Ben Mabrouk et al., 2009) use different wavelet estimators of the order of integration in the return and the volatility series of stock market indices. They show that the long memory property in stock returns is approximately associated with emerging markets rather than developed ones while strong evidence of long range dependence is found for all volatility series.

The wavelets based estimator outperforms the conventional OLS method in most cases of out-of-sample forecasting. Another interesting result is that the wavelets-based method has a better performance than a random walk, which means that wavelets-based regressor clearly extracts useful information about the underlying fundamental relationship.

(Roueff and Von Sachs, 2009) allow the long-memory parameter d to be varying over time. In their paper the authors adopt a semi-parametric approach for estimating the time-varying parameter d in order to avoid fitting a parametric model, such as ARFIMA, to the observed data. (Guegan and Zhiping, 2009) apply the Discrete Wavelet Packet Transform (DWPT) to estimate the time-varying parameters d of the locally stationary k-factor Gegenbauer process.
Ramsey has a series of pioneer papers on applications of wavelets in economics and finance. In his 1995 paper (Ramsey et al., 1995), the study focused on reveal the self-similarity in U.S. stock returns. The results are that stock market data are clearly complex, but there seems to be some evidence of non-randomness in the data, so they are more structured than mere representations of Brownian motion. Ramsey and Zhang examined the complex of financial data in their two papers (Ramsey and Zhang, 1994) and (Ramsey and Zhang, 1995). They simply visualize the S&P 500 index and foreign exchange by wavelet form dictionaries decomposition. They found that financial data are not random walks but very complex. There are bursts of high energy well localized and evidence of quasi-periodic se-

quence of Dirac delta function shocks to the system. And it appears that energy of the system is largely internal.

In another two papers, Ramsey and Lampart (Ramsey and Lampart, 1998a,b), examined the relationship between two economic variables in different scales, and the result is that the relationship between economic variables does vary significantly across scales. They also found that delay in the relationship between two variables might well be a function of the state space.

In the next paper (Ramsey, 1999), Ramsey further examined the relationship of two economic variables at each scale by study the distributional properties of the regression estimators and of the residuals in the context of the models. He confirmed the results in previous papers. (Davison et al., 1998) in another paper examine the different properties of the wavelet coefficients across scales. In this paper they studied twenty-one international commodity prices series, monthly data from 1960 to 1995. Their results are tentative.

(Karuppiah and Los, 2005) use wavelet multi-resolution analysis to analyze the nonstationarity and self-similarity of intraday Asian currency spot exchange rate. They use Lipschitz regularity exponent to measure the irregularity of foreign exchange at each scales. Their results are that foreign exchange markets exhibit nonlinear, non-normal dynamic structures, long term dependence, high kurtosis, and high degrees of non-informational trading. They also give some possible interpretations on the results.

(Norsworthy and Gorener, 2000) studies the relationship between individual asset returns and market returns in different scale and they found rather strong evidence that the market's influence on asset returns is principally in the high frequency movements of the asset returns.

Another paper (Norsworthy and Siregar, 2001) further examined the relationship between asset returns and market composite returns using intraday data with Kolmogorov-Smirnov test on the distributions of coefficients in each scale. Again

they found strong evidence that the differences are in high frequencies.
Another important application using wavelets in finance is financial time series forecasting, one of the ultimate goals of financial analysis. (Fryzlewicz, 2005) applies wavelet methods to process nonstationary financial time series for forecasting. He propose a new method for estimating time-varying second order quantities in the Locally Stationary Wavelet (LSW) framework, and provide an exploratory analysis of the daily FTSE 100 series using LSW toolbox. Another way of using wavelets for forecast is closely related with neural networks. Neural networks have been extensively used as a data mining method for short period forecasting. Wavelet method is an ideal decomposition, filtering, and de-noising technique for data preprocessing before they are input into neural networks.
(Triantafyllopoulos and Nason, 2009) show that the Locally Stationary Wavelet (LSW) process can be decomposed into a sum of signals, each of which follows a moving average process with time-varying parameters. They then show that such moving average processes are equivalent to state space models with stochastic design components. Using a simple simulation step, They propose a heuristic method of estimating the above state space models and then they apply the methodology to foreign exchange rates data.

Recently, (Ben Ammou et al., 2010) study the dynamics of the stock returns within the French stock market. Wavelet-based thresholding techniques are applied to the stock price series in order to obtain a set of explanatory variables that are practically noise-free. The Principal Components Regression (PCR) is then carried out on the new set of regressors.

Many new achievements have been made in statistical applications using wavelet methods. But potential applications using wavelets in financial economics have not been well developed. Most existing papers using wavelets in economics are stay in either macro level or on the data-driven side. In depth analysis based on finance theory using wavelets is still in its infancy. This study is an effort to fill in the gap between engineering and financial economics. Potential wavelets applications in financial economics include:

- Filter out "noise" traders (wavelets shrinkaging)
- Separation of Short term and long run performances (time-scale decomposition)
- Distinguish between market and individual behaviors
- Event study (isolate and help to model adaptation to shocks)
- De-correlation
- Non-stationary structure change (locate discontinuities)
- Forecasting Some advantages using wavelet methods
- Robustness of procedure (erroneous assumptions) (no parametric tests of procedures)
- Flexibility of regression fit (imprecise model formulations)
- Ability to handle complex relationships
- Efficiency of the estimators (few data points)
- Simplicity of implementation
- Ability to deal with non-stationarity of the stochastic innovations that inevitably are involved with economic and financial time series.

CHAPTER 3

Wavelets for Variance-Covariance Estimation

3.1 Introduction

First important use for the Discrete Wavelet Transform (DWT) and its variant, Maximal Overlap DWT (MODWT), is to decompose the simple variance of a time series on scale-by-scale basis.
Second, the wavelet variance is closely related to the concept of Spectral Density Function (SDF) and offers a simple summary of the SDF in cases where this function has a fairly simple structure. Third, the wavelet variance is a useful substitute for the variance of process for certain process with infinite variance.

This chapter is motivated by two interesting questions:

1. Correlation analysis;

2. Detecting and locating change point.

The wavelet covariance is shown to decompose the covariance between two stationary processes on scale by scale basis. The wavelet cross-covariance and cross-correlation are also defined in order to perform a more through scale by scale analysis of bivariate time series.

3.2 Wavelet Variance

Wavelet transform is able to analyze the variance of a stochastic process and decompose it into components that are associated to different time scales. In particular, given a stationary stochastic process X with variance σ_X^2 and defined the level j wavelet variance $\sigma_X^2(\tau_j)$, the following relationship holds

$$\sum_{j=0}^{\infty} \sigma_X^2(\tau_j) = \sigma_X^2$$

where $\sigma_X^2(\tau_j)$ represent the contribution to the total variability of the process due to changes at scale $\tau_j = 2^{j-1}$. This relationship says that wavelet variance decomposes the variance of a series into variances associated to different time scales[1]. By definition, the (time independent) wavelet variance for scale τ_j , $\sigma_X^2(\tau_j)$, is defined to be the variance of the j-level wavelet coefficients

$$\sigma_X^2(\tau_j) = var(\widetilde{w}_{j,t})$$

The scale τ_j is associated with the frequency interval $[1/2^{j+l}, 1/2^j]$, and we can use this property to obtain an approximate relation between the wavelet variance and the SDF of X_t via

$$\sigma_X^2(\tau_j) \approx 2 \int_{1/2^{j+l}}^{1/2^j} S_X(f)df \tag{3.1}$$

The factor of 2 is needed because the spectrum is an even function of frequency over the interval $[-1/2, 1/2]$. One way to visualize the wavelet variance is through

[1] The wavelet variance decomposes the variance of certain stochastic processes with respect to the scale $\tau_j = 2^{j-1}$ just as the spectral density decompose the variance of the original series with respect to frequency f, that is $\sum_{j=0}^{\infty} \sigma_X^2(\tau_j) = \sigma_X^2 = \int_{-1/2}^{1/2} S_X(f)df$

Equation 3.1 where the wavelet variance is a piecewise constant function with respect to frequency (see Genacy et al., 2002, p 239).

3.2.1 Estimation of the Wavelet Variance

Let a dyadic length $N = 2^J$ realization $(X_0, X_1, \ldots, X_{N-1})$ of a stochastic process $\{X_t\}$ and apply the partial DWT of order $J_p \leq J$ to produce the length N vector of wavelet coefficients to. As shown in (Percival, 1995), provided that $N - Lj \geq 0$, an unbiased estimator of the wavelet variance based on the DWT may be obtained, using

$$\sigma_X^2(\tau_j) = \frac{1}{2\tau_j N_j} \sum_{t=L'_j}^{N/2^j-1} w_{j,t}^2 \tag{3.2}$$

where $L'_j = (L-2)(1-2^{-j})$ is the number of DWT coefficients computed using the boundary, hence $N_j = N/2^j - L'_j$ is the number of wavelet coefficients at scale τ_j unaffected by the boundary and $w_{j,t} = \sum_{k=0}^{L_j-1} \widetilde{h}_{j,k} X_{t-k}, \quad t = \ldots, -1, 0, 1, \ldots$.

The restrictions of DWT, i.e. a sample size multiple of 2^J and sensitivity to circular shifts due to the down sampling approach, are overcome by the Maximal Overlap DWT (MODWT) which applies to any sample and is translation invariant, at the cost of giving up orthogonality. The Maximal Overlap Discrete Wavelet Transform (MODWT) is a non-orthogonal variant of the classical Discrete Wavelet Transform that unlike the orthogonal discrete wavelet transform, is translation invariant, as shifts in the signal do not change the pattern of coefficients.
An unbiased estimator of the wavelet variance based on the MODWT can be found as

$$\widehat{\sigma}_X^2(\tau_j) = \frac{1}{M_j} \sum_{t=L_j}^{N-1} \widetilde{w}_{j,t}^2 \quad j = 1, 2, \ldots, J. \tag{3.3}$$

(Percival, 1995) showed that the estimator $\widehat{\mathbf{v}}_X^2(\tau_j)$ is asymptotically Gaussian distributed with mean $\sigma_X^2(\tau_j)$ and variance $S_{\varpi,j}(0)/M_j$, where $S_{\varpi,j}(0)$ is the SDF of the scale τ_j squared wavelet coefficients $\varpi_{j,t} = W^2_{j,t}$ evaluated at frequency zero and is given by

$$S_{\varpi,j}(0) = \int_{-1/2}^{1/2} S_j^2(f)df \tag{3.4}$$

Thus, to a good approximation for large $M_j = N - L_j + 1$, we have

$$\frac{M_j^{1/2}(\widehat{\sigma}^2_X(\tau_j) - \sigma^2_X(\tau_j))}{S_{\varpi,j}(0)^{1/2}} \stackrel{d}{=} N(0,1) \tag{3.5}$$

To interpret the square integrability condition in Equation 3.4, we note that, $\{\overline{W}_{j,t}\}$ is a stationary process.

It has an Auto-Covariance Sequence (ACVS), say $\{s_{j,\tau}\}$. If the ACVS satisfies the square summability condition, then the ACVS and the SDF form a Fourier transform pair $\{s_{j,\tau}\} \leftrightarrow S_j(.)$. Parseval's theorem [2] then yields

$$\sum_{\tau=-\infty}^{\infty} s^2_{j,\tau} = \int_{-1/2}^{1/2} S_j^2(f)df$$

Thus, the square integrability condition holds as long as the ACVS for $\widetilde{w}_{j,t}$ dies down fast enough so that it is square summable.

3.2.2 Confidence Intervals for the Wavelet Variance

Following the above, let $\Phi^{-1}(p)$ represent the $p \times 100\%$ percentage point for the standard normal distribution. The symmetric nature of the Gaussian distribution implies that, for $0 \leq p \leq 1/2$,

$$\mathbf{P}[-\Phi^{-1}(1-p) \leq Z \leq \Phi^{-1}(1-p)] = 1 - 2p$$

For large N,

$$\left[\widehat{\sigma}^2_X(\tau_j) - \Phi^{-1}(1-p)\left(\frac{2S_{\varpi,j}(0)}{M_j}\right)^{1/2} ; \; \widehat{\sigma}^2_X(\tau_j) + \Phi^{-1}(1-p)\left(\frac{2S_{\varpi,j}(0)}{M_j}\right)^{1/2}\right] \tag{3.6}$$

[2]Perseval's theorem: $\sum_{t=-\infty}^{\infty} |a_t|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |A(f)|^2 df$ where $\{a_t\}$ is an infinite real or complex valued sequence and $A(f) = \sum_{t=-\infty}^{\infty} a_t \exp(ift)$

with probability $1-2p$, includes the true value of $\sigma^2_X(\tau_j)$ and constitutes a $100\times(1-2p)\%$ confidence interval for $\sigma^2_X(\tau_j)$ (Percival and Walden, 2000, pages: 311,312).
To use the above confidence interval, we must estimate $S_{\varpi,j}(0)$, which is the integral of $S_j^2(.)$. Since a periodogram can estimate large values of an SDF with relatively little biais, we can juste use the periodogram ($\widehat{S}_j^{(\mathrm{P})}$)as our estimator of $S_j(.)$:

$$\widehat{S}_j^{(\mathrm{P})}(f) = \frac{1}{M_j}\left|\sum_{t=L_j-1}^{N-1} \widetilde{W}_{j,t}e^{-i2\pi ft}\right|^2 \tag{3.7}$$

for large N and $0<|f|<1/2$, it follows

$$\frac{2\widehat{S}_j^{(\mathrm{P})}(f)}{S_j(f)} \stackrel{d}{=} \chi^2(2)$$

where $\chi^2(2)$ is a χ^2 random variable with two degree of freedom.

An approximately unbiased estimator of $S_{\varpi,j}(0)$ for large M_j is given by

$$\widehat{S}_{\varpi,j}(0) = \frac{1}{2}\int_{-1/2}^{1/2}\left[\widehat{S}_j^{(\mathrm{P})}(f)\right]^2 df \tag{3.8}$$

3.2.3 Exemple: Oil price volatility

In this example, we apply the MODWT-based wavelet variance to Oil price volatility data from December 17, 2003, to November 08, 2007.
Figure 3.1 displays the MODWT-based wavelet variance (the straight lines) with a 95% approximate CI of the estimate (represented by the upper, U, and lower, L, lines).
The main result emerging from wavelet variance analysis is a tendency for wavelet variance to decrease as the wavelet scale increases (there is an approximate linear relationship between the $\widehat{\sigma}^{\mathbf{2}}_X(\tau_j)$ and the wavelet scale τ_j indicating the potential for long memory in the volatility series. (Genacy et al., 2002).
Since the MODWT decomposes a series by local and scale, we may partition the

wavelet variance according to where we believe a change in variance occurs. In section section 3.3, we present a specific test to detect and locate changes in variance.

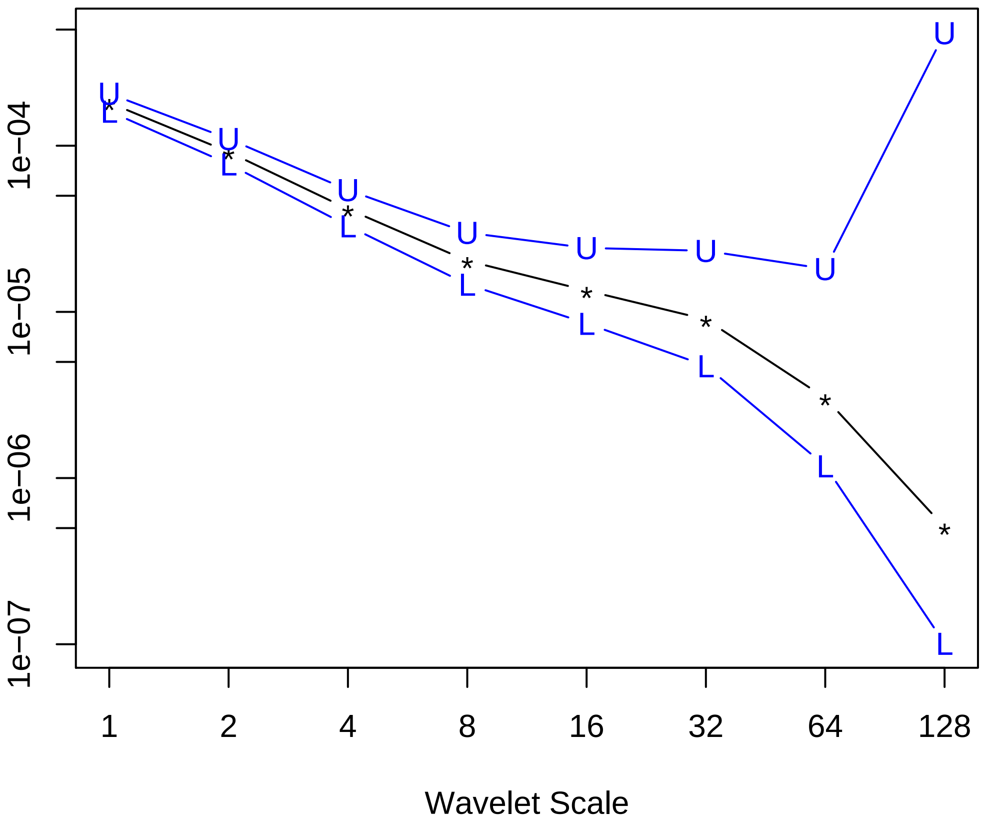

Figure 3.1: Wavelet variance of Oil price volatility

3.3 Testing for homogeneity

Whitcher, Byers, Guttorp and Percival (Whitcher et al., 1999) developed a framework for applying a test for homogeneity of variance on a scale-by-scale basis to long-memory processes. A good summary of the procedure is located in Gencay, Selcuk and Whitcher (Genacy et al., 2002, Section 7.3). The test relies on the usual econometric assumption that the crystals of coefficients, $w_{j,t}$ for scale j at time t have zero mean and variance $\sigma_X^2(\tau_j)$.

3.3.1 The test statistic

If $Y_0, \ldots, Y_{N-1}$ constitutes a portion of an Fractionally Differenced (FD) process with long memory parameter $0 < d < 1/2$, and with possibly nonzero mean, then each sequence of wavelet coefficients $W_{j,t}$ for Y_t is approximately a simple from a zero mean white noise process. This enables us to formulate our test for homogeneity of variance using wavelet coefficients for FD processes, as follows:
Let $X_0, \ldots, X_{N-1}$ be a time series that can be regarded as a sequence of independent Gaussian (normal) random variables with zero means and variances $\sigma_0^2, \ldots, \sigma_{N-1}^2$. We would like to test the hypothesis

$$H_0 : \sigma_0^2 = \ldots = \sigma_{N-1}^2 \tag{3.9}$$

A test statistic that can discriminate between Equation 3.9 and a variety of alternative hypotheses (such as $H_1 : \sigma_0^2 = \ldots = \sigma_k^2 \neq \sigma_{k+1}^2 = \ldots = \sigma_{N-1}^2$ where k is an unknown change point) is the Normalized Cumulative Sums of Squares (NCSS) test statistic D, which has previously been investigated by, (Inclan and Tiao, 1994).
Let

$$\mathcal{P}_k = \frac{\sum_{j=1}^{k} X_j^2}{\sum_{j=1}^{N} X_j^2}; \;\; k = 1, 2, \ldots, N-1. \tag{3.10}$$

denote the normalized partial energy sequence of X_t.
The desired test statistic is given by $D = \max(D^+, D^-)$ where

$$D^+ = \max_{0 \leq k \leq N-2} \left(\frac{k}{N-1} - \mathcal{P}_k\right) \quad \text{and} \quad D^- = \max_{0 \leq k \leq N-2} \left(\mathcal{P}_k - \frac{k-1}{N-1}\right) \tag{3.11}$$

In order to perform hypothesis testing here, we require knowledge of the distribution of D under H_0. (Inclan and Tiao, 1994) proved that, for large $\widetilde{N}_j$ and

$x_p > 0$,

$$\mathbf{P}(D \leq x_p) \approx 1 + 2\sum_{l=0}^{\infty}(-1)^l \exp(-l^2 \widetilde{N}_j x_p^2) \tag{3.12}$$

The statistic $\mathcal{P}_k$, as formulated in Equation 3.10, is measuring the accumulation of variance in the time series as a function of time. This quantity is then compared with a line at a 45° angle, where the maximum vertical deviation from this line is recorded as our test statistic D.

Critical levels for D under the null hypothesis can be readily obtained through Monte Carlo simulations for arbitrary sample size. The procedure presented here is to generate empirical critical values for the cumulative sum of squares test statistic D applied to the first scale of wavelet coefficients from the Haar wavelet filter. The procedure is as follows (Whitcher et al., 1999):

1. Generate a length N realization of a Gaussian white noise process;
2. Compute the DWT down to scale J, using the Haar wavelet filter;
3. Discard all wavelet coefficients at each scale that make explicit use of the periodic boundary conditions;
4. Compute the test statistic D for all scales based upon the remaining wavelet coefficients; and
5. Calculate the value of D such that $100\%(1-\alpha)$ simulated values are greater than that value. This is the $(1-\alpha)$-th quantile.

3.3.2 Locating a Variance Change

Once an observed time series has been tested and the null hypothesis rejected at some scale ,τ_j, we have succeeded in detecting a significant change in variance somewhere in the series. It may be of interest to estimate where this change occurred, but the DWT has poor time-resolution ; especially as j increases. The

MODWT provides a useful alternative to more precisely determine the location of a variance change after it has been detected. Through its superior time resolution, where each wavelet coefficient is associated with each t regardless of scale, the location of a variance change may be associated with a specific observation in the original time series.

Let

$$\widetilde{\mathcal{P}}_k = \frac{\sum_{t=L_j-1}^{k} \widetilde{w}_{j,t}^2}{\sum_{t=L_j-1}^{N-1} w_{j,t}^2} \quad k = L_j - 1, \ldots, N-2. \tag{3.13}$$

and define $\widetilde{D} = \max(\widetilde{D}^+, \widetilde{D}^-)$, where

$$\widetilde{D}^+ = \max\left(\frac{k - L_j + 2}{N - L_j} - \widetilde{P}_k\right) \qquad \widetilde{D}^- = \max\left(\widetilde{P}_k - \frac{k - L_j + 1}{N - 1}\right)$$

The estimated location of the variance change $\widehat{k}$ is the point at which $\widetilde{D}$ is achieved. (Whitcher et al., 1998) showed that $\widetilde{D}$ can accurately locate a change of variance when applied to wavelet coefficients from long-memory processes. The accuracy of this method improves as the ratio of the variances before and after $\widetilde{k}$ moves away from unity. When the null hypothesis is rejected over several scales, the estimated location associated with the lowest level of the wavelet transform (highest frequency interval) is recommended.

3.3.3 Application to Oil Price

Here we apply the methodology developed in this section to daily Oil Price from December, 17, 2003 to November, 08, 2007. An obvious increase in volatility $v_t = |r_t|$ is observed in the latter half of the time series.

Figure 3.2 shows the results of testing the daily Oil price volatility series (absolute returns) for a single change in variance at an unknown time. The bottom row is the original Oil Price volatility series, and the following three rows are the Normalized Cumulative Sum of Squares ($NCSS$) of the first three levels of its wavelet decomposition. These three levels are associate with changes in longer and longer

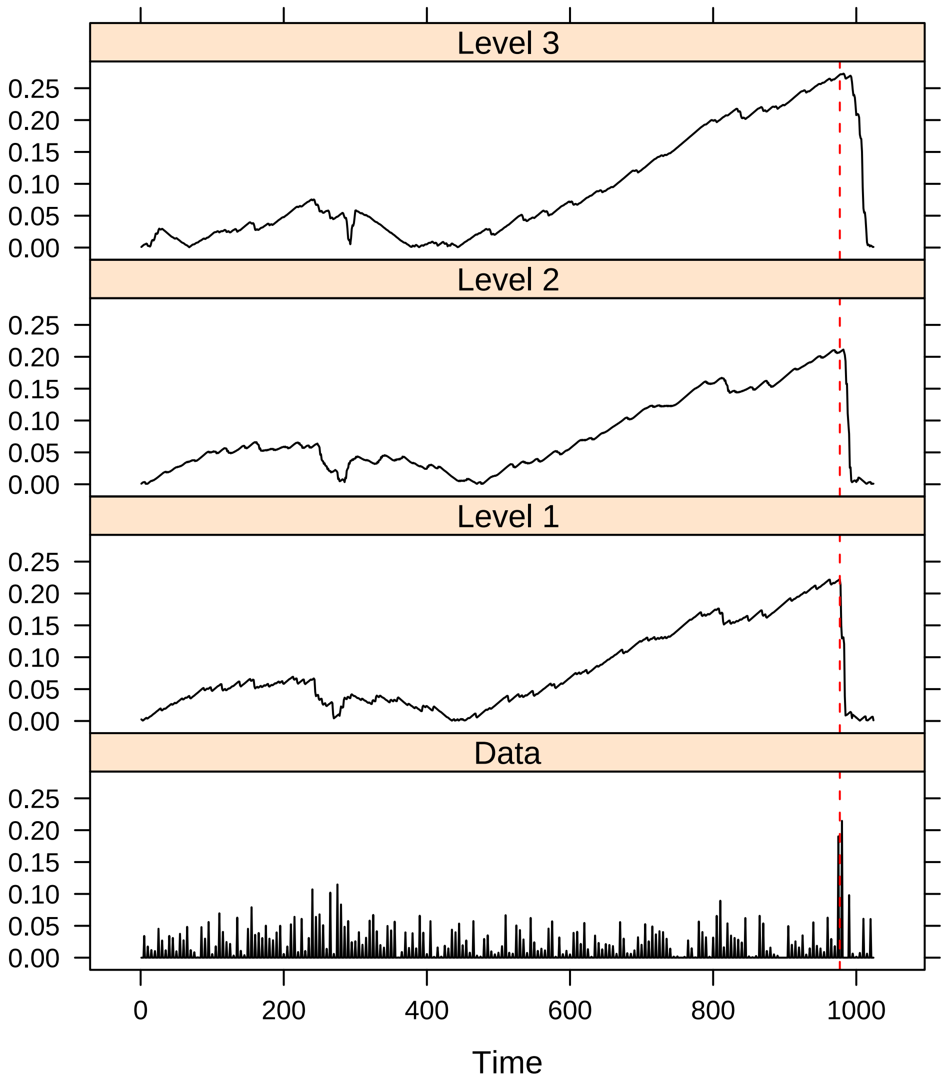

Figure 3.2: Oil price volatility (Bottom panel) along with the NCSS for its wavelet decomposition.

timescales; specifically - from bottom to top - changes on the order of one, two, and four days. The dashed vertical line denotes the location of the maximum for the scale one day wavelet coefficients (observation 977).

The null hypothesis of constant variance is rejected for the first three scales of the wavelet transform. The normalized cumulative sum of squares is displayed for the first three levels of wavelet coefficients in Figure 3.2. Since the level 1 coefficients (second row from the bottom in Figure 3.2) are associated with the highest frequencies, we use its maximum as the estimated time of variance change.

We now want to formally test for homogeneity of variance using the test statistic D in Section section 3.3. Testing through the DWT is necessary because we cannot assume the volatility v_t has mean zero and is uncorrelated. The approximate decorrelation property of the DWT allows us to apply the test statistic to each subseries of wavelet coefficients. The values of D for the first four scales are provided in Table Table 3.1, along with critical values computed via Monte Carlo simulations.

			Critical levels		
τ_j	N_j	D	10%	5%	1%
1	509	0.2606806	0.07672936	0.0851379	0.1020280
2	251	0.1180120	0.10924478	0.1212283	0.1452973
3	122	0.4642792	0.15186692	0.1691397	0.2025723
4	58	0.2793638	0.21834421	0.2425481	0.2945787

Table 3.1: Results of testing oil price volatility for homogeneity of variance using the LA(8) wavelet filter critical values determined by computer simulation.

Since D for τ_1 =1 day is greater than any of the tabulated critical values, we can reject the null hypothesis of homogeneity of variance at all three critical levels. For τ_2 =2 days, we can not reject the null hypothesis at 5% and 1% level of significance.

Table 3.2 shows the results of applying this test to the Oil Price return series at 10%, 5% and 1% level of significance respectively.

Scale	10%	5%	1%
d1	FALSE	FALSE	FALSE
d2	FALSE	TRUE	TRUE
d3	FALSE	FALSE	FALSE
d4	FALSE	FALSE	TRUE

Table 3.2: Testing for Homogeneity

3.4 Detection and location of multiple variance changes

In practice, a given time series may exhibit more than one change in variance. A natural approach is to test the entire series first, split at a detected change point, and repeat until no change points are found. This is known as a binary segmentation procedure studied by (Vostrikova, 1980), who proved its consistency. (Inclan and Tiao, 1994) used the binary segmentation procedure with the test statistic D to detect and locate changes of variance in sequences of Independent and Identically distributed (IID) Gaussian random variables. (Chen and Gupta, 1997) used a combination of this binary segmentation procedure with the Schwarz Information Criterion (SIC) to detect and locate changes of variance also in sequences of IID Gaussian random variables.

For detection, we consider the iterated cumulative sums of squares (ICSS) algorithm for the identification of multiple variance change points in sequence data, (Whitcher et al., 1998) adapted the ICSS algorithm to discrete wavelet transforms (DWT) and to maximal overlap discrete wavelet transforms (MODWT).

3.5 Wavelet Covariance and Cross-Covariance

In this section we expend wavelet methodology to easily handle two observed time series. The wavelet covariance is the covariance between the scale τ_j wavelet coefficients from a bivariate time series and the wavelet correlation is the correlation between the scale τ_j wavelet coefficients from a bivariate time series. If we introduce a lag k between the two series, we obtain the wavelet cross-covariance and wavelet cross-correlation as naturel extension to their classical times series analogs.

Let $X_t = (x_{1,t}, x_{2,t})$ be a bivariate stochastic process with univariate spectra $S_1(f)$ and $S_1(f)$, respectively, and let $W_{j,t} = (w_{1,j,t}, w_{2,j,t}$ be the scale τ_j wavelet coefficients computed from X_t. The wavelet covariance of $(x_{1,t}, x_{2,t})$ for scale τ_j is defined as

$$\gamma_X(\tau_j) = \frac{1}{2\tau_j} \operatorname{Cov}(w_{1,j,t}, w_{2,j,t}) \tag{3.14}$$

If the length of the wavelet filter is sufficient to eliminate any non-stationary features in both $x_{l,t}$ and $x_{2,t}$, then we may simply define the wavelet covariance via the expectation $\gamma_X(\tau_j) = \mathbb{E}(w_{1,j,t}, w_{2,j,t})/(2\tau_j)$ since we are guaranteed to have mean zero wavelet coefficients.

As with the wavelet variance for univariate processes, the wavelet covariance decomposes the covariance between two stochastic processes on a scale-by-scale basis (Whitcher, 1998).

$$\sum_{j=1}^{\infty} \gamma_X(\tau_j) = \operatorname{Cov}(w_{1,j,t}, w_{2,j,t}) \tag{3.15}$$

By introducing an integer lag k between $w_{l,j,t}$ and $w_{2,j,t}$, we establish a definition for the wavelet cross-covariance

$$\gamma_{X,k}(\tau_j) = \frac{1}{2\tau_j} \operatorname{Cov}(w_{1,j,t}, w_{2,j,t+k}) \tag{3.16}$$

3.5.1 Estimation

Suppose $\mathbf{X} = (X_0, X_1, \ldots, X_{N-1})$ is a N realizations of the bivariate process X_t and apply a partial MODWT of order $j_p < \log_2(T)$ to each univariate process $x_{1,t}$ and $x_{2,t}$, thus producing J length N vectors of MODWT coefficients

$$\begin{aligned}\widetilde{W}_j &= (\widetilde{W}_{j,0}, \widetilde{W}_{j,1}, \ldots, \widetilde{W}_{j,N-1}) \quad j = 1, 2, \ldots, J \\ &= ((\widetilde{w}_{1,j,0}, \widetilde{w}_{2,j,0}), (\widetilde{w}_{1,j,1}, \widetilde{w}_{2,j,1}), \ldots, (\widetilde{w}_{1,j,N/2^j-1}, \widetilde{w}_{2,j,N/2^j-1}))\end{aligned}$$

and the length N vector of MODWT scaling coefficients

$$\begin{aligned}\widetilde{V}_J &= (\widetilde{V}_{J,0}, \widetilde{V}_{J,1}, \ldots, \widetilde{V}_{J,N-1}) \\ &= ((\widetilde{v}_{1,J,0}, \widetilde{v}_{2,J,0}), (\widetilde{v}_{1,J,1}, \widetilde{v}_{2,J,1}), \ldots, (\widetilde{v}_{1,J,N/2^j-1}, \widetilde{v}_{2,J,N/2^j-1}))\end{aligned}$$

An unbiased estimator of the wavelet covariance based upon the MODWT is given by

$$\widetilde{\gamma}_X(\tau_j) = \frac{1}{\widetilde{N}_j} \sum_{k=L_j-1}^{N-1} \widetilde{w}_{1,j,k} \widetilde{w}_{2,j,k} \tag{3.17}$$

where $\widetilde{N}_j = N - L_j - 1$. The estimator $\widetilde{\gamma}_X(\tau_j)$ is asymptotically Gaussian distributed with mean $\gamma_X(\tau_j)$ and variance $S_{\varpi,j}(0)/\widetilde{N}_j$, where $S_{\varpi,j}(0)/\widetilde{N}_j$ is the Spectral Density Function (SDF) of the product of the scale τ_j wavelet coefficients $\varpi_{j,t} = w_{1,j,k} w_{2,j,k}$ evaluated at zero frequency and is given by (see Whitcher, 1998).

$$S_{\varpi,j}(0) = \int_{-1/2}^{1/2} S_{1,j}(f) S_{2,j}(f) df + \int_{-1/2}^{1/2} S_W^2(f) df \tag{3.18}$$

where $S_{1,j}(f)$ and $S_{1,j}(f)$ denote, respectively, the SDFs for $w_{1,j,t}$ and $w_{2,j,k}$, while $S_W(f)$ is the cross spectrum between $w_{1,j,t}$ and $w_{2,j,k}$. This result holds when $W_{j,t}$ is a bivariate Gaussian stationary process.

Estimation of the wavelet cross-covariance follows directly from the biased estimator of the usual cross-covariance. For $N > L_j$, a biased estimator of the wavelet cross-covariance based on the MODWT is given by

$$\widetilde{\gamma}_{X,k}(\tau_j) = \begin{cases} \widetilde{N}_j^{-1} \sum_{t=L_j-1}^{N-k-1} \widetilde{w}_{1,j,t}\widetilde{w}_{2,j,t}, & k = 0, 1, \ldots, \widetilde{N}_j - 1 \\ \widetilde{N}_j^{-1} \sum_{t=L_j-k}^{N-1} \widetilde{w}_{1,j,t}\widetilde{w}_{2,j,t}, & k = -1, -2, \ldots, -(\widetilde{N}_j - 1) \\ 0, & \text{otherwise.} \end{cases} \tag{3.19}$$

3.5.2 Confidence Intervals

Confidence intervals for the wavelet covariance and cross-covariance follow from the large-sample result in the previous section.

We proceed by first estimating the variance of $\widetilde{\gamma}_X(\tau_j)$ using periodogram-based estimators.

If $\mathbf{X}$ is a length N realization of a bivariate Gaussian process, then the SDF for the product of the scale τ_j wavelet coefficients $S_{\varpi,j}(0)$ is given by Equation 3.18 and is a function of the auto-spectra and cross spectrum of the scale τ_j wavelet coefficients.

(Whitcher, 1998) propose to estimate these spectral quantities using the periodogram $\widehat{S}^P_{1,j}(f)$, $\widehat{S}^P_{2,j}(f)$ and cross-periodogram $\widehat{S}^P_{W,j}(f)$ to produce an estimator for $\widehat{S}_{\varpi,j}(0)$.

Parseval's relation allows us to formulate an alternative representation for $\widehat{S}^G_{\varpi,j}(0)$ that uses only the auto-covariance and cross-covariance sequences instead of the auto-spectra and cross spectrum.

$$\int_{-1/2}^{1/2} \widehat{S}^P_{1,j}(f)\widehat{S}^P_{2,j}(f)df = \sum_{k=-(\widetilde{N}_j-1)}^{\widetilde{N}_j-1} \widetilde{\gamma}_{1,j,k}\widetilde{\gamma}_{2,j,k}, \tag{3.20}$$

where $\widetilde{\gamma}_{1,j,k}$ and $\widetilde{\gamma}_{2,j,k}$ are biased estimators of the Auto-Covariance Sequence (ACVS) for $w_{l,j,t}$ and $w_{2,j,t}$. The second integral in Equation 3.20, substituting the cross-periodogram for the cross spectrum, may be re-expressed via

$$\int_{-1/2}^{1/2} \left[\widehat{S}^{P}_{W,j}(f)\right]^2 df = \sum_{k=-(\widetilde{N}_j-1)}^{\widetilde{N}_j-1} \widetilde{\gamma}^2_{W,j,k} \tag{3.21}$$

where $\widetilde{\gamma}_{W,j,k}$ is the biased sample cross-covariance sequence (CCVS) between the scale τ_j wavelet coefficients $w_{l,j,t}$ and $w_{2,j,t}$.

Under the assumption of Gaussianity, (Whitcher, 1998) propose an unbiased estimator for $\widehat{S}_{\varpi,j}(0)$:

$$\widehat{S}^{G}_{\varpi,j}(0) = \frac{\widetilde{\gamma}_{1,j,0}\widetilde{\gamma}_{2,j,0}}{2} + \sum_{k=1}^{\widetilde{N}_j-1} \widetilde{\gamma}_{1,j,k}\widetilde{\gamma}_{2,j,k} + \frac{1}{2}\sum_{k=-(\widetilde{N}_j-1)}^{\widetilde{N}_j-1} \widetilde{\gamma}^2_{W,j,k} \tag{3.22}$$

An approximate $(1-\alpha)$ Confidence Interval (CI) for $\gamma_X(\tau_j)$ can given by (for details see Percival and Walden, 2000):

$$\widetilde{\gamma}_X(\tau_j) = \pm Z_{\frac{\alpha}{2}} \left[\frac{\widehat{S}^{P}_{\varpi,j}(0)}{\widetilde{N}_j}\right]^2 \tag{3.23}$$

where $Z_{\frac{\alpha}{2}}$ is the upper $(1-\alpha/2)$ quantile of the standard normal distribution Φ.

3.5.3 Exemple: Daily Foreign Exchange rates

We perform a wavelet variance/covariance analysis of daily foreign exchange rates for the U.S. Dollar - Euro(USD-EURO) and U.S. Dollar-Japanese Yen(USD-JPY).

Here we investigate the returns series $r_t = \ln(pt) - \ln(pt + l)$.

A natural question to ask is how well these two series are associated with one another. The wavelet covariance is a measure of this association when comparing the two series at the same time point.

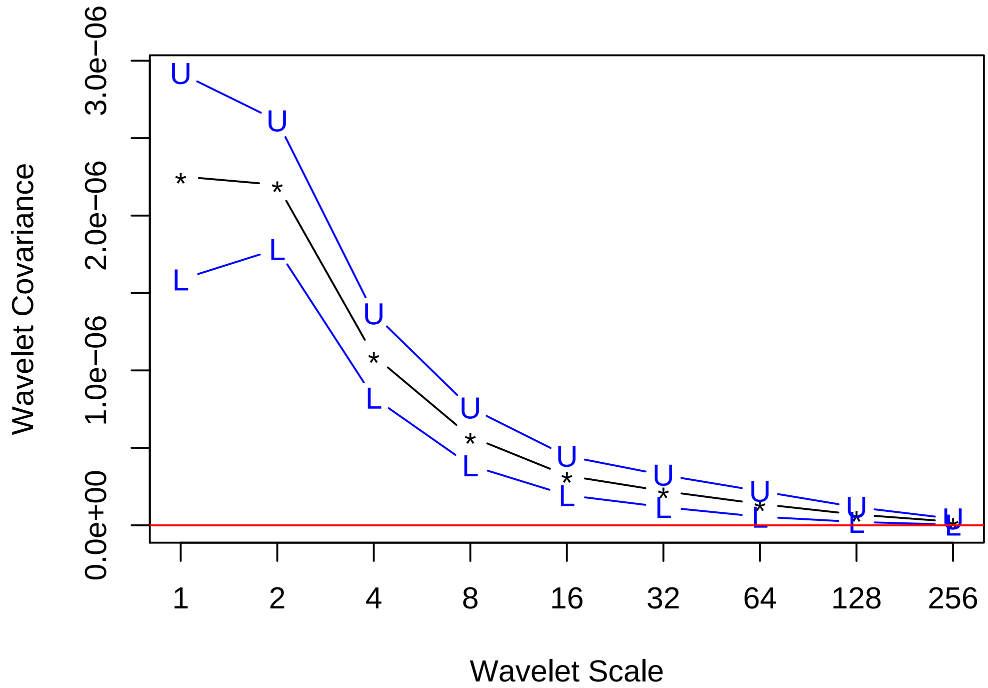

Figure 3.3: Wavelet covariance between the USD-EURO and USD-YEN exchange rate returns.

Figure 3.3 provides the MODWT-based wavelet covariance $\widetilde{\gamma}_X(\tau_j)$ with approximate confidence interval (The lines with the characters "L" and "U" denote the Lower and Upper bounds for the approximate 95% confidence interval respectively).

Wavelet covariances are positive and appear to be significantly different from zero (at the $\alpha = 0.05$ level of significance) except for the last two wavelet scales shown. Although there appears to be positive association between the two returns series, it is difficult to compare wavelet scales because of the differing variability exhibited by them. Standardizing by the variance of each series, at each scale, would be a simple way of overcoming this influence and make it possible to compare the magnitude of the association across scales. The wavelet correlation is therefore a natural quantity to use instead of the wavelet covariance.

3.6 Wavelet Correlation and Cross-Correlation

Although the wavelet covariance decompose the covariance between two stochastic processes on a scale-by-scale, in some situations it may be beneficial to normalize the wavelet covariance by the variability inherent in the observed wavelet coefficients.

Wavelet covariance at scale λ_j can be expressed as follows:

$$\gamma_{XY}(\tau_j) = Cov_{XY}(\tau_j) = \frac{1}{\widehat{N}_j} \sum_{k=L_j-1}^{N-1} \widehat{W}_{j,k}^{X} \widehat{W}_{j,k}^{Y}$$

Given that covariance does not take into account the variation of univariate time series, it is naturel to introduce wavelet correlation concept.

The wavelet correlation is simply made up of the wavelet covariance for $\{X_t, Y_t\}$, and wavelet variances for $\{X_t\}$ and $\{Y_t\}$. The MODWT estimator of wavelet correlation can be expressed as follows:

$$\widehat{\rho}_{X,Y} = \frac{Cov_{XY}(\tau_j)}{\widehat{\sigma}_X^2(\tau_j)\widehat{\sigma}_X^2(\lambda_j)}$$

As with the usual correlation coefficient between two random variables, $|\widehat{\rho}_{X,Y}(\tau_j)| \leq 1$. The wavelet correlation is analogous to its Fourier equivalent, the complex coherency (Genacy et al., 2002).

The wavelet cross-correlation decomposes the cross-correlation between two time series on a scale-by-scale basis thereby making it possible to see how the association between two time series changes as a function of time horizon. (Genacy et al., 2002) define the wavelet cross-correlation as:

$$\rho_{X,k}(\tau_j) = \frac{\gamma_{X,k}(\tau_j)}{\sigma_1(\tau_j)\sigma_2(\tau_j)} \tag{3.24}$$

where $\sigma_1^2(\tau_j)$ and $\sigma_2^2(\tau_j)$ are, respectively, the wavelet variances for $x_{1,t}$ and $x_{2,t}$ associated with scale τ_j and $\gamma_{X,k}(\tau_j)$ the wavelet covariance between $x_{1,t}$ and $x_{2,t-k}$

associated with scale τ_j.
Just as the usual cross-correlation is used to determine lead-lag relationships between two processes, the wavelet cross-correlation should be able to provide a lead-lag relationship on a scale by scale basis (Genacy et al., 2002).

3.7 Application

3.7.1 Data description and basic statistics

An analysis was conducted using monthly data for the interest rate of American Treasury securities at 3-month constant maturity provided by the Federal Reserve and the exchange rate between USD and EURO.
The closing S&P500 index is used as the indication of the stock price fluctuation. Empirical analysis covers the period form January 1990 to December 2008 providing a 228 observation in total. In this study, we investigate the returns series $r_t = log(p_t) - log(p_{t-1})$. Initially subsection 3.7.1 shows some brief summary

	min	max	mean	std.dev	skewness	kurtosis
Interest Rate	−1.85	0.30	−0.02	0.17	−7.37	68.06
Stock index	−0.06	0.08	0.00	0.02	0.14	0.16
Exchange Rate	−0.18	0.11	0.00	0.04	−0.88	2.04

Table 3.3: Descriptive Statistics for returns series

statistics for the returns series of interest rate, exchange rate and stock index respectively.
From the table, we make the following observations:

a) The mean of returns series is equal zero for all series.

b) Interest rate returns have higher standard deviation than exchange rate and stock index returns.

c) Monthly returns of interest rate tend to have high excess kurtosis.

Both series appear to have similar characteristics, in terms of mean and variation, but a more thorough description is available to use through a multi-scale analysis.

3.7.2 Multi-scale analysis:

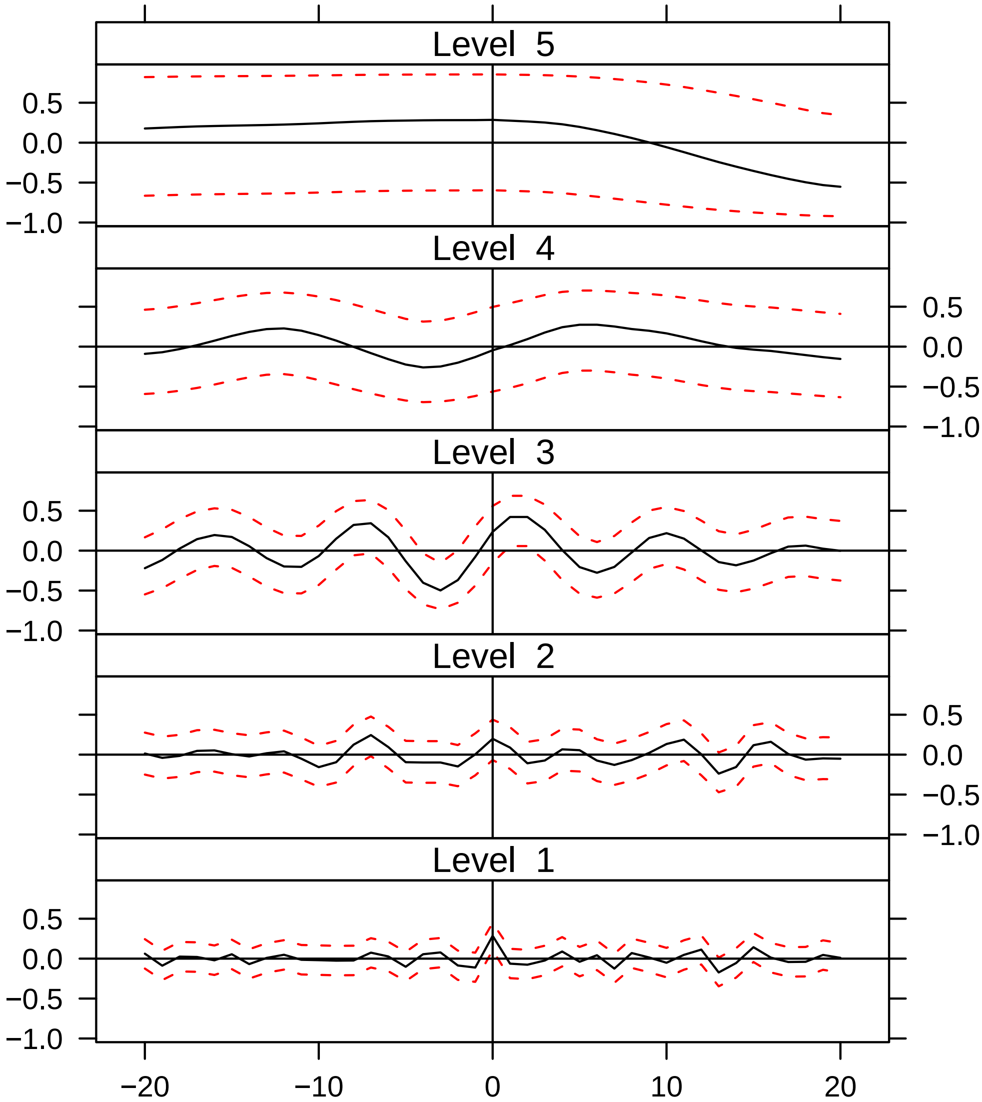

Figure 3.4: Wavelet cross-correlation between Interest rate and Exchange rate returns

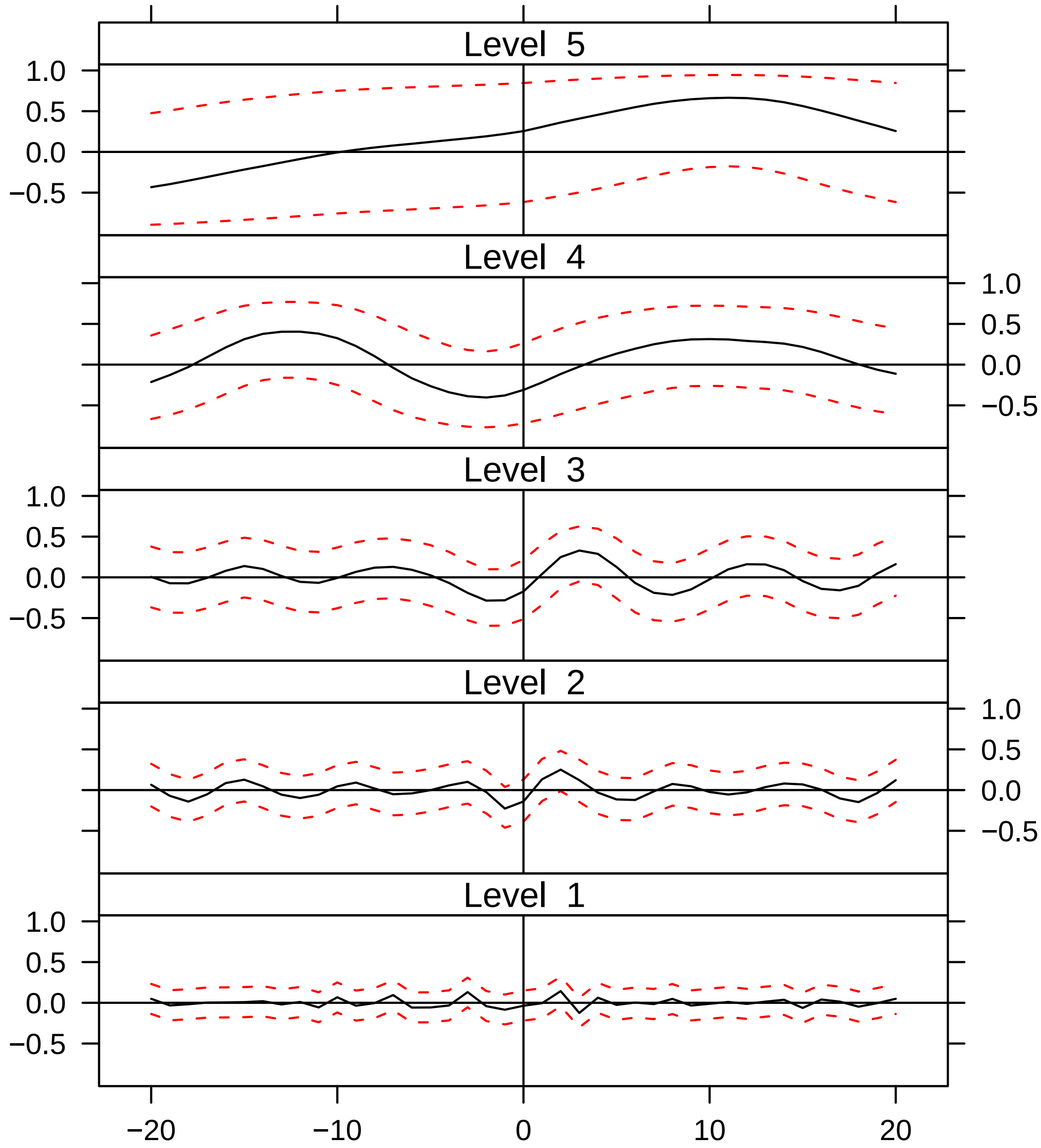

Figure 3.5: Wavelet cross-correlation between Interest rate and Stock index returns

We apply the Maximal Overlap Discrete Wavelet Transform to the monthly returns for the three series using the Daubechies (D) wavelet filter of length $L = 4$, that is $D(4)$, based on four non-zero coefficients (Daubechies, 1992), with periodic boundary conditions. The application of the translation invariant wavelet transform with a number of scales $J = 5$ produces five vectors of wavelet filter coefficients, that is w_5, w_4, w_3, w_2, w_1, and one vector of scaling coefficients, v_5.

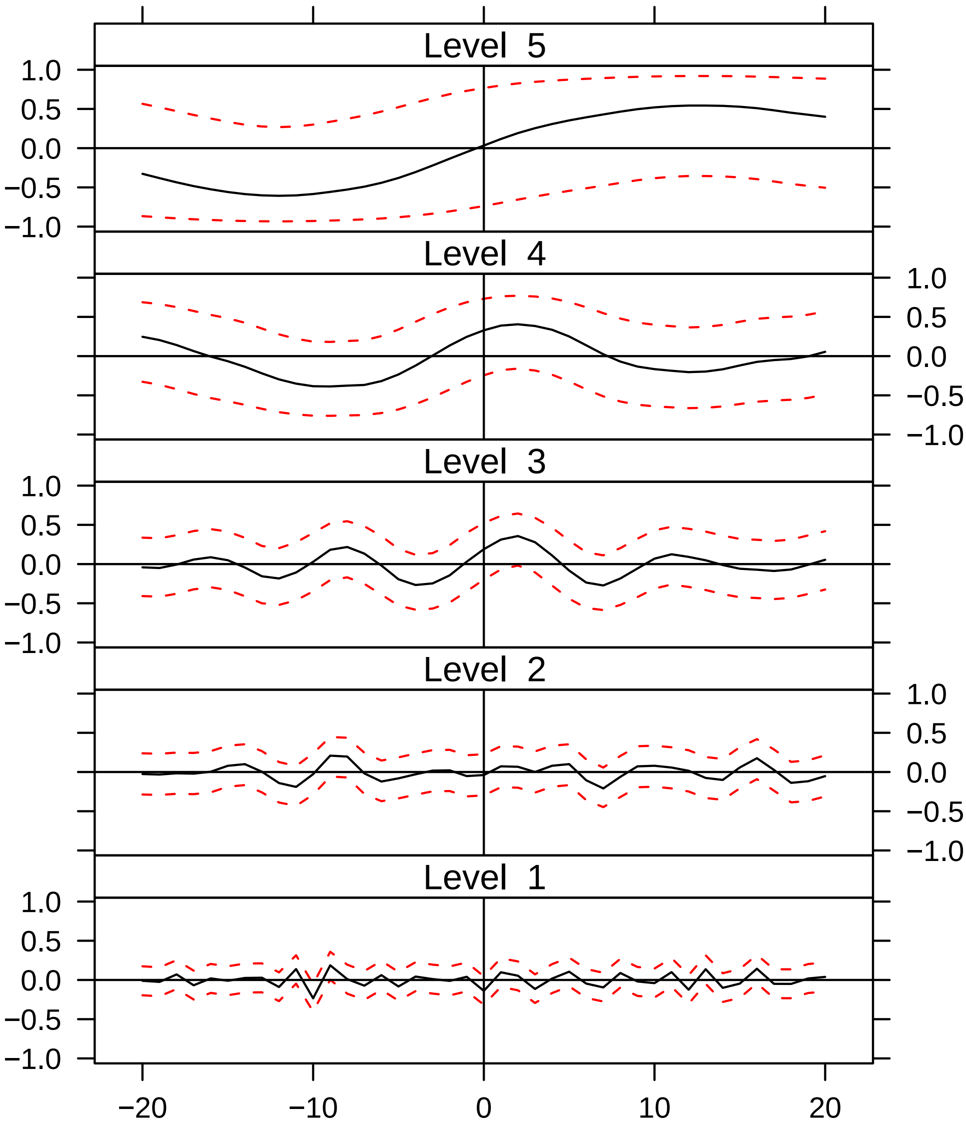

Figure 3.6: Wavelet cross-correlation between Exchange rate and Stock index returns

Since we use monthly data, the wavelet filter coefficients, $w_{5,k}, \ldots, w_{1,k}$, represent progressively finer scale deviations from the smooth behavior, and correspond to $32-64$, $16-32$, $8-16$, $4-8$ and $2-4$ months period, respectively.

In Figure 3.4, Figure 3.5 and Figure 3.6 we report the estimated wavelet cross-correlations coefficients and the corresponding approximate confidence intervals against time leads and lags for all scales between the interest rate returns and exchange rate returns, interest rate returns and stock index returns and between exchange rate returns and stock index returns respectively.

Figure 3.4 shows that, for all scales, the relationship between interest rate and exchange rate in generally not significantly different from zero at all leads and lags (slightly significant at scale 3). This means that interest rate returns and exchange rate returns in this period were independent and historical information of interest rate was not significantly predictive for exchange rate. In Figure 3.5, we report that at shortest scales, i.e. 1-3, the relationship between interest rate and stock index is not significantly different from zero. On the other hand, at the coarsest scales, i.e. 4-5, a significant positive relationship between the two series. Also, at scale 5, we underlying positive leading relationship between interest rate and stock returns, and we note the asymmetry in the wavelet cross-correlation sequence. At the fifth scale (associated with associations of 32 to 64 months), the interest rate is positively correlated with the stock return at a lead 10 months but not at a lag of 10 months.
This figure also shows that most of significant coefficients have positive values. This indicates that the interest rate appreciation (depreciation) was associated with a fall (rise) in stock index. The wavelet cross-correlation provides that exchange rate returns and stock index returns are independent at short horizons (high frequencies).
The Figure 3.6 also shows a significant relationship between the two series only at coarsest scale. We note also that significant coefficient have positive values at leads and negative values at lags (scale 5). This means the existence a relationship bidirectional between the two series at long horizons.

3.8 Conclusion

In this chapter, we have considered the use of wavelet in the analysis of variance, covariance, correlation and cross-correlation of time series.
(Genacy et al., 2002) introduced the concept of the wavelet variance, wavelet correlation and wavelet cross-correlation in terms of the discrete wavelet transform. We have used the DWT and MODWT to measure the variance and the association between two stationary processes and we have tested the homogeneity of variance of time series which exhibit long memory structure.

We have applied the notion of wavelet variance and testing homogeneity of variance to daily Oil Price volatility series from December, 17, 2003 to November, 08, 2007.
We showed that, the wavelet variance of this series has a tendency to decrease as the wavelet scale increase (there is an approximate linear relationship between the wavelet variance and the wavelet scale indicating the potential for long memory in the volatility series).
This means that the variance change over scale. The null hypothesis of constant variance is rejected for the first three scales of the wavelet transform.

The wavelet correlation and cross-correlation notions are applied to measure the association between two daily foreign exchange rate (USD/EURO and USD/YEN) and between interest rate, stock index and exchange rate.
We showed that the wavelet covariance between USD/EURO and USD/YEN is significantly positive for all wavelet scales.
We showed also, that for all scales, the relationship between interest rate and exchange rate is generally not significantly different from zero at all leads and lags. The relationship between interest rate and stock index is significantly positive at coarest scales (low frequencies). We reported a significant relationship between

exchange rate and stock index only at coarest scales. The significant coefficients have positive values at leads and negative values at lags. This means the existence a relationship bidirectional between the two series at long horizons.

CHAPTER 4

Wavelet and Long-Memory Process

4.1 Introduction

Long-memory phenomena have been observed frequently in finance, economics, hydrology, geophysics and many other fields. Generally speaking, a process is called long memory if its spectral density is unbounded at the origin. Consequently, the autocorrelation function decays at a hyperbolic rate such that the absolute values of the autocorrelations are not summable. Conventionally, the class of Fractionally Integrated Autoregressive Moving Average (ARFIMA) processes (Hosking, 1981) and (Granger and Joyeux, 1980) is used most widely in modeling time series with both long-memory and short-memory behaviors.

The corresponding Maximum Likelihood (ML) estimation (Sowell, 1992) is computationally intensive due to high-dimensional matrix inversion when dealing with large data sets. In contrast to the ARFIMA models constructed in the time domain, one alternative is to specify the spectral density semi-parametrically for small frequencies in the frequency domain. Accordingly, the corresponding inference for the long-memory parameter is derived in the frequency domain, for example the

GPH estimator (Geweke and Porter-Hudak, 1983), the Whittle estimator of (Fox and Taqqu, 1986) and the semi-parametric estimator of (Robinson, 1995). Since the semi-parametric model only characterizes the long-memory property of the underlying process, not the complete dependence structure of the underlying process, the implications on forecasting and predictions are limited.

It is well known that wavelet methods provide excellent tools for scale analysis of time series.

Wavelets are particularly powerful for analyzing long-memory properties due to self-similarity (Flandrin, 1992), (Masry, 1993) and (Wornell, 1993). In the literature about long memory, wavelet analysis has been used in simulating long-memory realizations and in estimating the long-memory parameter under a Fractionally Integrated (FI) process (McCoy and Walden, 1996) and (Jensen, 1999) and an ARFIMA process (Jensen, 2000). These studies proposed alternative likelihood-based estimation procedures for the conventional ARFIMA processes which essentially approximate the likelihood function in terms of the Discrete Wavelet Transform (DWT) of data in which the dependence between wavelet coefficients is ignored.

The advantage of using the wavelet approach is to reduce the computational order of calculating the likelihood, which is particularly useful for large data sets. Moreover, the wavelet MLE of the long-memory parameter has been found to be fairly robust to model specification in previous empirical studies. Therefore, when the main interest is only in estimating the long-memory parameter, the wavelet MLE can be viewed as a semi-parametric estimator for the long-memory parameter.

Motivated by the variance structure of DWT under a FI process and the success of wavelet MLE for estimating the long-memory parameter, we consider a new class of time series models in the wavelet domain in which the dependence structure is fully determined by the variances and covariances of wavelet coefficients at different scales. By imposing certain constraints on the variances (which are considered

as parameters in the new model) of wavelet coefficients at large scales, the new model can exhibit long memory. The constraints only rely on the behavior of the spectral density of a long-memory process near zero frequency. In other words, the new model is fairly semi-parametric in specifying the short-memory behavior which corresponds to the behavior of the spectral density away from zero frequency.
In contrast to previous studies focusing on an alternative likelihood-based estimation for ARFIMA processes using wavelet analysis, the primary interest here is to provide an alternative class of models which exhibit long memory. From the modeling point of view, the long-memory structure of the proposed model is similar but not identical to that of a FI model, because the way of constructing long memory in the proposed model is to match the variance structure of wavelet coefficients to that obtained from a FI process only in large scales. Moreover, the short-memory structure of the proposed model, which is fully determined by the variances of wavelet coefficients in small scales and the covariance of scaling function coefficients (parameters in the model), is different from those of ARFIMA models.
From the estimation point of view, since the short-memory structure in the proposed model is fairly semi-parametric, the estimator of long-memory parameter is expected to be robust. The wavelet MLE suggested by previous studies also has this advantage so that it can be viewed as a semi-parametric estimator when data are contaminated.
In this case, the model still provides sensible predictions that are not available under the conventional approach. For inference, maximum likelihood estimation is derived for the proposed long-memory wavelet model.

In this chapter we consider ways in which the Discrete Wavelet Transform (DWT) can be used in the analysis and synthesis of both stationary long memory processes and related nonstationary processes, all of which have Spectral Density Function (SDF) that plot approximatively as straight lines with negative slopes on log-frequency/log-power axes, at least over several octave of frequency f as f approaches 0.

4.2 Fractional Difference Processes

In the early 1980s, a family of models were developed to help analyze long- memory processes. (Granger and Joyeux, 1980) and (Hosking, 1981) introduced fractional $ARIMA$ models, which are a generalization of the standard $ARIMA(p, d, q)$ models defined by (Box and Jenkins, 1976).

Let x_t be a stochastic process whose d-th order backward difference

$$(1-L)^d X_t = \sum_{k=0}^{\infty} \begin{pmatrix} d \\ k \end{pmatrix} (-1)^k X_{t-k} = \varepsilon_t \tag{4.1}$$

is a stationary process, here d is a real number and $\begin{pmatrix} a \\ b \end{pmatrix}$ is defined by

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{a!}{b!(a-b)!} = \frac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)} \tag{4.2}$$

Where $\Gamma(\cdot)$ is a gamma density. If ε_t is a white noise process with variance σ_ε^2, then X_t is the simplest case of a fractional $ARIMA$ process, a fractional $ARIMA(0, d, 0)$ or Fractional Difference Process (FDP).

Now let X_t be a zero mean FDP(d) with $-1/2 < d < 1/2$. This process is stationary and invertible (Hosking, 1981). The auto-covariance sequence (ACVS) of X_t is defined to be

$$\gamma_\tau = \mathbb{E}(X_t X_t - \tau) = \frac{\sigma_\varepsilon^2 (-1)^\tau \Gamma(1-2d)}{\Gamma(1+\tau-d)\Gamma(1-\tau-d)} \tag{4.3}$$

which means the variance is given by

$$\text{Var}(X_t) = \gamma_0 = \frac{\sigma_\varepsilon^2 \Gamma(1-2d)}{\left[\Gamma(1-d)\right]^2} \tag{4.4}$$

The spectral density function (SDF) of X_t is

$$S_X(f) = \frac{\sigma_\varepsilon^2}{|2\sin(\pi f)|^{2d}} \quad \text{for} \quad -\frac{1}{2} < f < \frac{1}{2} \tag{4.5}$$

so that $S_X(f) \propto f^{-2d}$ approximately as $f \to 0$ and, thus, the SDF is approximately linear on the log scale. The spectra, plotted on a log-log scale, are approximately linear with slope varying with the fractional difference parameter d. When $0 < d < 1/2$, this SDF has an asymptote at frequency zero[1], in which case the process exhibits slowly decaying auto-covariances and provides a simple example of a long-memory process.

4.3 Wavelet transform of a long memory process

In this Section, we introduce the wavelet transform, and look at its properties in connection with long-memory processes.

Suppose $\{x_t\}$ is a long-memory process with long-memory parameter d, so its spectral density satisfies

$$S_x(f) \sim c_S f^{-2d} \text{ for } f \text{ near zero} \tag{4.6}$$

where $0 < d < 0.5$ and $c_S > 0$.

For notational simplicity, let $T = 2^p$ for some $p \in \mathbb{N}$ in the following context. The discrete wavelet transform (DWT) of $x = (x_1, \ldots, x_T)'$ satisfies

$$\mathbf{W} = \mathcal{W}x \tag{4.7}$$

where $\mathcal{W}$ is the matrix of the discrete wavelet transform associated with the selected wavelet filter (e.g., Haar wavelet) and the selected resolution (indicated by J).

The transformed vector $\mathbf{W}$ can be decomposed as

$$\mathbf{W} \equiv (d_{-1}, d_{-2}, \ldots, d_{-J}, s_{-J})' \tag{4.8}$$

[1] An asymptote in the spectrum at frequency zero is such that as $f \to O$, $S(f) \to \infty$.

in which $J \leq p$, $d_j = \{d_{j;k} \ : k = 1, \ldots, T^{2j}\}$ comprises the wavelet coefficients at the jth scale for $j = -1, \ldots, -J$, and $s_{-J} = \{s_{-J;k} \ : k = 1, \ldots, T^{2j}\}$ comprises the scaling-function coefficients. According to (Daubechies, 1992), the jth scale $(j = -1, \ldots, J)$ wavelet filter acts as an approximate band-pass filter with octave pass-band $[-2^{j+1}\pi; -2^{-j}] \cup [2^{j}\pi; 2^{j+1}]$ and the corresponding wavelet coefficients are approximately a bandpass representation of the original data. Moreover, the wavelet filter corresponding to the scaling-function coefficients s_{-J} acts as an approximate band-pass filter with octave pass-band $[-2^{-J}\pi; 2^{-J}\pi]$. In particular, when $J = p$, s_{-J} only contains the single scaling-function coefficient that corresponds to the sample mean of the observed series.

Given a zero-mean long-memory process with the power spectrum in Table 5.1, the wavelet coefficients and the scaling-function coefficients also have zero mean, i.e., $\mathbb{E}(d_j) = 0$, for $(j = -1, \ldots, J)$ and $\mathbb{E}(s_{-j}) = 0$. Moreover, the variances of the wavelet coefficients satisfy

$$\sigma_s^2 = Var(s_{-J,k}) \approx \frac{1}{2^{-J}\pi} \int_0^{2^{-J}\pi} S(f)df \approx \frac{2^J}{\pi} \int_0^{2^{-J}\pi} c_S f^{-2d} df = \frac{c_s}{1-2d} \pi^{-2d} 2^{2dJ}$$

$$\sigma_s^2 = Var(d_{J,k}) \approx \frac{1}{2^{j+1}\pi - 2^{j}\pi} \int_{2^{j}\pi}^{2^{j+1}\pi} S(f)df \approx \frac{c_s}{1-2d} \pi^{-2d} 2^{-2dj} \left(2^{1-2d} - 1\right)$$

for relatively large scale (i.e., $j = -p, -p+1, \ldots$). Similar results are derived by (McCoy and Walden, 1996) and (Jensen, 1999) for a FI(d) process, and by (Jensen, 2000) for an ARFIMA process.

It is well known that the DWT is an excellent decorrelator (Jensen, 1998) and therefore, the wavelet coefficients are nearly uncorrelated both within and between scales. In fact, this is the key property for having computational efficiency using wavelet MLE in which the wavelet coefficients are considered independent. For justifying this independent assumption, (Dijkerman and Mazumdar, 1994) showed that the correlations of wavelet coefficients decay exponentially fast between scales and hyperbolically fast within scales for a fractionally Brownian motion (fBm). Similarly, for a Fractionally Integrated Process, (Fan, 2003) showed that the within-

scale correlations decay hyperbolically and the decay rate increases as the length of wavelet filter increases. Moreover, the between-scale correlations decrease to zero uniformly across time lags as the length of wavelet filter increases.
Based on these theoretical properties about the DWT of a long-memory process, a model in the wavelet domain in which the dependence structure is characterized by the variances of wavelet coefficients at different scales, denoted by $\{\sigma_j^2 \ : j = -1, \dots, -J\}$ and the covariance matrix of scaling function coefficients, denoted by $\sigma_s \Gamma_s$ where Γ_s is a correlation matrix. By imposing constraints on σ_j^2 at large scales satisfying $\sigma_j^2 \propto 2^{-2dj}$ which is equivalent to Table 5.1, the proposed model exhibits long memory behavior with parameter d.

Similar to previous related studies, the correlations in wavelet coefficients within and between scales are assumed to be zero for ease of inference. The scaling function coefficients in s^{-J} preserve information of the spectral density on $[-2^{-J}\pi; 2^{-J}\pi]$ which represents low-frequency variations. Compared to wavelet coefficients, the scaling function coefficients are relatively smooth and the correlations between them cannot be ignored. The correlations in scaling function coefficients is taken into account in this model by assuming an AR(1) dependence structure. This choice is somehow arbitrary and more complex correlation structure can be assumed such as a higher-order AR or even a FI(d). But, in practice, we found that AR(1) works very well to approximate the dependence in scaling function coefficients for the various situations considered in our simulations. The precise model specification is given in the next section.

4.4 Wavelet Estimation of long-memory Process

4.4.1 Ordinary Least-Squares Estimation of Fractional Difference Processes

The approximate linear relationship between periodogram and Fourier frequencies (on a log-log scale) has been known for a long time now. Geweke and Porter-Hudak (Geweke and Porter-Hudak, 1983) first proposed regressing the log values of the periodogram on the log SDF to estimate the fractional differencing parameter d (we refer to this as the GPH estimator). Although very popular, the GPH estimator suffers from the poor asymptotic properties of the periodogram. Improvements to the GPH estimator have been suggested, such as restricting the number of frequencies used in the regression or using an alternative spectral estimator (smoothed or multi-taper).

Here we introduce a wavelet-based estimator of d also using ordinary least-squares (OLS) regression (Jensen, 2000). (Tkacz, 2001) has applied this technique to nominal interest rates in the United States and Canada.
For a vector of observations Y, the OLS model is formulated via

$$Y = X\beta + \varepsilon \tag{4.9}$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{1,1} \\ 1 & x_{2,1} \\ \vdots & \vdots \\ 1 & x_{J,1} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

are the length J vector of dependent observations, the $J \times 2$ dimensional model matrix and the length 2 parameter vector, respectively. The final vector

$$\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_J)'$$

where $'$ is the transpose operator, is the column of model errors with $\mathbb{E}(\varepsilon) = 0$ and $Var(\varepsilon) = \sigma_\varepsilon^2 \mathbf{I}_J$. The OLS estimator $\widehat{\beta}$ of β is

$$\widehat{\beta} = (X'X)^{-1} X'Y \tag{4.10}$$

and the covariance matrix of $\widehat{\beta}$ is given by

$$\Sigma_{\widehat{\beta}} = \sigma_\varepsilon^2 (X'X)^{-1} \tag{4.11}$$

For the remainder of this section, we are interested in estimating the slope parameter β_1 when the dependent observations are $y_j = log(\sigma_x^2(\tau_j))$ and the independent observations are $x_{j,1} = log(\tau_j)$ for $j = 1...., J$. Where $\sigma_x^2(\tau_j)$ is the wavelet variance of x at scale τ_j.

There is a clear linear relationship between the wavelet variance $\sigma_x^2(\tau_j))$ and scaleτ_j (on a log-log scale) for FDPs, as Figure 5.4 shows. (Jensen, 1999) proved that $\sigma_x^2(\tau_j) \to c\tau_j^{2d-1}$ as $j \to \infty$ for FDPs with $-1/2 < d < 1/2$. A reasonable regression model is therefore:

$$\ln(\sigma_x^2(\tau_j)) = \beta_0 + \beta_1 \ln(\tau_j) + \varepsilon_j \tag{4.12}$$

where $\beta_1 = 2d - 1$. Suppose that $Y_j = \ln(\sigma_x^2(\tau_j))$ and $x_j = \tau_j$ for $j = 1, \ldots, J$, the OLS estimators of β_1 is given by

$$\widehat{\beta}_1 = \frac{\sum_{j=1}^{J} (x_J - \overline{x}) Y_j}{\sum_{j=1}^{J} (x_J - \overline{x})^2} \tag{4.13}$$

where $\overline{Y}$ is the simple mean of Y. Hence, the OLS estimator of the fractional difference parameter is $\widehat{d} = (\widehat{\beta}_1 + 1)/2$. To determine the variance of our OLS estimator, we take the lower-right corner of $\Sigma_{\widehat{\beta}}$ in Equation 4.11; that is

$$Var(\widehat{\beta}_1) = \frac{\sigma_\varepsilon^2}{\sum_{j=1}^{J}(x_j - \overline{x})^2} \tag{4.14}$$

where the estimated variance of the model errors is given by (most easily expressed in matrix notation)

$$\sigma_\varepsilon^2 = \frac{(Y - X\widehat{\beta})'(Y - X\widehat{\beta})}{J - 2} \tag{4.15}$$

Basic properties of the variance tell us that $Var(\widehat{d}) = \frac{1}{4}Var(\widehat{\beta}_1)$.

To reduce the Mean Squared Error (MSE) of $\widehat{d}$, several authors have proposed a Weighted Least Squares (WLS) estimator of d instead of the OLS estimator (Abry et al., 1995, Abry and Veitch, 1998). The WLS procedure allows the model errors to have different variances, but they must still be uncorrelated. The estimation of $\widehat{d}_{wls}$ follows from the OLS procedure by simply introducing a diagonal matrix $(J \times J)$ of weights $\mathcal{C}$ such that:

$$\mathcal{C} = \text{diag}(\sigma_1^2, \ldots, \sigma_J^2) \tag{4.16}$$

The WLS estimator of β is given by

$$\widehat{\beta}_{\mathbf{WLS}} = (X'\mathcal{C}^{-1}X)^{-1}X'\mathcal{C}^{-1}Y \tag{4.17}$$

and the covariance matrix of $\widehat{\beta}_{\mathbf{WLS}}$ is given by $\Sigma_{\widehat{\beta}_{\mathbf{WLS}}} = (X'\mathcal{C}^{-1}X)^{-1}$. As before the WLS estimator of the fractional difference parameter is $d_{wls} = (\widehat{\beta}_{l,wls} + 1)/2$ and $Var(\widehat{d}_{wls}) = \frac{1}{4}Var(\beta_{l,wls})$. (Percival and Walden, 2000) showed that the WLS estimator reduces the MSE by a factor of two in comparison to the OLS estimator in simulation studies.

The possibility that the long-memory parameter d is not constant over time is an interesting generalization of the usual FDP.

(Veitch and Abry, 1999) developed a testing procedure for the time constancy of

d while Whitcher and Jensen (Whitcher and Jensen, 2000) proposed an OLS estimator for a nonstationary FDP, and Jensen and Whitcher (Jensen and Whitcher, 2000) applied the OLS-based estimator to a year of high-frequency foreign exchange rates. Parameter estimation for a nonstationary long memory time series model, through OLS or maximum likelihood, is in its infancy and should benefit greatly from wavelet-based methods.

4.4.2 Approximate Maximum Likelihood Estimation of Fractional Difference Processes

Wavelet-based Maximum Likelihood Estimation procedures have been investigated by McCoy and Walden (McCoy and Walden, 1996) and Jensen (Jensen, 1998, 2000). Although least squares estimation is popular because of its simplicity to program and compute, it produces much larger mean square errors when compared to maximum likelihood methods. The methodology presented here overcomes the difficulty of computing the exact likelihood by replacing the covariance matrix of the process with an approximation using the DWT. This is possible through the ability of the DWT to decorrelate long-memory processes.

If x is a length $N = 2^J$ FDP with mean zero and covariance matrix given by Σ_x, then we may write its likelihood as

$$L(d, \sigma_\varepsilon^2) = (2\pi)^{-N/2} \Big|\Sigma_x\Big|^{-1/2} \exp\Big[-\frac{1}{2} X' \Sigma_x^{-1} X \Big] \tag{4.18}$$

The quantity $\Big|\Sigma_x\Big|$ is the determinant of Σ_x. The maximum likelihood estimators (MLEs) of the parameters (d and σ_ε^2) are those quantities that maximize Equation 4.18. We now avoid the difficulties in computing the exact MLEs by using the approximate decorrelation of the DWT as applied to FDPs; that is,

$$\Sigma_x \approx \widehat{\Sigma}_x = \mathcal{W}' \Omega_N \mathcal{W} \tag{4.19}$$

where $\mathcal{W}$ is the orthonormal matrix defining the DWT (subsection 2.2.2) and Ω_N is a diagonal matrix containing the variances of DWT coefficients computed from FDPs; that is,

$$\Omega_N = \text{diag}\,(\underbrace{S_1, S_1, \ldots, S_1}_{N/2}, \underbrace{S_2, S_2, \ldots, S_2}_{N/4}, \ldots, \underbrace{S_j, S_j, \ldots, S_j}_{N/j}, \ldots, S_J, S_{J+1}) \quad (4.20)$$

where

$$S_j = 2^{j+1} \int_{1/2^{j+1}}^{1/2^j} \frac{\sigma_\varepsilon^2}{|2\sin(\pi f)|^2 d}\, df \;\; j = 1, 2, \ldots, J$$

and

$$S_{j+1} = 2^J \left(\frac{\sigma_\varepsilon^2 \Gamma(1-2d)}{\Gamma^2(1-d)} - \sum_{j=1}^{J} \frac{S_j}{2^j} \right)$$

The approximate likelihood function is now

$$\widehat{L}(d, \sigma_\varepsilon^2) = (2\pi)^{-N/2} \left| \widehat{\Sigma}_x \right|^{-1/2} \exp\left[-\frac{1}{2} X' \widehat{\Sigma}_x^{-1} X \right] \quad (4.21)$$

Hence, we want to find values of d and σ_ε^2 that minimize the log-likelihood function

$$\begin{aligned} \widehat{\mathcal{L}}(d, \sigma_\varepsilon^2) &= -2\ln\left(\widehat{L}(d, \sigma_\varepsilon^2)\right) - N\ln(2\pi) \\ &= \ln\left(\left|\widehat{\Sigma}_x\right|\right) + X'\widehat{\Sigma}_x^{-1} X \end{aligned} \quad (4.22)$$

where the innovations' variance is formerly estimated:

$$\widehat{\sigma}_\varepsilon^2 = \frac{1}{N} \left(\frac{\mathbf{V}_j' \mathbf{V}_j}{S_{J+1}'(d)} + \sum_{j=1}^{J} \frac{\mathbf{W}_j' \mathbf{W}_j}{S_J'(d)} \right)$$

here $S_J'(d) = \dfrac{S_j(d, \sigma_\varepsilon^2)}{S_\varepsilon^2} \quad j = 1, 2, \ldots, J.$

Minimizing numerically $\widehat{\mathcal{L}}(d, \sigma_\varepsilon^2)$ in this case over the parameter space $]0, 1/2[$[2] yields approximate MLE for the fractional differencing parameter d.

[2] Although d can vary in $]-1/2, 1/2[$, in practice, we will focus primarily on fractional difference parameters that are strictly positive.

4.5 Application to Stock Market Index:

4.5.1 Data description and statistical properties:

The data sets consists of daily closing price of six stock market index. There are respectively S&P500, CAC40, Nasdaq, Nekkei, Dow Jones and DJ Eurostoxx. These indices are used from Yahoo finance website.

The sample period spans from 15/04/2002 to 22/04/2002 for a total of 1770 daily observations[3]. We calculate the daily return, r_t,of the series as follows:

$$r_t = \ln(p_t - p_{t-1})$$

where p_t denotes the value of the index on day t. We focus on squared daily returns, r_t^2, as proxies for the volatility of the series. The series index, returns and squared returns are displayed in Figure 4.1, Figure 4.2 and Figure 4.3 respectively. The Figure 4.4 display the autocorrelation function of the squared returns.

Table 4.1 reports summary statistics for the daily returns over the sample period. The sample mean returns is null for all series. All analyzed series display significant kurtosis values greater than 3.

This suggests that fat-tailed leptokurtic distributions may characterize the returns of the series of study. Focusing on the skewness coefficients' estimates, we see that the SP500 and Nikkei index returns are slightly negatively skewed as compared to the Gaussian distribution whereas the rest series returns present positive skewness statistics. There are significant departures for normality as the series is skewed and leptokurtic.

[3]It is remarkable to note that the full data samples were trimmed so that the effective number of observations for the volatility series in hand is a power of 2. This is required in order to implement the wavelet OLS estimator

	SP500	CAC40	DJIA	Nasdaq	Nikkei	DJ.EuroStox
min	-0.095	-0.095	-0.082	-0.096	-0.121	-0.082
max	0.110	0.106	0.105	0.112	0.132	0.121
median	0.001	0.000	0.000	0.001	0.000	0.000
mean	0.000	0.000	0.000	0.000	0.000	0.000
std.dev	0.014	0.016	0.013	0.016	0.017	0.016
skewness	-0.118	0.107	0.157	0.027	-0.403	0.128
kurtosis	9.392	6.049	9.018	5.005	7.799	6.337
normtest.W	0.881	0.913	0.891	0.942	0.925	0.905
p-value	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.1: Summary statistics of Stock Market returns

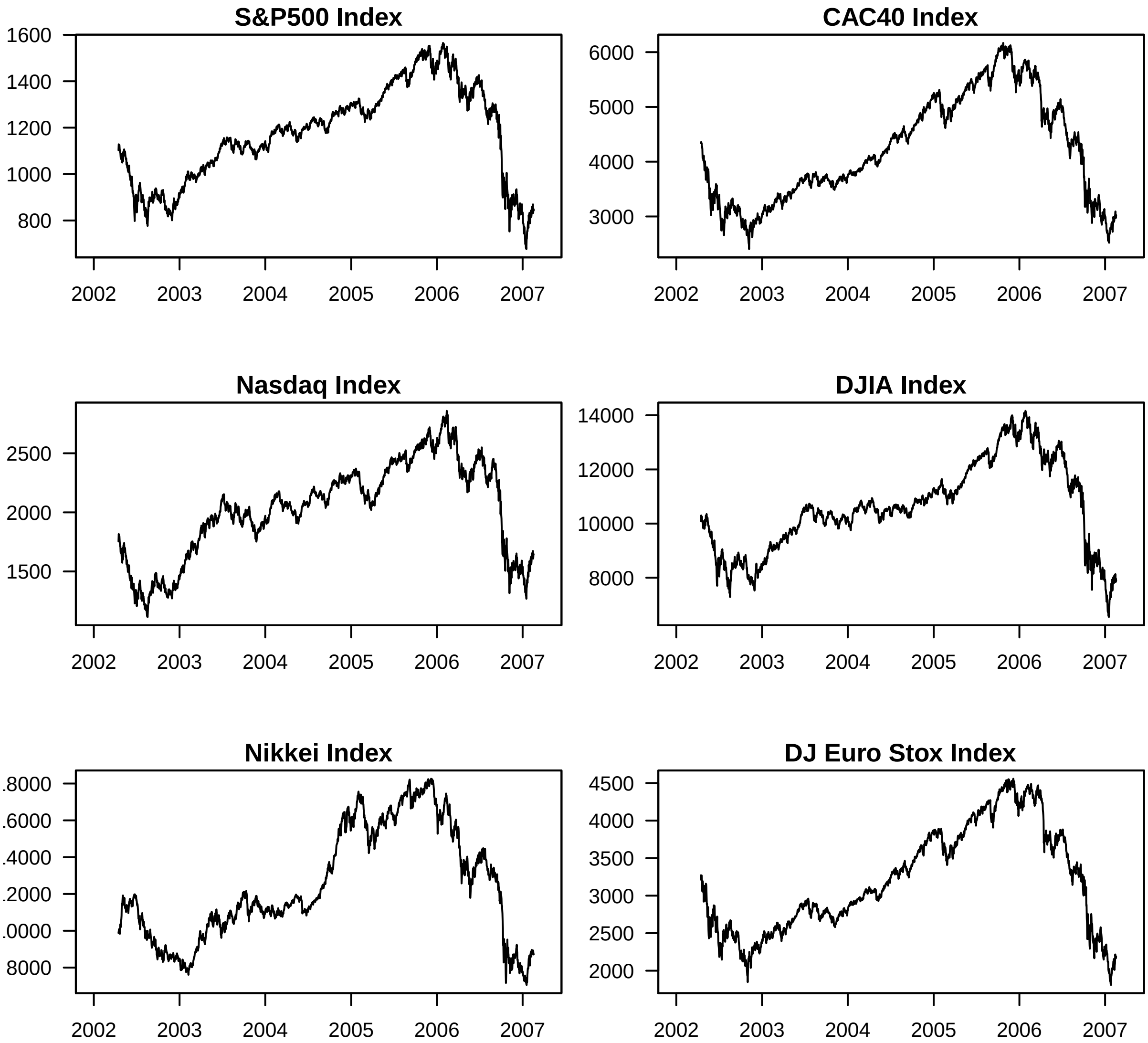

Figure 4.1: Stock Market Index Series

4.6 Empirical results:

In this section, we apply a wavelet estimator to our data.
As described in (Abry and Veitch, 1998), wavelet estimator is robust against additional trends. Unfortunately this robustness depends on the chosen underlying wavelet. We choose Haar wavelet and Daubechies wavelets of the order four and height to show this dependence.
We estimate the fractional differencing parameter d by three methods: Maximum Likelihood estimator (ML), GPH estimator and wavelet estimator (the Weighted

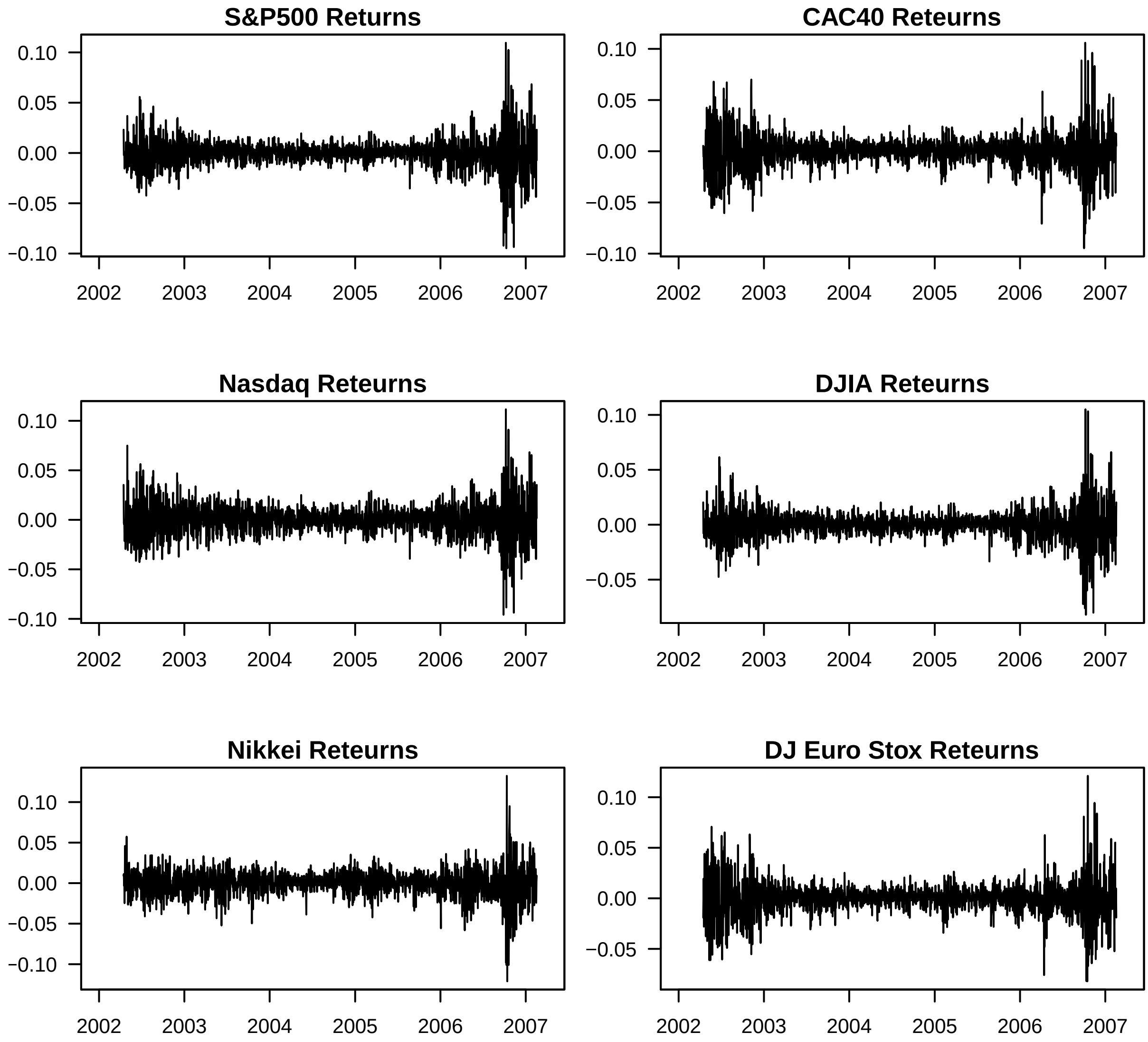

Figure 4.2: Stock Market Return Series

Least Squares Estimation (WLSE), the Maximum Likelihood Estimation (MLE) and the Ordinary Least Squares estimation (OLS) described previously). To ensure the robustness of our results, we consider a wide range of wavelet filters, namely the Haar wavelet, the Daubechies extremal phase filters $D(L)$ for $L = 4, 8$ and the Daubechies least asymmetric filters $LA(L)$ for $L = 4, 8$.

For the GPH estimator, we estimate the fractional differencing parameter with

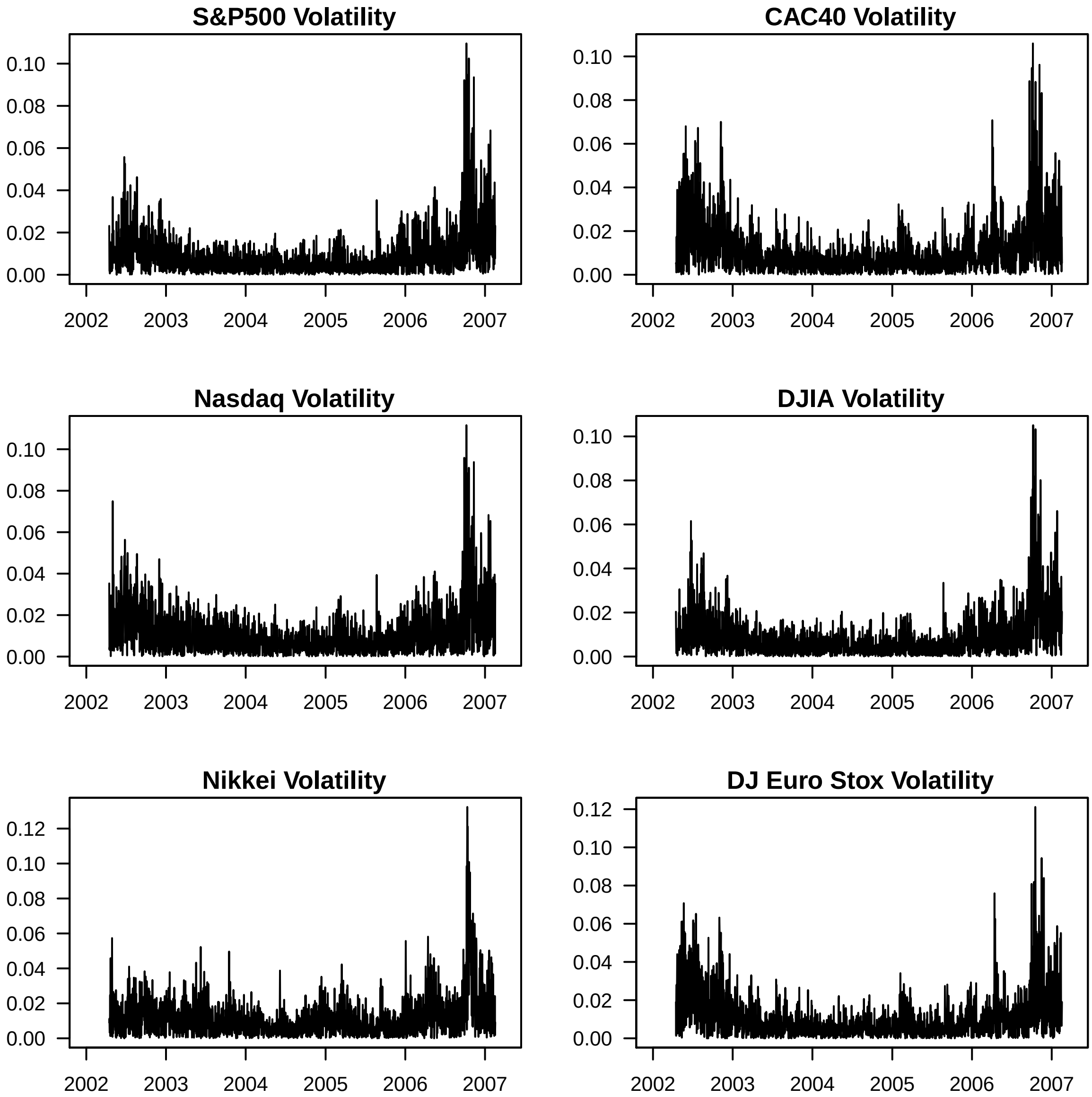

Figure 4.3: Stock Market Volatility Series

estimation windows of $T^{0.85}$. The null hypothesis of interest is whether the volatility series is integrated of order zero ($H_0 : d = 0$), versus the alternative of fractional integration ($H_1 : d \neq 0$). Estimates for the fractional integration parameter d along with absolute asymptotic t-statistics (in parentheses) for the null hypothesis $d = 0$, are shown in Table 4.2. According to the GPH results, the estimated fractional differencing parameters for all stock market volatility are significant at the 5% level pointing ***stationary long memory*** behavior with $0 < d < 0.5$.

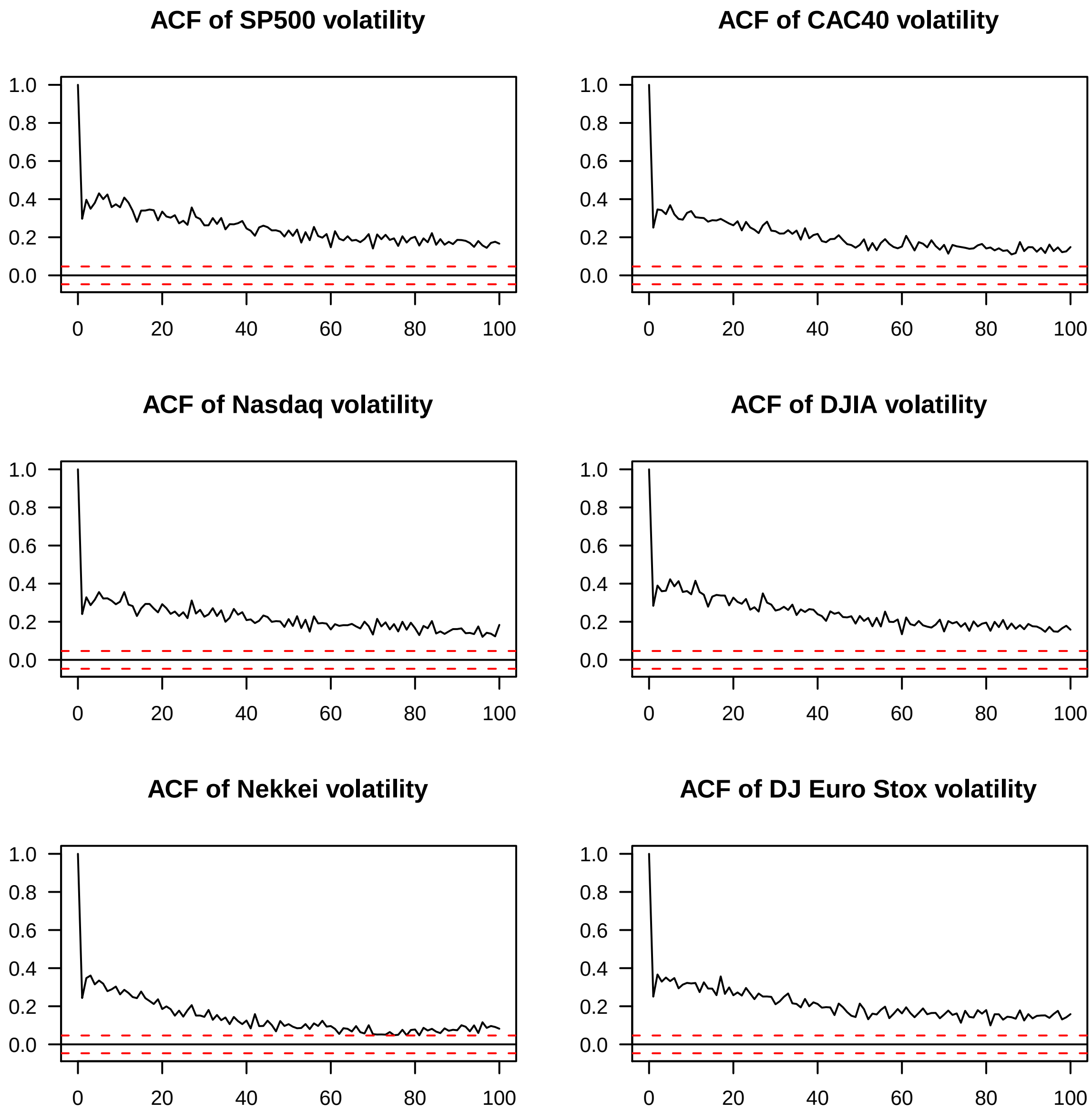

Figure 4.4: Autocorrelation of Volatility Series

	SP500	CAC40	DJIA	Nasdaq	Nikkei	DJ.EuroStox
MLE	0.225382	0.211825	0.221479	0.204983	0.215173	0.212984
GPH	0.202732	0.193055	0.188158	0.175591	0.291831	0.182148
	(7.108634)	(6.769296)	(6.597614)	(6.156960)	(10.232796)	(6.386859)

Table 4.2: GPH estimator of the differencing parameter. The value in parentheses are the t-statistics for the null hypothesis $d = 0$

We apply in the following a wavelet estimator as described previously to the six volatilities of stock returns.

The results of wavelet based estimator are given in Table 4.3.

As we see the wavelet based estimator gives evidence of long memory in the volatilities of the stocks. Because the estimates are in the same range as with the Maximum Likelihood method and the GPH method we have evidence that the observed dependencies are real and not spurious long memory.

4.7 Conclusion

Absolute daily returns of six stocks are considered.

All of them show evidence of long memory by using standard methodology. We prove whether is an artefact of trends or structural breaks or if there is real evidence of long memory. This is done by applying three methods for distinguishing trends and long memory. At first we used a maximum likelihood estimator. Afterwards a log-periodogram based methodology is used. At last we estimated the memory parameter with a wavelet estimator which is robust against trends. Absolute daily returns of SP500, CAC40, DJIA, Nasdaq, Nikkei, and DJ.Eurostox are considered. The ML based estimator gives values clearly greater zero for all data. For all of the considered stocks the standard GPH-estimator estimate similar values for the memory parameter.

		Haar	D(4)	D(8)	LA(4)	LA(8)
WLSE	SP500	0.2552534	0.1199371	-0.0603214	0.1199371	-0.0616800
	CAC40	0.2308362	0.1472446	0.0521677	0.1472446	0.0546592
	DJIA	0.2566247	0.1289205	-0.0487493	0.1289205	-0.0452805
	Nasdaq	0.2065951	0.0915240	-0.0520755	0.0915240	-0.0531718
	Nikkei	0.3047232	0.2538280	0.1517089	0.2538280	0.1447832
	DJ.Eurostox	0.2298510	0.1442468	0.0467442	0.1442468	0.0547381
MLE	SP500	0.3059438	0.1473141	-0.0247492	0.1473141	-0.0861251
	CAC40	0.2986800	0.2142526	0.0745283	0.2142526	0.0386858
	DJIA	0.3142016	0.1794389	-0.0285518	0.1794389	-0.0623330
	Nasdaq	0.2721162	0.1123735	-0.0090294	0.1123735	-0.1037307
	Nikkei	0.3336335	0.3211967	0.1776601	0.3211967	0.1596386
	DJ.Eurostox	0.3145928	0.1799988	0.1056316	0.1799988	0.0669937
OLS	SP500	0.2965362	0.0098224	-0.1149412	0.0443466	-0.0891698
	CAC40	0.2558679	0.0802783	0.1086934	0.1463964	0.1038756
	DJIA	0.2989605	-0.0261508	-0.1318246	0.0380215	-0.0926101
	Nasdaq	0.2902089	0.0092969	-0.1483582	-0.0318254	-0.1370387
	Nikkei	0.3791740	0.1419149	0.0384017	0.2012580	-0.0176337
	DJ.Eurostox	0.1795479	0.0716379	-0.0020495	0.1956864	0.0465343

Table 4.3: Wavelet estimator of the differencing parameter using two wavelet filter families (Haar and Daubechies).

CHAPTER 5

Causality between trade and economic growth: A wavelet filtering based analysis

5.1 Introduction

In this chapter, we use a wavelet filtering based approach to study the econometric relationship between exports, imports and Gross Domestic Product (GDP). We explore the interactions between these primary macroeconomic indicators in a co-integrating and vector autoregression framework and their dynamic causality under the Granger-Causality setup. The much studied relationship between these three primary indicators of the economy is explored with the help of the wavelet multi-resolution filtering technique. Instead of an analysis at the original series level, as is usually done, we first decompose the variables using wavelet decomposition technique at various scales of resolution and obtain relationship among components of the decomposed series matched to its scale. The analysis reveals interesting aspects of the interrelationship among the three fundamental macroeconomic variables.

The rest of the chapter is organized as following. In the following section we briefly describe the relationship between exports (imports) and GDP, while section 5.3 deals with the application of appropriate econometric tools on the wavelet filtered series. Finally, results of empirical study are presented and discussed in section 5.4.

5.2 Literature revue

The role of export to improve the growth potential of a country occupies the center stage in especially development literature where export promotion and increased openness gradually have replaced import substitution to enhance growth. This shift from import substitution to export promotion and increased openness implies as well a shift in the trade and industry policy from being highly import substituting and government controlled to become more liberalized and deregulated. This shift in policies has also been central in policy recommendations to developing countries concerning improvements of their growth potential. An increased openness to trade will enhance competition for firms producing for the international market. Such an environment generates incentives for an increased productivity and incentives for innovations as well as the possibility to pay higher wages in line with the increased productivity. Furthermore, an increased openness to trade is also central in international negotiations about trade and tariff barriers where trade theory suggests that all parties on aggregate will enhance their welfare position in relation to their autarky situation. An Export-Led Growth (ELG) hypothesis which states that exports are keys to promoting economic growth provides one of the answers to this fundamental question. There is a considerable literature that investigates the link and causation between exports and economic growth, but the conclusions still remain a subject of debate.

A number of empirical studies have documented a strong and positive relationship between export and economic growth including (Balassa, 1978), (Jung and Marshall, 1985), (Chow, 1987), (Ahmad, 1991), and (Moosa, 1999) among others. The results reveal evidence in support of the export-led growth hypothesis for various countries. Several studies have also shown that it is possible to have Growth-Led Exports (GLE) which has the reverse causal flow from economic growth to exports growth. In the GLE case, export expansion could be stimulated by productivity gains caused by increases in domestic levels of skilled-labor and technology (Bhagwati, 1988), (Krugman, 1984). The third alternative is that of Import-Led Growth (ILG) which suggests that economic growth could be driven primarily by growth in imports. Endogenous growth models show that imports can be a channel for long-run economic growth because it provides domestic firms with access to needed intermediate factors and foreign technology (Coe and Helpman, 1995). Growth in imports can serve as a medium for the transfer of growth-enhancing foreign R&D knowledge from developed to developing countries (Joy, 2001). In this chapter, we use the wavelet based filtering technique for exploring the long-run equilibrium and other relationship between the three important macroeconomic indicators, exports, imports and economic growth for Tunisia.

5.3 Econometric tools

In this section, we discuss the appropriate econometric tools can be used in the context of the problem of studying interrelationships among the exports, imports and GDP indicators.

5.3.1 Unit root test

A unit root test tests whether a time series variable is non-stationary using an autoregressive model. The most famous test is the Augmented Dickey-Fuller test (Dickey and Fuller, 1979). Another test is the Phillips-Perron test (Phillips and Perron, 1988). Both these tests use the existence of a unit root as the null hypothesis. Stationarity tests, on the other hand, are for the null that y_t is $I(0)$. The most commonly used stationarity test, the KPSS test, is due to Kwiatkowski, Phillips, Schmidt and Shin (1992) (Kwiatkowski et al., 1992).

4.3.1.1 Augmented Dickey Fuller test

The testing for unit root is formulated in the statistical hypothesis testing framework as a test of the null hypothesis H_0: series is non stationary, $I(1)$, against the alternative H_1: series is stationary, $I(0)$.

For general Autoregressive (AR) models the test is given by Dickey and Fuller (Dickey and Fuller, 1979). The test regresses the first difference of y_t, i.e., $\Delta y_t = y_t - y_{t-1}$ on the lag level y_{t-1}, an intercept (or possible trend) and enough lagged values of Δy_t, referred to as the ***augmenting lags***. This regression is defined as follow:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_t. \tag{5.1}$$

where α is a constant, β the coefficient on a time trend and p the lag order of the autoregressive process. Imposing the constraints $\alpha = 0$ and $\beta = 0$ corresponds to modeling a random walk and using the constraint $\beta = 0$ corresponds to modeling a random walk with a drift. Depending upon the structure of regression function, there are three cases of ADF, which are presented below:

i) Zero mean ADF:

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_t. \tag{5.2}$$

ii) Single mean ADF:

$$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_t. \tag{5.3}$$

iii) Time trend ADF:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_t. \tag{5.4}$$

In each of the above cases, the null hypothesis (H_0) of unit root boils down to testing the null $H_0 : \gamma = 0$ against the alternative hypothesis of $\gamma < 0$. The ADF t-statistic and normalized bias statistic are based on the least squares estimates of Equation 5.1 and are given by:

$$ADF_t = t_{\gamma=0} = \frac{\widehat{\gamma}}{SE(\widehat{\gamma})}$$

$$ADF_\tau = \frac{T\widehat{\gamma}}{1 - \sum_{i=1}^{p} \delta_i}$$

Where ADF_τ is computed it can be compared to the relevant critical value for the Dickey-Fuller Test. If the test statistic is greater (in absolute value) than the critical value, then the null hypothesis of $\gamma = 0$ is rejected and no unit root is present. The ADF t-statistic is then the usual t-statistic for testing $\gamma = 0$.

4.3.1.2 Phillips-Perron Test

An alternative to parametric setup, a non parametric test procedure, called the Phillips Perron (PP) test (**Phillips and Perron, 1988**) is also proposed in the literature. Asymptotic distribution of Phillips and Perron test is same as the ADF test. (**Phillips and Perron, 1988**) developed a number of unit root tests that have become popular in the analysis of financial time series. The Phillips-Perron (PP) unit root tests differ from the ADF tests mainly in how they deal with serial correlation and heteroscedasticity in the errors. In particular, where the ADF tests use a parametric auto-regression to approximate the ARMA structure of the errors in the test regression, the PP tests ignore any serial correlation in the test regression. The test regression for the PP tests is

$$\Delta y_t = \beta' \mathbf{D}_t + \gamma y_{t-1} + \varepsilon_t. \tag{5.5}$$

where D_t is a vector of deterministic terms (constant, trend etc.). The PP tests correct for any serial correlation and heteroscedasticity in the errors $u\varepsilon_t$ of the test regression by directly modifying the test statistics $t_{\gamma=0}$ and τ. These modified statistics, denoted Z_t and Z_τ, are given by

$$Z_t = \Big(\frac{\widehat{\sigma}^2}{\widehat{\lambda}^2}\Big)^{1/2} \times t_{\gamma=0} - \frac{1}{2}\Big(\frac{\widehat{\lambda}^2 - \widehat{\sigma}^2}{\widehat{\lambda}^2}\Big) \times \Big(\frac{T \times SE(\widehat{\gamma})}{\widehat{\sigma}^2}\Big) \tag{5.6}$$

$$Z_\tau = T\widehat{\gamma} - \frac{1}{2}\frac{T^2 SE(\widehat{\gamma})}{\widehat{\sigma}^2}(\widehat{\lambda}^2 - \widehat{\sigma}^2) \tag{5.7}$$

The terms $\widehat{\sigma}^2$ and $\widehat{\lambda}^2$ are consistent estimates of the variance parameters

$$\sigma^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E(\varepsilon_t^2)$$

$$\lambda^2 = \lim_{T\to\infty} \sum_{t=1}^{T} E(T^{-1} S_T^2)$$

where $S_T^2 = \sum_{t=1}^{T} \varepsilon_t$. Under the null hypothesis that $\gamma = 0$, the PP Z_t and Z_τ statistics have the same asymptotic distributions as the ADF t-statistic and normalized bias statistics. One advantage of the PP tests over the ADF tests is that the PP tests are robust to general forms of heteroscedasticity in the error term ε_t. Another advantage is that the user does not have to specify a lag length for the test regression.

4.3.1.3 KPSS test

The ADF and PP unit root tests are for the null hypothesis that a time series y_t is $I(1)$. Stationarity tests, on the other hand, are for the null that y_t is $I(0)$. The most commonly used stationarity test, the KPSS test, is due to Kwiatkowski, Phillips, Schmidt and Shin (Kwiatkowski et al., 1992). They derive their test by starting with the model

$$y_t = \beta' \mathbf{D}_t + \mu_t + \varepsilon_t \tag{5.8}$$

$$\mu_t = \mu_{t-1} + u_t, \quad u_t \sim WN(0, \sigma_u^2) \tag{5.9}$$

where D_t contains deterministic components (constant or constant plus time trend), ε_t is $I(0)$ and may be heteroscedastic. Notice that μ_t is a pure random walk with innovation variance σ_u^2 . The null hypothesis that y_t is $I(0)$ is formulated as $H0 : \sigma_u^2 = 0$, which implies that μ_t is a constant. Although not directly apparent, this null hypothesis also implies a unit moving average root in the ARMA representation of Δy_t. The KPSS test statistic is the Lagrange multiplier (LM) or score statistic for testing $\sigma_u^2 = 0$ against the alternative that $\sigma_u^2 > 0$ and is given by

$$KPSS = \frac{1}{T^2 \widehat{\lambda}^2} \sum_{t=1}^{T} \widehat{S}_t^2 \tag{5.10}$$

where $\widehat{S}_t = \sum_{j=1}^{t} \widehat{\varepsilon}_j$, $\widehat{\varepsilon}_t$ is the residual of a regression of y_t on $\mathbf{D}_t$ and $\widehat{\lambda}^2$ is a consistent estimate of the long-run variance of ε_t using $\widehat{\varepsilon}_t$. Under the null that y_t is $I(0)$, Kwiatkowski, Phillips, Schmidt and Shin show that $KPSS$ converges to a function of standard Brownian motion that depends on the form of the deterministic terms $\mathbf{D}_t$ but not their coefficient values β.

5.3.2 Cointegration analysis

The second stage involves testing for the existence of a long-run equilibrium relationship between real exports, real imports and GDP within a multivariate framework.

Further, we explore a co-integrating relationship among the variables. We first consider co-integration testing in a univariate time series setup. If a time series Y_t, with no deterministic component, can be represented by a stationary and invertible ARMA process after differencing d times, the series is said to be integrated of order d, i.e.,$Y_t \sim I(d)$. Furthermore, if all elements of a vector time series Y_t are $I(d)$ and there exists a vector $\beta \neq 0$ such that $\beta^T Y_t \sim I(d-b)$ for any $b > 0$, the vector process is said to be co-integrated CI(d, b), with β as the co-integrating vector. A special case of $b = d = 1$ is of importance in analysis of economic time series, as this implies long run equilibrium (stationary) relationship among the variables involved in co-integration. We can infer that in such a situation, in the long run, the $I(1)$ variables are "tied together" even though they might drift apart is the short run.

The standard procedure for testing co-integration is through Engel-Granger test (Granger, 1987). The test is a two-step procedure where if $X_1, X_2, \ldots, X_k$ are k $I(1)$ variables. Then first we find the OLS regression of say, X_1 on $(X_2, \ldots, X_k)$, i.e.,

$$X_{1t} = \alpha + \beta_1 X_{2t} + \ldots + \beta_{k-1} X_{kt} + \varepsilon_t$$

Next we apply stationarity test, like the ADF test on the estimated residuals and infer that $(X_1, X_2, ..., X_k)$ is a set of co-integrated variables if and only if the estimated residuals are stationary. The co-integrating vector in such a situation is $(1, -\beta_1, -\beta_2, \ldots, -\beta_{k-1})$. The long run equilibrium relationship between the variables being $X_1 = \alpha + \beta_1 X_2 + \ldots + \beta_{k-1} X_k$.

Existence of con-integration can also be tested under VAR setup. The VAR based con-integration testing is known as the Johansen's tests. Under the VAR setup, we consider the vector autoregressive formulation with stationary errors

$$Y_t = \mu + \sum_{i=1}^{p} \phi_i Y_{i-1} + \varepsilon_t$$

The first difference formulation of the above model is

$$\Delta Y_t = \mu + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + \Pi Y_{t-p} + \varepsilon_t$$

where

$$\Gamma_i = (\phi_1 + \phi_2 + \ldots + \phi_i) - I_k \,; \quad \text{for } i = 1, 2, \ldots, p-1$$

and

$$\Pi = (\phi_1 + \phi_2 + \ldots + \phi_i) - I_k$$

The matrix Π contains information on possible co-integrating relations between k elements of Y_t. Rank(Π) equals the number of independent co-integrating vectors of the system. Number of distinct co-integrating vectors can thus be obtained by checking the significance of the characteristic roots of Π. (Johansen, 1988) uses maximum likelihood based approach for testing the number of characteristic roots that are significantly different from zero.

Johansen's "Trace Test" procedure sequentially tests the following hypotheses:

$H_0^0 : \; r = 0$ vs $H_A^0 : r \; \geq 0$

$H_0^1 : \; r \leq 1$ vs $H_A^1 : r \; \geq 2$

$\vdots$

$H_0^{k-1} : \; r \leq k-1$ vs $H_A^{k-1} : r \; = k$

where r denotes the number of co-integrating vectors in the system.

Johansen's "Trace Statistics", for testing the rth hypothesis is given by

$$\lambda_{\text{trace}}(r_0) = -T \sum_{i=r_0+1}^{k} \log(1 - \widehat{\lambda}_i) \tag{5.11}$$

where T is the sample size and $\widehat{\lambda}_i$ are the estimated eigenvalues of the matrix Π. If rank(Π) $= r_0$ then $\widehat{\lambda}_{r_0+1}, \ldots, \widehat{\lambda}_k$ should all be close to zero and $\lambda_{\text{trace}}(r_0)$ should be small. In contrast, if rank(Π) $> r_0$ then some of $\widehat{\lambda}_{r_0+1}, \ldots, \widehat{\lambda}_k$ will be nonzero (but less than 1) and $\lambda_{\text{trace}}(r_0)$) should be large. The asymptotic null distribution of $\lambda_{\text{trace}}(r_0)$ is not chi-square but instead is a multivariate version of the Dickey-Fuller unit root distribution which depends on the dimension $n - r_0$ and the specification of the deterministic terms.

Then the "maximum eigenvalue" test statistic for testing the rth hypothesis is given by

$$\lambda_{\max}(r_0) = -T\log(1-\widehat{\lambda}_{r_0+1}) \tag{5.12}$$

As with the trace statistic, the asymptotic null distribution of $\lambda_{\max}(r_0)$ is not chi-square but instead is a complicated function of Brownian motion, which depends on the dimension $n - r_0$ and the specification of the deterministic terms.

5.3.3 Causality relationships

Causality denotes a necessary relationship between one event (called cause) and another event (called effect) which is the direct consequence of the first. In others words, whether one variable can help forecast another variable or not.

One way to address this question was proposed by (Granger, 1969) and popularized by Sims (Sims, 1972). Testing causality, in the Granger sense, involves using F-tests to test whether lagged information on a variable Y provides any statistically significant information about a variable X in the presence of lagged X. If not, then "Y does not Granger-cause X". Here, assuming a particular autoregressive lag length p, we estimate the following unrestricted equation by ordinary least squares (OLS):

$$X_t = \mu + \sum_{i=1}^{p}\alpha_i X_{t-i} + \sum_{i=1}^{p}\beta_i Y_{t-i} + u_t \tag{5.13}$$

The null hypothesis under this setup of causality testing is framed as

H_0 : Y does not Granger-cause X.

In terms of model (5.13), we are interested in testing the following hypothesis

H_0' : $\beta_1 = \beta_2 = \ldots = \beta_p = 0$.

This is the simultaneous testing of a subset of regression parameters and can be tested using usual F-statistic. It is worth noting that with lagged dependent variables, as in Granger-causality regressions, the test is valid only asymptotically. The test procedure can be extended to cover the causality situation involving groups of variables.

Next we discuss testing Granger-Causality under the VAR setup. Let a k component vector time series process Y_t be re-arranged and partitioned in subgroups Y_{1t} and Y_{2t} with dimensions k_1 and k_2, respectively ($k = k_1 + k_2$); i.e., $Y_t = (Y_{1t}^T Y_{2t}^T)$ (where Y^T denote the transpose of Y_t) with the corresponding white noise process $\varepsilon_t = (\varepsilon_{1t}^T \varepsilon_{2t}^T)$.
Consider the VAR(p) for Y_t, i.e.,

$$Y_t - \alpha = \sum_{i=1}^{p} \phi_i (Y_{t-i} - \alpha) + \varepsilon_t \qquad \text{or} \qquad \Phi(B)(Y_t - \alpha) = \varepsilon_t \tag{5.14}$$

where

$$\Phi(B) = I_k - \phi_1 B - \ldots - \phi_p B^p$$

Realize that, in terms of the partition of Y_t, we can write the $VAR(p)$ model for Y_t, with partitioned coefficients matrices as $\Phi_{ij}(B)$ for $i, j = 1, 2$ as follows:

$$\Phi(B) Y_t = \begin{pmatrix} \Phi_{11}(B) & \Phi_{12}(B) \\ \Phi_{21}(B) & \Phi_{22}(B) \end{pmatrix} \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \delta + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{5.15}$$

In the VAR setup, the variables in Y_{1t} are said to "Granger Cause" Y_{2t}, but Y_{2t} not to "Granger Cause" Y_{1t} only if $\Phi_{12}(B) = 0$. The implication of this model structure is that future values of the process Y_{1t} are influenced only by its own past and not by the past of Y_{2t}, whereas future values of Y_{2t} are influenced by the past of both Y_{1t} and Y_{2t}. On the other hand, if the future of Y_{1t} is not influenced by the past values of Y_{2t}, then it is more logical to model Y_{1t} separately from Y_{2t}.

5.4 Empirical analysis

5.4.1 Data description and descriptive statistic

For the present study, we have considered Tunisian macroeconomic time series data on exports, imports and GDP index. The data set,obtained from the National Statistics Institute of Tunisia, is quarterly and covers the periods 1961:1-2007:4. All the variables are in natural logarithms. The definitions of variables are the following: $X = \ln \text{Real Exports}$, $M = \ln \text{Real Imports}$ and $GDP = \ln \text{Gross Domestic Product}$.

	mean	std.dev	skewness	kurtosis	JB	p.value
Y	3.1171	0.0512	-0.5492	-1.0741	18.2965	0.0001
X	3.0664	0.0688	-0.5950	-1.0281	19.2133	0.0001
M	3.0659	0.0773	-0.7715	-0.8498	24.3055	0.0000

Table 5.1: Summary statistics

The descriptive statistic is given in the Table 5.1. From this table, the mean is not significantly different from zero for either series. Normality is tested with the Jarque-Bera test, distributed as $\chi^2(2)$ under the null hypothesis, so is strongly rejected for all series. Since rejection could be due to either excess kurtosis, or skewness. We report these statistics, separately in the Table 5.1.

The following graphs Figure 5.1 portray the evolution of real exports, real imports and real GDP during the period of study.

5.4.2 Wavelet decomposition

We first obtain a wavelet decomposition of the respective macroeconomic series. All the wavelet decompositions are done using a Daubechies extremal phase filters

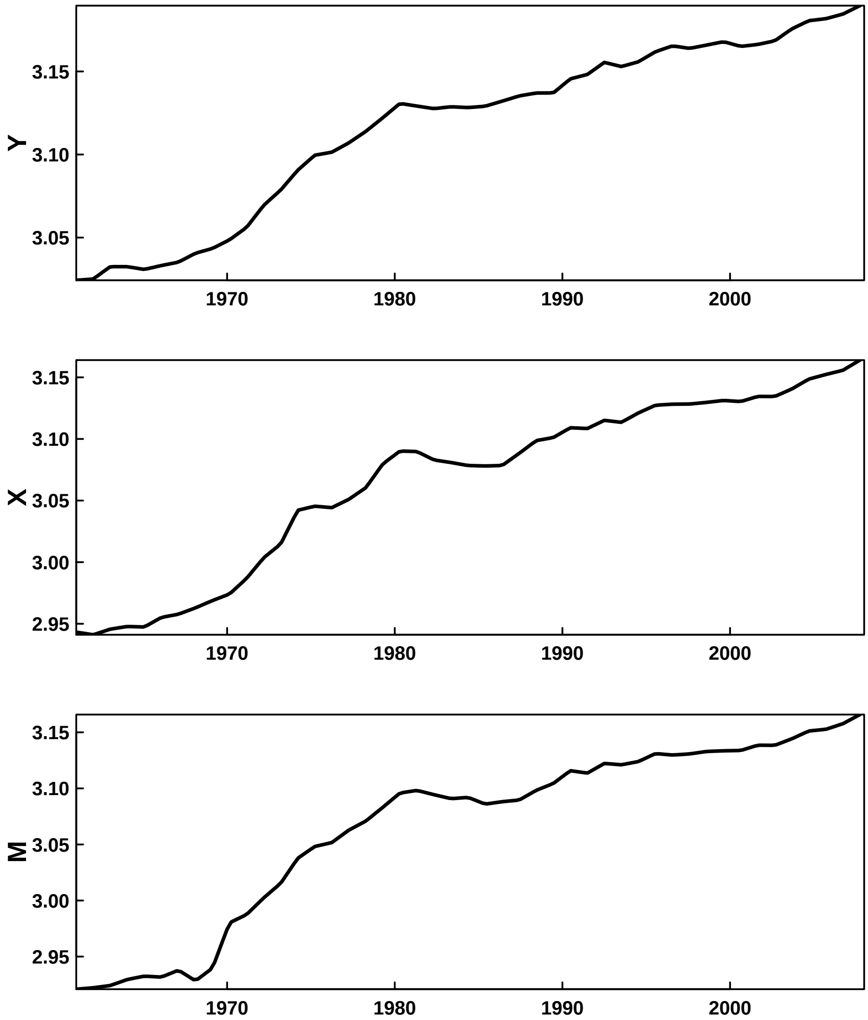

Figure 5.1: GDP, Exports and Imports series

of length $L = 4$, that is $D(4)$, based on four non-zero coefficients, with periodic boundary conditions. The application of the translation invariant wavelet transform with a number of scales $J = 3$ produce three vectors of details coefficients, that is d_1, d_2 and d_3,, and one vector of wavelet smooth. In figures Figure 5.2, Figure 5.3, and Figure 5.4 we report the wavelet decomposition of series.

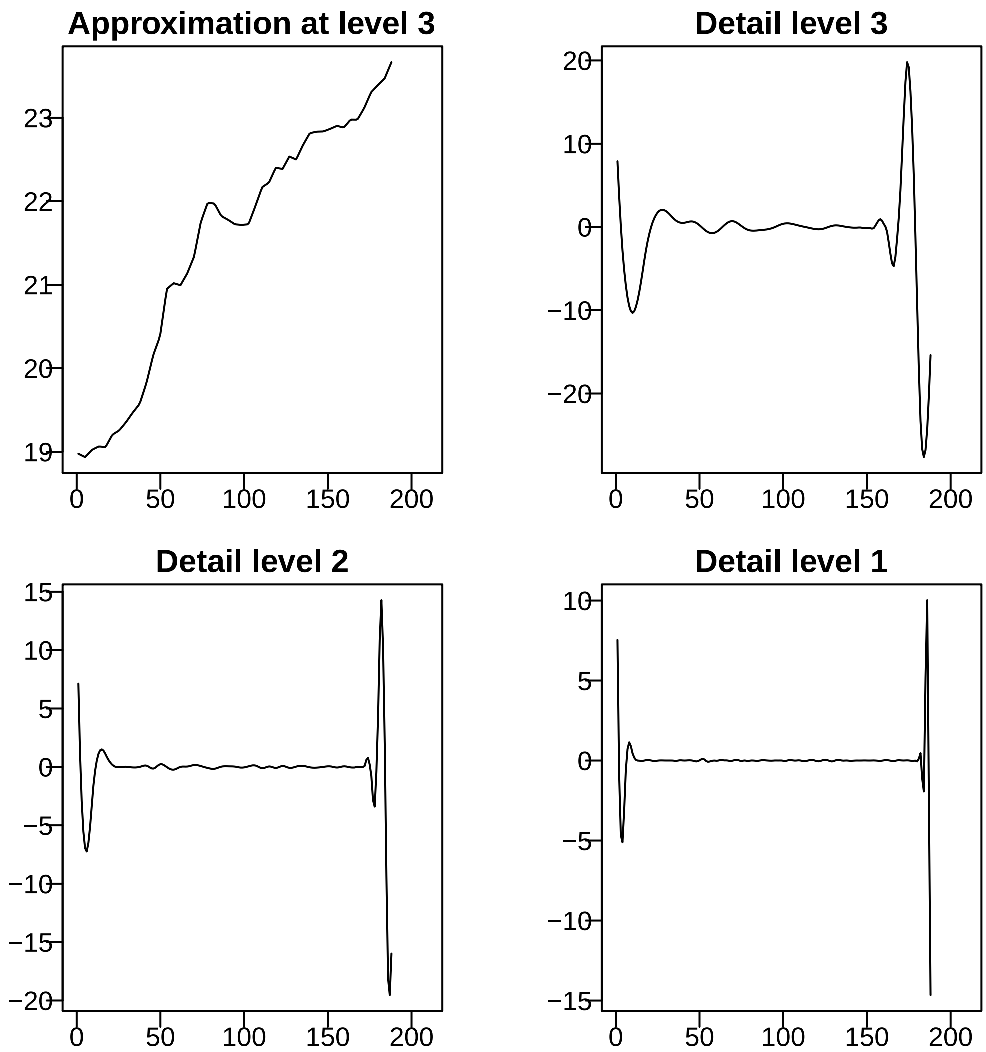

Figure 5.2: Wavelet decomposition of Exports series

5.4.3 Test results for unit roots

The tests are performed on all the component series obtained through wavelet decompositions of the original data. First we present the unit root test results and stationarity properties of the respective approximation series by performing the Augmented Dickey Fuller, Phillips-Perron and KPSS tests in top half of Table 5.2. We observe from this table that the respective approximation series are non stationary time series processes with respect to three test's statistics. In order to find out their order of integration, we further investigate the unit root test properties of their first difference series. From this test (bottom half of Table 5.2, we observe

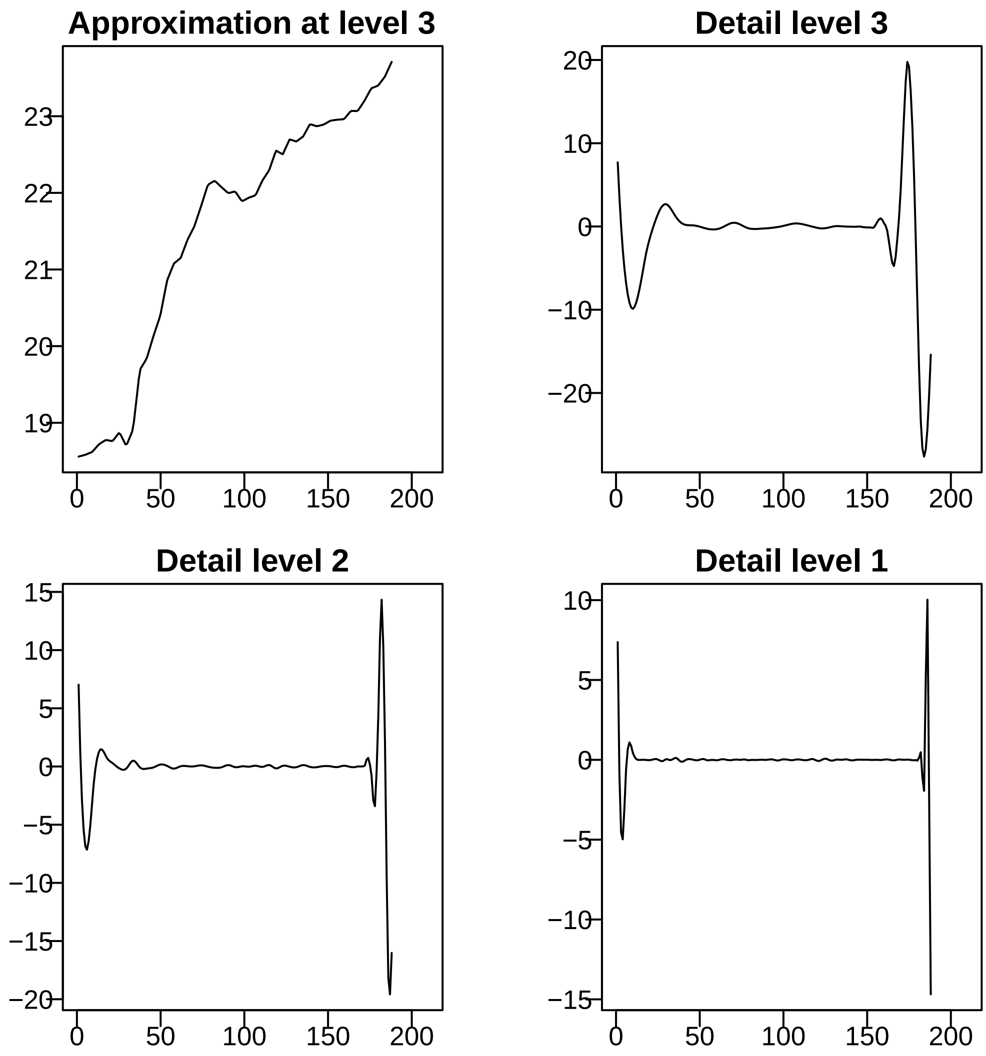

Figure 5.3: Wavelet decomposition of Imports series

that the first difference series of Exports and Imports are all stationary. Hence we conclude that the 4 level approximation series of Export, Import and GDP are all order one integrated, i.e. $I(1)$ time series processes.

5.4.4 Cointegration test results

The cointegration tests were performed utilizing the Johansen (Johansen, 1991) and (Johansen, 1995) methodology. The Johansen methodology is a generalization of the Dickey-Fuller test. Two likelihood ratio tests, λ_{max} and λ_{trace}, were used to test the hypotheses regarding the number of cointegrating vectors.

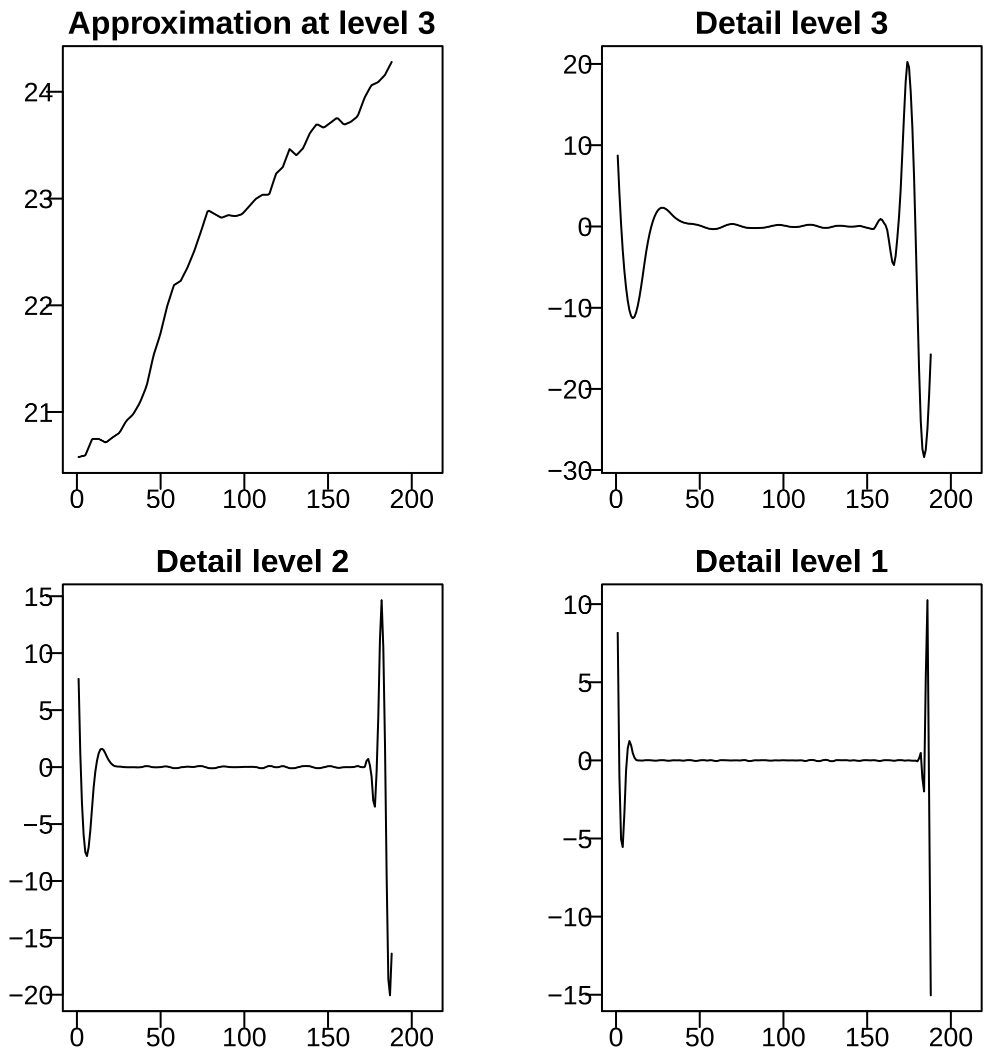

Figure 5.4: Wavelet decomposition of GDP series

Before implementing the Johansen procedure for co-integration analysis, the autoregression order of the VAR in Equation 5.14 has to be correctly specified. Therefore, to select the correct specification, we based our decision on the Schwarz' Bayesian information criterion (BIC) and selected $p = 3$.

We next proceed to testing for co-integration using the maximum likelihood approach developed by (Johansen, 1991) and (Johansen, 1995). The results of testing for the number of co-integrating vectors are reported in Table 5.4, which presents the maximum eigenvalue and the trace statistics, the 5% critical values as well as the corresponding λ values. The maximum eigenvalue tests for at most r co-

	ADF Test		Philips Perron Test		KPSS Test	
	Statistic (Lag)	p-value	Statistic (Lag)	p-value	Statistic (Lag)	p-value
X	-2.08588 (5)	0.48641	-1.96583 (4)	0.95592	4.47958 (3)	0.01000
M	-1.91339 (5)	0.56832	-1.50448 (4)	0.97650	4.31119 (3)	0.01000
Y	-1.85945 (5)	0.59555	-1.43931 (4)	0.97571	4.53364 (3)	0.01000
ΔX	-3.42928 (5)	0.05154	-28.65515 (4)	0.01000	0.37662 (3)	0.08723
ΔM	-3.89190 (5)	0.01589	-31.51380 (4)	0.01000	0.56147 (3)	0.02782
ΔY	-3.20722 (5)	0.08855	-28.38377 (4)	0.01000	0.51762 (3)	0.03770

Table 5.2: Unit root test results of level 4 approximation series and their first difference series

		ADF Test		Philips Perron Test		KPSS Test	
		Statistic (Lag)	p-value	Statistic (Lag)	p-value	Statistic (Lag)	p-value
Detail 1	X	-7.24741 (5)	0.0100	-106.19603 (4)	0.0100	0.02307 (3)	0.10
	M	-7.22903 (5)	0.0100	-105.57160 (4)	0.0100	0.02310 (3)	0.10
	Y	-7.21588 (5)	0.0100	-107.69627 (4)	0.0100	0.02301 (3)	0.10
Detail 2	X	-9.03592 (5)	0.0100	-52.18433 (4)	0.0100	0.05973 (3)	0.10
	M	-9.05793 (5)	0.0100	-52.12218 (4)	0.0100	0.05949 (3)	0.10
	Y	-9.06295 (5)	0.0100	-52.53184 (4)	0.0100	0.06004 (3)	0.10
Detail 3	X	-3.82256 (5)	0.0193	-23.60265 (4)	0.0288	0.12859 (3)	0.10
	M	-3.80145 (5)	0.0203	-23.54744 (4)	0.0293	0.12840 (3)	0.10
	Y	-3.83209 (5)	0.0188	-23.66561 (4)	0.0283	0.12948 (3)	0.10

Table 5.3: Unit root test statistics of the details time series

integrating vectors against the alternative of exactly $r + 1$, and the trace tests for at most r co-integrating vectors against an alternative of at least $r + 1$ vectors. (Johansen, 1991) discussed the likelihood based co-integration theory for such a model without constant terms. It turns out, however, that the presence of a con-

		Trace test		Max test	
H_0	λ	Statistic	p-value	Statistic	p-value
$r = 0$	0,12389	34,855	0,0112	24,469	0,0142
$r \leq 1$	0,038899	10,387	0,2566	7,3401	0,4585
$r \leq 2$	0,016333	3,0466	0,0809	3,0466	0,0809

Table 5.4: Johansen's test for multiple cointegration vectors. **Note that the p-values are computed via the approximations by (Doornik, 1998).**

stant in the non-stationary part of the data generating process plays a crucial role in the statistical analysis and for the interpretation of the model. In particular, (Johansen, 1991) proved that the asymptotic distribution of the test statistics and estimators in an error-correction model is not invariant to the assumption made about the constant term. (Osterwald-Lenum, 1992) also proved that the distribution of the trace statistic could be affected by the presence of a constant term. Thus, we used the (Johansen, 1995) likelihood ratio test to test whether the model should contain a constant term. The results of this test indicate that a model with an unrestricted constant should be adopted for the analysis. Both the maximum eigenvalue and trace statistics indicate the existence of a unique long-run relationship among the three endogenous variables included in the analysis. The estimates of the β and α vectors are presented in Table 5.5 below.
At this stage we observe that the adjustment coefficient on M is small and insignificant. Using a likelihood ratio test confirmed that real imports are weakly exogenous and, therefore, the model was re-estimated under the restriction that the coefficient on M is zero. The estimates of the α and β vectors after imposing a weak exogeneity restriction on M are presented in Table 5.6.

Using the results in Table 5.6, the long-run relationship between the three endogenous variables is given by:

$$Y_t = 0.82001X_t - 0.068618M_t \tag{5.16}$$

Variable	Y	X	M
β	1	-0.85127	0.098715
	–	(-11.34180)	(1.486179)
α	-0.0090580	0.061435	-0.0054243
	(-0.6683)	(2.843)	(-0.2145)

Table 5.5: The β and α vectors. Note that the cointegrating vector is normalized on Y and the figures in parentheses are t-statistics.

Variable	Y	X	M
β	1	-0.82001	0.068618
	–	(-11.31438)	(1.067121)
α	-0.011637	0.057971	–
	(-0.8546)	(2.663)	–

Table 5.6: The β and α vectors after imposing a weak exogeneity restriction on M. Note that the cointegrating vector is normalized on Y and the figures in parentheses are t-statistics.

The main conclusions from the results in Table 5.6 can be summarized as follows: Real exports has a significant positive effect on GDP as expected and both variables are adjusting to their long-run equilibrium relationship. This means that, in the short-run, changes in the real exports and changes in GDP are influenced by the long-run relationship that exists between them. Real imports has a insignificant negative effect on GDP.

5.4.5 Test results for Granger Non-causality

Testing for Granger non-causality in the context of stable VAR models involves testing whether some parameters of the model are jointly zero. In the past such testing has involved a standard F-test in a regression context.

The recent literature has shown that the conventional testing procedure for Granger non-causality using the F-statistic has size and power problems because of its dependence on the pretesting for co-integration (Zapata and Rambaldi, 1997). A much more accurate and simpler procedure was proposed independently by (Dolado and Lutkepohl, 1996) and (Toda and Yamamoto, 1995) and is known as the "augmented VAR approach". As shown by (Zapata and Rambaldi, 1997), the augmented VAR testing procedure is very simple to compute and is independent of the co-integration properties of the data. For the purposes of this chapter, the procedure is applied as follows:

1. Since the VAR model contains three lags and since the highest order of integration in the data is one, we first estimate a VAR in levels with four lags, then
2. We test jointly that the first three lags of the relevant variable are zero using a Wald test, which has a χ^2 distribution.

Table 5.7 reports the results of testing for Granger non-causality between GDP, Exports expansions and Imports expansions. The results suggest that there are clear indications of existence of a bi-directional causality between Exports and GDP (both the ELG and GLE hypotheses are supported by the data at the 5% level of significance.). Nevertheless, there is no relevant causality between import and export growths at 10% level of significance.

Consequently, we can conclude that, in Tunisia, there is a simultaneous cause and effect between economic growth and export growth. This simultaneity arises from the fact that, at the initial stages of development, economic growth promotes exports, but along the way exports start generating the capital needed for further economic growth.

Null hypothesis	Statistics	p-value
Exports does not Granger-cause GDP	4.2963	0.2312
Imports does not Granger-cause GDP	2.8741	0.4114
GDP does not Granger-cause Exports	3.9581	0.26602
GDP does not Granger-cause Imports	0.7346	0.8650
Imports does not Granger-cause Exports	7.3567	0.0613
Exports does not Granger-cause Imports	6.2653	0.0994

Table 5.7: Test results for Granger non-causality

5.5 Conclusion

In this chapter, we used wavelet based filtering technique for establishing relationships among macroeconomic indicators of exports, imports and economic growth.

An accurate analysis of co-integrated long-run equilibrium or causal relationships among these macroeconomic indicators is important.
We consider the case of the Tunisian economy, explore and extract interesting relationships using wavelet technique:
Using quarterly data over the time period 1961:1-2007:4, after filtering the series using wavelet technique, we have analyzed the time series properties of the exports, imports and economic growth variables in order to determine appropriate functional form for testing the ELG hypothesis.
The study find that, the exports, imports and GDP are co-integrated.
From the causality tests we have seen that there exist a bi-directional relationship between the Exports and GDP, no relevant causality between import and export growths at 10% level of significance and a bi-directional relationship between import and economic growths.

CHAPTER 6

Conclusion and Prospects

This thesis lies with the scope of the whole topic of time series analysis. The principal tool used is the wavelet transforms. It is a mathematical tool recently developed which has been proved be efficient compared to classical methods as Fourier Analysis which was the basic tool in signal analysis during a long period.

In the first chapter, we briefly summarize the main contributions and then move to discuss possible directions for future research.

In a second chapter, we recalled the necessary fundamental notions of wavelet analysis such as multi-resolution analysis with some statistical applications such as regression, filtering,...

In a third chapter, we have used the wavelet transform to analyse the variance and correlation between two time series. The wavelet variance and wavelet correlation decompose the variance of a time series and correlation between two time series on scale by scale basis respectively. In the same chapter, we have tested the homogeneity of variance for a time series which exhibit a long memory structure using wavelet transforms.

Chapter 4 is devoted to the study of fractional processes. In the study of the estimation of long memory parameters, there exists mainly three methods through wavelet techniques: The Ordinary Least Squares (OLS) method, the Maximum Likelihood Estimator (MLE) and the Weighted Least Squares Estimator (WLSE). These methods are applicable due to the ability of the Discrete Wavelet Transform (DWT) to decorrelate long memory processes. We have described these methods and applied them to estimate long memory parameters of six stock market indices.

Finally, in chapter 5, we have applied a wavelet filtering technique to study the econometric relationship between exports, imports expansions and economic growth in Tunisia.

Using quarterly data over the time period 1961:1-2007:4 and after filtering the series using Daubechies Extremal Phase Wavelet Filter with length $L = 4$, we have analyzed the the time series properties.

We find that, the three variables considered are co-integrated.
From the causality tests, we have seen that there exists a feedback relationship between the exports and Growth Domestic Product (GDP), no relevant causality between exports and imports, whereas there exists a bi-directional causality between imports and economic growth.

Future research

Some challenging problems remain for future research:
The first problem is to provide a test for homogeneity of covariance. A cumulative sum of squares test statistic may be useful for investigating departures from a constant association between two time series. More work is needed to formulate

specific statistical hypotheses.

Another problem is to provide an approximate likelihood estimation method for the locally stationary processes. For the generalized fractional integrated model, there exist one wavelet-based estimation: the wavelet-based OLS method (Whitcher and Jensen, 2000).

Final conclusions

In this thesis we have attempt to provide a new useful methodology in order to analyse problems in time series analysis where techniques were previously lacking. The two primary areas being detecting and locating change of variance analysis in time series with long memory structure using wavelet transforms, and extending wavelet analysis of variance for univariate time series to a wavelet analysis of co-variance for bivariate time series; i.e, introducing the wavelet cross-covariance and cross-correlation.

The wavelet transforms are a powerful mathematical tool, enabling staticians to examine much more complicated processes by separating features on scale by scale basis. However, it cannot be applied to problems without caution. Which wavelet filter to use is a very important issue. To help select an appropriate wavelet filter, several of various lengths should be applied to the data and visually analysis to help detect the potential leakage of low frequency features throughout the multi-resolution analysis. Most importantly, how the shape of the underlying wavelet filter matches the physical processes from where the data was sampled? Keeping these issues in mind, there should be great variety of problems where wavelet analysis is beneficial.

Bibliography

Abry, P., Gonçalves, P., and Flandrin, P. (1995). Wavelets, spectrum analysis and 1/f processes. *Wavelets and statistics*.

Abry, P. and Veitch, D. (1998). Wavelet analysis of long-range-dependent traffic. *Transactions on Information Theory*, 44:2–15.

Ahmad, J A, C. K. (1991). Causality between exports and economic growth. *Economics Letters*, 37:243–248.

Antoniadis, A. (1991). *Multiresolution analysis and wavelets for density extimation.* Irvine, University of California.

Antoniadis, A., Grégoire, G., and Nason, G. (1999). Density and hazard rate estimation for right-censored data by using wavelet methods. *Journal Of The Royal Statistical Society Series B*, 61(1):63–84.

Balassa, B. (1978). Exports and economic growth: further evidence. *Journal of Development Economics*, (5):181.

Ben Ammou, S., Kacem, Z., and Haouas, N. (2010). Risk reduction using wavelets for denoising principal-components regression models. *Journal of Risk Finance*, 11(2):180 – 203.

Ben Mabrouk, A., Kortas, H., and Ben Ammou, S. (2009). Wavelet estimators for long memory in stock markets. *International Journal of Theoretical and Applied Finance (IJTAF)*, 12(03):297–317.

Bhagwati, J, N. (1988). *Protectionism.* MIT Press, Cambridge, Massachusetts.

Box, G, E. and Jenkins, G, M. (1976). *Time series Analysis.* Holden Day, San Francisco.

Campbell, J., Lo, A., and MacKinlay, A. (1997). *The Econometrics of Financial Markets.* Princeton University Press, Princeton.

Chen, J. and Gupta, A. (1997). Testing and locating variance changepoints with application to stock prices. *Journal of the American Statistical Association*, 92:739–747.

Chow, P, C. Y. (1987). Causality between export growth and industrial performance: Evidence from the nic's. *Journal of Development Economics*, 26:55–63.

Coe, T, D. and Helpman, E. (1995). International r&d spillovers. *European Economic Review*, 39:859–887.

Coifman, R, R. and Donoho, D, L. (1995). *Translation-invariant wavelet denoising: Wavelets and Statistics.* Lecture Notes in Statistics. Springer-Verlag, New York.

Daubechies, I. (1992). *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics.

Daubechies, I., Guskov, I., Schröder, P., and Sweldens, W. (1999). Wavelets on irregular point sets. *Phil. Trans. R. Soc. Lond. A*, 357:2397–2413.

Davison, R., Labys, W, C., and Lesourd, J, B. (1998). Wavelet analysis of commodity price behavior. *Journal of Computational Economics*, 11:103–128.

Dickey, D, A. and Fuller, W, A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–431.

Dijkerman, R. and Mazumdar, R. (1994). On the correlation structure of the wavelet coefficients of fractional brownian motion. *IEEE Trans. Inform. Theory*, 40:1609–1612.

Dolado, J, J. and Lutkepohl, H. (1996). Making wald test work for cointegrated var systems. *Econometric Reviews*, 15:369–386.

Donoho, D, L. (1992). Interpolating wavelet transforms. Technical report, Department of Statistics, Stanford University, Stanford. Tech. rept.

Donoho, D, L. (1993). Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data. *American Mathematical Society*, 47:173–205.

Donoho, D, L. (1995). De-noising by soft thresholding. *Transactions on Information Theory*, 41(3):613–627.

Donoho, D, L. and Johnstone, I, M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–55.

Donoho, D, L. and Johnstone, I, M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *J, Am, Statist. Ass*, (90):1200–1224.

Donoho, D, L. and Johnstone, I, M. (1998). Minimax estimation via wavelet shrinkage. *The Annals of Statistics*, 26(3):879–921.

Donoho, D, L. and Johnstone, I, M. (1999). Asymptotic minimaxity of wavelet estimatiors with sampled data. *Statistica Sinica*, 9(1):1–32.

Donoho, D, L. and Yu, T, P. Y. (1998). *Nonlinear 'wavelet transforms' based on median- thresholding.* Stanford University.

Doornik, J, A. (1998). Approximations to the asymptotic distribution of cointegration tests. *Journal of Economic Surveys*, 12:573–93.

Doukhan, P. (1988). Formes de toeplitz associées à une analyse multiéchelle. *C. R. Acad. Sci. Paris. Sér. A*, 306:663–666.

Doukhan, P. (1990). Déviations quadratique d'estimateurs de densité par projections orthogonales. *C. R. Acad. Sci. Paris. Sér. A*, 310:425–430.

Dutilleux, P. (1989). An implementation of the "algorithme à trous" to compute the wavelet transform. In Combes, J. M., Grossmann, A., and Tchamitchian, P., editors, *Wavelets, Time-Frequency Methods and Phase Space*, pages 298–304. Springer-Verlag.

Enders, W. (1995). *Applied Econometric Time Series*. John Wiley & Sons.

Fan, Y. (2003). On the approximate de-correlation property of the discrete wavelet transform for fractionally differenced processes. *IEEE Transactions on Information Theory*, 49:516–521.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional brownian motion. *IEEE Trans. Inform. Theory*, 38(2, part 2):910–917.

Fox, R. and Taqqu, Murad, S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary gaussian time series. *The Annals of Statistics*, 14(2):517–532.

Fryzlewicz, P. (2005). Modelling and forecasting financial log-returns as locally stationary wavelet processes. *Journal of applied statistics*, 32(5):503–528.

Gao, H, Y. (1993). *Wavelet Estimation of Spectral Densities in Time Series Analysis*. PhD thesis, PhD Dissertation, Department of Statistics, University of California, Berkeley.

Genacy, R., Selcuk, F., and Whitcher, B. (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press, New York.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4:221–238.

Granger, C. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, (1):15 – 29.

Granger, C, W. J. (1969). Investigating the causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–238.

Granger, C, W. J. (1987). Coi-integration and error correction: representation, estimation and testing. *Econometrica*, 55:251–276.

Guegan, D. and Zhiping, L. (2009). Wavelet method for locally stationary seasonal long memory processes. Documents de travail du centre d'economie de la sorbonne, Université Panthéon-Sorbonne (Paris 1).

Hosking, J, R. M. (1981). Fractional differencing. *Biometrika*, 68:165–176.

Inclan, C. and Tiao, G. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89(427):913–923.

Jensen, M. (1997). Making wavelets in finance. *Financial Engineering News*, 1:1–10.

Jensen, M. (1999). Using wavelets to obtain a consistent ordinary least squares estimator of the long-memory parameter. *Journal of Forecasting*, 18:17–32.

Jensen, M. J. (1998). An approximate wavelet mle of short and long memory parameters. *Studies in Nonlinear Dynamics and Econometrics*, 3.

Jensen, M. J. (2000). An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets. *Journal of Economic Dynamics and Control*, 24(3):361–387.

Jensen, M. J. and Whitcher, B. (2000). Time-varying long-memory in volatility: Detection and estimation with wavelets. Technical report.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3):231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–80.

Johansen, S. (1995). *Likelihood Based Inferences in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Johnstone, I. and Silverman, B. W. (1997). Wavelet threshold estimators for data with correlated noise. *J. R. Statist. Soc.*, Vol B, 52(2):319–351.

Johnstone, I. and Silverman, B. W. (1998). Empirical bayes approaches to mixture problems and wavelet regression. Technical report.

Joy, M. (2001). Imported machinery and growth in ldcs. *Journal of Development Economics*, 65(1):209–224.

Jung, Woo, S. and Marshall, Peyton, J. (1985). Exports, growth and causality in developing countries. *Journal of Development Economics*, 18(1):1–12.

Karuppiah, J. and Los, C. A. (2005). Wavelet multiresolution analysis of high-frequency asian fx rates, summer 1997. *International Review of Financial Analysis*, 14(2):211–246.

Krugman, P. (1984). *Monopolistic Competition in International Trade*, chapter Import protection as export promotion. Oxford University Press.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3):159–178.

Mallat, S, G. and Hwang, W, L. (1992). Singularity detection and processing with wavelets. *IEEE Trans. Inform. Theory*, 38(2):617–643.

Mallat, S. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693.

Masry, E. (1993). Strong consistency and rates for deconvolution of multivariate densities of stationary process. *Stochastic Process Appl.*, 47:53–74.

McCoy, E. and Walden, A. (1996). Wavelet analysis and synthesis of stationary long-memory processes. *Journal of Computational and Graphical Statistics*, 5(1):26–56.

Moosa, I. (1999). Is the export-led growth hypothesis valid for australia? *Applied Economics*, 31(7):903–06.

Nason, G. and Silverman, B. (1995). The stationary wavelet transform and some statistical applications. pages 281–300. Springer-Verlag.

Nason, G, P., Sapatinas, T., and Sawczenko, A. (1998). Statistical modelling of time series using non-decimated wavelet representations.

Nason, G, P. and von Sachs, R. (1999). Wavelets in time series analysis. *Philosophical Transactions of the Royal Society of London A*, 357:2511–2526.

Norsworthy, J, L. D. and Gorener, R. (2000). Wavelet-based analysis of time series: An export from engineering to finance. IEEE International Engineering Management Society Conference. Albuquerque, New Mexico.

Norsworthy, J, L. D. and Siregar, D. (2001). Managing information in financial markets: A comparison of intraday distributions of trades on good. In *Proceedings of the 2001 Portland International Conference on Management Engineering and Technology*. Portland Oregon USA.

Osterwald-Lenum, M. (1992). A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics: four cases. *Oxford Bulletin of Economics and Statisitics*, 54:461–72.

Percival, D. (1995). On estimation of the wavelet variance. *Biometrika*, 82:619–631.

Percival, D. B. and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press.

Phillips, P, C. B. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75:335–346.

Ramsey, J. (1999). The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London*, A(357):2593–2606.

Ramsey, J. and Lampart, C. (1998a). The decomposition of economic relationships by timescale using wavelets: Money and income. *Macroeconomic Dynamics*, 2(1):49–71.

Ramsey, J., Usikov, D., and Zaslavskiy, G. (1995). An analysis of u.s. stock price behavior using wavelets. *Fractals*.

Ramsey, J. B. and Lampart, C. (1998b). The decomposition of economic relationships by time scale using wavelets: Expenditure and income. *Studies in Nonlinear Dynamics and Econometrics*, 3(1):23–42.

Ramsey, J. B. and Zhang, Z. (1994). The application of wave form dictionaries to stock market index data. Working Papers 94-05.

Ramsey, J. B. and Zhang, Z. (1995). The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance*, 4:341–372.

Robinson, P, M. (1995). Gaussian semiparametric estimation of long-range dependence. *The Annals of Statistics*, 23:1630–1661.

Roueff, F. and Von Sachs, R. (2009). Locally stationary long memory estimation. Technical report, INRIA a CCSD electronic archive server based on P.A.O.L [http://hal.inria.fr/oai/oai.php] (France).

Shensa, M, J. (1992). Wedding the a trous and mallat algorithms. *IEEE Trans. Signal Process*, 40(10):2464–2482.

Silverman, B. W. (2000). *Wavelets: The Key to Intermittent Information?*, chapter Wavelets in statistics: beyong the standard assumptions, pages 71–87. Oxford University Press.

Sims, C. A. (1972). Money, income, and causality. *American Economic Review*, 62(4):540–52.

Sowell, F. (1992). Modeling long-run behavior with the fractional arima model. *Journal of Monetary Economics*, 29(2):277–302.

Stein, C. M. (1995). Inadmissibility of the usual estimatior for the mean of a multivariate normal distribution. In *Proc. Third Berkeley Symp. Math. Statist. Probab.* Berkeley,University of California.

Tewfik, A. and Kim, M. (1992). Correlation structure of the discrete wavelet coefficients of fractional brownian motion. *IEEE Transactions on Information Theory*, 38:904 –909.

Tkacz, G. (2001). Estimating the fractional order of integration of interest rates using a wavelet ols estimator. *Studies in Nonlinear Dynamics and Econometrics*, 5(1):19–32.

Toda, H, Y. and Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66:225–250.

Triantafyllopoulos, K. and Nason, G. (2009). A note on state space representations of locally stationary wavelet time series. *Statistics & Probability Letters*, 79(1):50–54.

Veitch, D. N. and Abry, P. (1999). A statistical test for the time constancy of scaling exponents. *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, 49:2325–2334.

Vostrikova, L, Y. (1980). Detection of changes of mean functions in random processes. *Theory Probab. Appl.*, 25:867–869.

Walter, G. G. (1992). Approximation of the delta function by wavelets. *J. Approx. Theory*, 71(3):329–343.

Whitcher, B. (1998). *Assessing Nonstationary Time Series Using Wavelets*. PhD thesis, University of Washington, Department of Statistics.

Whitcher, B., Byers, S., Guttorp, P., and Percival, D. (1998). Testing for homogeneity of variance in time series: Long memory, wavelets and the nile river. Technical report, Water Resour. Res.

Whitcher, B., Guttorp, P., and Percival, D. (1999). Multiscale detection and location of multiple variance changes in the presence of long memory. *Journal of Statistical Computation and Simulation*, 68:65–88.

Whitcher, B. and Jensen, M. (2000). Wavelet estimation of a local long memory parameter. *Exploration Geophysics*, 31:89–98.

Wornell, G, W. (1993). wavelet based representations for the $1/f$ family of fractal processes. In *Proceedings of the IEEE*, volume 81, pages 1428–1450.

Zapata, H, O. and Rambaldi, A, N. (1997). Monte carlo evidence on cointegration and causation. *Oxford Bulletin of Economics and Statistics*, 59:285–298.

Zheng, G., Starck, J., Campbell, J., and Murtagh, F. (1999). The wavelet transform for filtering financial data streams. *Journal of Computational Intelligence in Finance*, 7(3):18–35.