department of economics
"S. Cognetti de Martiis"

Via Po, 53 – 10124 Torino (Italy)
Tel. (+39) 011 6704043 - Fax (+39) 011 6703895
URL: http//www.de.unito.it

WORKING PAPER SERIES

OLD AND NEW SPECTRAL TECHNIQUES FOR ECONOMIC TIME SERIES

Lisa Sella

Dipartimento di Economia "S. Cognetti de Martiis"

Working paper No. 09/2008

Università di Torino

OLD AND NEW SPECTRAL TECHNIQUES FOR ECONOMIC TIME SERIES

Lisa Sella

May, 2008

e-mail address: lisa.sella@unito.it

Abstract. This methodological paper reviews different spectral techniques well suitable to the analysis of economic time series. While econometric time series analysis is generally yielded in the time domain, these techniques propose a complementary approach based on the frequency domain. Spectral decomposition and time series reconstruction provide a precise quantitative and formal description of the main oscillatory components of a series: thus, it is possible to formally identify trends, low-frequency components, business cycles, seasonalities, etc. Since recent developments in spectral techniques allow to manage even with short noisy dataset, nonstationary processes, non purely periodic components these tools could be applied on economic datasets more widely than they nowadays are.

Keywords. Time series analysis; Spectral Methods; Frequency domain; Singular Spectrum Analysis; Time series decomposition; Denoising.

JEL Classification. C15; C49; C8; E32

1. Introduction

Dynamical properties of the economic systems are generally inquired in the time-domain by analyzing the projections of interesting functions of time onto the phase space. However, information obtained by the time-domain analysis could be effectively supplemented by a frequency-domain approach, i.e. spectral analysis (Granger and Hatanaka, 1964; Priestley, 1981; Press et al., 1986; Medio, 1992; Pollock, 2008). This methodology has been developed in scientific fields such as digital signal processing, oceanography, meteorology, and so on, and lies on the remark that most regular behaviour of a time series is periodic; thus, the periodic components embedded in the analyzed series can be determined by computing their periods, amplitude, and phase (Ghil et al., 2002). The main assumption underlying this use of time- and frequency-domain analysis techniques to inquire global system dynamics is that the series *X(t)* we actually observe is an effective realization of the underlying stochastic data generating process *X*. Thus, with some care the inference on specific series can be extended to the whole system.

Therefore, spectral tools are particularly attractive for applied economic inquiries, since they allow to investigate crucial issues like trend-cycles separation, fundamental cycles extraction, seasonal variability, sectorial contribution to growth, denoising, and so on (Lisi and Medio, 1997; Higo and Nakada, 1998; Atesoglu and Vilasuso, 1999; Baxter and King, 1999; A'Hearan and Woitek, 2001; Croux et al., 2001; Gerace, 2002; Aadland, 2005). All these critical economic topics have found some answers within the time-domain econometric analysis framework, but our discipline would have great benefit from a more diffuse and complementary use of the frequency-domain analysis. In fact, the high descriptive power of such techniques allows a rigorous definition, quantification, and extraction of long, medium, and short term components of a time series, thus favouring the inspection of both cyclical phenomena and lead-lag relations (Iacobucci, 2003). This feature is particularly attractive whenever some precise and adjustable description of economic fluctuations through formal decomposition is needed and it could considerably help to enlighten both causal relations and feedbacks among the economic variables.

Clearly, the phenomenological nature of such inquiry does not need any *a priori* modelling, thus allowing the researcher to inspect the dataset in an inductive manner. Anyhow, some spectral techniques call for covariance stationary series, whose first and second moments do not depend on time. Since such requirement is not always satisfied in time series, a relaxation of this assumption is desirable. Thus, some new spectral tools have been developed to analyze nonstationary time series: among them Singular Spectrum Analysis (SSA) is an innovative flexible data-adaptive method allowing spectral decomposition in short, noisy, and chaotic time series without particular characteristics. This is especially appreciable in economics, where most series are short and noisy. Notwithstanding evidence of chaos in economic datasets has not been confirmed yet, Broomhead and King (1986) demonstrate that SSA works well even with mildly nonlinear data, as economic series effectively are (Neftci, 1984; Brock, 1986; Neftci and McNevin, 1986; Brock and Sayers, 1988; Frank and Stengos, 1988; Serletis, 1996; Acemoglu and Scott, 1997; Altissimo and Violante, 1998; Stanca, 1999; Stock and Watson, 1999; Brock, 2000; McConnell and Perez-Quiros, 2000; Piselli, 2004). Moreover, SSA is successful at different time scales, allowing the researcher to inspect the series at various resolutions. Thus, a more widespread application of spectral tools in economics is desirable to both strengthen and improve the findings on the time-domain analysis side.

Going deeper into the topic, the typical spectral analysis tool is the power spectrum, which records the contribution to the process total variance of components belonging to different frequency bands. Thus, power spectral analysis is generally applied to inspect the fraction of total information carried by different frequencies (Granger, 1966; Medio, 1992). In fact, the *Wiener-Khinchin theorem* implies that the autocorrelation function and the power spectrum of a function of time $X(t)$ contain both the same information about X. Then, from the *Parseval theorem* the total power of a function of time is the same in both the time and frequency domain. However, among the drawbacks of power spectrum analysis there is the difficulty to distinguish between quasi-periodic and chaotic signals, as much as between low- and high-dimensional chaotic attractors; moreover, some information contained in the original signal is usually lost.

The techniques surveyed in section 2 concern some power spectrum estimation methods, which are consequently subject to the above drawbacks; on the contrary, methods in sections 4-6 allow in addiction to reconstruct the main components of the series, thus providing a rigorous determination of their periodic and/or quasi-periodic characteristics.

This paper is a methodological survey, thus the emphasis is on spectral methods and their characteristics rather than on the existent empirical applications: the aim is to introduce the reader to the advantages and weaknesses of some frequency-domain techniques, highlighting the attempts to improve these tools whenever some features of the series compromise the statistical reliability of the analysis. Note that in this framework robustness is achieved just when the analyses performed with different methods lead to similar results.

The next sections are organized as follows. Section 2 provides a brief review of classical spectral estimation, highlighting some feature of the standard Fourier approach solved by successive spectral tools. Section 3 surveys the Maximum Entropy Method, which underlies the close connection between power spectrum and auto-regressive processes. In section 4 the problems of signal reconstruction and precise localization of purely periodic and narrowband components is addressed. The SSA method, diffusely explained in section 5, is a data-adaptive tool allowing a reliable signal reconstruction and a precise quantification of trend, low-frequency components, dominant cycles, and seasonal movements. Some drawbacks of the method are solved through Monte Carlo simulations (section 6), which implement testing procedures against different types of noise processes. Finally, section 7 provides a simple description of the wavelet method, which performs a double decomposition of the signal into the time/frequency space and consequently proves very useful in the analysis of multiscale nonstationary processes. Section 8 concludes.

2. Brief review of classical spectral estimation

Following the classical Fourier interpretation, signals can be represented as linear superpositions of sinusoidal modes, each with its own frequency, amplitude, and phase. Thus, a *p-periodic function* $f_p(t)$ can be expressed by

$$f_p(t) = \sum_{k=-\infty}^{+\infty} a_k \exp(2\pi ikt/p), \tag{1}$$

where $i = \sqrt{-1}, |t| < \frac{p}{2}$. The complex coefficients a_k are determined by

$$a_k = \frac{1}{p}\int_{-p/2}^{p/2} f_p(t)\exp(-2\pi ikt/p)dt, \tag{2}$$

i.e. each a_k represents the contribution of the *k*-th frequency component to the periodic function (1).

An extension to non-periodic functions treats them as infinitely periodic, i.e. $p \to \infty$. With appropriate transformations of (1) and (2) we obtain the so-called *Fourier pair*, i.e. both a frequency- and time-domain representation of the underlying process respectively giving its amplitude and phase as a function of frequency ω, and its values as a function of time t:

$$F(\omega) = \int_{-\infty}^{+\infty} \exp(-2\pi i\omega t) f(t)dt \text{ and } f(t) = \int_{-\infty}^{+\infty} \exp(2\pi i\omega t) F(\omega)d\omega .$$

This pair represents the basis for time series interpretation in spectral analysis.

Ghil et al. (2002, p. 16) underlie how

“a very irregular motion possesses a smooth and continuous spectrum, which indicates that all frequencies in a given band are excited by such process. On the other hand, a purely periodic or quasi-periodic process is described by a single line or a (finite) number of lines in the frequency domain. Between these two extremes, nonlinear deterministic but `chaotic´ processes can have spectral peaks superimposed on a continuous and wiggly background”.

Thus, our aim is to distinguish among different spectral components in the analysis of economic nonlinear time series, in order to single out the main cyclicalities and trend characterizing the economic system. This is fundamental for the investigation of many economic phenomena, from macroeconomics to finance, form sectorial studies to environmental economics, and so on. In fact, with proper spectral methods we could inductively enhance our empirical comprehensions of the structure and cyclical behaviour of the series at different time scales, together with some causal and feedback relations among economic variables (cf. e.g. Granger and Hatanaka, 1964; Granger, 1969; Klotz and Neal, 1973; Iacobucci, 2003). As an example, Pollock (2008) reflects on the nature of the business cycle and its eventual relationship with trend by means of some different detrending methods based on the discrete Fourier transform, underlying that “a clear understanding of the business cycle can be achieved only in the light of its spectral analysis” (Pollock, 2008, p.2). On the same line, Wang

(1999) proposes an innovative measure of persistence in economic time series based on the discrete Fourier transform of their whole spectrum, rather than on statistics at the zero frequency only. Concerning financial applications, e.g. Chiarella and El-Hassan (1997) and Takahashi and Takehara (2008) apply Fourier transform algorithms in order to evaluate currency options and derivative security prices.

2.1 Blackman-Tukey windowed correlogram analysis

Classical spectral techniques generally estimate the power spectrum of a series by its *periodogram*, i.e. the squared modulus of its direct discrete Fourier transform. However, power leakage (see below in the section) induces systematic distortions and other problems with the estimate variance produce inconsistency.

Thus, Blackman and Tukey (1958) derived an indirect non-parametric method based on the *Wiener-Khinchin theorem*, which states that the lag autocorrelation function of a time series and its power spectrum are Fourier transforms of each other, i.e. respectively the time- and frequency-domain representations of the same process. By a windowed fast Fourier transform, the Blackman-Tukey method (BT) provides the series *correlogram*, i.e. a power spectrum estimate through its autocorrelation function. This algorithm is particularly advantageous with respect to other classical techniques since it reduces the estimate variance and bias by weighting over non-overlapping bins in the frequency domain (Chatfield, 1984). The resulting power spectrum is

$$P(\omega) = \frac{1}{m}\left|\sum_{j=0}^{m-1} \rho_j w_j \exp(i\omega j)\right|^2 ,$$

where m represents the window length, i.e. the maximum lag considered, ρ_j the autocorrelation function, and w_j the windowing function. Both the windowing function and the window length are crucial issues, since a proper choice of the filter is fundamental for consistence and low bias (Percival and Walden, 1993). Moreover, in the application of this technique there is a clear trade-off between resolution and the estimate variance: the lower m, the higher the number of independent samples we can extract from the series, the lower the estimate variance; but on the other side, too narrow windows may loose relevant features of the signal.

Ghil et al. (2002) underlie that BT is quite an efficient estimate of the continuous part of the spectrum, while pure sinusoidal components are hardly detected because of low resolution. Obviously this constitutes an heavy drawback in economic applications, since we are mainly interested in identifying cyclical regularities of the series, i.e. their periodic or quasi-periodic components. Thus, in the early 80s David Thomson developed the Multi-Taper Method (MTM, section 4 below), an extension of the BT algorithm actually maximizing resolution by averaging the series along a special set of optimal functions (Thomson, 1982).

Concerning economic variables, in a pioneering work Granger (1966) observes that they possess a typical spectral shape: regardless of the time series length and the size of the filtering window, their spectrum shows a large bump at the zero frequency band, which is quickly but smoothly reabsorbed in the next bands. This is due to *leakage* phenomena in spectral estimation of time series showing a large peak at some frequency: in fact, some of the power associated with such frequency actually leaks into the estimate of neighbouring frequency bands. In most economic time series leakage is due to the fact that trend usually represents by far the largest portion of the total variance, i.e. a large bump in the power spectrum.

Thus, apart from the technical drawbacks of classical spectral estimation mentioned in section 1, an extension to recent spectral reconstruction techniques helps the economist to curb spectral leakage, detrending the series and defining the features of its main fluctuations (sections 4-6 below).

3. Maximum Entropy Method (MEM)

Another way to solve the shortcoming of resolution in spectral estimation methods consists in exploiting the close connections among Information Theory, entropy measures, and nonlinear complex systems analysis (Golan, 2002). In fact, MEM consists in a high-resolution estimate of the power spectrum by a stepwise extrapolation of the corresponding autocorrelation function. At each step the estimate is carried out in order to maximize the information entropy of the autocorrelation probability density function, i.e. the measure of uncertainty associated to the autocorrelation function (Burg, 1967; van den Bos, 1971).

Applied to a time series *X(t)*, MEM consists in finding the maximizing joint probability density subject to the time series constraints, since conceptually all information embedded in the series represents a constraint. Jaynes (1982) deals with two types of information set particularly widespread in economic applications: the means of different quantities, and the estimated autocovariances. In this last case Jaynes' formulation allows both to derive the optimal interpolation of missing values and to predict autocovariances not belonging to the *information gathering region I*, which is an important aspect for economic forecasting.

The MEM estimated spectrum is

$$P(\omega) = \frac{1}{\sum_{k \in I} \lambda_k \exp(i\omega k)}, |\omega| \leq \pi, \quad (3)$$

representing the smoothest and fail-safe spectrum consistent with the data (Burg, 1967). In this formula λ_k represents the Lagrange multiplier associated to the *k*-th constraint of the optimization problem. An important feature of (3) is that the more λ_k tends to zero, the less contributing to the power spectrum structure is the corresponding *k*-th datum.

Thus, the relevant dataset in MEM is the significant independent information, while redundancies are automatically dropped out by owning zero potential. As a consequence, superimposing superfluous constraints does not change the final result (Jaynes, 1982). This is particularly interesting for econometric applications, since it drops out the overspecification problems usually compromising the performance of statistical techniques. Accordingly, Lagrange multipliers signal each datum potential measuring the importance of the respective constraint, i.e. the informative power carried out by the corresponding datum. In addiction, we can perform significance tests considering that the most true spectrum should be close to the MEM prediction, since by definition the great majority of all possible spectra share the information it contains. The tests are performed comparing the data entropy with the maximum entropy derived imposing the null hypothesis: a large discrepancy represents an evidence against the null.

3.1 MEM and Autoregressive models (AR)

Particularly interesting is the fact that the ME spectrum in equation (3) has the same analytical form of the corresponding *AR(m)*[1], where *m* represents the maximum lag of the relevant data, i.e. $\lambda_{m+1} \approx 0$ (Jaynes, 1982). Thus, by the ME principle the optimal spectrum estimate corresponds to an AR model: this circumstance represents a deep link between spectral estimation and autoregressive stochastic processes, which are particularly recurrent in economic modelling. Percival and Walden (1993) show that for an *AR(m)* process with generalized white noise disturbance

$$X_{N+m} = a_m X_{N+m-1} + ... + a_1 X_N + \eta_N , \eta_N \sim GWN(0, \sigma^2)$$

the associated power spectrum is

$$P_X(\omega) = \frac{\sigma^2}{\left|1 + \sum_{k=1}^{m} a_k \exp(i\omega k)\right|}.$$

Thus, MEM performs best whenever estimating line frequencies for time series actually generated by an AR process (Ghil et al., 2002). On the contrary, noisy data considerably complicate MEM estimation, since the anomalies often bias the analysis towards some small unrepresentative subclass of all the possible processes consistent with the dataset. In addiction, the number of detected spectral peaks generally increases with the MEM order *m*: various authors suggested to use information criteria to help the correct choice of *m* (Akaike, 1969; Akaike, 1974; Haykin and Kessler, 1983), but they often tend to either under or overestimate the proper order. Thus, Penland et al. (1991) propose a preliminary denoising of the time series through SSA (see section 5.2): this procedure should clean the series and facilitate MEM peak detection.

Summing up, MEM is a parametric method boiling down in the search of the AR process better fitting the original time series behaviour. Since the robustness of an empirical analysis generally comes from homogeneous results in both parametric and non-parametric methods, MEM is often used jointly with other spectral estimation techniques.

Concerning some existent economic applications, Callen et al. (1985) use MEM in order to evaluate the predictive power of AR models of spot rates with respect to random walk models: this technique is preferred to standard

Box-Jenkins' since it neither arbitrarily truncates the data in the time domain, nor imposes periodic extensions in the frequency domain, thus relieving distortions in structural change detections. Vinod (2006) elaborates a ME bootstrapping algorithm to avoid both shape-destroying transformations of time series like detrending, differencing, etc., and structural change and unit root tests with complicated asymptotics, while Wu and Stengos (2005) propose partially adaptive estimators via ME densities. Among other MEM applications to economics, Marchand (1985) evaluates the impact of monetary activities upon the Canadian regional housing markets, Paris and Howitt (1998) solve ill-posed production problems recovering flexible cost functions from very limited datasets, Paris (2001) proposes ME estimators unaffected by multicollinearity, which uniformly dominate the Ordinary Least Squares one, while Peeters (2004) compares generalized ME and Full Information Maximum Likelihood procedures allowing in addiction to treat the heterogeneity of observations.

4. Multi-Taper Method (MTM)

From this section on we introduce some spectral tools capable to both perform power spectrum estimation and to reconstruct the series dominant frequencies.

First of all, MTM is a non-parametric technique generally applied to time series characterized by both broadband and line spectral components, i.e.

respectively continuous parts and pure sinusoids (Thomson, 1982; Park, 1992; Percival and Walden, 1993). Thus, peak detection is no more limited to purely periodic components, which are generally just a small part of time series components.

Then, MTM provides more stable spectral estimates than BT (see section 2), since now the variance is further reduced by averaging the correlogram over a small set of orthogonal windows, the so-called *eigentapers*, rather than using just a single filter (Thomson, 1990b). This method is also more heuristic than BT (Box and Jenkins, 1970), since the optimal eigentapers are now the solution of a variational problem minimizing the finite-length power leakage outside of a proper frequency band, which can cause artificial high power detections at frequencies different from those of the true peaks (Ghil et al., 2002).

Since merely the first eigentapers effectively reduce spectral leakage (Slepian, 1978), in MTM the classical trade-off resolution vs. stability concerns the K eigenfunctions actually adopted and the choice of both the bandwidth and the effective number of tapers. Spectral estimates are performed firstly premultiplying the data by the K orthogonal eigentapers, thus obtaining K tapered series $\{X(t)w_k(t):\ t=1,\dots,N\}$; then averaging the individual spectra $S_k(\omega) \equiv |Y_k(\omega)|^2$, where $Y_k(\omega)$ is the k-th tapered series discrete Fourier transform. The resulting estimated power spectrum

$$S(\omega) = \frac{\sum_{k=1}^{K} \mu_k |Y_k(\omega)|^2}{\sum_{k=1}^{K} \mu_k}$$

detects line components until the chosen frequency band. Differently from previous methods, MTM directly determines the amplitude of each line in the spectrum even in processes with high noise background. This is particularly relevant, since it allows the quantitative reconstruction of periodic components even in highly noisy contexts, like the ones generally occurring in economics.

Assuming that the analyzed time series is the sum of a sinusoid of frequency ω_0 and amplitude A plus a noise background made up of other negligible sinusoids and white noise, i.e.

$$X(t) = A \exp(2\pi i \omega_0 t) + \eta(t),$$

a least squares fit in the frequency domain yields

$$\hat{A}(\omega_0) = \frac{\sum_{k=1}^{K} U_k^*(0) Y_k(\omega_0)}{\sum_{k=1}^{K} |U_k(0)|^2},$$

where $U_k(\omega)$ is the discrete Fourier transform of the k-th eigentaper, while the asterisk denotes complex conjugation. Related hypothesis testing allows both to significantly detect low-amplitude harmonic oscillations in short time series, and to reject large amplitude periodicities when the *F test* fails, since the value of this statistic does not depend on the magnitude of $A(\omega)$. This is a valuable improvement with respect to classical methods, whose F

values depended on the estimated amplitude (Jenkins and Watts, 1968). Kendall and Stuart (1977) define statistical confidence intervals for the estimated amplitude based on the ratio of the variance captured by the filtered time series to the residual variance. For hypothesis testing the null could be $A=0$, i.e. $X(t)$ is a white noise process.

The Thomson (1982)'s MTM conventional assumption that $X(t)$ consists of a finite superposition of separate, purely periodic, fixed-amplitude components and (locally) white noise is quite restrictive, since quasi-periodic and other components are widely diffuse in time series. Thus, in standard MTM a continuous spectrum (e.g. that of coloured noise) is resolved into spurious lines incorrectly associated to arbitrary frequencies. However, since presumably most narrowband variability is not associated with strictly harmonic purely periodic behaviour, Mann and Lees (1996) propose to combine conventional harmonic analysis, which works well just in the case of low signal-to-noise ratios (MacDonald, 1989), with additional criteria allowing to detect significant narrowband quasi-oscillatory components as well.

Thus, Mann and Lees' innovative procedure bears to distinguish among harmonic, anharmonic, and background noise components; such feature is particularly valuable in signal reconstruction (see section 4.1). This new method tests all the detected peaks against a red noise background null hypothesis, whose variance and lag-one autocorrelation are directly estimated from the data. First of all the spectral background is robustly estimated by minimizing the misfit between the AR(1) red noise empirical spectrum and the adaptively weighted multitaper spectrum, which is smoothed to become insensitive to outliers (Ghil et al., 2002). Then, the significance levels against the estimated noise are determined; finally, a reshaped spectrum isolates narrowband, possibly intermittent, amplitude-and-phase-modulated oscillations from harmonic phase-coherent sinusoids. Finally, note that such algorithm allows to filter out long timescales, which is not necessarily a desirable property for the analysis of the economic systems. On the other side, an highly desirable property is the possibility to reconstruct the signal dominating components, i.e. the ones which are significantly different from the data-adapted background noise.

4.1 Signal reconstruction

Similarly to SSA techniques (see section 5), it is possible to reconstruct in the time domain the signal components detected by MTM. However, while in this case the information comes from a frequency-domain decomposition, in SSA the reconstruction is carried out in the lag domain (Ghil et al., 2002). Signal reconstruction in the time domain is particularly useful for economic purposes, since it allows the researcher to focalize on each spectral component, defining the characteristics of trend, low-frequency fluctuations, business cycles, and seasonal oscillations. Thus, time series analysis gains through spectral reconstruction the ability to quantitatively inspect periodicity, phase, and amplitude of all the significant cyclicalities of the original signal above a properly chosen background noise.

In MTM, the discrete-time reconstructed signal centred in ω_0 is given by

$$\tilde{X}(N\Delta t) = \Re\{A_N \exp(-2\pi i \omega_0 N\Delta t)\},$$

where A_n is the complex envelope determined from a problem involving both the complex amplitudes of each of the K MTM-eigenspectra and appropriate boundary conditions (Park, 1992). This reconstruction method represents and improvement of classical demodulation methods (Bloomfield, 1976).

Among existent MTM economic applications, Movshuk (2003) tries to assess the distortive effect of detrending methods in demand analysis, Harris and Poskitt (2004) identify the cointegration rank of partially non-stationary processes, Gan and Zhang (2005) explore the effect of thick markets on local unemployment rate fluctuations.

5. Singular Spectrum Analysis (SSA)

SSA is a fully non-parametric method especially well suiting for short, noisy, and chaotic time series (Vautard and Ghil, 1989; Vautard et al., 1992). It is a frequency-domain statistical technique related to the Empirical Orthogonal Functions (EOFs) analysis (North et al., 1982): assuming time series *stationarity* and examining just enough realizations of a process *X(t)*, we can learn some information about its probability distribution (Ghil and Mo, 1991).

One of SSA main advantages with respect to classical Fourier methods is its ability to detect oscillations *modulated* both in amplitude and phase (Allen and Smith, 1996). Thus, the original signal is no more simply decomposed into periodic sine and cosine functions, but rather into data adaptive waves possibly exhibiting non-constant amplitude and/or phase.

Initially introduced as a data analysis technique in digital signal processing, SSA has been applied to several scientific fields, including financial and macroeconomic time series analysis. In this peculiar case, in fact, the problem of a substantial lack of observations, which e.g. traditionally induced eminent scholars to consider business cycles as inherently unpredictable (Fisher, 1925; Slutsky, 1937), could be at least partially solved, since SSA well applies to short time series. However, the method is not conceived to build models, but rather to identify information about the time series deterministic and stochastic parts (Ormerod and Campbell, 1997). In particular, SSA should both accurately forecast the short-term system evolution and capture its long-term features, highlighting some of the system peculiar properties, such as its degree of randomness (Gershenfeld and Weigend, 1993).

As mentioned in the introduction, one main problem with SSA application to economic data could concern their not verified chaotic nature, since many studies have detected nonlinearities in macroeconomic time series, but no evidence of chaos has been clearly confirmed yet (Neftci, 1984; Brock, 1986; Neftci and McNevin, 1986; Brock and Sayers, 1988; Frank and Stengos, 1988; Serletis, 1996; Acemoglu and Scott, 1997; Altissimo and Violante, 1998; Stanca, 1999; Stock and Watson, 1999; Brock, 2000; McConnell and Perez-Quiros, 2000; Piselli, 2004). However, Broomhead and King (1986) show that SSA works well even with mildly nonlinear noisy data, as economic time series seem to be (Lisi and Medio, 1997; Ormerod and Campbell, 1997).

As an example, Ormerod and Campbell (1997) apply SSA to UK and US quarterly GDP growth rates. They find no underlying structure in the UK dataset, while some small sign of regularity is detected for the US, although not so clear as to contradict the hypothesis of no predictability for the business cycle. However, Palm (1997) observes that this ambiguous result could be essentially due to GDP aggregation, which clearly tends to flatten the series power spectrum.

Other economic applications exploit the filtering characteristics of SSA: Lisi and Medio (1997) use its denoising features to improve their out-of-sample short-term predictions of the exchange rate. On the financial side Thomakos et al. (2002) apply SSA to daily realized futures volatility. They clearly describe such tool as a non-parametric method particularly useful in decomposing volatility into economically significant components, succeeding in identifying both a trend and some significant cycles in their S&P500 and Eurodollar series.

5.1 Technical introduction

Since no *a priori* structure is imposed on the data, SSA provides both qualitative and quantitative information about the deterministic and stochastic parts of the underlying system without requiring prior knowledge of its characteristics (Broomhead and King, 1986; Vautard and Ghil, 1989). SSA principal goal consists in the identification and distinction of pure signal from noise in the analysis of nonlinear dynamics, both in univariate and multivariate contexts (*single-channel* and *multi-channel SSA*, respectively); here we cover just the single-channel method, since the multi-channel is simply an extension of it (Lisi and Medio, 1997).

Strictly speaking, SSA is a *linear* method providing the orthogonal decomposition of the time series lag-covariance matrix, i.e. a projection of its principal components onto the vector space of the time series delay-coordinates (Vautard and Ghil, 1989; Vautard et al., 1992). Intuitively, SSA provides a decomposition of the time series into orthogonal modes characterized by different dynamical properties. To identify periodic or quasi-periodic activity in the signal Vautard and Ghil (1989) use the eigenvalue spectrum, i.e. a plot of the time series eigenvalues ranked by order with the correspondent confidence bars: the *near-equality* of an eigenvalue pair in phase quadrature *may be associated* with an oscillatory movement (Ghil and Mo, 1991). Intuitively, each eigenvalue represents the fraction of total variance explained by the associated component: if two components explain more or less the same variance and their modes are in phase quadrature, they may represent an oscillatory pair. Thus, with some additional techniques we can identify both slow modes, which presumably represent trends, and either regular or irregular oscillations from a background noise and/or uninteresting processes (Vivaldo, 2007). This is clearly innovative with respect to the classical *Fourier analysis*, where the original series is projected onto a basis of simple sine and cosine functions (Stein and Weiss, 1971): now irregular modulated fluctuations can be treated as well. Moreover, the fully data-adaptive framework of the method allows to manage nonstationary time series too, thus avoiding the additional noise carried by the eventual transformation of an integrated series. With respect to MEM, SSA appreciably allows to reconstruct in the time domain any signal decomposition onto the chosen vector space, independently of its significance with respect to a null hypothesis background noise.

Going deeper into technical details, applying a sliding *m*-window to an N_T-length time series we obtain a sequence of $N=N_T-m+1$ vectors [2] $\{\boldsymbol{x}_i \in \Re^m | i = 1,2,...,N\}$ in the embedding space $\Re^m$, i.e. the space of all the *m*-elements patterns (Broomhead and King [1986]). Then we define a linear map $\boldsymbol{X} : \Re^m \rightarrow \Re^N$ called *trajectory matrix*, i.e.

$$\boldsymbol{X} = N^{-1/2} \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_N' \end{bmatrix}.$$

The triple *(X,* $\Re^N$ *,* $\Re^m$*)* can be analyzed through singular system analysis methods such as SSA (Broomhead and King, 1986). However, there are several aspects an accurate SSA must tackle over.

5.2 Dimensionality and Principal Components Analysis

An important issue in SSA is *dimensionality* since the number of degrees of freedom, i.e. the number of modes effectively containing significant power, is crucial in distinguishing chaotic from stochastic systems (Packard et al., 1980).

In economics, for example, the concept of *deterministic chaos* proves particularly useful (Brock and Sayers, 1988; Medio, 1992), although a reliable empirical verification is quite difficult to implement. Moreover, since signals are finite-sampled and generally perturbed by noise, the way we manage dimensionality is fundamental to recognize our system dynamics. A possible approach consists in computing the number of significant mutually-orthogonal directions of the reconstructed attractor, i.e. the dataset *statistical dimension S*. However, as long as noise and finite sampling affect the system, this dimension is not an invariant characteristic (Vautard and Ghil, 1989).

S is directly connected with Principal Components Analysis (PCA) and SSA. In fact, from a statistical point of view the dynamics of the data points covering the embedded attractor can be linearly described by its principal axes. This issue concerns the dimension of the reconstructed phase space: in economic terms, it corresponds to identifying a small number of variables controlling the dynamics of our system (Medio, 1992). However, most directions in the embedding space are noise-dominated and can thus be neglected without significant information loss (Vautard and Ghil, 1989).

Following PCA, the phase space principal directions are determined solving the eigenvalue problem for the sampled time series embedded in $\Re^m$. A more detailed exposition can help the reader to make some parallels with more familiar econometric tools. As Broomhead and King (1986) explain developing the *Takens' method of delays* (Takens, 1981), the set of orthonormalized vectors $\{\boldsymbol{s}_i \in \Re^N\}$ producing a set of linearly independent vectors in $\Re^m$ (when applied to the trajectory matrix $\boldsymbol{X}$) is generally a subset of the complete orthonormal basis $\{\boldsymbol{c}_i | i = 1,2,...,m\}$ for the embedding space. Thus, by construction

$$\boldsymbol{s}_i'\boldsymbol{X} = \sigma_i \boldsymbol{c}_i', \tag{4}$$

where $\{\sigma_i\}$ is a set of real constants for normalization. Since $\{\boldsymbol{c}_i\}$ is orthonormal, it follows that

$$\boldsymbol{s}_i'\boldsymbol{X}\boldsymbol{X}'\boldsymbol{s}_j = \sigma_i \sigma_j \delta_{ij}, \tag{5}$$

where δ_{ij} is the Kronecker's delta[3]. The *structure matrix* containing the correlations between all the pairs of patterns composing $\boldsymbol{X}$, i.e. $\boldsymbol{\Theta} = \boldsymbol{X}\boldsymbol{X}'$, is real, symmetric, and non-negative definite; thus, its eigenvectors form a complete orthonormal basis for $\Re^N$. Equation (5) is solved by the eigenvectors of $\boldsymbol{\Theta}$, i.e.

$$\boldsymbol{\Theta s}_i = \sigma_i^2 \boldsymbol{s}_i , \qquad (6)$$

being $\{\sigma_i^2\}$ the corresponding eigenvalues. Equation (4) guarantees that there are at least m non-zero σ_i.

Looking at an inverse relation of (4), i.e. $\boldsymbol{Xc}_i = \sigma_i \boldsymbol{s}_i$, we can observe how SSA is based on a *lag-covariance matrix* orthogonal decomposition (see section 4.1). In fact, since $\{\boldsymbol{s}_i\}$ is an orthogonal set, from $\boldsymbol{c}'_j \boldsymbol{X'Xc}_i = \sigma_i \sigma_j \delta_{ij}$ we can derive the *eigenvalue equation*

$$\boldsymbol{\Xi c}_i = \sigma_i^2 \boldsymbol{c}_i , \qquad (7)$$

with $\boldsymbol{\Xi} = \boldsymbol{X'X}$ representing the *lag-covariance matrix*[4] of the process *X(t)* sampled in our N_T-length time series.

If we normalize the eigenvectors $\boldsymbol{E}^l$ of equation (7) such that $\sum_{l=1}^{m} \boldsymbol{E}_i^l \boldsymbol{E}_j^l = \delta_{ij}, 1 \le i \le m, 1 \le j \le m$, the *spectral decomposition formula* holds, thus allowing to factorize the diagonal matrix $\boldsymbol{\Xi}$ into a canonical form, i.e. $\boldsymbol{\Xi}_{ij} = \xi(i-j) = \sum_{l=1}^{m} \sigma_l^2 \boldsymbol{E}_i^l \boldsymbol{E}_j^l, 1 \le i \le m, 1 \le j \le m$. The eigenvectors $\boldsymbol{E}^l$ represent the lagged sequences of length m providing the new orthonormal basis onto which the signal is decomposed (Vautard et al., 1992). They are generally called Empirical Orthogonal Functions and come directly from the data, while the corresponding eigenvalues $\{\sigma^2{}_l\}$ represent the so called Principal Components, i.e. the fraction of total variance explained in the dataset by each orthogonal direction.

Coming back to the connection between statistical dimension *S* and PCA, the square roots of the above eigenvalues are called *singular values*, and their set is the *singular spectrum*. The dataset noise level can be identified by the flat-floor tail of the singular spectrum when singular values are sorted in decreasing order. S identifies the number of singular values above the noise floor, i.e. the number of significant components characterizing the dataset (Vautard and Ghil, 1989).

5.3 Noise-reduction strategies

Coming back to the drawbacks of noise in spectral analysis, SSA provides an interesting denoising tool for its effectiveness in determining the series statistical dimension: then, denoised series can be reconstructed in the time domain by neglecting all the spectral components belonging to the noise tail of the eigenvalue spectrum; thus, spectral estimation can be performed avoiding many problems due to noise.

Following Broomhead and King (1986), we analyze the effects of noise on spectral techniques considering the projection of the trajectory matrix onto the orthonormal basis $\{\boldsymbol{c}_i\}$, i.e. $\boldsymbol{XC}$. Thus, $(\boldsymbol{XC})'(\boldsymbol{XC}) = \boldsymbol{\Sigma}^2$, where $\boldsymbol{\Sigma} = diag(\sigma_1, \sigma_2, ..., \sigma_m)$; since each σ_i^2 is the mean square projection onto the covariance matrix $\boldsymbol{X'X}$, $\{\boldsymbol{c}_i\}$ gives the directions and $\{\sigma_i\}$ the lengths of the principal axes of the embedded ellipsoid. Unfortunately, noise tends to obscure system dynamics especially along those directions associated with small σ_i.

In the case of monotonic and rapidly decreasing-to-0 spectra the simplest noise reduction strategy is a fixed low-pass filter. However, for non monotonic spectra exhibiting peaks over a wide range of frequencies the issue is more complicate (Vautard et al., 1992).

A plausible alternative consists in measuring an experimental noise time series and its root mean square projections onto the orthonormal basis $\{\boldsymbol{c}_i\}$ in order to compute the *signal-to-noise ratio* associated with each *singular vector* of $\boldsymbol{X}$, i.e. the vectors of $\boldsymbol{C}$ and the eigenvectors of $\boldsymbol{\Theta}$. That way we are implicitly assuming that the realization of our N_T-length finite process $X(t)$ is the sum of a dynamical process $Y(t)$ and some external noise process. Then, we separate the original signal from its experimentally-measured noise component, thus allowing the rejection of a portion of noise actually affecting the signal. The main danger is to remove parts of the relevant dynamics eventually hidden in the detected noise.

However, since such experimental noise measurement is not always possible, other denoising techniques have been developed. For example, SSA is considered a powerful tool for signal reconstruction from noisy data (Vautard et al., 1992). Since Vautard and Ghil (1989) demonstrate that the last PCs of $X(t)$ are noise-dominated, Vautard et al. (1992) develop a systematic method to detect the eigenvalue spectrum break actually identifying the noise floor. They define a *noise-reduction ratio* $n(p)$ using the p Reconstructed Components (RCs) derived from the data (see section 5.6), i.e.

$$n(p) = \frac{\sum_{i=1}^{N_T} (y_i - \sum_{k=1}^{p} x_i^k)^2}{\sum_{i=1}^{N_T} (y_i - x_i)^2}, \qquad (8)$$

where the numerator represents the noise associated with the first p RCs, while the denominator the total noise. Thus, the average optimum noise reduction occurs when (8) is minimized with respect to p, which reveals the dataset statistical dimension. However, since real data do not generally provide several realizations of the same process, *Monte Carlo simulations* try to make up for this lack: different techniques have been developed, such as the *surrogate-data method* and the *Monte Carlo SSA* (see section 5.6).

5.4 Robustness and statistical stability of the eigenvalues

Another crucial aspect for the reliability of SSA is the statistical stability of the eigenvalues: various methods have been developed to estimate the statistical confidence of the eigenvectors. In principle, since SSA is based on estimates of the lagged autocovariances, the eigenvalues should converge at least as well as the Blackman-Tukey estimation (Vautard et al., 1992). Ghil and Mo (1991) estimate the eigenvalue error by the formula

$$\delta\sigma_k = (2/N_{dof})^{1/2} \sigma_k , \qquad (9)$$

where $N_{dof} = (N_T/m) - 1$ is the number of degrees of freedom associated with the m-length window. The authors show that such error is proportional to the eigenspectrum: in order to reduce (9) the dataset sampling interval should be smaller than the autocorrelation length of the field.

However, if $\delta\sigma_k$ is at least comparable to the spacing between σ_k and a neighbouring eigenvalue, then the sampling error for the associated EOF is comparable to the size of the neighbouring EOF, thus producing degenerate multiplets. As a consequence, in such case the eigenvectors derived from the sample actually represent a random mixture of the true eigenvectors, and the ambiguity in choosing the proper linear combination of EOFs increases the sampling error (North et al., 1982).

Summing up, numerical and sampling errors do not guarantee that eigenvalue pairs effectively represent oscillatory movements. In fact, the eigenvalues whose error bars significantly overlap with those of the noise spectrum are highly suspected of being themselves part of the noise. However, the more noisy is the series, the less significant pairs tend to emerge from the noise floor. The problem could be partially solved by properly increasing *m*, since the spectrum tail consequently flattens (Vautard et al., 1992). However, the widening of the window length is not always practicable because of the drawbacks listed in the next section.

5.5 Choice of the window length m

In SSA the choice of the window length is crucial, since the method cannot solve periods longer than *m* (Vautard and Ghil, 1989; Vautard et al., 1992): thus, the higher periodicity we want to look for, the larger *m* we need. It is exactly the *flexibility* in the choice of *m* which makes SSA more successful at different time scales than other spectral methods; however, a proper decision must evaluate some drawbacks.

In fact, the selection of *m* represents a compromise between information and statistical confidence (Ghil and Mo, 1991). In order to avoid the last values of the *autocovariance function* being dominated by statistical error, the choice set is generally limited to $m \leq \frac{N_T}{3}$. However, when *m* is considerably larger than the lifetime of the oscillation[5] we are inspecting, the analysis suffers from the so called *Gibbs' effect*, i.e. the overshooting of the eigenfunction series at jump discontinuities. This effect tends both to reduce the amplitude of the fit within the spell and to produce artificial periodicity off the spell, which is then smoothed out (Vautard et al., 1992).

Moreover, if *m* is too small neighbouring spectral peaks tend to coalesce because of the coarse resolution, while high resolutions, i.e. large *m*, are likely to split broad peaks into several components in neighbouring frequencies.

Practically, Vautard et al. (1992) suggest to select *m* between the oscillation period we want to analyze and the average lifetime of its spells. However, since the latter quantity is not *a priori* known, the authors suggest that SSA is usually successful for periods in the range $\left(\frac{m}{5}, m\right)$.

5.6 Principal and Reconstructed Components

Since SSA comes directly from Principal Components Analysis (see section 5.2), we can think of it in terms of Principal and Reconstructed Components (PCs and RCs).

First of all, the *k*-th PC is the coefficient of the projection of the original series onto the *k*-th *Empirical Orthogonal Function* (EOF), i.e. the *k*-th eigenvector $\boldsymbol{E}^k$ of the cross-covariance matrix $\boldsymbol{\Xi}$:

$$a_i^k = \sum_{j=1}^{m} x_{i+j} E_j^k , 0 \leq i \leq N_T - m .$$

Thus, PCs are reciprocally orthogonal [6] *N*-length processes providing *weighted moving-averages* of the process *X(t)*. In fact, since PCs are the coefficients of the linear combination of any EOFs subset minimizing the least-squares distance between the resulting fit and the original series over the chosen window, from a spectral point of view EOFs are data-adaptive moving-average filters. Vautard et al. (1992) demonstrate that the sum of the PCs spectra is identical to the original series power spectrum. This result is particularly interesting, since it underlies the completely linear nature of SSA, which valuably simplifies the analysis of nonlinear time series through linear tools.

Now we have introduced all the elements to present the *Karhunen-Loéve expansion* of x_i, which is the SSA corner stone:

$$x_{i+j} = \sum_{k=1}^{m} a_i^k E_j^k , 1 \leq j \leq m, 0 \leq i \leq N_T - m . \qquad (10)$$

It illustrates how each observation of the series *X(t)* is linearly decomposed to analyze the information carried by each eigenvector of the orthonormal basis, which is determined by sliding the chosen *m*-length weighting-moving-average window.

Thus, PCs are filtered versions of the original series unfortunately not performing unique expansions: in fact, since equation (10) depends on combinations of *i* and *j*, there are clearly multiple ways to reconstruct most of the signal components. Moreover, this algorithm reconstructs just an *N*-length signal, instead of the N_T-length one we need to make comparisons with the original data.

The problem could be managed by means of RCs: considering a subset *A* of *K* eigenelements, the associated PCs are combined to form the *partial reconstruction Y(t)* of the original series, which is the solution of the least-squares problem

$$\min \sum_{i=0}^{N_T-m} \sum_{j=1}^{m} (y_{i+j} - \sum_{k \in A} a_i^k E_j^k)^2 \ ;$$

thus, the augmented version of the optimal series is the closest in a least-squares sense to the projection of *X(t)* onto the subset of EOFs belonging to *A* (Vautard et al., 1992).

It is important to notice that the resulting RCs are characterized by purely additive properties: the *k*-th RC for $k \in A$ is denoted by $x^k(t)$, and the associated partial reconstruction is $Y(t) = \sum_{k \in A} x^k(t)$. Consequently, *X(t)* can be expanded as the linear sum of its RCs, i.e.

$$X(t) = \sum_{k=1}^{m} x^k(t) \,. \quad (11)$$

This property is particularly useful in economic applications, since it allows to combine the effects of significant fluctuations in order to inspect the overall dynamics and eventually some kind of feedback.

As an example, one of the most puzzling issues in the analysis of economic cycles deals with the trend-cycle separation and the nature of eventual feedbacks between the overall trend and the cyclical characteristics of the economy. Thus, a method allowing the time-domain reconstruction of significant economic fluctuations and their simple linear summation could shed some new light on controversial issues of that kind.

However, despite the linear summation in (11), the transform between $X(t)$ and $x^k(t)$ is nonlinear, since the relation between each EOF and the original series is actually nonlinear: this fact allows a proper decomposition of nonlinear time series. Finally, notwithstanding RCs are correlated even at zero-lag, they allow both component reconstructions on the whole time span and a precise localization of short oscillations spells, thus representing a fundamental improvement with respect to PCs.

5.7 Interpretation of the eigenelements

This section clears some concepts for the correct interpretation of SSA results. In particular, for economic (and other) applications it is fundamental to correctly denoise the series, identify significant oscillations, and distinguish between trend and ultra-low frequencies. The following subsections deal with all these subjects.

5.7.1 Trend and ultra-low frequencies

A basic assumption in SSA is the *covariance stationarity of the data generating process X(t)*, i.e. the time invariance of its first and second moments: this assumption is essential, since SSA is based on the estimate of the process lag-covariance matrix from the dataset. Notably, Vautard and Ghil (1989) hypothesise that lag-covariances of components outside the dataset time span may be estimated through the available observations, thus suggesting new forecasting and backcasting tools. However, these procedures are plausible just when assuming stationarity of the generating process on the observed timescale (Allen and Smith, 1996, p.3376). Nevertheless, the single N_T-length realization $\{X(t)\}$ we actually observe may appear nonstationary if the stationary process X is essentially characterized by periods longer than N_T (Vautard et al., 1992). Thus, if we do not dispose of several realizations and checks for stationarity, we cannot effectively distinguish between trends and ultra-low frequencies.

In order to solve the *impasse*, Ghil and Vautard (1991) show how SSA provides itself a fully non-parametric data-adaptive detrending method performing as a low-pass filter. Moreover, Vautard et al. (1992) develop a systematic data-adaptive algorithm to remove trend and ultra-low frequencies based on the *Kendall nonparametric test for global trend identification* (Kendall and Stuart, 1968). Note that a sizable trend should appear in the first few PCs of the series, since it accounts for most of its variance; however, since finite length may produce artificial trends, after the first detrending a new test is performed on the detrended series and so on, until the reconstructed series shows no significant trend at all.

5.7.2 Eigenvalues pairs

As we said in section 5.1, nearly-equality of two successive eigenvalues and phase quadrature of their associated EOFs may jointly identify a fundamental oscillation. Ghil and Mo (1991) characterize nearly-equality as

$$\left|\sigma_{k+1} - \sigma_k\right| \le \min\left\{\delta\sigma_k, \delta\sigma_{k+1}\right\},$$

where $\delta\sigma_k$ is defined by (9); on the contrary, phase quadrature is analyzed through the lag-correlations of the corresponding PC pairs.

However, since we usually lack of reliable estimators for lag-correlations, Vautard et al. (1992) propose two substitutive criteria based on spectral properties of the eigenvectors. First of all, oscillatory pairs of two successive eigenelements *(k,k+1)* must be spectrally localized around the same frequency, i.e. the quantity $\delta f_k = \left|f_k - f_{k+1}\right|$ has to be small. Secondly, since oscillations are generally identified by high spectral peaks, the frequency f^* resolving the oscillatory pair must present a reconstruction filter whose *response function* explains at least 2/3 of the process variance at f^*[7].

Notably, these criteria are not very sensitive to changes in the window length. However, they may sometimes be misleading. In fact, Allen and

Smith (1996) show some cases where the existence of trends and/or the intentional suppression of some cycles produce spurious EOF pairs. Then, if the noise is mistakenly identified by eigenvalue pairs included in the SSA-based noise reducing filter, any subsequent analysis is clearly biased. Thus, in order to reduce the probability of type I errors, i.e. the rejection of a null hypothesis when it is actually true, Monte Carlo experiments are sometimes performed (see section 6).

5.7.3 Eigenvectors

A crucial point in SSA concerns the interpretation of the EOFs, in order to assign them a clear meaning in the analysis of system dynamics; this issue is directly related to the assumptions made upon noise.

As previously explained, the detection of oscillatory EOF pairs may often correspond to nontrivial signals; however, this analysis is neither immediate nor unambiguous: if we do not adopt proper devices, nontrivial oscillatory pairs could be falsely detected even in pure noise processes. Moreover, since in conventional SSA both high- and low-ranked EOFs tend to pair up, some authors consider the stability of a pair with respect to variations in the window length m as an evidence for the significance of the corresponding oscillation. However, Allen and Smith (1996) notice that this kind of stability does not always identify nontrivial signals. In fact, the red noise eigenspectrum typically shows more power than the average at certain low frequencies: thus, the EOF pairing tends to occur over different m even if the generating process is a pure noise.

Coming back to the meaning of SSA eigenvectors, in pure deterministic signals EOFs with m' non-zero eigenelements define the linear m'-subspace where the attractor lies; clearly, increasing the sample size N_T they converge to the principal axes of the true attractor (Allen and Smith, 1996).

A similar interpretation is possible for the first p highest-ranked EOFs coming from series affected by white noise. In such cases the *eigenspectrum truncation* is a standard practice, since clearly when $p \geq m'$ the whole variance of the significant signal is projected onto the p selected EOFs, while just a p/m fraction of noise is projected, since the typical feature of white noise processes concerns the projection of equal variance onto all the series EOFs. Consequently, the method provides an m/p enhancement of the signal-to-noise ratio.

Unfortunately, this issue becomes more complicated when the series is affected by coloured noise: since the lag-covariance matrix is no more a scalar multiple of the identity matrix but depends on the signal-to-noise ratio and on the lag-covariance matrices of both the true signal and the noise, algorithms providing the eigenspectrum truncation are generally no longer correct and the eigenvalues rank-order is unreliable (Allen and Smith, 1996). This is clearly a serious problem for the reliability of the SSA analysis, since its standard procedure does not provide an effective denoising tool against the very widespread coloured noise processes. In order to solve this drawback, some authors propose a Monte Carlo approach to SSA (see section 6).

Finally, SSA eigenvectors may be interpreted as data-adaptive *moving-average filters* for noise. Thus, according to Broomhead et al. (1987), the significant signal can be separated from noise by exploiting the EOFs topological features: in fact, deterministic eigenvectors show a regular shape, while stochastic ones have noisy profile. However, an overestimation of the number of significant eigenvalues is not too a dangerous mistake, since it simply corresponds to leaving some noise in the data (Medio, 1992).

6. Monte Carlo SSA and the surrogate data method

As previously underlined, standard SSA may be misleading in the case of either low signal-to-noise ratio or coloured background noise. In fact, they could make slow modes being incorrectly interpreted as nontrivial signals, and/or nontrivial signals actually embedded in coloured noise being neglected if their variance is smaller than that of the noise slow modes. Thus, Monte Carlo (MC) approaches has been developed, allowing to distinguish a given time series from any well-defined process, including noise backgrounds (Ghil and Vautard, 1991; Vautard et al., 1992; Allen and Smith, 1996; Paluš and Novotná, 2004; Paluš and Novotná, 2006).

In particular, Allen and Smith (1996) propose a MC method to isolate red noise processes eventually embedded in the dataset, but a generalized version can be applied to any kind of coloured noise. The method implements the so called *surrogate data method* (Theiler et al., 1992), which creates surrogate series and estimates their lag-covariance matrix distribution, thus defining their respective projections onto the interested EOFs in order to test whether the original data coefficients are significantly different from those of a data-adaptive noise process.

More clearly, surrogate data are different realizations of the hypothesized noise, since MC-SSA is based on a *null hypothesis (NH) approach*. Thus, we firstly assume some NH noise process and generate the relative surrogate realizations to test the significance of the detected SSA components with respect to the null. Secondly, both the eigenvalues coming from the dataset and the surrogate data bars are plotted against the dominant frequency of the corresponding EOF. Note that the surrogate data bars represent the *a%* confidence intervals of the corresponding surrogate eigenvalues, signalling that *a%* surrogate realizations exhibit a *k*-th eigenvalue lying between the *k*-th bar extremes. Once the problem of selecting a single dominant frequency is solved[8], a sort of coarsely discretized power spectrum is obtained. Finally, the data eigenspectrum and the surrogate data bars are compared frequency-by-frequency: when a data eigenvalue falls above the corresponding surrogate bar, the related EOF accounts for more power than expected under the null, thus signalling an eventually nontrivial signal (Allen and Smith, 1996).

Unfortunately, such occurrence is necessary but not sufficient to detect nontrivial signals: since our NH test does actually correspond to *m* mini-tests, the probability of a by-chance excursion of at least one eigenvalue above the corresponding *a%*-confidence-level surrogate bar is generally higher than *(100 – a)%* (cf. Thomson, 1990a). Allen and Smith (1996) suggest a *two-pass Monte Carlo approach* to directly estimate such by-chance probability: after computing the distribution of the surrogate data lag-covariance matrix and performing a second pass trough the ensemble, the probability of a given number of excursions above a predetermined percentile is directly estimated from its relative frequency in an element of the ensemble.

Pointing out the features of MC-SSA, we observe that this procedure is fundamental when either the number of significant EOFs is small or the NH is particularly complicate. However, since the parameters of the underlying noise process are often unknown, we must apply *maximum likelihood tests* on them whenever failing to reject the NH (cf. Allen and Smith, 1996). In addiction, when dealing with *hybrid NHs* testing whether the residuals from some nontrivial signal identification are attributable to a chosen noise process, the *signal-reconstruction approach* described in section 5.6 can be applied to generate composite signal-plus-noise surrogates. Unfortunately, this technique does not yield unbiased estimates of noise parameters[9], while

SSA-based reconstructions are poor when applied to irregularly sampled and/or heteroskedastic data. However, it is extremely valuable the potential of the method to deal with complicate and hybrid NHs, since it highly rises the confidence of the researcher in identifying nontrivial signals and noise backgrounds, which is the core of the spectral methods described in the paper.

Coming back to the relevant risks of standard SSA, its full data-adaptive framework is clearly a precious advantage with respect to other spectral tools, but it may be hazardous when detecting unknown signals. In fact, in this case artificial variance compression and similar effects raise the probability of type I error. Thus, Allen and Smith (1996) suggest an improved Monte Carlo algorithm which allows both to retain the data-adaptive properties when extracting already detected signals, and to avoid the above mentioned shortcoming when no signal has been identified yet. The core assumption is that any NH is true until otherwise established. So, the method projects both the data and the surrogates onto the EOFs expected under the NH, rather than deriving them from the data themselves. Then, when a signal is detected, it is analyzed and reconstructed through standard SSA. The risk of such technique is to miss relevant signals which do not exactly align with the NH EOFs, but its great advantage is a more precise quantification of the type I error probability, which is the worst error type in econometrics.

To conclude, a crucial point when applying MC-SSA is that surrogates and data have to be identically treated in order to perform correct statistical testing[10]: a careful analysis should stop only when the NH is no longer rejectable.

6.1 Enhanced MC-SSA

A main trouble with MC-SSA is the assumption that relevant signals linearly add to the specified noise background. Thus, in order to be detected, a signal must show significantly greater variance in its characteristic frequency band than the NH (Paluš and Novotná, 2004; Paluš and Novotná, 2006). However, the signal of interest has often more complicated origins, which cannot always be correctly analyzed through standard MC-SSA. Paluš and Novotná (1998) propose an enhanced MC-SSA method, which tests in addiction the dynamical properties of the SSA modes against the surrogate ones, thus allowing to detect interesting dynamical modes independently of their relative variance.

The enhanced method evaluates the uncertainty embedded in the series by applying some notions from both the theory of stochastic processes and the information theory. As explained in section 3, the uncertainty of a stochastic variable is measured by its *entropy* (Cover and Thomas, 1991). Since economic time series can be considered as single realizations of the underlying stochastic process, i.e. sequences of stochastic variables, the uncertainty they carry in can be measured by entropy. Moreover, since time series generally show the temporal evolution of not completely random systems, the *entropy rate*, i.e. the rate at which they forget information about their previous states, can be thought of as an important quantitative characterization of the system temporal complexity (Paluš and Novotná, 2004).

Since the possibility to compute entropy rates from experimental data with the well-known *Kolmogorov-Sinai Entropy* (KSE) is very limited (Cover and Thomas, 1991), Paluš (1996) proposes to use the so called C*oarse-grained Entropy Rate* (CER). CER is based on *mutual information* $I(x;x_\tau)$, which is the average amount of information about X_τ contained in the τ-1 variables $X_1, X_2, \ldots, X_{\tau-1}$. The time series $X(t)$ is a realization of the ergodic and stationary stochastic process X. CERs are not supposed to estimate the exact entropy rates because of their dependence on particular experimental and numerical setups, but rather to produce measures of the regularity and predictability of the analyzed series in a relative sense. It means that different datasets can be compared by their CERs if both they were observed in the same experimental conditions and CERs were estimated using the same numerical parameters.

The enhanced MC-SSA is implemented through six steps (Paluš and Novotná, 2004; Paluš and Novotná, 2006):

1. the standard MC-SSA is performed, identifying the frequency bin with the highest power;

2. a NH benchmark model is fitted on the analyzed series, and the residuals computed;

3. surrogate data are generated with the above model using randomly permutated residuals as innovations;

4. each surrogate realization undergoes SSA, finding for each frequency bin the surrogate eigenvalues distribution at the chosen confidence level;

5. for each frequency bin the data eigenvalues are compared with the surrogate bars: if an eigenvalue lies outside the respective surrogate range, the NH is rejected;

6. regularity indexes are computed for both data- and surrogate-SSA modes; then, they are statistically tested with procedures analogous to the eigenvalue test.

Summing up, the rejection of a NH always suggests that something has been neglected in the benchmark model. However, a rejection based on the eigenvalues hints a different covariance structure than the NH noise; on the contrary, a rejection based on the regularity index suggests that the dataset contains some dynamically interesting signal characterized by higher regularity and better predictability than the NH model.

7. Wavelet analysis technique

For the application of SSA and derived techniques the analyzed series do not necessarily need to be stationary, since these methods manage in taking into account the distortive effects of finite-length on stationarity.

Nonetheless, an approach particularly successful in dealing with realizations of *multiscale nonstationary processes*, which show nonstationary power at many different frequencies, is the wavelet analysis technique (Daubechies, 1990; Ramsey, 2002; Schleicher, 2002; Crowley, 2007). It attempts to solve the typical problem of classical analysis methods concerning the localization of oscillatory movements in time and frequency, the first involving the choice of the filtering window size, the second the eventual localization of the dominant frequencies of periodic signals.

The innovative feature of this method with respect to the techniques surveyed above consists in the *simultaneous* decomposition of time series into the time/frequency space, thus gaining information on both the amplitude of any eventual periodic signal, and its variation through time (Torrence and Compo, 1998). Thus, the approach involves a transform from one-dimensional time series images to two dimensional time-frequency ones.

It is important to underlie that an high precision in time localization in the high-frequency band requires a trade-off in terms of a reduced frequency resolution, and vice versa (cf. the uncertainty principle: Chui, 1992; Lau and Weng, 1995).

The core of wavelet parametric method consists in sliding window functions of particular shapes (either real or complex) along the analyzed series, in order to obtain a time series of the projection amplitudes. Among the most common wavelet functions there are *Morlet*, *Paul*, the *Mexican hat*, and the *derivative of Gaussian* (DOG) functions (see Farge, 1992; Lau and Weng, 1995; Torrence and Compo, 1998; Crowley, 2007). A main advantage of this method is the possibility to vary the wavelet scale during the same analysis, changing its width. It basically works as a bandpass filter with known response function, the *wavelet function* $\psi(\eta)$, where η is a non-dimensional time parameter; consequently, the method can reconstruct time series by inverse filtering.

The *wavelet transform* (WT) is a sort of filter defined as the inner product of the wavelet function with the original *N*-length time series *X(t)*, i.e.

$$W_n(s) = \sum_{n=0}^{N-1} x_n \psi^* \left[\frac{(n'-n)\delta t}{s} \right], \tag{12}$$

where s represents the wavelet scale, n the localized time index, δt the constant sampling time, and the asterisk the complex conjugate, i.e. the number obtained changing the sign to the imaginary part of a complex number.

Since the WT is generally complex, both its amplitude $|W_n(s)|$ and phase $[\Im\{W_n(s)\}/\Re\{W_n(s)\}]$ can be identified, where $\Im\{W_n(s)\}$ and $\Re\{W_n(s)\}$ respectively represent the WT imaginary and real parts.

Evaluating (12) for various s along n, a two-dimensional variability image is derived by plotting the wavelet amplitude and phase. The wavelet power spectrum, i.e. the *spectrogram* or *scalogram*, contains information about the relative power of the signal at a certain scale and time. Qiu and Er (1995) demonstrate a result particularly useful in applications: the bias of the wavelet spectrogram in noisy signals depends just on the power associated with noise, thus being independent of time, scale, and the wavelet function; on the contrary, the variance of the wavelet spectrogram depends on the spectrogram of the signal component, which is a function of time and scale. More important for hypothesis testing, Torrence and Compo (1998) demonstrate that the wavelet power is distributed like a χ_2^2 . Thus, these results suggest a fruitful use of wavelet techniques in time series analysis.

Concerning economic applications, Ramsey (2000) suggests to implement these tools for multiple purposes: first of all as an exploratory device to enlighten the dynamics of frequency components in economic and financial datasets; to explore the relationships among economic variables at a disaggregate-scale level; to forecast series by scale, in order to analyze both their global and local features; and finally, to deal with local inhomogeneity.

The next subsections deal with some introductory technical details about the wavelet computation algorithm, its variance decomposition procedures and some statistical hypotheses testing. For a recent complete review of the method addressed to economists refer to Crowley (2007).

7.1 Computation algorithm

The WT can be implemented in the frequency domain through the Fast Fourier Transform (FFT), which is an algorithm for Fourier transform (see section 2) allowing the simultaneous computation of all the *N* points of the series. In the *Fourier domain* the WT (12) becomes

$$W_n(s) = \sum_{k=0}^{N-1} \hat{x}_k \hat{\psi}^*(s\omega_k)\exp(i\omega_k n\delta t) \ , \tag{13}$$

where ω is the frequency and the hat (^) indicates the FFT algorithm, which in the case of *X(t)* is calculated by

$$\hat{x}_k = \frac{1}{N}\sum_{n=0}^{N-1} x_n \exp(2\pi ikn/N) . \tag{14}$$

Note that in order to make the scalogram comparable with different power spectra, the wavelet function requires a normalization (Torrence and Compo, 1998).

Then, there are six steps to compute the WT and perform the decomposition of the series into the time/frequency space:

1. choose a mother wavelet, e.g. Morlet, DOG, etc.;
2. compute the mother wavelet FFT;
3. compute the time series FFT;
4. find all the scales, which are power-of-two multiples of the smallest resolvable scale s_0 (Torrence and Compo, 1998; Lau and Weng, 1995);
5. for each scale:

- compute the daughter wavelet Fourier transform at that scale through the formula $\hat{\psi}(s\omega_k) = \left(\frac{2\pi s}{\delta t}\right)^{1/2} \hat{\psi}_0(s\omega_k)$;
- normalize dividing it by the total wavelet standard deviation;
- multiply the result by the time series FFT (14);
- compute the inverse transform back to real space using (13);

6. make a contour plot.

A main problem with the computation of the WT in the Fourier space is that the time series is thus supposed periodic. An *escamotage* to avoid the signal at one end of the series being wrapped around to the other end consists in padding one time series end with zeros.

7.2 Statistical NH testing

Standard statistical testing on the WT involves the series scalogram. In fact, whenever a peak exceeds the chosen background Fourier spectrum, its related features are likely to be true with a certain confidence level. However, the wavelet testing procedure is a bit different from the previous ones, since it can involve both local and global characteristics of the series.

More in details, Torrence and Compo (1998) show that the NH of interest can be tested by assuming a proper background *mean spectrum*, since the *local wavelet spectrum*, i.e. the time/frequency spectrogram, follows the mean Fourier spectrum, i.e. the spectrum obtained by time-averaging the power map at each duration. Since the distribution of the Fourier power spectrum is given by

$$\frac{N|\hat{x}_k|^2}{2\sigma^2} \xrightarrow[d]{} \frac{1}{2} P_k \chi_2^2 \ ,$$

where P_k represents the mean spectrum at the Fourier frequency k, the corresponding distribution for the *local wavelet spectrogram* at each time n and scale s is given by

$$\frac{|W_n(s)|^2}{\sigma^2} \xrightarrow[d]{} \frac{1}{2} P_k \chi_2^2 \ . \tag{15}$$

In (15) P_k represents the mean spectrum at the Fourier frequency k corresponding to the wavelet scale s: apart from this relation, the distribution of the local spectrogram is notably independent of the chosen wavelet function.

As in classical Fourier analysis, smoothing the wavelet spectrum can desirably enhance the confidence in regions of significant power; however, the spectrum can now be smoothed by averaging either over time (*global wavelet spectrum*) or over scale (Torrence and Compo, 1998). In the time domain the technique is particularly useful as it succeeds both in the global and local frameworks, since the *global* wavelet spectrum is an unbiased and consistent estimator of the true spectrum (Percival, 1995), and the NH spectrum can be measured also in the *local* context (Kestin et al., 1998). On the frequency side, the scale-averaged spectrogram corresponds to a time series of the dataset average variance over a certain band: this property can be used to examine eventual frequency modulations.

At last, the confidence contours are plotted at the desired level. The *confidence interval*, i.e. the probability that the true wavelet power lies within a certain interval, is derived by (15):

$$\frac{2}{\chi_2^2(p/2)}|W_n(s)|^2 \leq W_n^{*2}(s) \leq \frac{2}{\chi_2^2(1-p/2)}|W_n(s)|^2,$$

where $W_n^{*2}(s)$ represents the true wavelet power, and p the desired significance.

7.3 Variance decomposition through wavelet variance estimation

The wavelet variance decomposition is particularly useful to analyze stochastic processes X, since it allows to enlighten the global variance contribution of components associated with different scales (Percival, 1995), i.e.

$$\sum_{j=0}^{\infty} v_X^2(2^j) = \text{var}(X_t).$$

Roughly speaking, $v_X^2(\lambda)$ is a measure of the variation of a weighted average of the process X with bandwidth λ: the plot $v_X^2(\lambda)$ versus λ indicates which scales do mainly contribute to the global variability. Notice that the wavelet variance provides a way of regularising the spectrum, since it summarizes spectral information using one value per octave frequency band (Percival, 1995).

A shortcoming of the wavelet variance is due to its bias towards *power-law processes*, i.e. this method tends to identify power laws in the data even when other models would be more explicative (Percival and Guttorp, 1994). In order to avoid this drawback, a complementary spectral analysis is highly recommended.

Two estimators are generally used: the *Discrete Wavelet Transform* (DWT) and the *Maximal-Overlap Estimator* (MOE). DWT constructs an unbiased and consistent estimator for the variance by discretely sampling the wavelet, which acts as a filter for scale λ; on the contrary, MOE is based on the *overlapping* technique, which is generally used to reduce the estimate variance (Percival, 1995). It consists in subdividing sampled data into K subintervals, calculating each transform, and then averaging over them to obtain more accurate estimates (Medio, 1992). An analytical description of WTE and MOE confidence intervals is derived in Percival (1995). Nevertheless, notwithstanding the wavelet filter for scale λ can be regarded as a band-pass filter, the asymptotic relative efficiency of WTE with respect to MOE is always less than unity and can even approach 0.5 (cf. Percival, 1995), thus suggesting the use of MOE.

Summing up, the wavelet method provides a valid alternative to the SSA and MC-SSA analyses in the case of multiscale nonstationary processes, supplying a double decomposition in the time/frequency domain and supporting reach procedures for hypotheses testing.

8. Conclusions

This work is a methodological review of some classical and more recent spectral techniques for time series analysis: their features suggest that a more widespread use in economics and finance would be particularly profitable to enlighten economic system dynamics. In fact, while most traditional econometric time series analysis substantially lies in the time domain, spectral analysis is yielded in the frequency domain. The idea dates back to the 1960s, when some advances in spectral estimation algorithms and the increasing necessity of deeper insights in time series structure suggested the application of spectral methods to macroeconomic time series (cf. e.g. Granger and Hatanaka, 1964; Nerlove, 1964; Granger, 1966). The main aim of these techniques is to detect low, medium, and high frequency components carrying the most information in the time series, thus providing precise filtering methods, the identification of the signal dominant cycles, trend-cycle separation, business cycles extraction, and the analysis of co-movements among different series.

The present work surveys some recent univariate methods for spectral analysis and time series reconstruction. Since the robustness of a result is generally obtained when different spectral techniques produce quite homogeneous outcomes, it is particularly important to study and use a differentiated set of spectral tools. From the above survey the highly diversified nature of these techniques emerges: from classical Fourier methods, which succeed in identifying purely periodic signals, to more recent MEM, which allows to separate periodic harmonic components from quasi-periodic ones; from SSA and MC-SSA, which provide a data adaptive detrending tool and support both the reconstruction of signal components from the eigenvalue decomposition and some testing procedures for different null hypotheses noise processes, to wavelet analysis, which is particularly useful in the analysis of nonstationary processes. Our inquiry does not include cross-spectral analysis, which is a valuable tool for the investigation of co-movement among series (for a recent survey cf. Iacobucci, 2003).

Clearly, all these techniques hold a good potential for economic application both in micro- and macro-analysis, since they allow to gain deep quantitative insights into any noisy, nonlinear and short time series.

Acknowledgments

I would like to thank Gianna Vivaldo, for the precious and illuminating introduction to spectral analysis techniques, and Prof. Vittorio Valli, for his patient supervision. This work has been financed by Progetto Alfieri (CRT Foundation) and LABORatorio Revelli, thanks to Prof. Bruno Contini's kind availability.

Notes

[1] More precisely, power spectra are Fourier transforms of the corresponding lag autocorrelation functions, which determine *Wiener prediction filters*, whose coefficients can always be interpreted as AR coefficients (Jaynes, 1982).

[2] Two SSA algorithms are generally applied, differing for their window characteristics: here we essentially refer to Broomhead and King (1986)'s one, whose window stops when either the beginning or the end of the series is reached, i.e. $N=N_T-m+1$. On the contrary, Vautard

and Ghil (1989) develop a sliding-off-the-ends window algorithm, where $N=N_T+m-1$ and the lag-covariance matrix contains some missing value.

[3] Kronecker's delta is a 2-variable function (e.g. *i* and *j*) assuming value 1 if $i=j$ and 0 otherwise.

[4] For each $\xi(j)=\frac{1}{N}\sum_{i=1}^{N-j}x_i x_{i+j}$ the *Yule-Walker estimate* is used. However, both the estimation algorithms by Broomhead and King (1986) and Vautard and Ghil (1989) explained in note 2 are biased, depending both on N_T and $\boldsymbol{X}$. Nevertheless, the latter provides a significant additional noise reduction when applied to short series, being subjected to more bias but less variance than the first (Allen and Smith, 1996).

[5] Note that sometimes the frequency band we are interested in is associated with the least-unstable periodic orbit embedded in a strange attractor, which can intermittently attract our system trajectories, thus generating oscillations with strongly variable amplitude (Vivaldo, 2007). In such cases a simple decomposition of the signal onto sine and cosine functions looses most information; thus, data-adaptive decomposition is strongly recommended.

[6] Note that orthogonality does not imply independence, which in SSA holds just at zero-lag.

[7] The response function ρ of a subset A of eigenelements at frequency f is defined as $\rho_A(f)=\frac{1}{m}\sum_{k\in A}\left|\widetilde{\boldsymbol{E}}^k(f)\right|^2$, where the argument of the absolute value is the *reduced Fourier transform* of the corresponding eigenvector $\boldsymbol{E}^k$. The criterion suggested by Vautard et al. (1992) is based on the analysis of the response function $\rho_k(f^*)+\rho_{k+1}(f^*)$, which must be close to 1 because of the orthogonality constraint imposed on the EOFs.

[8] Since SSA EOFs are not pure sinusoids, their association to a single frequency is quite problematic. Allen and Smith (1996) propose to identify as dominant the frequency maximizing the squared correlation with the corresponding sinusoid, while Vautard et al. (1992) apply the reduced Fourier transform.

[9] This is valid for both large and short series: in the first case, when the noise is added to the RCs its variance is distributed over all frequencies, including the signal ones; in the other case, SSA-based reconstructions tend to be over-fitted near the series endpoint (Allen and Smith, 1996).

[10] A critical issue concerns the way we deal with *artificial variance compression* (Allen and Smith, 1996): in standard SSA it arises since the algorithm makes EOFs maximizing the variance accounted for in the data by the smallest possible number of patterns. Thus, when analyzing a pure noise segment the highest ranked EOF artificially accounts for an improbably high variance relative to the variance it accounts for in an arbitrary series generated by the same noise process, and vice versa. Then, the *Florida-Milwaukee-UCLA test* based on the eigenspectrum shape was implemented (Elsner and Tsonis, 1994; Elsner, 1995): it does no more project each surrogate lag-covariance matrix onto the data EOFs, but on a new EOF basis for each surrogate realization. In such way the artificial variance compression effect is present in both the data and the surrogate eigenspectra. However, since such test compares the overall shape of the ranked eigenspectra in both the data and the surrogates, it may be misleading when the rank-order is misleading, as could happen in standard SSA.

References

Aadland, D. (2005) Detrending time-aggregated data. *Economic Letters* 89: 287-93.

Acemoglu, D. and Scott, A. (1997) Asymmetric business cycles: theory and time-series evidence. *Journal of Monetary Economics* 40(3): 501-33.

A'Hearn, B. and Woitek, U. (2001) More international evidence on the historical properties of business cycles. *Journal of Monetary Economics* 47: 321-46.

Akaike, H. (1969) Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics* 21: 243-47.

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Authomatic Control*19: 716-23.

Allen, M. R. and Smith, L. A. (1996) Monte Carlo SSA: detecting irregular oscillations in the presence of colored noise. *Journal of Climate* 9: 3373-404.

Altissimo, F. and Violante, G. L. (1998) Nonlinear VAR: some theory and an application to US GNP and unemployment. *Temi di Discussione del Servizio Studi Banca d'Italia* 338.

Atesoglu, H. S. and Vilasuso, J. (1999) A band spectral analysis of exports and economic growth in the united states. *Review of International Economics* 7(1): 140-52.

Baxter, M. and King, R. G. (1999) Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics* 81: 575-93.

Blackman, R. B. and Tukey, J. V. (1985) *The Measurement of Power Spectra from the Point of View of Communication Engineering.* New York: Dover Publications.

Bloomfield, P. (1976) *Fourier Analysis of Time Series: an Introduction.* New York: John Wiley.

Box, G. E. P., Jenkins, G. M. (1970) *Time Series Analysis: Forecasting and Control.* San Diego, CA: Holden-Day.

Brock, W. A. (1986) Distinguish random and deterministic systems: abridged version. *Journal of Economic Theory* 40: 168-95.

Brock, W. A. (2000) Whither nonlinear? *Journal of Economic Dynamics and Control* 24: 663-78.

Brock, W. A. and Sayers, C. L. (1988) Is the business cycle characterized by deterministic chaos? *Journal of Monetary Economics* 22: 71-90.

Broomhead, D. S., Jones, R. and King, G. P. (1987) Topological dimension and local coordinates from time series data. *Journal of Physics A* 20: L563-69.

Broomhead, D. S. and King, G.P. (1986) Extracting qualitative dynamics from experimental data. *Physica D* 20: 217-36.

Burg, J. P. (1967) Maximum entropy spectral analysis. *37th Annual International Meeting Society of Exploration Geophysicists.*

Callen, J. L., Kwan, C. C. Y., Yip, P. C. Y. (1985) Foreign-exchange rate dynamics: an empirical study using maximum entropy spectral analysis. *Journal of Business and Economic Statistics* 3(2): 149-55.

Chatfield, C. (1984) *The Analysis of Time Series: an Introduction.* New York: Chapman and Hall.

Chiarella, C., El-Hassan, N. (1997) Evaluation of derivative security prices in the heath-jarrow-morton framework as path integrals using fast fourier transform techniques. Working Paper 72, School of Finance and Economics, University of Technology, Sydney, Australia.

Chui, C. K. (1992) *An Introduction to Wavelets.* San Diego, CA: Academic Press, Harcourt Brace Jovanovich.

Cover, T. M. and Thomas, J. A. (1991) *Elements of Information Theory.* New York: J. Wiley & Sons.

Croux, C., Forni, M. and Reichlin, L. (2001) A measure of comovement for economic variables: theory and empirics. *Review of Economics and Statistics* 83(2): 232-41.

Crowley, P. M. (2007) A guide to wavelets for economists. *Journal of Economic Surveys* 21(2): 207-67.

Daubechies, I. (1990) The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory* 36(5): 961-1004.

Elsner, J. B. (1995) Significance tests for SSA. *Proceedings of the 19th Climate Diagnostic Workshop College Park MD CAC/NOAA U.S. Dept. of Commerce* 187-90.

Elsner, J. B. and Tsonis, A. A. (1994) Low-frequency oscillations. *Nature* 372: 507-08.

Farge, M. (1992) Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics* 24: 395-457.

Fisher, I. (1925) Our unstable dollar and the so-called business cycle. *Journal of the American Statistical Association* 20: 179-202.

Frank, M. Z. and Stengos, T. (1988) Some evidence concerning macroeconomic chaos. *Journal of Monetary Economics* 22: 423-38.

Gan, L., Zhang, Q. (2005) The thick market effect on local unemployment rate fluctuations. Working Paper 11248, NBER, Cambridge, MA.

Gardiner, C. W. (1983) *Handbook of Stochastic Methods. For Physics, Chemistry and the Natural Sciences.* Berlin Heidelberg: Springer-Verlag.

Gerace, M. P. (2002) US military expenditures and economic growth: some evidence from spectral methods. *Defence and Peace Economics* 13(1): 1-11.

Gershenfeld, N. A. and Weigend, A. S. (1994) The future of time series: learning and understanding. In N. A. Gershenfeld and A. S. Weigend (eds.) *Time Series Prediction: Forecasting the Future and Understanding the Past.* USA: Perseus Books Publishing.

Ghil, M., Allen, M. R., Dettinger, M. R., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F. and Yiou, P.

(2002) Advanced spectral methods for climatic time series. *Reviews of Geophysics* 40(1): 1-41.

Ghil, M. and Mo, K. (1991) Interseasonal oscillations in the global atmosphere. part I: northern hemisphere and tropics. *Journal of the Atmospheric Sciences* 48(5): 752-79.

Ghil, M. and Vautard, R. (1991) Interdecadal oscillations and the warming trend in global temperature time series. *Nature* 350: 324-27.

Golan, A. (2002) Information and entropy econometrics - editor's view. *Journal of Econometrics* 107(1-2): 1-15.

Granger, C. W. J. (1966) The typical spectral shape of an economic variable. *Econometrica* 34(1): 150-61.

Granger, C. W. J. (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3): 424-38.

Granger, C. W. J. and Hatanaka, M. (1964) *Spectral Analysis of Economic Time Series.* Princeton, NJ: Princeton University Press.

Harris, D., Poskitt, D. S. (2004) Determination of cointegrating rank in partially non-stationary processes via a generalized von-Neumann criterion. *The Econometrics Journal* 7(1): 191-217.

Haykin, S., Kessler, S. (1983) Prediction-error filtering and maximum-entropy spectral estimation. In S. Haykin (ed.) *Nonlinear Methods of Spectral Analysis.* New York: Springer-Verlag.

Higo, M. and Nakada, S. K. (1998) How can we extract a fundamental trend from an economic time-series? *Monetary and Economic Studies* December

Iacobucci, A. (2003) Spectral analysis for economic time series. *OFCE Working Papers* 2003-07.

Jaynes, E. T. (1982) On the rationale of maximum-entropy methods. *Proceedings of the IEEE* 70(7): 939-52.

Jenkins, G. M., Watts, D. G. (1968) *Spectral Analysis and Its Applications.* San Francisco, CA: Holden Day.

Kendall, M. G. and Stuart, A. (1968) *The Advanced Theory of Statistics.* London: Griffin.

Kestin, T. S., Karoly, D. J., Yano, J. I. and Rayner, N. A. (1998) Time frequency variability of ENSO and stochastic simulation. *Journal of Climate* 11(9): 2258-72.

Klotz, B. P., Neal, L. (1973) Spectral and cross-spectral analysis of the long swing hypothesis. *Review of Economics and Statistics* 55(3): 291-98.

Lau, K.-M. and Weng, H. (1995) Climate signal detection using wavelet transform: how to make a time series sing. *Bulletin of American Meteorological Society* 76(12): 2391-401.

Lisi, F. and Medio, A. (1997) Is a random walk the best exchange rate predictor? *International Journal of Forecasting* 13: 255-67.

MacDonald, G. J. (1989) Spectral analysis of time series generated by nonlinear processes. *Review of Geophysics* 27: 449-69.

Mann, M. E. and Lees, J. M. (1996) Robust estimation of background noise and signal detection in climatic time series. *Climatic Change* 33: 409-45.
Marchand, C. (1985) The impact of monetary activity upon regional housing markets: a maximum entropy spectral analysis. *Environment and Planning A* 17(7): 889-904.
McConnell, M. M. and Perez-Quiros, G. (2000) Output fluctuations in the United States: what has changed since the early 1980's? *American Economic Review* 90(5): 1464-76.
Medio, A. (1992) *Chaotic Dynamics - Theory and Applications to Economics.* Cambridge: Cambridge University Press.
Movshuk, O. (2003) Does the choice of detrending method matter in demand analysis? *Japan and the World Economy* 15(3): 341-59.
Neftci, S. N. (1984) Are economic time series asymmetric over the business cycle? *Journal of Political Economy* 92: 307-28.
Neftci, S. N. and McNevin, B. (1986) Some evidence on the non-linearity of economic time series: 1890-1981. *Working Paper of the C.V. Starr Center for Applied Economics* 86/26.
Nerlove, M. (1964) Spectral analysis of seasonal adjustment procedures. *Econometrica* 32(3): 241-86.
North, G. R., Bell, T. L., Cahalan, R. F. and Moeng, F. J. (1982) Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review* 110: 699-706.
Ormerod, P. and Campbell, M. (1997) Predictability and economic time-series. In C. Heij, J. M. Schumacher, B. Hanzon, C. Praagman (eds.) *System Dynamics in Economic and Financial Models.* John Wiley & Sons.
Packard, N. H., Crutchfield, J. P., Farmer, J. D. and Shaw, R. S. (1980) Geometry from a time series. *Physical Review Letters* 45(9): 712-16.
Palm, F. C. (1997) Comments. In C. Heij, J. M. Schumacher, B. Hanzon, C. Praagman (eds.) *System Dynamics in Economic and Financial Models.* John Wiley & Sons.
Paluš, M. (1996) Coarse-grained entropy rates for characterization of complex time series. *Physica D* 93: 64-77.
Paluš, M. and Novotná, D. (1998) Detecting modes with nontrivial dynamics embedded in colored noise: enhanced Monte Carlo SSA and the case of climate oscillations. *Physics Letters A* 248: 191-202.
Paluš, M. and Novotná, D. (2004) Enhanced Monte Carlo Singular System Analysis and detection of period 7.8 years oscillatory modes in the monthly NAO index and temperature records. *Nonlinear Processes in Geophysics* 11: 721-29.
Paluš, M. and Novotná, D. (2006) Quasi-biennial oscillations extracted from the monthly NAO index and temperature records are phase-synchronized. *Nonlinear Processes in Geophysics* 13: 287-96.
Paris, Q. (2001) Multicollinearity and maximum entropy estimators. *Economics Bullettin* 3(11): 1-9.

Paris, Q., Howitt, R. E. (1998) An analysis of ill-posed production problems using maximum entropy. *American Journal of Agricultural Economics* 80(1): 124-38.
Park, J. (1992) Envelope estimation for quasi-periodic geophysical signals in noise: a multitaper approach. In A. T. Walden, P. Guttorp (eds.) *Statistics in the Environmental and Earth Sciences.* London: Edward Arnold.
Peeters, L. M. K. (2004) Estimating a random-coefficients sample-selection model using generalized maximum entropy. *Economics Letters* 84(1): 87-92.
Penland, C., Ghil, M. and Weickmann, K. (1991) Adaptive filtering and maximum entropy spectra, with application to changes in atmospheric angular momentum. *Journal of Geophysical Research* 96: 22659-71.
Percival, D. B. and Guttorp, P. (1994) Long-memory processes, the Allan variance and wavelets. In E. Foufoula-Georgiou, P. Kumar (eds.) *Wavelets in Geophysics.* New York: Academic Press.
Percival D. B., Walden A. T. (1993) *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques.* Cambridge: Cambridge University Press.
Percival, D. P. (1995) On estimation of the wavelet variance. *Biometrika* 82: 619-31.
Piselli, P. (2004) Business cycle non-linearities and productivity shocks. *Temi di Discussione del Servizio Studi Banca d'Italia* 516.
Pollock, D. S. G. (2008) The frequency analysis of the business cycle. Working Paper 08/12, University of Leicester, Leicester, UK.
Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. (1986) *Numerical Recipes: the Art of Scientific Computing.* New York: Cambridge University Press.
Priestley, M. B. (1981) *Spectral Analysis and Time Series.* London: London Academic Press.
Qiu, L. and Er, M. H. (1995) Wavelet spectrogram of noisy signals. *International Journal of Electronics* 79(5): 665-77.
Ramsey, J. (2002) Wavelets in economics and finance: past and future. *Studies in Nonlinear Dynamics and Econometrics* 6(3): 1-27.
Schleicher, C. (2002) An introduction to wavelets for economists. Working Paper 2002-3, Bank of Canada, Ottawa, Canada.
Serletis, A. (1996) Is there chaos in economic time series? *Canadian Journal of Economics* 29(Special Issue): S210-12.
Slepian, D. (1978) Prolate spheroidal wave functions, Fourier analysis and uncertainty. V - The discrete case. *Bell System Technical Journal* 57: 1371-430.
Slutsky, E. (1937) The summation of random causes as the source of cyclical processes. *Econometrica* 5(2): 105-46.
Stanca, L. M. (1999) Asymmetries and nonlinearities in Italian macroeconomic fluctuations. *Applied Economics* 31: 483-91.

Stein, E. M. and Weiss, G. (1971) *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton, NJ: Princeton University Press.

Stock, J. H. and Watson, M. W. (1999) Business cycle fluctuations in US macroeconomic time series. In J. B. Taylor and M. Woodford (eds.) *Handbook of Macroeconomics. Vol. 1A.* Amsterdam: Elsevier.

Takahashi, A., Takehara, K. (2008) Fourier transform method with an asymptotic expansion approach: an application to currency options. *International Journal of Theoretical and Applied Finance* forthcoming.

Takens, F. (1981) Detecting strange attractors in turbulence. In D. A. Rand, L.-S. Young (eds.) *Lecture Notes in Mathematics* Berlin: Springer.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J. D. (1992) Testing for nonlinearity in time series: the method of surrogate data. *Physica D* 58: 77-94.

Thomakos, D. D., Wang, T. and Wille, L. T. (2002) Modelling daily realized futures volatility with singular spectrum analysis. *Physica A* 312: 505-19.

Thomson, D. J. (1982) Spectrum estimation and harmonic analysis. *Proceedings of the IEEE* 70: 1055-92.

Thomson, D. J. (1990a) Time series analysis of Holocene climate data. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 330(1615): 601-16.

Thomson, D. J. (1990b) Quadratic-inverse spectrum estimates: applications to paleoclimatology. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 332: 539-97.

Torrence, C. and Compo, G. P. (1998) A practical guide to wavelet analysis. *Bulletin of American Meteorological Society* 79(1): 61-78.

van den Bos, A. (1971) Alternative interpretation of maximum entropy spectral analysis. *IEEE Transactions on Information Theory* 17: 493-94.

Vautard, R. and Ghil, M. (1989) Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D* 35: 395-424.

Vautard, R., Yiou, P. and Ghil, M. (1992) Singular-spectrum analysis: a toolkit for short, noisy chaotic signals. *Physica D* 58: 95-126.

Vinod, H. D. (2006) Maximum entropy ensembles for time series inference in economics. *Journal of Asian Economics* 17(6): 955-78.

Vivaldo, G. (2007) Private communication.

Walker, G. (1931) On periodicity in series of related terms. *Proceedings of the Royal Society of London A* 131(818): 518-32.

Wang, P. (1999) Spectral analysis of economic time series behaviour. Working Paper 9914, Manchester School of Management, Manchester, UK.

Whitney, H. (1936) Differentiable manifolds. *Annals of Mathematics* 37(3): 645-80.

Wu, X., Stengos, T. (2005) Partially adaptive estimation via the maximum entropy densities. *Econometrics Journal* 8: 352-66.

Yule, G. U. (1927) On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London A* 226: 267-98.