Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2012, Article ID 819503, 14 pages
doi:10.1155/2012/819503

Research Article

Analysis of Stock Market Indices with Multidimensional Scaling and Wavelets

J. Tenreiro Machado,[1] Fernando B. Duarte,[2] and Gonçalo Monteiro Duarte[3]

[1] *Department of Electrical Engineering, Institute of Engineering, 4200-072 Porto, Portugal*
[2] *Faculty of Engineering and Natural Sciences, Lusofona University, 1749-024 Lisboa, Portugal*
[3] *Faculty of Economics and Management, Lusofona University, 1749-024 Lisboa, Portugal*

Correspondence should be addressed to J. Tenreiro Machado, jtm@isep.ipp.pt

Received 20 January 2012; Accepted 12 March 2012

Academic Editor: Katica R. (Stevanovic) Hedrih

Stock market indices (SMIs) are important measures of financial and economical performance. Considerable research efforts during the last years demonstrated that these signals have a chaotic nature and require sophisticated mathematical tools for analyzing their characteristics. Classical methods, such as the Fourier transform, reveal considerable limitations in discriminating different periods of time. This paper studies the dynamics of SMI by combining the wavelet transform and the multidimensional scaling (MDS). Six continuous wavelets are tested for analyzing the information content of the stock signals. In a first phase, the real Shannon wavelet is adopted for performing the evaluation of the SMI dynamics, while their comparison is visualized by means of the MDS. In a second phase, the other wavelets are also tested, and the corresponding MDS plots are analyzed.

1. Introduction

Economical indices measure segments of the stock market and are normally used to benchmark the performance of stock portfolios. This paper proposes a method for analyzing the correlations embedded in international stock markets. The study of the international stock markets may have different leitmotifs. Economic motivations to identify the main factors which affect the behavior of stock markets across different exchanges and countries. Statistical motivations to visualize correlations in order to suggest some potentially plausible parameter relations and restrictions. The understanding of such relations would be helpful to the design of good portfolios [1, 2].

The financial time series are inherently noisy, nonstationary, and deterministic chaotic, that is to say the distribution of financial time series is changing over the time. The noise component is due to the unavailability of complete information from the signal behaviour to capture the dependency between past and future values.

The complexity of the problem motivated the adoption of the wavelets for the study of the stock market indices (SMIs) [3, 4]. A *wave* is usually defined as an oscillating function of time, such as sinusoid. Fourier analysis is wave analysis. It expands signals in terms of sinusoids (or, equivalently, complex exponentials) which has proven to be valuable in mathematics and engineering, especially for periodic, time-invariant, or stationary phenomena. A *wavelet* is a "small wave," which has its energy concentrated in time to give a tool for the analysis of transient, nonstationary, or time-varying phenomena. The wavelet transform allow users to establish a compromise between precision in the frequency and time domains. Several types of continuous wavelets are tested and, based on the emerging patterns, the real Shannon wavelet is considered as the best one for the analysis. The wavelet charts depict complex patterns and, due to the large number of cases, a comparison index is performed. Based on the similarity measure, the multidimensional scaling (MDS) visualization tool is adopted. MDS is a data analysis technique for depicting the similarity or dissimilarity of data. MDS is used to represent (dis)similarity data between objects by a variety of distance models. The term similarity is used to indicate the degree of likeness between two objects, while dissimilarity indicates the degree of unlikeness. MDS represents a set of objects as points in a multidimensional space in such a way that the points corresponding to similar objects are located close together, while those corresponding to dissimilar objects are located far apart. The researcher then attempts to make sense of the derived object configuration by identifying meaningful regions and/or directions in the space [5–9].

The remainder of this paper is organized as follows. Section 2 introduces the financial indices, the fundamental concepts adopted in the study, and the methodology of analysis. Section 3 analyzes the market stocks indices using wavelets. Section 4 presents a MDS analyzes based on wavelets. Finally, Section 5 draws the main conclusions.

2. Financial Indices: Fourier and Wavelet Transforms

Our data consist of the n daily close values of $S = 33$ stock markets, listed in Table 1, from January 2, 2000 up to December 31, 2009, to be denoted by $x_i(t)$, where $t = 1, \ldots, n$ represent time and $i = 1, \ldots, S$. These specific stock markets were chosen because they are considered to be representative of the reality. The inclusion of more indexes would lead to confusion and, therefore, become counterproductive.

The data is obtained from data provided by Yahoo Finance web site [10] and measures indices in local currencies.

For example, Figure 1 depicts the time evolution of daily closing prices of the six stock markets versus time (in years). The charts exhibit the well-known noisy and chaotic characteristics. Each signal $x_i(t)$ is complex and difficult to analyze in the time domain. Therefore, for highlighting the characteristics of $x_i(t)$ is required the application of adequate signal processing tools. In the sequel is analyzed the performance of the Fourier and wavelets transforms.

2.1. Fourier Analysis

For each signal index, the corresponding Fourier transform (FT) is calculated, according to:

$$\mathcal{F}\{x(t)\} = \int_{-\infty}^{+\infty} x(t) e^{-j\Omega t} dt, \quad j^2 = -1, \tag{2.1}$$

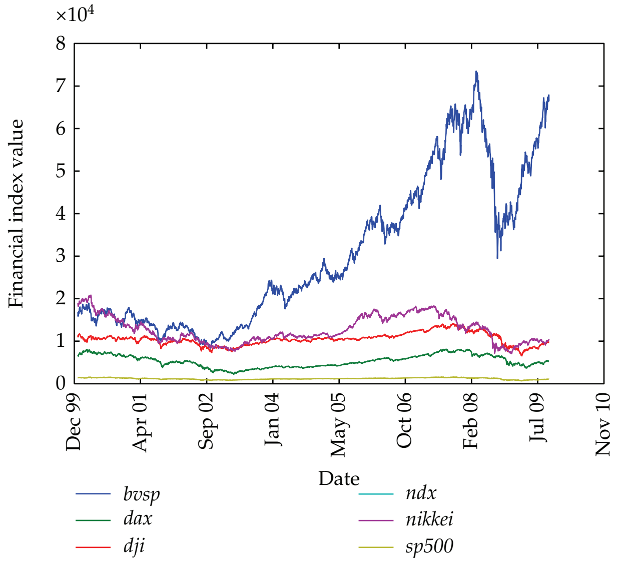


Figure 1: Time series for the {*bvsp, dax, dji, ndx, nikkei, sp500*} indices from January 2000 up to December 2009.

where $\mathcal{F}$ is the Fourier operator, $x(t)$ is the index value, t is time, and Ω is the angular frequency.

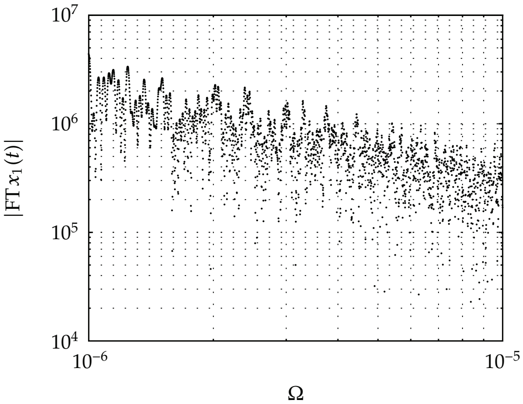
Figure 2 shows the $|\mathcal{F}\{x_k(t)\}|$ *versus* Ω for the $\{bvsp, dax, dji, ndx, nikkei, sp500\}$ indices. The charts for the other SMI are of the same type and are not represented. It is well known that this tool "dilutes" the signal time information leading only to a global representation. Therefore, since the signals may be not stationary, accessing limited periods of time is problematic.

2.2. Wavelet Analysis

The continuous wavelet transform [11–13] is defined as

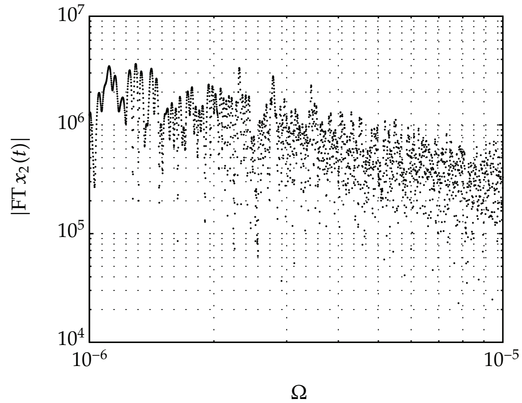
$$[W_\psi x(t)](a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^* \frac{t-b}{a} dt, \quad a > 0, \tag{2.2}$$

where the symbol $*$ denotes the complex conjugate, the parameters (a, b) represent the dyadic dilation and the dyadic position, respectively, and ψ is a function called the mother wavelet. The mother wavelet is the source for generating daughter wavelets, which are simply the translated and scaled versions of the mother wavelet. Often the parameter a is interpreted qualitatively as the inverse of the frequency of the Fourier analysis. The wavelet transform is often compared with the Fourier transform, in which signals are represented as a sum of sinusoids. The main difference is that wavelets are localized in both time and frequency, whereas the standard Fourier transform is only localized in frequency. Wavelets give a better signal representation using multiresolution analysis, with balanced resolution at any time and frequency.



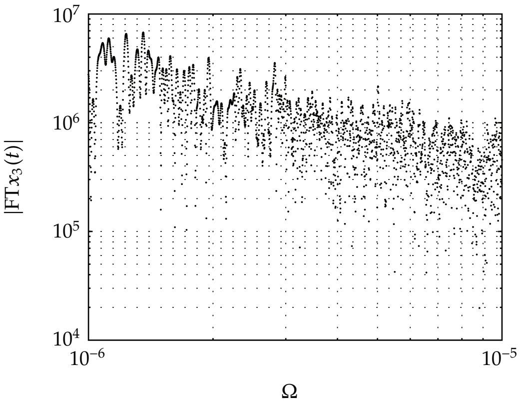
bvsp

(a)



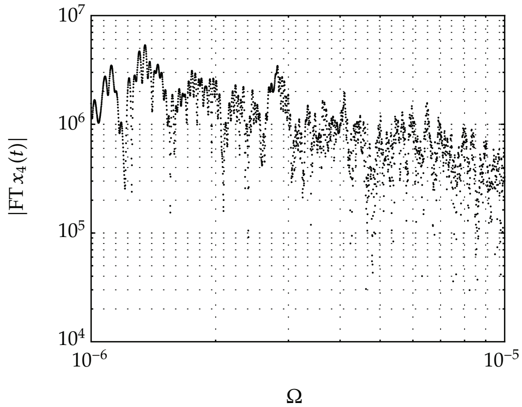
dax

(b)



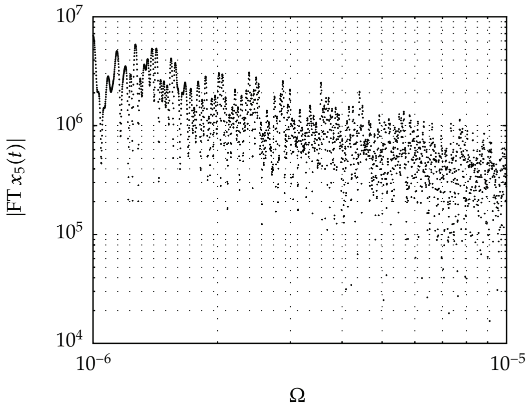
dji

(c)



ndx

(d)



nikkei

(e)

Figure 2: $|\mathcal{F}\{x_k(t)\}|$ *versus* Ω for the $\{bvsp, dax, dji, ndx, nikkei, sp500\}$ indices.

Table 1: Thirty-three stock markets.

i	Stock market index	Abbreviation	Country
1	Dutch Euronext Amsterdam	aex	The Netherlands
2	All Ordinaries equities market	aord	Australia
3	Index of the Vienna Bourse	atx	Austria
4	EURONEXT BEL-20	bfx	Belgium
5	Bombay Stock Exchange Index	bse	India
6	San Paulo (Brazil) Stock	bvsp	Brazil
7	Budapest Stock Exchange	bux	Hungary
8	Cotation Assisée en Continu	cac	France
9	Egypt CMA Index	ccsi	Egypt
10	Dow Jones Industrial	dji	USA
11	Deutscher Aktien Index	dax	Germany
12	Footsie	ftse	United Kingdom
13	Stock market index in Hong Kong	hsi	Hong Kong
14	Iberia Index	ibex	Spain
15	Jakarta Stock Exchange	jkse	Indonesia
16	Bursa Malaysia	klse	Malaysia
17	Stock Market index of South Korea	ks11	South Korea
18	Argentina Merval Index	ks11	Argentina
19	Oslo Børs All Share Index	oseax	Norway
20	Italian Bourse	mibtel	Italy
21	Bolsa Mexicana de Valores	mxx	Mexico
22	NASDAQ	ndx	USA
23	Tokyo Stock Exchange	nikkei	Japan
24	New York Stock Exchange	nya	USA
25	Stock exchange of Portugal	psi20	Portugal
26	Standard & Poor's	sp500	USA
27	Shanghai Stock Exchange	ssec	China
28	Swiss Market Index	ssmi	Switzerland
29	Straits Times Index	sti	Singapore
30	Tel Aviv 100 Index	ta100	Israel
31	Toronto Stock Exchange	tsx	Canada
32	Taiwan Stock Exchange	twii	Taiwan
33	South Africa Index	zadowd	South Africa

Wavelet transforms are classified into discrete wavelet transforms (DW) and continuous wavelet transforms (CW). Note that both DW and CW are continuous-time (analog) transforms. They can be used to represent continuous-time (analog) signals. CW operates over every possible scale and translation whereas DW uses a specific subset of scale and translation values or representation grid [14–17].

In this paper are investigated three real and three complex valued CWs, namely, the Haar, Ricker (also called Mexican hat), Shannon, Hermitian hat, Shannon complex, and Morlet wavelets, denoted by {HW, RW, SW, HHW, SCW, MW}, defined by the expressions:

$$\text{Haar: } \psi(t) = \begin{cases} 1, & 0 \le t < \frac{1}{2} \\ -1, & \frac{1}{2} \le t \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\text{Shannon real: } \psi(t) = \frac{1}{\pi t}[\sin(2\pi t) - \sin(\pi t)]$$

$$\text{Ricker (Mexican hat): } \psi(t) = \frac{2}{\sqrt{3\sigma}\pi^{1/4}}\left(1 - \frac{t^2}{\sigma^2}\right)e^{-t^2/2\sigma^2} \tag{2.3}$$

$$\text{Hermitian hat: } \psi(t) = \frac{2}{\sqrt{5}\pi^{1/4}}\left(1 - t^2 + jt\right)e^{t^2/2}$$

$$\text{Shannon complex: } \psi(t) = \frac{\sin(\pi t)}{\pi t}e^{-j2\pi t}$$

$$\text{Morlet: } \psi(t) = c_\sigma \pi^{-1/4} e^{-t^2/2}\left(e^{j\sigma t} - e^{-1/2\sigma^2}\right), \qquad c_\sigma = \left(1 + e^{-\sigma^2} - 2e^{-3/4\sigma^2}\right)^{-1/2}.$$

Tackling the financial data through wavelets leads to a considerable volume of information. Therefore, for condensing the results of the wavelet charts a similarity measure r between two plots is developed in the next section. This index allows the construction of a symmetrical correlation matrix $\mathbf{R}$ comparing all cases. Based on the matrix it is then possible to use visualization tools for establishing a graphical locus of the thirty-three stock markets.

The multidimensional scaling (MDS) visualization tool assigns a point to each item in a multidimensional space and then arranges the the "cloud" of points in a low-dimensional space in order to reproduce the observed similarities.

Bearing these ideas in mind, for the stock market analysis, in the next section, is adopted (i) the set of thirty-three SMI listed in Table 1 (ii) the CWs for the signal analysis, (iii) the six continuous wavelets {HW, RW, SW, HHW, SCW, MW} defined in (2.3), (iv) the measure of similarity between wavelet charts using an appropriate index, and (v) the adoption of the visualization technique for obtaining a graphical output.

3. Wavelet Analysis of the SMI

For each of the thirty three index signals, $x_i(t)$, $i = 1, \ldots, S$, the wavelet transform $[W_\psi x(t)](a, b)$ is calculated. The results of the wavelet analysis depend on the mother function ψ to be adopted. Therefore, before comparing all indices, a preliminary evaluation is developed in order to characterize of each function ψ for the analysis of the indices signal. In this line of thought, Figure 3 depicts the absolute value of the wavelet for the *dji* index and the six functions listed in (2.3), where it is adopted that $\sigma = 1$ and $\sigma = 5$ for the RW and the MW, respectively. Furthermore, it is established a "time step" $dt = 1$ for the sequence base increment along the index, and the parameters (a, b) are considered to vary from zero up to the maximum length of signal index. In the charts, we must take care with the results at the limits of the intervals of variation of the parameters (a, b) due to truncation effects.

We verify that we get very different charts for each function ψ, that is, we conclude that the observation lens provided by ψ hide or reveal different signal characteristics and that some may be better adapted to this than others. The HW seems to be the "worst" probably

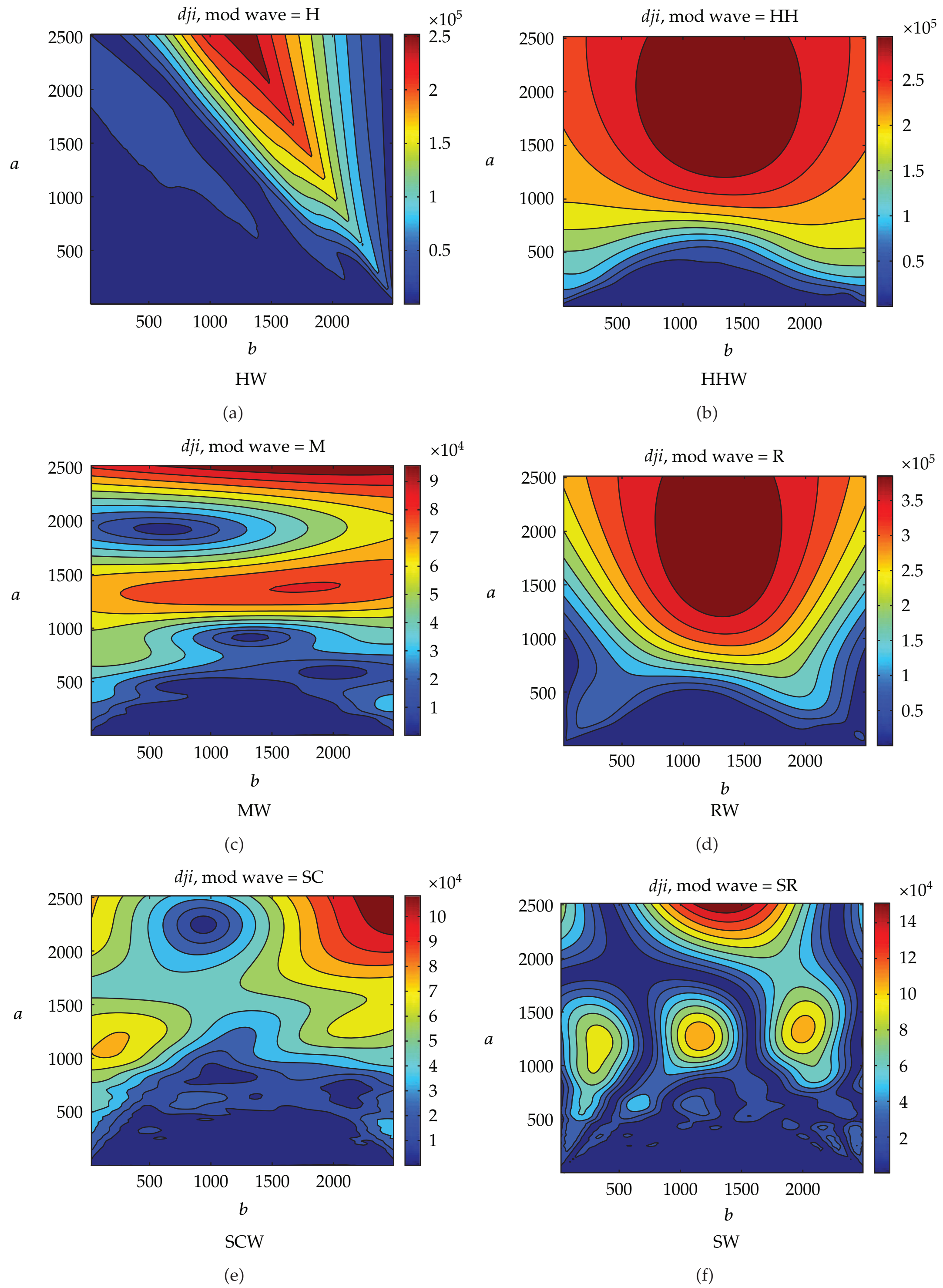


Figure 3: Absolute value of the six wavelets {HW, RW, SW, HHW, SCW, MW} listed in (2.3) *versus* (a, b) for the *dji* index, with $dt = 1$.

because it is more adapted to digital signals, while in the present case we have a different type of time evolution. We observe that the RW has similarities to HHW, and, identically, the SCW to the MW. In these four cases, we observe a pattern for a in the middle of the interval and for low values of b. Particularly, the RW seems to present an slight higher level of detail with the presence of three objects. The SW seems to be a "good" wavelet, depicting a clear emergence of three objects. Other SMIs were tested leading to similar observations. Therefore, the SW is adopted for developing a first phase of exploratory analysis of the thirty three SMIs. Figure 4 depicts the absolute value of the SW of the $\{bvsp, dax, dji, ndx, nikkei, sp500\}$ indices, where for the sake of completeness, dji is repeated.

We verify that the charts are distinct from each other but reveal similarities namely, the emergence of several objects for low values of b. While for most cases, we have three objects some charts present several levels according with b following a logic of different scales of resolution.

4. MDS Analysis Based on Wavelets

After calculating the wavelet transform, it is possible to compare visually the plots and to establish a qualitative grouping by similarities. Nevertheless, it is preferable to define a quantitative measure avoiding subjective assessments. It should be noted that wavelets results are complex-valued. Figure 3 represents only the absolute value, but this approach is frequently adopted because it is simple and produces good results. In order to emphasize the shape of the wavelet plots, it was decided to normalize the charts, by converting the a and b axes into the interval $[0, 1]$ and by rescaling the wavelet absolute values so that the total volume becomes one. In other words, for each plot is considered for the x and y scale axes the values $a/\max(a)$ and $b/\max(b)$, for the z axis the values $\|[W_\psi x(t)](a,b)\| / \iint_{(a,b)} \|[W_\psi x(t)](a,b)\| da\, db$. Each plot can now be interpreted as a probability density function, and for comparing the normalized plots it is adopted the measure:

$$r_{ij} = \left[\left(\mu_{a,i} - \mu_{a,j}\right)^2 + \left(\sigma_{a,i} - \sigma_{a,j}\right)^2 + \left(\mu_{b,i} - \mu_{b,j}\right)^2 + \left(\sigma_{b,i} - \sigma_{b,j}\right)^2\right]^{1/2}, \quad i, j = 1, \ldots, 33, \tag{4.1}$$

where the symbols μ and σ represent the arithmetic average and standard deviation, (a, b) are the wavelet parameters, and i, j are listed in the set of financial indices. Based on the r_{ij} index, it is now possible to calculate a $\mathbf{R}_{33\times33}$ symmetric matrix of distances in the sense of (4.1) and to use a visualization tool for mapping the indices characteristics.

In order to reveal possible relationships between the SMIs the MDS technique is used. In this perspective, several MDS criteria are tested. The Sammon criterion revealed good results and is adopted in this work [18, 19].

4.1. Using MDS and the SW

Figure 5 depicts the two-(a) and three-(b) dimensional maps generated by the MDS [6, 8, 20–23].

Usually the MDS output quality is evaluated by the Sheppard and stress diagrams. The first plots the distances *versus* the original dissimilarities, and the second plots a measure of the mapping difficulty (called stress) *versus* the number of dimensions in the MDS

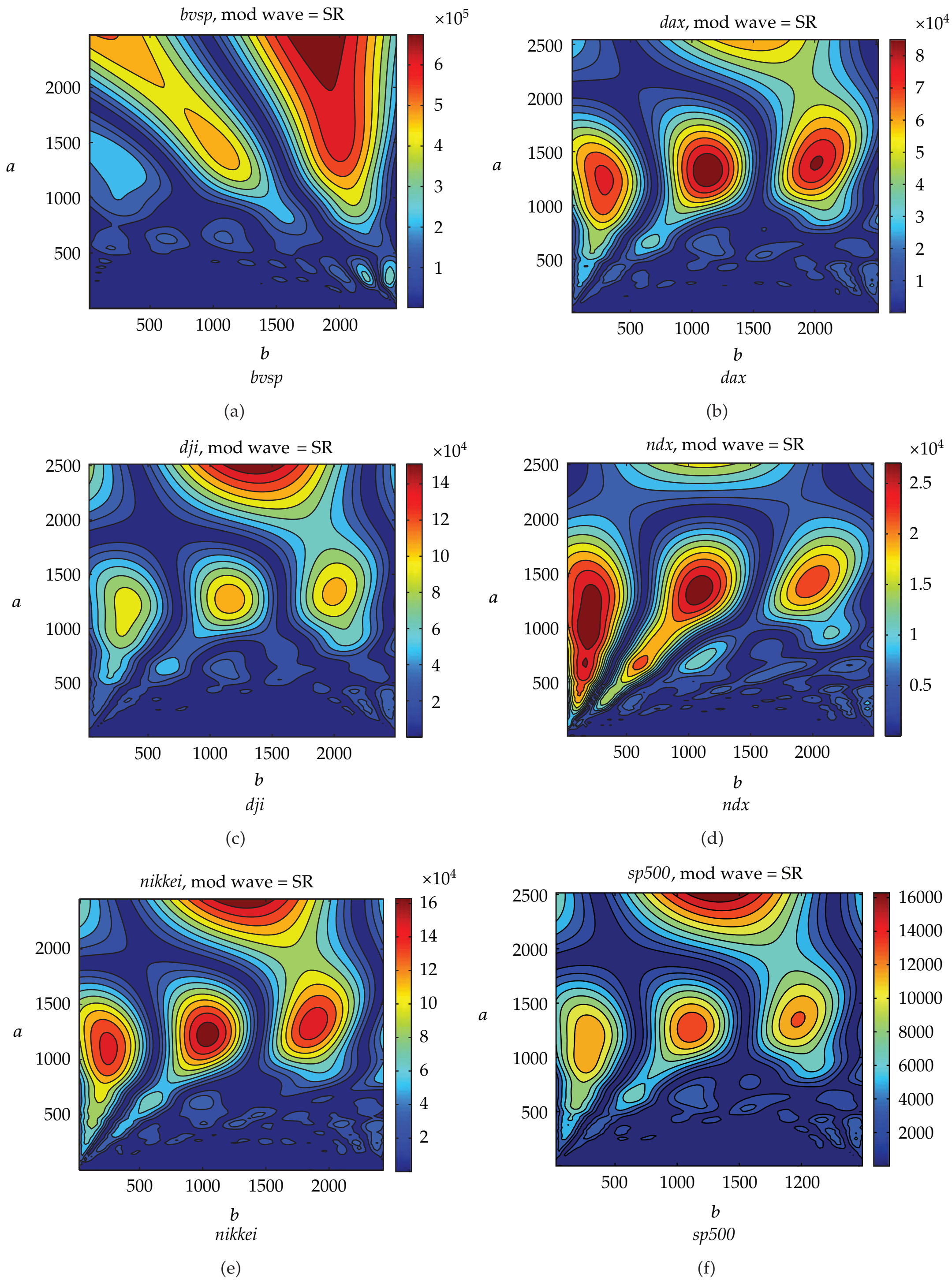
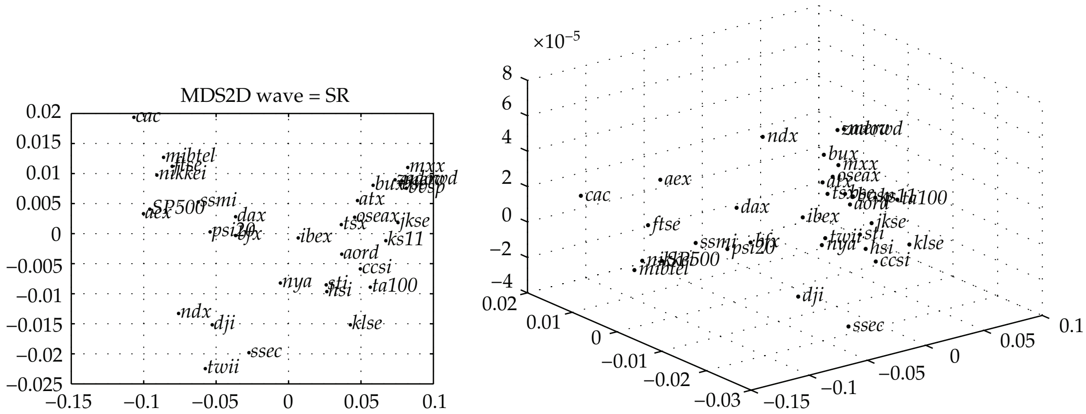


Figure 4: Absolute value of the SW *versus* (a, b) for the indices $\{bvsp, dax, dji, ndx, nikkei, sp500\}$, with $dt = 1$.



(a) (b)

Figure 5: Two-(a) and three-(b) dimensional MDS map of the thirty-three indices.

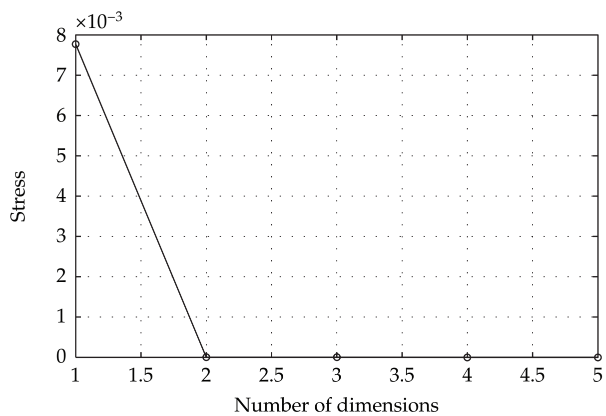


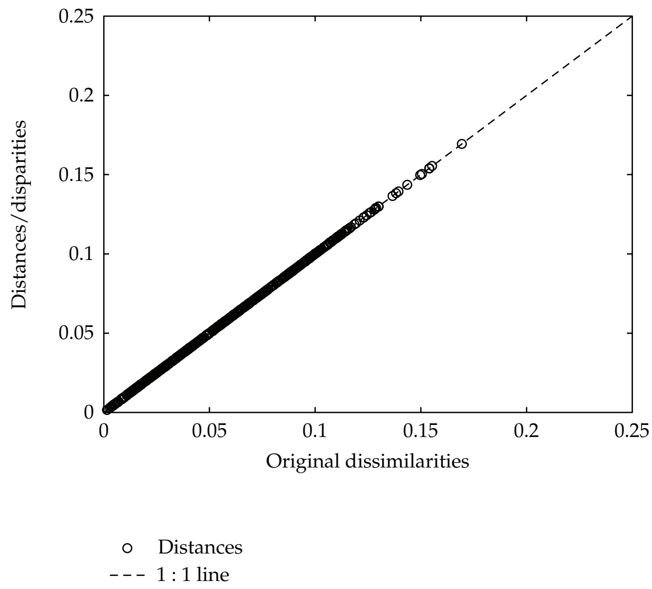
Figure 6: Stress value *versus* number of plotting dimensions of the MDS map using SW.

representation. Obviously the closer to the 45-degree line, in the first case, or the lower the values, in the second case, the better is the MDS map. In this perspective, Figures 6 and 7 show the corresponding stress and Sheppard diagrams using SW, respectively, demonstrating a good fit of the two- and three-dimensional MDS maps [24].

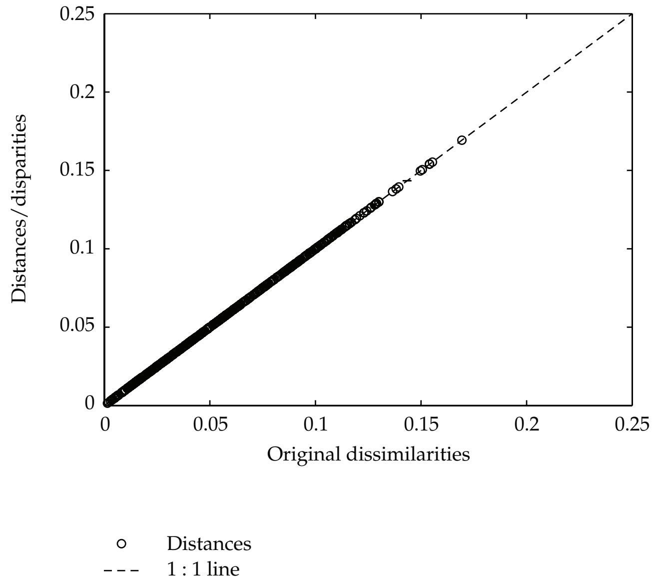
The aim of the MDS technique is to project the high-dimensional information into a low-dimensional space while preserving a good accuracy. Therefore, two-dimensional plots are sufficient, which will ease the comparison.

4.2. Using MDS and the HW, RW, HHW, SCW, and MW

We tested the SW for calculating the matrix of distances and plotting the MDS chart. The choice of the SW was merely based on a visual characteristics of the wavelet plots that revealed structured features. It is well known that the choice of a particular wavelet depends heavily on the application. In spite of the efforts that have been done in automating this choice, the fact is that often the best method is simply to experiment all wavelets. Having



(a)



(b)

Figure 7: Shepard diagram of the 2D-MDS (a) and 3D-MDS (b) using SW.

these ideas in mind, Figure 8 shows the two-dimensional MDS plots for all wavelets (the SW MDS map is repeated for easing the comparison).

Comparing the six MDS plots, we verify that we get a distinct chart for each wavelet. This result is usual since MDS depends on the analysis index, which in our case corresponds to the type of wavelet (2.3) and the comparison measure (4.1). Furthermore, MDS plots must be interpreted only on the basis of clustering of points since they are insensitive to translations and rotations. Therefore, we observe that the MDS chart based on the HW has

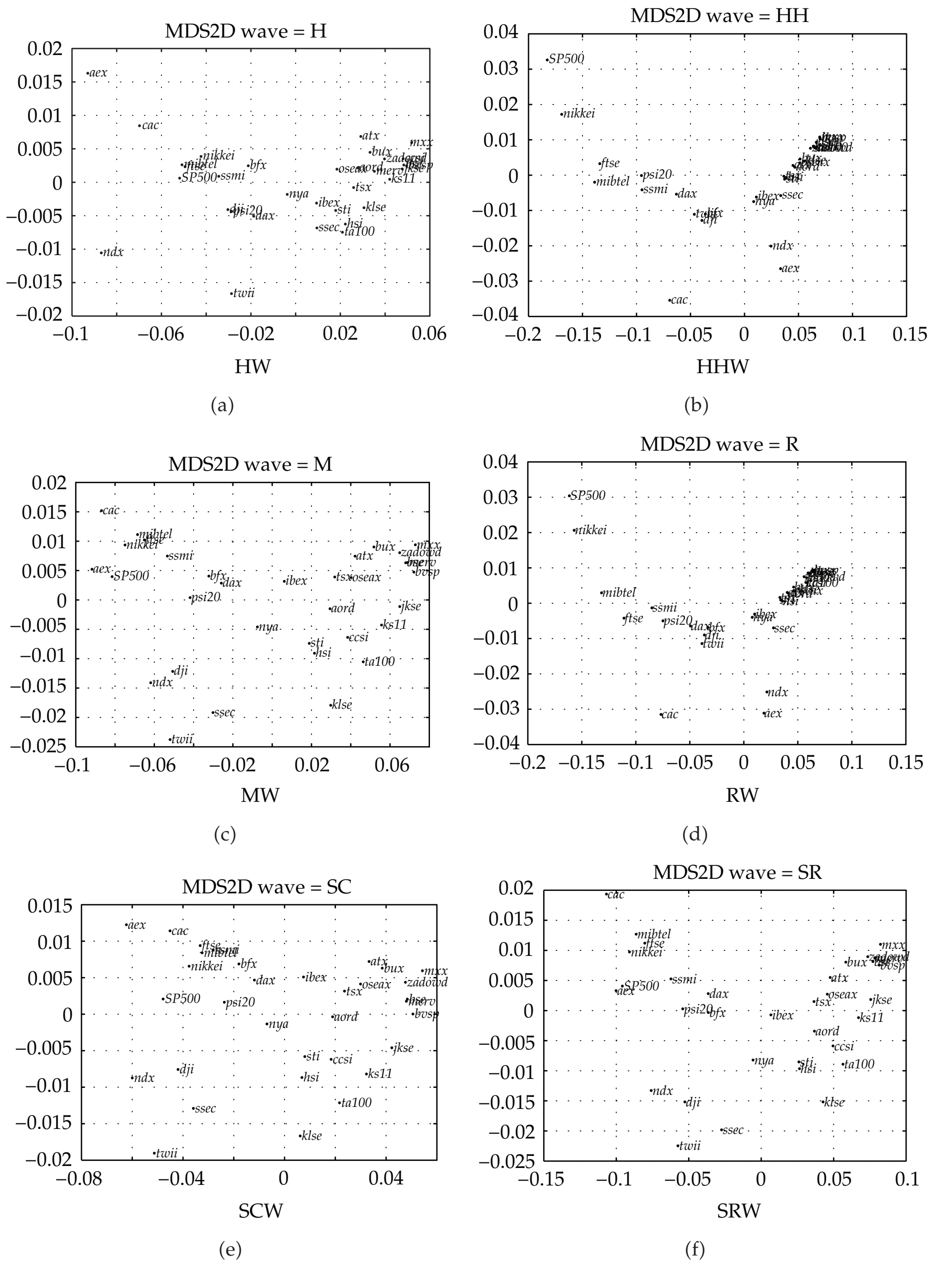


Figure 8: Two-dimensional MDS map, of the thirty-three indices, with the six wavelets {HW, HHW, MW, RW, SCW, SRW} listed in (2.3).

the most different features. Again, this result is common since it is known that the HW is more adapted to analyse digital signals. In what concerns the other five charts, the classification of the "best" one is more complicated since the clusters are not so different. Often MDS users prefer to have simple images, that is, with some clusters (that make sense for the application) but without an heavy overlapping of points. In this perspective, the SW is a good choice in the authors opinion, but SMI experts working in different areas may have distinct empirical choices. In fact, the adopting possibility of distinct measures based on the same methodology is one of the key features of the MDS scheme.

Considering the SW, there are several empirical conclusions we can draw from the MDS graph (Figure 8), and we will mention just a few here. We can clearly observe that there seem to emerge clusters, which show similar behavior [19]. Hence, there does not seem to be a single behavior market, but perhaps there are several important behaviors markets according their characteristics. We can say that $\{ndx, dax, psi20, aex, ssec, mibtel, cac\}\{atx, mxx, bux, zadowd, merv, bse, ccsi, bvsp, kse\}$ form separate clusters while $\{ibex, ssmi, ftse, nikkei, twii, hsi, bfx, tsx, sti, nya, klse, sp500, sji\}$ seem to form the "center cluster." Furthermore, the indices such as $\{aord\}$, $\{ks11\}$, $\{oseax\}$, and $\{ta100\}$ are separated from those groups.

5. Conclusion

It seems that there are many distinct analogies between the dynamics of complex physical and financial systems. This information can be analyzed with tools usually adopted in dynamical systems and signal processing. In this paper was studied the evolution of financial indices by means of continuous wavelets. The application to the thirty-three SMI by means of six different wavelets revealed the dynamical characteristics. After comparing the distinct possibilities, the Shannon real wavelet was adopted for guiding the study. For comparing the results, an index inspired in probability theory was defined, and the MDS visualization technique was applied. The charts lead to the emergence of patterns and clusters capable of being interpreted and compared. Having established the processing methodology the analysis was repeated for the other wavelets, and the results were compared.

Acknowledgments

This work is supported by FEDER Funds through the "Programa Operacional Factores de Competitividade—COMPETE" program and by National Funds through FCT "Fundação para a Ciência e a Tecnologia" under the projects: FCOMP-01-0124-FEDER-PEst-OE/EEI/UI0760/2011.

References

[1] R. R. Nigmatullin, "Universal distribution function for the strongly-correlated fluctuations: general way for description of different random sequences," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 3, pp. 637–647, 2010.

[2] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. Eugene Stanley, "Econophysics: financial time series from a statistical physics point of view," *Physica A*, vol. 279, no. 1, pp. 443–456, 2000.

[3] R. Gençay, F. Selçuk, and B. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press, New York, NY, USA, 2002.

[4] A. Sharkasi, H. J. Ruskin, and M. Crane, "Interrelationships among international stock market indices: Europe, Asia and the Americas," *International Journal of Theoretical and Applied Finance*, vol. 8, no. 5, pp. 603–622, 2005.

[5] I. Borg and P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, Springer, New York, NY, USA, 2005.

[6] T. Cox and M. Cox, *Multidimensional Scaling*, Chapman & Hall, Washigton, DC, USA, 2001.

[7] J. B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, vol. 29, pp. 1–27, 1964.
[8] J. Kruskal and M. Wish, *Multidimensional Scaling*, Sage Publications, Newbury Park, Calif, USA, 1978.
[9] J. W. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Transactions on Computers*, vol. 18, no. 5, pp. 401–409, 1969.
[10] http://finance.yahoo.com.
[11] S. R. D. R. Jaffard and Y. Meyer, *Wavelets: Tools for Science & Technology*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 2001.
[12] G. G. Walter, *Wavelets and Other Orthogonal Systems*, CRC, New York, NY, USA, 2000.
[13] G. G. Walter, *Wavelets and Signal Processing: An Application-Based Introduction*, Springer, New York, NY, USA, 2005.
[14] A. Cohen and R. D. Ryan, *Wavelets and Multiscale Signal Processing*, Chapman & Hall, London, UK, 1995.
[15] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms*, Prentice Hall, New York, NY, USA, 1998.
[16] C. K. Chui, *Wavelets: A Mathematical Tool for Signal Processing*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 1997.
[17] Y. Nievergelt, *Wavelets Made Easy*, Birkhäuser, Boston, Mass, USA, 1999.
[18] F. B. Duarte, J. A. Tenreiro MacHado, and G. Monteiro Duarte, "Dynamics of the Dow Jones and the NASDAQ stock indexes," *Nonlinear Dynamics*, vol. 61, no. 4, pp. 691–705, 2010.
[19] J. T. MacHado, G. M. Duarte, and F. B. Duarte, "Identifying economic periods and crisis with the multidimensional scaling," *Nonlinear Dynamics*, vol. 63, no. 4, pp. 611–622, 2011.
[20] A. Buja, D. F. Swayne, M. L. Littman, N. Dean, H. Hofmann, and L. Chen, "Data visualization with multidimensional scaling," *Journal of Computational and Graphical Statistics*, vol. 17, no. 2, pp. 444–472, 2008.
[21] S. Nirenberg and P. E. Latham, "Decoding neuronal spike trains: how important are correlations?" *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 12, pp. 7348–7353, 2003.
[22] J. O. Ramsay, "Some small sample results for maximum likelihood estimation in multidimensional scaling," *Psychometrika*, vol. 45, no. 1, pp. 139–144, 1980.
[23] J. Woelfel and G. A. Barnett, "Multidimensional scaling in Riemann space," *Quality and Quantity*, vol. 16, no. 6, pp. 469–491, 1982.
[24] R. N. Shepard, "The analysis of proximities: multidimensional scaling with an unknown distance function. I," *Psychometrika*, vol. 27, pp. 219–246, 1962.