Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Alternative Methodology for Turning-Point Detection in Business Cycle : A Wavelet Approach

Peter Martey ADDO, Monica BILLIO, Dominique GUEGAN

2012.23

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/bandeau-haut/documents-de-travail/
ISSN : 1955-611X

Alternative Methodology for Turning-Point Detection in Business Cycle: A Wavelet Approach

Peter Martey ADDO*

European Doctorate in Economics–Erasmus Mundus (EDEEM)
Université Paris1 Panthéon-Sorbonne, MSE-CES UMR8174 , 106-113 boulevard de l'hopital, 75013, Paris, France
Ca'Foscari University of Venice, 30121, Venice, Italy
email: pkaddo2000@yahoo.com

Monica BILLIO

Department of Economics, Ca'Foscari University of Venice, 30121, Venice, Italy
email: billio@unive.it

Dominique GUÉGAN

CES - Centre d'économie de la Sorbonne - CNRS : UMR8174 - Université Paris I - Panthéon Sorbonne,
EEP-PSE - Ecole d'Économie de Paris - Paris School of Economics, France
email: dguegan@univ-paris1.fr

Abstract

We provide a signal modality analysis to characterize and detect nonlinearity schemes in the US Industrial Production Index time series. The analysis is achieved by using the recently proposed 'delay vector variance' (DVV) method, which examines local predictability of a signal in the phase space to detect the presence of determinism and nonlinearity in a time series. Optimal embedding parameters used in the DVV analysis are obtained via a differential entropy based method using wavelet-based surrogates. A complex Morlet wavelet is employed to detect and characterize the US business cycle. A comprehensive analysis of the feasibility of this approach is provided. Our results coincide with the business cycles peaks and troughs dates published by the National Bureau of Economic Research (NBER).

Keywords: Nonlinearity analysis, Surrogates, Delay vector variance (DVV) method, Wavelets, Business Cycle, Embedding parameters
JEL: C14, C22, C40, E32

1. Introduction

In general, performing a nonlinearity analysis in a modeling or signal processing context can lead to a significant improvement of the quality of the results, since it facilitates the selection of

*Correspondence to: Peter Martey Addo, Université Paris1 Panthéon-Sorbonne, MSE-CES UMR8174 , 106-113 boulevard de l'hopital, 75013, Paris, France, Email: pkaddo2000@yahoo.com

Documents de Travail du Centre d'Economie de la Sorbonne - 2012.23

appropriate processing methods, suggested by the data itself. In real-world applications of economic time series analysis, the process underlying the generated signal, which is the time series, are a priori unknown. These signals usually contain both linear and nonlinear, as well as deterministic and stochastic components, yet it is a common practice to model such processes using suboptimal, but mathematically tractable models. In the field of biomedical signal processing, e.g., the analysis of heart rate variability, electrocardiogram, hand tremor, and electroencephalogram, the presence or absence of nonlinearity often conveys information concerning the health condition of a subject (for an overview, see Hegger R. and Schreiber, 1999). In some modern machine learning and signal processing applications, especially biomedical and environmental ones, the information about the linear, nonlinear, deterministic or stochastic nature of a signal conveys important information about the underlying signal generation mechanism. There has been an increasing concerns on the forecasting performance of some nonlinear models in modelling economic variables. Nonlinear models often provide superior in-sample fit, but rather poor out-of-sample forecast performance (Stock and Watson, 1999). In cases were the nonlinearity is suprious or relevant for only a small part of the observations, the use of nonlinear models will lead to forecast failure (see Terasvirta, 2011). It is, therefore, essential to investigate the intrinsic dynamical properties of economic time series in terms of its deterministic/stochastic and nonlinear/linear components reveals important information that otherwise remains not clear in using conventional linear methods of time series analysis.

Several methods for detecting nonlinear nature of a signal have been proposed over the past few years. The classic ones include the 'deterministic versus stochastic' (DVS) plots (Casdagli, 1994), the Correlation Exponent and 'δ-ϵ' method (Kaplan, 1994). For our purpose, it is desirable to have a method which is straightforward to visualize, and which facilitates the analysis of predictability, which is a core notion in online learning. In this paper, we adopt to the recently proposed phase space based 'delay vector variance' (DVV) method (Gautama T., 2004a), for signal characterisation, which is more suitable for signal processing application because it examines the nonlinear and deterministic signal behaviour at the same time. This method has been used for the qualitative assessment of machine learning algorithms, analysis of functional magnetic resonance imaging (fMRI) data, as well as analysing nonlinear structures in brain electrical activity and heart rate variability (HRV) (Gautama T., 2004b). Optimal embedding parameters will be obtained using a differential entropy based method proposed in Gautama T., 2003, which allows for simultaneous determination of both the embedding dimension and time lag. Surrogate generation used in this study will be based on a recently refined Iterative Amplitude Adjusted Fourier Transform (iAAFT) using a wavelet-based approach, denoted WiAAFT (Keylock, 2006).

Wavelet analysis has successfully been applied in a great variety of applications like signal filtering and denoising, data compression, imagine processing and also pattern recognition. The application of wavelet transform analysis in science and engineering really began to take off at the beginning of the 1990s, with a rapid growth in the numbers of researchers turning their attention to wavelet analysis during that decade. The wavelet transforms has the ability to perform local analysis of a time series revealing how the different periodic components of the time series change over time. The maximum overlap discrete wavelet transform (MODWT) has commonly been used by some economists (Whitcher and Percival, 2000, Gallegati and Gallegati, 2007, Gallegati, 2008). The MODWT can be seen as a kind of compromise between the discrete wavelet transform (DWT) and the continuous wavelet transform (CWT); it is a redundant transform, because while it is efficient with the frequency parameters it is not selective with the time parameters. The CWT, unlike the DWT, gives us a large freedom in selecting our wavelets and yields outputs that makes it much easier to interpret. The continuous wavelet transform has

emerged as the most favoured tool by researchers as it does not contain the cross terms inherent in the Wigner-Ville transform and Choi-Williams distribution (CWD) methods while possessing frequency-dependent windowing which allows for arbitrarily high resolution of the high frequency signal components. As such, we recommend the use of CWT to analyze the time series to discover patterns or hidden information.

The choice of the wavelet is important and will depend on the particular application one has in mind. For instance, if a researcher is concerned with information about cycles then complex wavelets serves as a necessary and better choice. We need complex numbers to gather information about the phase, which, in turn, tells us the position in the cycle of the time-series as a function of frequency. There are many continuous wavelets to choose from; however, by far the most popular are the Mexican hat wavelet and the Morlet wavelet. In this work, we employ a complex Morlet wavelet which satisfies these requirements and has optimal joint time-frequency concentration (Aguiar-Conraria and Soares, 2011), meaning that it is the wavelet which provides the best possible compromise in these two dimensions.

In this paper, we provide a novel procedure for analysing US Industrial Production Index data, based on the DVV method and then we employ wavelet analysis to detect US business cycle. Our results is consistent with business cycle peaks and troughs dates published by the National Bureau of Economic Research (NBER). The paper is organised as follows: In section 2, we discuss wavelets and wavelet transforms and then give an overview on recent types of surrogate generation. An entropy-based method for determining the embedding parameters of the phase-space of a time series is presented. We then provide the 'Delay Vector Variance' methodology with an illustration. In section 3, we present a comprehensive analysis of the feasibility of this approach to analyse the US Business cycle.

2. Background on Methodology: Wavelet Analysis and 'Delay Vector Variance' Method

In this section, we first introduce some notation and operators that will be referred to for the rest of this paper. The concept of wavelet analysis and our choice of analysing wavelet is presented in section 2.2. Surrogate generation methodology and differential entropy based method for determining optimal embedding parameters are then presented in section 2.3 and section 2.4 respectively. Lastly, we present, in section 2.5, an overview of the 'delay vector variance' method with illustrations.

2.1. Notation and Operators

Consider the mapping

$$\Gamma : \mathcal{L}^2(\mathbb{R}) \longrightarrow \mathcal{L}^2(\mathbb{R})$$

Let $f \in \mathcal{L}^2(\mathbb{R})$, $\alpha, \beta \in \mathbb{R}$ and $s \in \mathbb{R}^+$ where $\mathbb{R}^+ := \{t \in \mathbb{R} : t > 0\}$. Unless otherwise stated, the complex conjugate of $z \in \mathbb{C}$ is denoted $\bar{z}$ and the magnitude of z is denoted $|z|$. The symbol i will represent the square root of -1, i.e., $i^2 = -1$. We present in table 1 some notations and operators that will be often referred to in this manuscript.

2.2. Wavelet and Wavelet Analysis

It is a time-frequency signal analysis method which offers simultaneous interpretation of the signal in both time and frequency allowing local, transient or intermittent components to be elucidated. These components are often not clear due to the averaging inherent within spectral only methods like the fast Fourier transform (FFT).

Operator, Γ	Notation, Γf	Output	Inverse, $\Gamma * f = \Gamma^{-1} f$	Fourier transform, $\widehat{(\Gamma f)}$
Dilation	$(D_s f)(t)$	$\frac{1}{s^{1/2}} f(\frac{t}{s})$	$D_{s^{-1}} f$	$D_{s^{-1}} \hat{f}$
Involution	$\tilde{f}$	$\bar{f}(-t)$	$\tilde{f}$	$\tilde{\hat{f}}$
Translation	$(\tau_\alpha f)(t)$	$f(t-\alpha)$	$\tau_{-\alpha} f$	$e_{-a} \hat{f}$
Modulation	$(e_\alpha f)(t)$	$e^{i2\pi\alpha t} f(t)$	$e_{-\alpha} f$	$\tau_\alpha \hat{f}$
Reflection	$(Rf)(t)$	$f(-t)$	Rf	$R\hat{f}$

Table 1: Notations and Operators

A Wavelet is a function, ψ, which has a small concentrated burst of finite energy in the time domain and exihibits some oscillation in time. This function must be in the space of measurable functions that are absolutely and squared-integrable, i.e. $\psi \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$, to ensure that the Fourier transform[1] of ψ is well-defined and ψ is a finite energy signal. A single wavelet function generates a family of wavelets by dilating (stretching and contracting) and translating (moving along the time axis) itself over a continuum of dilation and translation values. If ψ is a wavelet analysing function then the set $\{\tau_t D_s \psi\}$ of all the dilated (by $s \neq 0$) and translated (by t) versions of ψ is that wavelet family generated by ψ. Dilation in time by contracting values of scale ($s > 1$) corresponds to stretching dilation in the frequency domain.

The basic concept in wavelet transforms is the projection of data onto a basis of wavelet functions in order to separate large-scale and fine-scale information (Bruce L, 2002). Thus, the signal is decomposed into a series of shifted and scaled versions of a mother wavelet function to make possible the analysis of the signal at different scales and resolutions. For reconstruction of a signal, it is necessary that ψ be such that $\{\tau_t D_s \psi\}$ span a large enough space of interest.

- Thus, every signal f of interest should be representable as a linear combination of dilated and translated versions of ψ.
- Knowing all the inner products $\{\langle f, \tau_t D_s \psi \rangle\}$, the signal should be recoverable.

The wavelet ψ is assumed to satisfy the admissibility condition,

$$C_{adm,\psi} = \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \tag{1}$$

where $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(\xi)$,

$$\hat{\psi}(\omega) = \int_{\mathbb{R}} \psi(\xi) e^{-i\omega\xi} d\xi \tag{2}$$

The admissibility condition (1) implies

$$\hat{\psi}(0) = \int_{\mathbb{R}} \psi(\xi) d\xi = 0 \tag{3}$$

[1]Given $\psi \in \mathcal{L}^1$, the Fourier transform of ψ, $\hat{\psi}(\gamma) = \int \psi(t) e^{-i2\pi\gamma t} dt$ is well-defined since the integral converges absolutely.

For s restricted to $\mathbb{R}^+$, the condition (1) becomes

$$C_{adm_+,\psi} = \int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty. \tag{4}$$

This means that the wavelet has no zero-frequency component. The value of the admissibility constant, $C_{adm,\psi}$ or $C_{adm_+,\psi}$ depends on the chosen wavelet. This property allows for an effective localization in both time and frequency, contrary to the Fourier transform, which decomposes the signal in term of sines and cosines, i.e. infinite duration waves.

There are essentially two distinct classes of Wavelet transforms: the continuous wavelet transform and the discrete wavelet transform. We refer the reader to Addison, 2005; Walden and Percival, 2000 for a review on Wavelet transforms. In this work, we mainly employ the Maximal Overlap Discrete Wavelet Transform (MODWT) and a complex continuous wavelet transform.

2.2.1. Maximal Overlap Discrete Wavelet Transform

The Maximal Overlap Discrete Wavelet Transform (MODWT), also related to notions of 'cycle spinning' and 'wavelet frames', is with the basic idea of downsampling values removed from discrete wavelet transform. The MODWT unlike the conventional discrete wavelet transform (DWT), is non-orthogonal and highly redundant, and is defined naturally for all sample sizes, N (Walden and Percival, 2000). Given an integer J such that $2^J < N$, where N is the number of data points, the original time series represented by the vector $X(n)$, where $n = 1, 2, \cdots, N$, can be decomposed on a hierarchy of time scales by details, $D_j(n)$, and a smooth part, $S_J(n)$, that shift along with X:

$$X(n) = S_J(n) + \sum_{j=1}^{J} D_j(n) \tag{5}$$

with $S_j(n)$ generated by the recursive relationship

$$S_{j-1}(n) = S_J(n) + D_J(n). \tag{6}$$

The MODWT details $D_j(n)$ represent changes on a scale of $\tau = 2^{j-1}$, while the $S_j(n)$ represents the smooth or approximation wavelet averages on a scale of $\tau_J = 2^{J-1}$. Gallegati and Gallegati, 2007 employed this wavelet transform to investigate the issue of moderation of volatility in G-7 economies and also to detect the importance of the various explanations of the moderation.

2.2.2. Continuous Wavelet Transform (CWT)

The continuous wavelet transform (CWT) differs from the more traditional short time Fourier transform (STFT) by allowing arbitrarily high localization in time of high frequency signal features. The CWT permits for the isolation of the high frequency features due to it's variable window width related to the scale of observation. In particular, the CWT is not limited to using sinusoidal analysing functions but allows for a large selection of localized waveforms that can be employed as long as they satisfy predefined mathematical criteria (described below).

Let $\mathcal{H}$ be a Hilbert space, the CWT may be described as a mapping parameterized by a function ψ

$$C_\psi : \mathcal{H} \longrightarrow C_\psi(\mathcal{H}). \tag{7}$$

The CWT of a one-dimensional function $f \in \mathcal{L}^2(\mathbb{R})$ is given by

$$\begin{aligned} C_\psi : \mathcal{L}^2(\mathbb{R}) &\longrightarrow C_\psi(\mathcal{L}^2(\mathbb{R})) \\ f &\mapsto \langle f, \tau_t D_s \psi \rangle_{\mathcal{L}^2(\mathbb{R})} \end{aligned} \tag{8}$$

where $\tau_t D_s \psi$ is a dilated (by s) and translated (by t) version of ψ given as

$$(\tau_t D_s \psi)(\xi) = \frac{1}{|s|^{\frac{1}{2}}} \psi(\frac{\xi - t}{s}) \tag{9}$$

Thus, the CWT of one-dimensional signal f is a two-dimensional function of the real variables time t, and scale $s \neq 0$. For a given ψ, the CWT may be thought of in terms of the representation of a signal with respect to the wavelet family generated by ψ, that is, all it's translated and dilated versions. The CWT may be written as

$$(C_\psi f)(t, s) := \langle f, \tau_t D_s \psi \rangle \tag{10}$$

For each point (t, s) in the time-scale plan, the wavelet transform assigns a (complex) numerical value to a signal f which describes how much f like a translated by t and scaled by s version of ψ.

The CWT of a signal f is defined as

$$(C_\psi f)(t, s) = \frac{1}{|s|^{\frac{1}{2}}} \int_{\mathbb{R}} f(\xi) \bar{\psi}(\frac{\xi - t}{s}) d\xi \tag{11}$$

where $\bar{\psi}(\xi)$ is the complex conjugate of the analysing wavelet function $\psi(\xi)$. Given that ψ is chosen with enough time-frequency localization[2], the CWT gives a gives a picture of the time-frequency characteristics of the function f over the whole time-scale plane $\mathbb{R} \times (\mathbb{R} \backslash \{0\})$. When $C_{adm,\psi} < \infty$, it is possible to find the inverse continuous transformation via the relation known as *Calderón's reproducing identity*[3],

$$f(\xi) = \frac{1}{C_{adm,\psi}} \int_{\mathbb{R}^2} \langle f, \tau_t D_s \psi \rangle \tau_t D_s \psi(\xi) \frac{1}{s^2} ds dt. \tag{12}$$

and if s restricted in $\mathbb{R}^+$, then the *Calderón's reproducing identity* takes the form

$$f(\xi) = \frac{1}{C_{adm_+,\psi}} \int_{-\infty}^{\infty} \int_0^{\infty} \langle f, \tau_t D_s \psi \rangle \tau_t D_s \psi(\xi) \frac{1}{s^2} ds dt. \tag{13}$$

Let α and β be arbitrary real numbers and f, f_1, and f_2 be arbitrary functions in $\mathcal{L}^2(\mathbb{R})$ The CWT, C_ψ, with respect to ψ statisfies the following conditions:

1. Linearity
   - $(C_\psi(\alpha f_1 + \beta f_2))(t, s) = \alpha(C_\psi f_1)(t, s) + \beta(C_\psi f_2)(t, s)$

[2]The time-frequency concentrated functions, denoted $TF(\mathbb{R})$, is a space of complex-valued finite energy functions defined on the real line that decay faster than $\frac{1}{t}$ simultaneously in the time and frequency domains. This is defined explicitly as $TF(\mathbb{R}) := \{\varphi \in \mathcal{L}^2(\mathbb{R} : |\varphi(t)| < \eta(1 + |t|)^{-(1+\varepsilon)} \quad and \quad |\hat{\varphi}(\gamma)| < \eta(1 + |\gamma|)^{-(1+\varepsilon)} for \quad \eta < \infty, \varepsilon > 0\}$

[3]This identity is also called the resolution identity.

2. Time Invariance
   - $(C_\psi(\tau_\beta f))(t, s) = (C_\psi f)(t - \beta, s)$
3. Dilation
   - $(C_\psi(D_\alpha f))(t, s) = (C_\psi f)(\alpha t, \alpha^{-1} s)$
4. Negative Scales
   - $C_\psi f(t, -s) = (C_\psi R f)(-t, s)$

The time invariance property of the CWT implies that the wavelet transform of a time-delayed version of a signal is a time-delayed version of its wavelet transform. This serves as an important property in terms of pattern recognition. This nice property is not readily obtained in the case of Discrete wavelet transforms (Addison, 2005; Walden and Percival, 2000).

The contribution to the signal energy at the specific scale s and location t is given by

$$\mathcal{E}(t, s) = |C_\psi|^2 \tag{14}$$

which is a two-dimensional wavelet energy density function known as the scalogram. The wavelet transform C_ψ corresponding to a complex wavelet is also complex valued. The transform can be separated into two categories:

- Real part $\mathcal{R}\{C_\psi\}$ and Imaginary part $\mathcal{I}\{C_\psi\}$
- Modulus (or Amplitude), $|C_\psi|$ and phase (or phase-angle), $\Phi(t, s)$,

which can be obtained using the relation :

$$C_\psi = |C_\psi| e^{i\Phi(t,s)} \quad \textit{and} \quad \Phi(t, s) = \arctan\left(\frac{\mathcal{I}\{C_\psi\}}{\mathcal{R}\{C_\psi\}}\right). \tag{15}$$

In this work, we employ a complex wavelet to in order to separate the phase and amplitude information. In particular, the phase information will be useful in detecting and explaining the cycles in the data.

2.2.3. Choice of Wavelet

The *Morlet wavelet* is the most popular complex wavelet used in practice. A complex Morlet wavelet[4] (Teolis, 1998) is defined by

$$\psi(\xi) = \frac{1}{\sqrt{\pi f_b}} e^{i2\pi f_c \xi - \frac{\xi^2}{f_b}} \tag{16}$$

depending on two parameters: f_b and f_c, which corresponds to a bandwidth parameter and a wavelet center frequency respectively. The Fourier transform of ψ is

$$\hat{\psi}(\zeta) = e^{-\pi^2 f_b (\zeta - f_c)^2}, \tag{17}$$

[4]The complete Morlet wavelet can also be defined as $\psi(t) = \frac{1}{\pi^{1/4}}(e^{iw_0 t} - e^{-\frac{w_0^2}{2}})e^{-\frac{t^2}{2}}$ where w_0 is the central frequency of the mother wavelet. The second term in the brackets is known as the correction term, as it corrects for the non-zero mean of the complex sinusoid of the first term. In practice it becomes negligible for values of $w_0 > 5$.

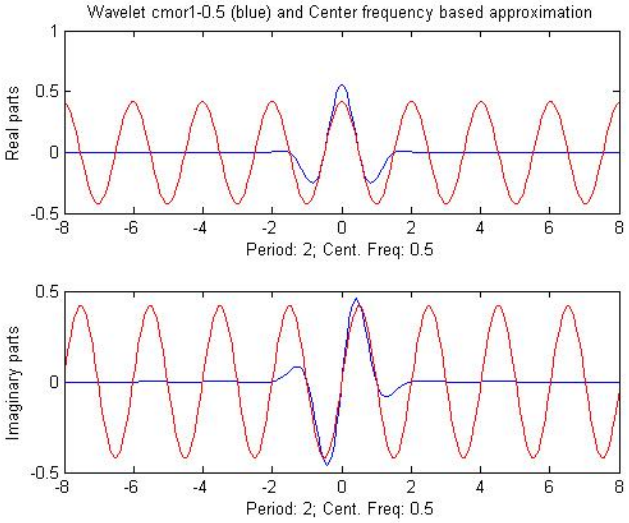


Figure 1: Complex Morlet wavelet with $f_b = 1$ and $f_c = 0.5$

which is well-defined since $\psi \in \mathcal{L}^1(\mathbb{R})$. It can easily be shown that the Morlet wavelet (16) is a modulated gaussian function and involutive, i.e. $\psi = \tilde{\psi}$. The Fourier transform $\hat{\psi}$ has a maximum value of 1 which occurs at f_c, since

$$\|\psi\|_1 := \int |\psi| = 1.$$

This wavelet has an optimal joint time-frequency concentration since it has an exponential decay in both time and frequency domains, meaning that it is the wavelet which provides the best possible compromise in these two dimensions. In addition, it is infinitely regular, complex-valued and yields an exactly reconstruction of the signal after the decomposition via CWT.

In this work, the wavelet that best detect the US business is the complex Morlet wavelet with $f_b = 1$ and $f_c = 0.5$. In this case, the Morlet wavelet becomes

$$\psi(\xi) = \frac{1}{\sqrt{\pi}} e^{i\pi\xi - \xi^2}, \tag{18}$$

which we will often refer to as Morlet wavelet. The nature of our choice of wavelet function and the associated center frequency is displayed in figure 1.

2.3. Surrogate generation

Surrogate time series, or 'surrogate' for short, is non-parametric randomised linear version of the original data which preserves the linear properties of the original data. For identification of nonlinear/linear behavior in a given time series, the null hypothesis that the original data conform to a linear Gaussian stochastic process is formulated. An established method for generating constrained surrogates conforming to the properties of a linear Gaussian process is the Iterative Amplitude Adjusted Fourier Transform (iAAFT), which has become quite popular (Teolis, 1998; Schreiber and Schmitz, 1996, 2000; Kugiumtzis, 1999).

For $\{|x_k|\}$, the Fourier amplitude spectrum of the original time series, X, and $\{c_k\}$, the sorted version of the original time series, at every iteration j, two series are generated: (i) $r^{(j)}$, which has the same signal distribution as the original and (ii) $X^{(j)}$, with the same amplitude spectrum as the original. Starting with $r^{(0)}$, a random permutation of the original time series, the iAAFT method is given as follows:

- compute the phase spectrum of $r^{(j-1)} \rightarrow \{\phi_k\}$,
- $X^{(j)}$ is the inverse transform of $\{|x_k|exp(i\phi_k)\}$, and
- $r^{(j)}$ is obtained by rank ordering $X^{(j)}$ so as to match $\{c_k\}$.

The iteration stops when the difference between $\{|x_k|\}$ and the amplitude spectrum of $r^{(j)}$ stops decreasing. This type of surrogate time series retains the signal distribution and amplitude spectrum of the original time series, and takes into account a possibly nonlinear and static observation function due to the measurement process. The method uses a fixed point iteration algorithm for achieving this, for the details of which we refer to Schreiber and Schmitz, 1996, 2000.

Wavelet-based surrogate generation is a fairly new method of constructing surrogate for hypothesis testing of nonlinearity which applies a wavelet decomposition of the time series. The main difference between fourier transform and wavelet transform is that the former is only localized in frequency, whereas the latter is localized both in time and frequency. The idea of a wavelet representation is an orthogonal decomposition across a hierarchy of temporal and spatial scales by a set of wavelet and scaling functions.

The iAAFT-method has recently been refined using a wavelet-based approach, denoted by WiAAFT (Keylock, 2006), that provides for *constrained realizations*[5] of surrogate data that resembles the original data while preserving the local mean and variance as well as the power spectrum and distribution of the original except for randomizing the nonlinear properties of the signal. The WiAAFT-procedure follows the iAAFT-algorithm but uses the Maximal Overlap Discrete Wavelet Transform (MODWT) where the iAAFT-procedure is applied to each set of wavelet detail coefficients $D_j(n)$ over the dyadic scales 2^{j-1} for $j = 1, \cdots, J$, i.e., each set of $D_j(n)$ is considered as a time series of its own.

2.4. Optimal Embedding Parameters

In the context of signal processing, an established method for visualising an attractor of an underlying nonlinear dynamical signal is by means of time delay embedding (Hegger R. and Schreiber, 1999). By time-delay embedding, the original time series $\{x_k\}$ is represented in the so-called 'phase space' by a set of delay vectors[6] (DVs) of a given embedding dimension, m, and time lag, $\tau : x(k) = [x_{k-\tau}, \cdots, x_{k-m\tau}]$. Gautama T., 2003 proposed a differential entropy based method for determining the optimal embedding parameters of a signal. The main advantage of this method is that a signal measure is simultaneously used for optimising both the embedding dimension and time lag. We provide below an overview of the procedure:

[5]These are surrogate realizations that are generated from the original data to conform to certain properties of the original data, e.g., their linear properties, i.e., mean, standard deviation, distribution, power spectrum and autocorrelation function (Schreiber and Schmitz, 1996, 2000).

[6]Time delay embedding is an established method for visualising an attractor of the underlying nonlinear dynamical signal, when processing signals with structure (Hegger R. and Schreiber, 1999).

The "Entropy Ratio" is defined as

$$R_{ent}(m,\tau) = I(m,\tau) + \frac{m \ln N}{N}, \tag{19}$$

where N is the number of delay vectors, which is kept constant for all values of m and τ under consideration,

$$I(m,\tau) = \frac{H(x,m,\tau)}{\langle H(x_{s,i},m,\tau)\rangle_i} \tag{20}$$

where x is the signal, $x_{s,i}$ $i = 1,\cdots,T_s$ surrogates of the signal x, $\langle\cdot\rangle_i$ denotes the average over i, $H(x,m,\tau)$ denotes the differential entropies estimated for time delay embedded versions of a time series, x, which an inverse measure of the *structure* in the phase space. Gautama T., 2003 proposed to use the Kozachenko-Leonenko (K-L) estimate (Leonenko and Kozachenko, 1987) of the differential entropy given by

$$H(x) = \sum_{j=1}^{T} \ln(T\rho_j) + \ln 2 + C_E \tag{21}$$

where T is the number of samples in the data set, ρ_j is the Euclidean distance of the j-th delay vector to its nearest neighbour, and $C_E(\approx 0.5772)$ is the Euler constant. This ratio criterion requires a time series to display a clear structure in the phase space. Thus, for time series with no clear structure, the method will not yield a clear minimum, and a different approach needs to be adopted, possibly one that does not rely on a phase space representation. When this method is applied directly to a time series exhibiting strong serial correlations, it yields embedding parameters which have a preference for $\tau_{opt} = 1$. In order to ensure robustness of this method to the dimensionality and serial correlations of a time series, Gautama T., 2003 suggested to use the iAAFT method for surrogate generation since it retains within the surrogate both signal distribution and approximately the autocorrelation structure of the original signal. In this Paper, we opt to use wavelet-based surrogate generation method, WiAAFT by in Keylock, 2006, for reasons already discussed in the previous section.

2.5. 'Delay Vector Variance' method

The Characterisation of signal nonlinearities, which emerged in physics in the mid-1990s, have been successfully applied in predicting survival in heart failure cases and also adopted in practical engineering applications (Ho A et al., 1997; Chambers and Mandic, 2001). The 'delay vector variance' (DVV) method (Gautama T., 2004a) is a recently proposed phase space based method for signal characterisation. It is more suitable for signal processing application because it examines the nonlinear and deterministic signal behaviour at the same time. The algorithm is summarized below:

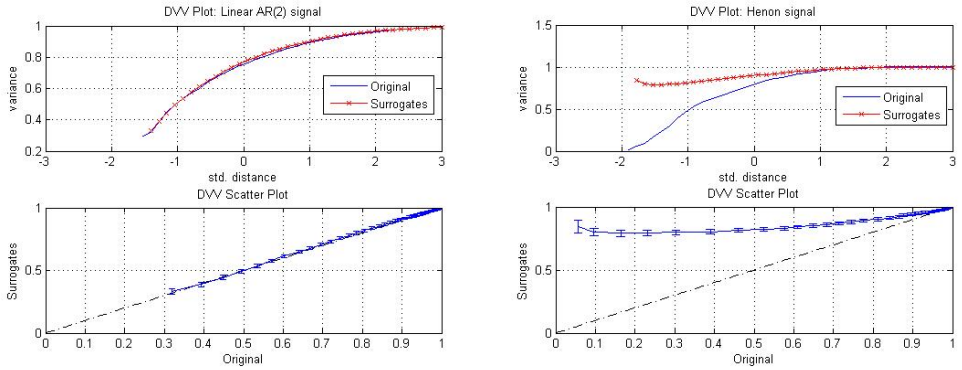
- For an optimal embedding dimension m and time lag τ, generate delay vector (DV): $x(k) = [x_{k-\tau},\cdots,x_{k-m\tau}]$ and corresponding target x_k
- The mean μ_d and standard deviation, σ_d, are computed over all pairwise distances between DVs, $\|x(i) - x(j)\|$ for $i \neq j$.

- The sets Ω_k are generated such that $\Omega_k = \{x(i) | \|x(k) - x(i)\| \leq \varrho_d\}$, i.e., sets which consist of all DVs that lie closer to $x(k)$ than a certain distance ϱ_d, taken from the interval $[\min\{0, \mu_d - n_d\sigma_d\}; \mu_d + n_d\sigma_d]$, e.g., uniformly spaced, where n_d is a parameter controlling the span over which to perform the DVV analysis.

- For every set Ω_k, the variance of the corresponding targets, σ_k^2, is computed. The average over all sets Ω_k, normalised by the variance of the time series, σ_x^2, yields the *target variance*, σ^{*2} :

$$\sigma^{*2}(\varrho_d) = \frac{\frac{1}{N}\sum_{k=1}^{N}\sigma_k^2(\varrho_d)}{\sigma_x^2} \tag{22}$$

  where N denotes the total number of sets $\Omega_k(\varrho_d)$



(a) AR(2) signal

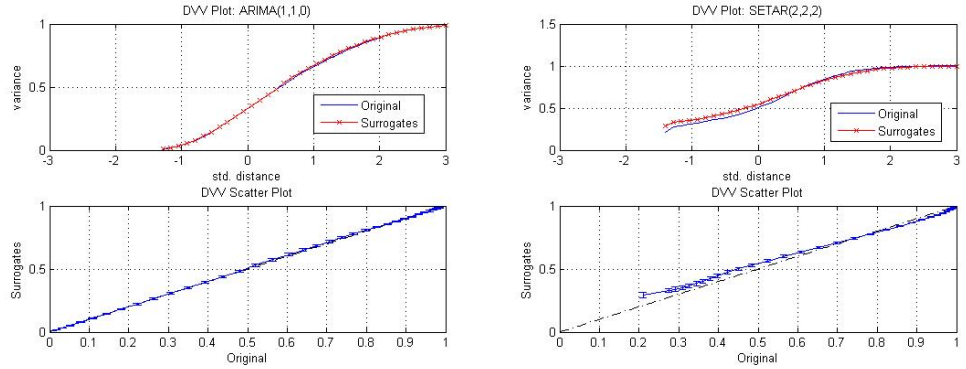
(b) Henon signal

Figure 2: Nonlinear and deterministic nature of signals. The first row of Diagrams 2a and 2b are DVV plots for a linear benchmark signal: AR(2) signal and a nonlinear benchmark signal: Henon signal, where the red line with crosses denotes the DVV plot for the average of 25 WiAAFT-based surrogates while the blue line denotes that for the original signal. The second row of Diagrams 2a and 2b denote the DVV scatter diagrams for those two signals, where error bars denote the standard deviation of the target variances of surrogates.

As a result of the standardisation of the distance axis, the resulting *DVV plot* are easy to interpret, as illustrated in the first row of Figure 2a and Figure 2b. The minimum target variance, which corresponds to the lowest point of the curve, is a measure for the amount of noise which is present in the time series. The presence of a strong deterministic component will lead to small target variances for small spans, n_d. At the extreme right, the DVV plots smoothly converge to unity, as illustrated in Figure 2a and Figure 2b. The reason behind this is that for maximum spans, all DVs belong to the same set, and the variance of the targets is equal to the variance of the time series. In the following step, the linear or nonlinear nature of the time series is examined by performing DVV analysis on both the original and 25 WiAAFT surrogate time series. Due to the standardisation of the distance axis, these plots can be conveniently combined within a scatter diagram, where the horizontal axis corresponds to the DVV plot of the original time series, and

the vertical to that of the surrogate time series. If the surrogate time series yield DVV plots similar to that of the original time series, as illustrated by the first row of Figure 2a, the DVV scatter diagram coincides with the bisector line, and the original time series is judged to be linear, as shown in second row of Figure 2a. If not, as illustrated by first row of Figure 2b, the DVV scatter diagram will deviate from the bisector line and the original time series is judged to be nonlinear, as depicted in the second row of Figure 2b.

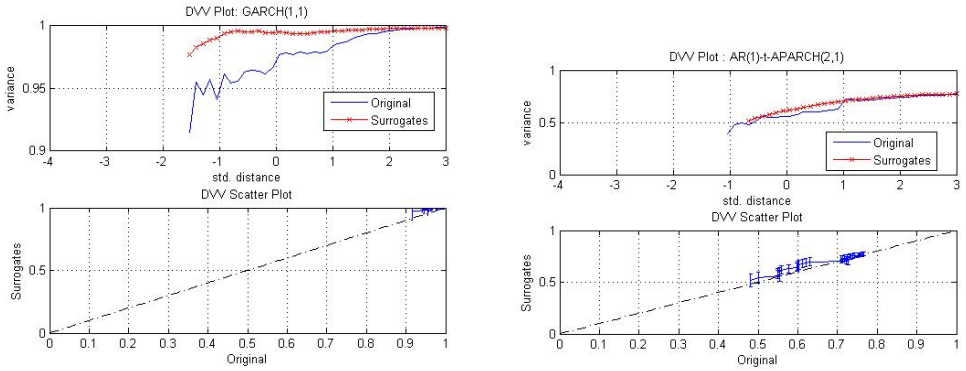
In Figure 3 and Figure 4, we provide the structure of the DVV analysis on some simulated processes such as: a self-exciting threshold autoregressive process (SETAR), linear autoregressive integrated moving average (ARIMA) signal, a Generalised autoregressive conditional heteroskadastic process (GARCH), and a signal with a mean equation as Autoregressive (AR) process and the innovations generated from a skewed Student-t APARCH (asymmetric power autoregressive conditional heteroskadastic) process. The interpretation of the DVV analysis is not different from the previous illustration.



(a) DVV analysis on ARIMA(1,1,0) signal

(b) DVV analysis on SETAR(2,2,2) signal

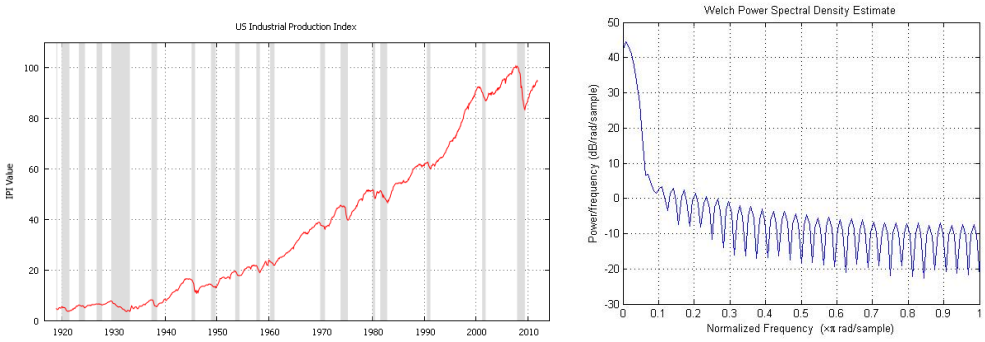
Figure 3: DVV analysis on ARIMA and SETAR signals



(a) DVV analysis on GARCH(1,1) signal

(b) DVV analysis on AR(1)-t-APARCH(2,1) signal

Figure 4: DVV analysis on GARCH and AR(1)-t-APARCH(2,1) signals



(a) Time Domain (b) Frequency Domain

Figure 5: US Industrial Production Index (IPI) time seris. Diagram 5a is the plot of the monthly IPI series for the period: 1919:01 - 2011:11 ($n = 1115$), where the shaded regions corresponds to NBER business cycles peaks and troughs from 1920. Diagram 5b is the plot of the IPI in frequency domain.

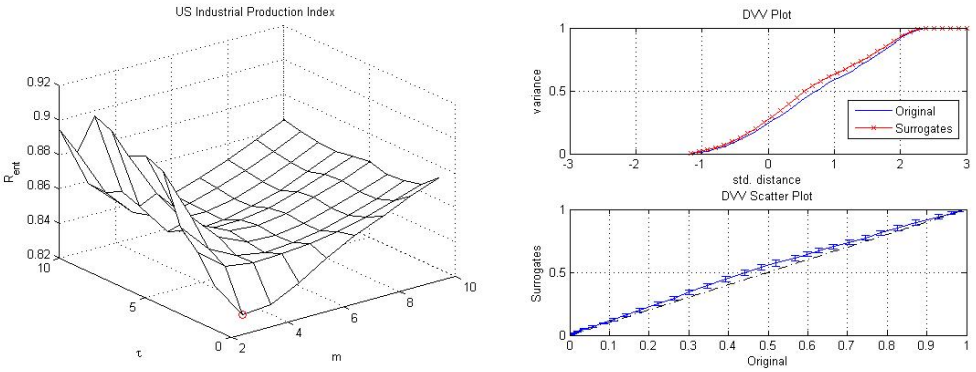
3. Data Analysis

It is now well-known that the United States and all other modern industrial economies experience significant swings in economic activity. In this section, we perform analysis to characterize and detect nonlinear schemes for the US business cycle. The monthly US Industrial Production Index (IPI) time series[7] spanning over the period January,1919 to November,2011 is considered for our analysis. Figure 5a is the plot of the monthly IPI series for the period: 1919:01 - 2011:11, implying 1115 observations, where the shaded regions corresponds to NBER[8] business cycles peaks and troughs from 1920. Figure 5b is the plot of the IPI spectral density representation. Firstly, we characterize the *nature* of the time series using the DVV method with WiAAFT surrogates and then employ complex Morlet wavelet to discover the cycles or hidden nonlinear information in the data. Coefficient plots from the wavelet analysis is displayed in Figure 8 were the the business cycles peaks and troughs dates are represented by blue peaks. Alternative representation of the Coefficient plot is Figure 7 with the blue regions of the Angle coefficients plot corresponding to the business cycles peaks and troughs dates.

We now give a comprehensive analysis of the data under consideration. To begin with, we opted for the differential-entropy based method (Gautama T., 2003) to determine the optimal embedding parameters, i.e., the embedding dimension, m, and the time lag, τ, for the DVV method with WiAAFT surrogates. The optimal embedding parameters obtained for the IPI series are $m = 3$ and $\tau = 1$ with an *entropy ratio*, $R_{ent}(m, \tau) = 0.8240$, indicated as an open circle in the diagram with a clear structure in Figure 6a. This results indicates the presence of time correlations, in the time series, implying a higher degree of structure, thus, a lower amount of disorder. Based on this, we obtained the DVV scatter diagram for the IPI data, as illustrated in the diagram in the second row of Figure 6b. The DVV analysis of the time series, figure 5a, indicates a presence

[7]The data can be downloaded from Federal Reserve Bank of St. Louis
[8]National Bureau of Economic Research

of a strong deterministic component governing the dynamics since the minimum target variance, $\sigma^{*2}_{min} \approx 0$ for small spans. This can be observed on the DVV plot on the first row of Figure 6b. From the second row of Figure 6b, the DVV scatter diagram does not coincide with bisector line, indicating its nonlinear nature eventhough it approaches the bisector. The DVV analysis suggests that the time series under consideration exhibits weak nonlinearity with deterministic characteristics. This results indicates that the nonlinearity might be relevant for only a small part of the observations. Based on the DVV-analysis, a linear autoregressive (AR) model applied to this data is likely to yield superior point forecasts compared to other nonlinear models like smooth transition autoregressive (STAR) models and Markov switching model which allows for regime specific mean and variances with autoregressive parameters kept constant. We suggest[9] the use of self-exciting threshold autoregressive (SETAR) models in capturing the dynamics of the underlying time series. This is based on analysis comparing both in-sample and out-of-sample forecasting performance of linear Autoregressive (AR) model, STAR models and self-exciting threshold autoregressive (SETAR) models applied to the time series under consideration.



(a) Differential-Entropy based method

(b) DVV analysis of the US IPI time series

Figure 6: The *nature* of the US IPI series

In the following step, we perform wavelet analysis on the data using wavelets of the form in equation (16) at different bandwidths f_b and center frequency f_c. The Morlet wavelet that best detect cycles and hidden information in the data is given in equation (18). The colormap used in the coefficient plots and scalogram plot ranges from blue to red, and passes through the colors cyan, yellow, and orange. Looking at the *blue peaks* on the phase-angle plot of the Coefficient plots in Figure 8, we discover the US business cycle peaks and troughs dates. The modulus plots on the Coefficient plots in Figure 7 and Figure 8 corresponds to the amplitude of the cycle. Our results on the US business cycle peaks and troughs dates coincides with the published dates from the National Bureau of Economic Research (NBER). The wavelet energy at time periods are displayed on the scalogram in Figure 9a and the pseudo-frequency corresponding to scales are displayed in Figure 10.

We are able to discover all the NBER business cycles peaks and troughs from 1920, which

[9]The results are not reported in this paper since it is beyond the scope of this work. However, unreported results are available from authors upon request.

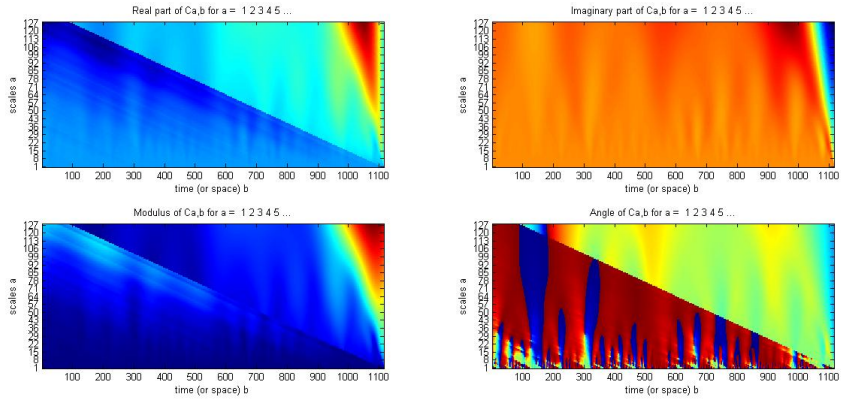


Figure 7: Complex Morlet wavelet transform coefficients plots: First row is the Real and Imaginary parts and the second row represents Modulus and Angle coefficients. The colormap ranges from blue to red, and passes through the colors cyan, yellow, and orange. The blue peaks on the Angle Coefficient plot corresponds to the monthly dates for peaks and troughs of U.S. business cycles from 1920. The corresponding amplitudes can be read from the Modulus plot.

is reported on Table 2. The Wall Street Crash of 1929, followed by the Great Depression[10] of the 1930s - the largest and most important economic depression in the 20th century - are well captured on the phase-angle coefficient plot in the figure 8 for time period (128 - 235). The corresponding amplitude and energy of the Great depression are represented by the cyan color on the modulus plot in Figure 8 and the scalogram in Figure 9a respectively. The three recessions between 1973 and 1982: the oil crisis - oil prices soared, causing the stock market crash are shown on the blue peaks of the phase-angle coefficient plot in the figure 8 for the time periods (659-676), (733-739), (751-767). Furthermore, the bursting of dot-com bubble - speculations concerning internet companies crashed is also detected for the time periods (987-995). We can observe from Figure 9a that the scalogram reveals high energy of 0.02% at time periods (1068 - 1115) corresponding to the Financial crisis of 2007-2011, followed by the late 2000s recession and the 2010 European sovereign debt crisis. In order to compare the late-2000s financial crisis with the Great Depression of the 1930s, we perform the wavelet analysis considering the period of January, 1919 to December, 1939. The US business cycle peaks and troughs dates for this period were clearly detected as shown on the Coefficient plots in Figure 11. Looking at Figure 9b, we clearly observe high energy levels on the interval $(0.02\% - 0.04\%)$ for period of the Great Depression of the 1930s. This results, as observed from the scalograms in Figure 9, provides evidence to support the claim made by many economists that the late-2000s financial crisis, also known as the Global Financial Crisis (GFC) is the worst financial crisis since the Great Depression of the 1930s.

[10]The Great Depression in the United States, where, at its nadir in 1933, 25 percent of all workers and 37 percent of all nonfarm workers were completely out of work (Gene Smiley: Rethinking the Great Depression (American Ways Series)

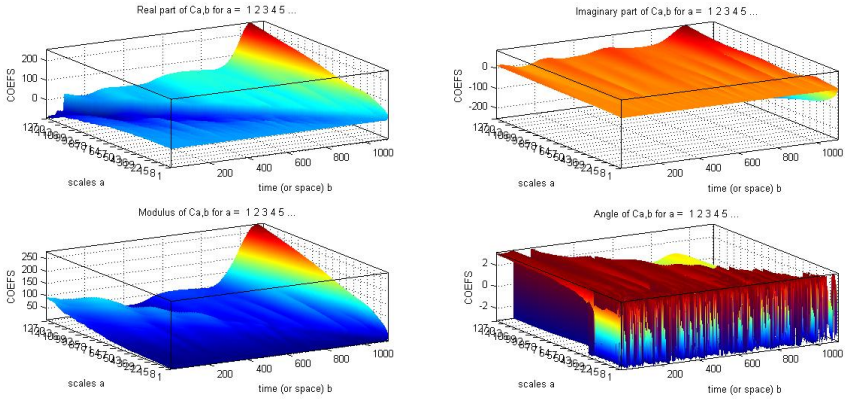
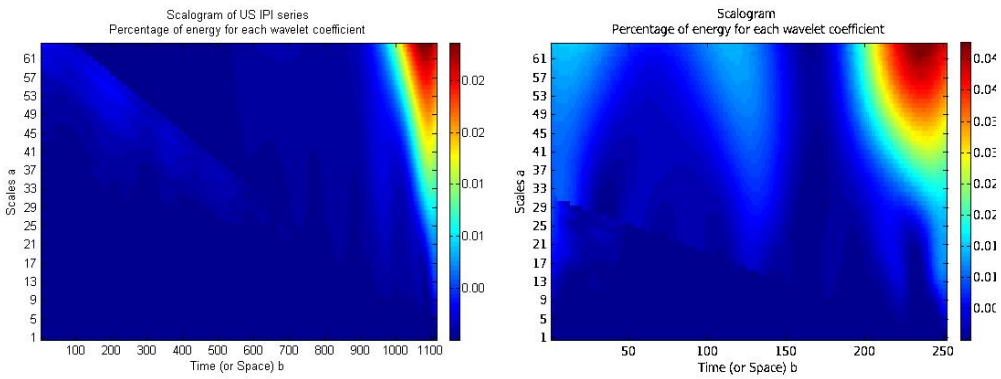


Figure 8: Complex Morlet wavelet transform coefficients plots: First row is the Real and Imaginary parts and the second row represents Modulus and Angle coefficients. The colormap ranges from blue to red, and passes through the colors cyan, yellow, and orange. The blue regions on the Angle Coefficient plot corresponds to the monthly dates for peaks and troughs of U.S. business cycles from 1920. The corresponding amplitudes can be read from the Modulus plot.



(a) Scalogram from January, 1919 to November, 2011.

(b) Scalogram from January, 1919 to December, 1939.

Figure 9: The colormap of the scalograms ranges from blue to red, and passes through the colors cyan, yellow, and orange. The bar on the left-hand side of the scalogram plot indicates the percentage of energy for each wavelet coefficient. The scalogram in figure 9a reveals high energy of 0.02%, at time periods (1068 - 1115) corresponding to the period of late-2000s financial crisis, also known as the Global Financial Crisis. Higher energy levels on the interval (0.02% – 0.04%) can be clearly observed, in figure 9b for the Great Depression of the 1930s.

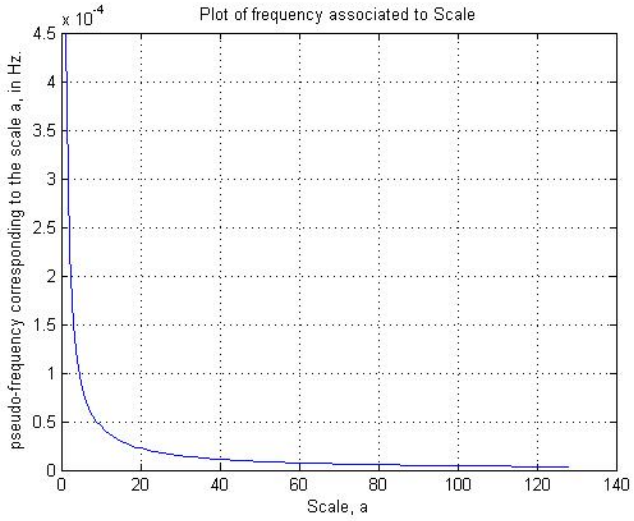


Figure 10: The pseudo-frequency associated to scale, in Hertz (Hz)

Business Cycle dates		
Peak	Trough	Time Index
1920:01	1921:07	13-31
1923:05	1924:07	53-67
1926:10	1927:11	94-107
1929:08	1933:03	128-171
1937:05	1938:06	221-235
1945:02	1945:10	314-322
1948:11	1949:10	359-370
1953:07	1954:05	415-425
1957:08	1958:04	464-472
1960:04	1961:02	496-506
1969:12	1970:11	613-623
1973:11	1975:04	659 -676
1980:01	1980:07	733-739
1981:07	1982:11	751-767
1990:07	1991:03	859-867
2001:03	2001:11	987-995
2007:12	2009:06	1068-1086

Table 2: Business Cycle Peaks and Troughs in the United States, 1920-2009. The peak and trough dates, in the format YYYY:MM, represent the start and end of “episodes” of some sort. (see http://www.nber.org/cycles.html)

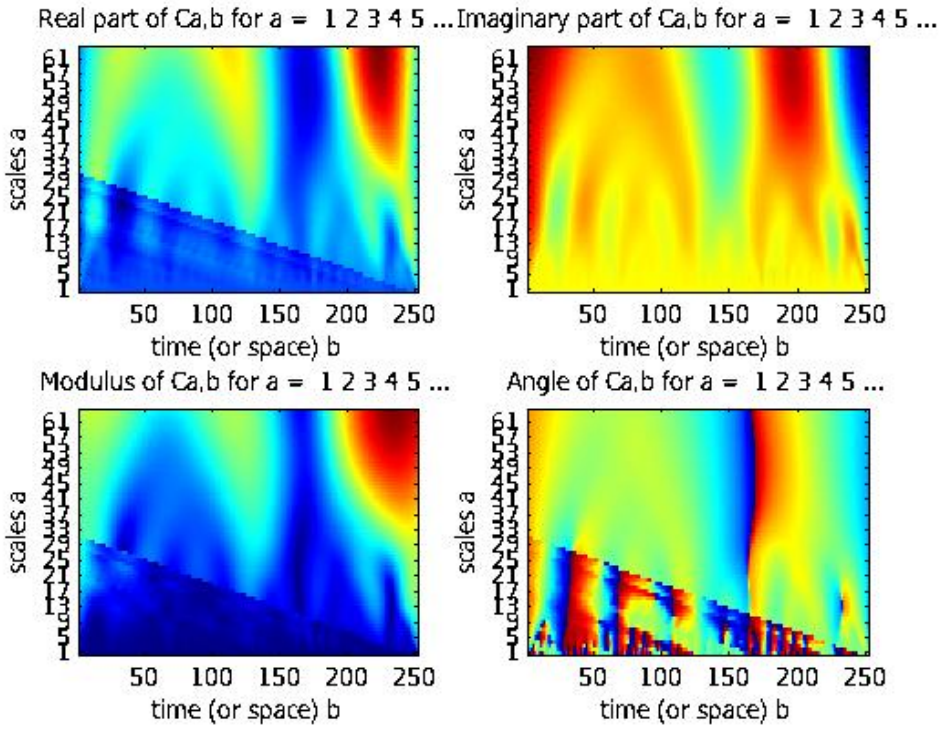


Figure 11: Complex Morlet wavelet transform coefficients plots: First row is the Real and Imaginary parts and the second row represents Modulus and Angle coefficients. The colormap ranges from blue to red, and passes through the colors cyan, yellow, and orange. The blue regions on the Angle Coefficient plot corresponds to the monthly dates for peaks and troughs of U.S. business cycles from 1920 to 1939. The corresponding amplitudes can be read from the Modulus plot.

4. Conclusion

We have utilised the *Delay Vector Variance*(DVV) method and Wavelet analysis to characterize the *nature* of the US Industrial Production Index data. It has been shown that the DVV method and the Complex Morlet Wavelet can serve as an alternative way to detect and characterize nonlinearity schemes in the US business cycle. In particular, we have shown that the wavelet approach serves as an alternative methodology for turning-point detection in business cycle analysis.

References

Addison, P. S., 2005. Wavelet transforms and ecg: a review. Physiological Measurement 26, 155–199.

Aguiar-Conraria, Soares, 2011. The continuous wavelet transform: A primer, nIPE – WP 16.

Bruce L, Koger C, J. L., 2002. Dimentionality reduction of hyperspectral data using discrete wavelet transform feature extraction. IEEE Transactions on Geoscience and Remote Sensing 40, 2331–2338.

Casdagli, M. C. . W. A. S., 1994. Exploring the continuum between deterministic and stochastic modelling, in time series prediction: Forecasting the future and understanding the past. Reading, MA: Addison-Wesley, 347–367.

Chambers, D., Mandic, J., 2001. Recurrent neural networks for prediction: learning algorithms architecture and stability. Chichester, UK: Wiley.

Gallegati, M., 2008. Wavelet analysis of stock returns and aggregate economic activity. Computational Statistics and Data Analysis 52, 3061–3074.

Gallegati, M., Gallegati, M., 2007. Wavelet variance analysis of output in g-7 countries. Studies in Nonlinear Dynamics and Econometrics 11 (3), 6.

Gautama T., Mandic D.P., V. H. M., 2003. A differential entropy based method for determining the optimal embedding parameters of a signal. In Proceedings of ICASSP 2003, Hong Kong IV, 29–32.

Gautama T., Mandic D.P., V. H. M., 2004a. The delay vector variance method for detecting determinism and nonlinearity in time series. Physica D 190 (3–4), 167–176.

Gautama T., Mandic D.P., V. H. M., 2004b. A novel method for determining the nature of time series. IEEE Transactions on Biomedical Engineering 51, 728–736.

Hegger R., H. K., Schreiber, T., 1999. Practical implementation of nonlinear time series methods: The tisean package. Chaos 9, 413–435.

Ho A, K. Moody, G. P., C. Mietus, J. Larson, M. L., Goldberger, D., 1997. Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. Circulation 96, 842–848.

Kaplan, D., 1994. Exceptional events as evidence for determinism. Physica D 73 (1), 38–48.

Keylock, C., 2006. Constrained surrogate time series with preservation of the mean and variance structure. Physical Review E 73, 036707.

Kugiumtzis, D., 1999. Test your surrogate data before you test for nonlinearity. Physics Review E 60, 2808–2816.

Leonenko, N., Kozachenko, L., 1987. Sample estimate of the entropy of a random vector. Problems of Information Transmission 23, 95–101.

Schreiber, T., Schmitz, A., 1996. Improved surrogate data for nonlinearity tests. Physics Review Lett. 77, 635–638.

Schreiber, T., Schmitz, A., 2000. Surrogate time series. Physica D 142, 346–382.

Stock, M. W., Watson, J. H., 1999. Forecasting inflation. Journal of Monetary Economics 44 (2), 293–335.

Teolis, A., 1998. Computational signal processing with wavelets. Birkhauser.

Terasvirta, T., September 2011. Modelling and forecasting nonlinear economic time series, wGEM workshop, European Central Bank.

Walden, D., Percival, A., 2000. Wavelet Methods for Time Series Analysis. Cambridge: Cambridge University Press.

Whitcher, B. G. P., Percival, D., 2000. Wavelet analysis of covariance with application to atmospheric time series. Journal of Geophysical Research 105, 14941–14962.