BRNO UNIVERSITY OF TECHNOLOGY

Faculty of Electrical Engineering
and Communication

DOCTORAL THESIS

Brno, 2022 Ing. Eva Klejmová

BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF RADIO ELECTRONICS
ÚSTAV RADIOELEKTRONIKY

SELECTED ASPECTS OF STATISTICAL SIGNIFICANCE TESTING IN TIME-FREQUENCY ANALYSIS
VYBRANÉ ASPEKTY TESTOVÁNÍ STATISTICKÉ VÝZNAMNOSTI V ČASOVĚ-FREKVENČNÍ ANALÝZE

DOCTORAL THESIS
DIZERTAČNÍ PRÁCE

AUTHOR — Ing. Eva Klejmová
AUTOR PRÁCE

SUPERVISOR — doc. RNDr. Jitka Poměnková, Ph.D.
ŠKOLITEL

BRNO 2022

ABSTRACT

This doctoral thesis is focused on analyses and assessment of the quality of the frequency and time-frequency transform of the data and the formulation of recommendations for working with such methods. When using these methods, the question arises of how to evaluate which components of the spectrogram are statistically significant and which are not. In this thesis, we analyze the properties of standard statistical significance tests. We discuss their advantages and disadvantages taking into account the heteroskedastic character of data. Based on our experiments we propose two types of improved testing methods that reduce the negatives standard tests. The final step is creating a framework for data filtering using our proposed methods.

KEYWORDS

spectrogram, time-frequency analysis, wavelet transform, Fourier transform, autoregressive process, significance testing, co-movement filtering

ABSTRAKT

Přeložená dizertační práce se zabývá analýzou a posouzením kvality odhadu frekvenční a časově-frekvenční transformace dat a formulaci doporučení pro práci s metodami. Při použití těchto metod vyvstává otázka, jak vyhodnotit, které složky spektrogramu jsou statisticky významné a které nikoli. V této práci analyzujeme vlastnosti standardních testů statistické významnosti. Diskutujeme o jejich výhodách a nevýhodách s ohledem na heteroskedastický charakter dat. Na základě našich experimentů jsou v práci navrženy dva typy testovacích metod, které snižují negativní aspekty standardních testů. Práce jen zakončena vytvořením rámce pro filtrování dat pomocí námi navržených metod.

KLÍČOVÁ SLOVA

spektrogram, časově frekvenční analýza, vlnková transformace, Fourierova transformace, autoregresivní proces, testování významnosti, filtrování společného pohybu

Author's Declaration

Author:	Ing. Eva Klejmova
Author's ID:	115199
Paper type:	Doctoral thesis
Academic year:	2021/22
Topic:	Selected Aspects of Statistical Significance Testing in Time-Frequency Analysis

I declare that I have written this paper independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the paper and listed in the comprehensive bibliography at the end of the paper.

As the author, I furthermore declare that, with respect to the creation of this paper, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll. of the Czech Republic, Section 2, Head VI, Part 4.

Brno ....

author's signature*

*The author signs only in the printed version.

ACKNOWLEDGEMENT

I would like to thank the supervisor of my thesis, doc. RNDr. Jitka Poměnková, Ph.D. for her patience, recommendations, suggestions, valuable comments and generous help. I would like to thank everyone, who kept me motivated to finish this work, my friends, my coworkers and my family. Finally, I would like to thank E. Ctiborová for her brilliant idea to buy gummy bears in bulk.

Contents

List of Figures

List of Tables

Introduction

The need to analyze the data can be found across a variety of scientific disciplines. Despite the diversity of disciplines, it is a common goal to obtain the maximum information from data analysis to help solve the tasks set. Concerning the scientific area, such data are given as observations in the form of time series of input signals. The standard analytical instruments are given in the time or frequency domain. Linking of both approaches giving us a more compact view can be done via time-frequency techniques. The combination of time and frequency tools provides a more efficient means of data analysis, allowing us a deep look into the signal structure.

The graphical representation of time-frequency analysis is a spectrogram. Its estimation can be done via several parametric or nonparametric methods. The most used are short-time Fourier transform, estimation via the time-frequency varying autoregressive process, and wavelet transform. While the periodogram is the classic estimator for stationary signals, multiple windows or short-time Fourier transformation can be useful for non-stationary signals. The time-frequency varying autoregressive process is a simplification of the general autoregressive moving average model. The signals can be corrupted by noise which can affect the precision of instantaneous frequency; therefore is good to investigate several types of analyses methods to reach the required precision.

The key aspect of time-frequency analysis is the precision of the estimated spectrogram. For further processing and filtering of data, it is appropriate to specify which components of the spectrogram are statistically significant and which are not. There are several test methods for this purpose. One of the most used methods is based on the identified distribution of background noise, and several requirements need to be met for its appropriate usage. Other methods work with the use of geometric and topological changes or simulations of background noise. The obtained selection of significant regions can then be used for further description and filtering of the data, taking into account the objectives of the analysis, either in time, frequency, or time-frequency domain.

Considering the current progress in the field and the gap in current research, this work is focused on describing the framework of analysis from the use of individual methods, through statistical significance testing of their results, and finally filtering of data based on these significance tests. We can divide this thesis into several parts. The first part deals with the description and the usage of basic methods for time-frequency analysis. The second part deals with the description of commonly used significance test methods and the definition of their weaknesses. In the third part, methods are proposed that eliminate these weaknesses. In the last part, these proposed methods are used to filter the required data components.

1 State of the Art

The need to describe and analyze the input signal for further use occurs across all scientific disciplines, from technical to social sciences. In terms of approach, we can define the analysis in the time domain (TD), frequency domain (FD), and time-frequency domain (TFD). Fundamental analysis can be performed in the time domain. Such an analysis deals with the changes in a signal over a span of time, i.e. variation of the amplitude of the signal with time. In contrast, the frequency domain describes the behavior of the signal across a given frequency band concerning a range of frequencies and can include information on phase shift. It is possible to use time-domain techniques or frequency domain techniques separately; however, their ability to capture the frequency behavior of the analyzed time series with respect to the time is somewhat limited. The combination of time and frequency tools provides a more efficient means of statistical analysis, reflecting the fact that the time-frequency analysis of input signal is an instrument that has been used in interdisciplinary analysis for a long time.

Time-frequency (TF) techniques are an instrumental approach, reflecting both the time and frequency behavior of input time series. These approaches predominates in the last decade in many fields of science. It is a useful instrument in natural sciences [1–4], engineering [5, 6], biology or medicine [7–9] or social and economic sciences.

The time-frequency representation of the signal can be estimated via several approaches. The most common method is Short Time Fourier Transform (STFT). The periodogram or its modification, such as the multiple window method using Slepian sequences [9] can also be used. We can also use estimation via the time-frequency varying Autoregressive Process (TFAR) [10], wavelet analysis (CWT) [11,12] or alternatively Modified empirical mode decomposition method [13]. While the periodogram belongs to the group of the classic estimator for stationary signals, multiple windows or STFT can be a valuable instrument for non-stationary signals [9,14–16]. As Jiang and Mahadevan [17] wrote, the advantage of the wavelet analysis is that it can capture the features of non-stationarity signal due to the simultaneous time-frequency decomposition of inputs. The TFAR process is a simplification of the general AutoRegressive Moving Average (ARMA) model.

Among the advantages of Fourier transform and its derivatives, we can include low computational complexity and a wide range of software and hardware implementations with a selection of optimal parameters that provide satisfactory results. AR process used for estimation of signal spectrum representation provides fair results, especially in very short signals when STFT tends to fail. For longer signals, it provides good results [18]. In such cases, the variance of insignificant cyclical

components that usually take the character of noise has a lower level than in the case of STFT. This advantage can be useful when we investigate thresholding such as in [19]. The time-varying representation of the AR process provides a more complex view compared to a simple spectrum estimate in the frequency domain only. It has time and frequency resolution corresponding to the size of the window and the size of window overlap, which must be selected. In such a way, it is similar to the STFT method. Unfortunately, the disadvantage of the method is its accuracy which strongly depends on the selection of optimal lag order. Therefore, it is good to investigate various optimization criteria for its optimal selection. Another disadvantage is that there are not many existing implementations; most are only on a software level. Continuous wavelet transform is a relatively new method compared to Fourier transform. As pointed out in [20] or [17], the wavelets have several advantages. It is applicable to non-stationary data. It also has the ability to uncover the latent process with changing cyclical patterns. Such features are typical for an economic time series. Additionally, the wavelet analysis has very good time resolution, and there is no need to optimize the parameters. There is only discussion about the mother wavelet and the scale selection.

The need to validate the estimated model arises with the application of TF methods on real values with respect to the application area (engineering, medicine, etc.). This leads to the significance testing [21, 22]. The fundamental work in this field can be found in Torrence and Compo [23]. This paper presents the comparison of the windowed Fourier transform to the wavelets. The authors also focus on the relationship between wavelet scale and Fourier frequency and the choice of an appropriate wavelet basis function. The proposed statistical significance test is given for wavelet power spectra and is based on theoretical derivation for white and red noise processes.

Motivated by the work of Torrence and Compo [23], Ge [24] proposes significance testing of wavelet power and wavelet power spectrum. He derived the sampling distributions for the power spectrum of a Gaussian White Noise (GWN). And also for the wavelet power of GWN. He proved that the results given by [23] are numerically accurate when if the sampling period factor is incorporated. Ge [25] uses the same methodological approach for wavelet cross-spectrum and linear coherence. However, one of the disadvantages of this test is that it takes into account the variance of the entire signal. In specific cases where the data exhibit highly variable volatility, the variance of the whole signal may not be sufficiently descriptive. The question is then how this affects the accuracy of the test and how to interpret the results.

A similar approach to Torrence and Campo can be found in the work of Schulte et al. [21]. They use geometric and topological methods for assigning contiguous significance regions of significant wavelet coefficient with respect to selected noise

models with application to climatic data. Also, James and Fleming [22] use the Torrence and Compo approach to identify significant spatial scales of pattern and spatial boundaries in geo-science.

Model validation of structural dynamics example is proposed by Jiang and Mahadevan [17]. They investigate simulation-based predictions of structural response on the virtually generated data. The authors use testing with the help of Monte Carlo simulations to infer whether the model prediction and experimental observation represent two coherent processes. Wang et al. [26] present another point of view by introducing the general sequential Monte Carlo method to estimate the probability density function and to optimize wavelet transform for extracting bearing fault features.

Given the above methods, we found that the literature insufficiently describes several areas. One of these areas is the effect of the character of the data on the significance testing of time-frequency methods. This character can manifest itself through structural changes in the data that lead to changes in volatility. This raises the question of how to interpret the results of standard tests in this case. A related insufficiently described area is how TFA can be used to provide additional information that will contribute to subsequent signal analysis.

Most standard methods have been designed to work with technical data. The physical nature of these types of data is usually known; their description is available, their behavior and their content are known, and there is knowledge of what their deterministic components may look like. However, there are scientific areas where factors, that are often unpredictable, may affect or change the character of the data. This problem is typical, for example, for economic data, where due to various events (economic shocks, crises, pandemics, etc.) diverse structural changes may arise, such as in trend, volatility, growth, etc.

Given the above observations, this work will focus on the analysis of economic, technical, and simulated data.

2 Dissertation Objectives

This dissertation thesis deals with analyses and assessment of the quality of the frequency and time-frequency transform and with the formulation of recommendations for working with such methods. We take into account how much a priori information will help to obtain maximum information about the data.

We researched literature and resources and evaluated current progress and gaps in this field. We found out that the literature does not deal with the influence of the data character on the significance testing of the time-frequency methods. Most of the literature focuses on the technical area, where the application of methods and interpretation of results is facilitated by knowledge or information about the data character. In some cases, such as the selected photonic Doppler velocimetry data set, these are rather experimental data, therefore, a priori information may be more general.

The different types of data in terms of nature are economic data, which are less informative in terms of technical analysis. The process and mechanism of this data generation are influenced by factors such as unexpected events, economic shocks, psychological factors, etc., which are difficult to predict and simulate. This may appear as structural changes in the data, to which standard methods may not respond correctly in all cases. For the TFD application and subsequent testing, the question then arises as to how data with structural changes can be analyzed to obtain relevant results. The third type of data is simulated data used to verify standard and designed methods.

We focus on the issue of statistical testing of data mentioned above in order to verify the standard methods and to propose methods for cases where the data volatility is changing in time. Based on these we defined the following objectives of the thesis.

Objective I. *Is it possible to use and combine different characteristics of individual TFA methods to obtain relevant spectrogram?*
In Chapter 6 we propose an approach to incorporate advantages and suppress disadvantages of individual methods in order to bring out significant components and suppress noise.

Objective II. *How can we modify standard tests to eliminate/reduce their disadvantages and shortcomings in case of data with changing volatility?*
In Chapters 8 and 9 we use knowledge gained in Chapter 7 to propose SAB and LAB testing approach and recommendation for their usage in such case.

Objective III. *How can we use these modified test methods for subsequent data filtering?*

In some cases, it is useful to work selected spectral components represented in the time domain. Especially in the case of evaluation of time-series co-movement. Therefore, in Chapters 10 and 11 we propose co-movement filtering as an instrument for obtaining co-movement indicator. In its construction and application, the expertise gained from Objective II. is used.

3 Methodological Background

This chapter is divided in four sections which provide methodical background of well known methods used in following chapters. The first section introduces basic tools for time-frequency analysis. The second provides background for co-movement measures, the third describes the singular value decomposition, and the last section contains selected aspects of kernel smoothing.

3.1 Time-Frequency Analysis (TFA)

As mentioned in Chapter 1 the time-frequency techniques are nowadays common instrument for analysis of different input signals or time series. Wide range of its usage can be found across different scientific disciplines. Estimating signal spectrograms can be done via several parametric or non-parametric methods. The most commonly used non-parametric methods are Fourier transform and its derivatives, wavelet transform or multiple window method. One of the most used parametric methods is the autoregressive process. All these approaches are shortly described in following subchapters.

3.1.1 Fourier Transform (FT)

One of the most common methods used for spectrum estimation is the Fourier transform (FT) and its modifications. Given the character of input signals used in this work, the focus is on discrete-time representation. If the input signal $s(n)$ has an infinite length, then the discrete-time Fourier transform (DTFT) is used [10].

$$S_{\mathrm{DTFT}}(f) = \sum_{n=-\infty}^{\infty} s(n)\mathrm{e}^{-\mathrm{j}2\pi fn}. \tag{3.1}$$

If an input signal $s(n)$ is a discrete time series of N elements, then the discrete Fourier transform (DFT) is used. It can be defined as [27]

$$S_{\mathrm{DFT}}(f) = \sum_{n=0}^{N-1} s(n)\mathrm{e}^{-\mathrm{j}2\pi fn/N}. \tag{3.2}$$

The transform form frequency domain back to time domain is called an inverse discrete Fourier transform and can be defined as [27]

$$s(n) = \frac{1}{N}\sum_{n=0}^{N-1} S_{\mathrm{DFT}}(f)\mathrm{e}^{\mathrm{j}2\pi fn/N}. \tag{3.3}$$

A slight modification of this method is short time Fourier transform. In this modification a sliding observation window is used. The individual spectrum estimations are then sorted in time and can be plotted as spectrogram. The usual mathematical definition of the STFT is [27]

$$S_{\mathrm{STFT}}(m,f) = \sum_{n=0}^{N-1} s(n)g(n-m)\mathrm{e}^{-\mathrm{j}2\pi fn/N} \tag{3.4}$$

where $s(n)$ is the input signal, $g(n)$ is the window function. The window function is symmetric discrete signal with unit norm $||g|| = 1$. Commonly used types of windows are Rectangle, Hamming, Gaussian, Hanning, etc. (for details see [27]). The magnitude squared of resulting spectrogram $S_{\mathrm{STFT}}(m,f)$ is called power spectral density.

3.1.2 Wavelet Transform (WT)

Continuous wavelet transform (CWT) is usually defined as the integral of a signal with respect to the wavelet function. It can be described as integral of analyzed signal with base function (mother wavelet) [27]:

$$S_{\mathrm{CWT}}(a,\tau) = \int_{-\infty}^{\infty} s(u)\frac{1}{\sqrt{a}}\psi^*\left(\frac{u-\tau}{a}\right)\mathrm{d}u, a > 0, \tau \in R \tag{3.5}$$

where the "$*$" denotes complex conjugation, $s(n)$ is the time series, $\psi^*\left(\frac{u-\tau}{a}\right)$ is a scaled version of the mother wavelet, τ denotes the time shift, and a denotes the scale (or frequency). We can also define WT of a discrete time series $s(n)$ as the convolution of $s(n)$ with a scaled and translated version of the wavelet function $\psi(\cdot)$ [23]

$$S_{\mathrm{CWT}}(a,\tau) = \sum_{n=0}^{N-1} s(n)\psi^*\left(\frac{n-\tau}{a}\right)\delta n \tag{3.6}$$

where δn is the time step, a is the wavelet scale and τ is the time shift.

To be the invertible transform, basis (mother wavelets) functions must be mutually orthogonal, have zero mean value and limited to finite time interval [27]. That is

$$\begin{aligned} &i) \quad \int_{-\infty}^{\infty}\psi\left(\tfrac{n-\tau}{a}\right)\mathrm{d}n = 0, \\ &ii) \quad \int_{-\infty}^{\infty}\psi\psi^*\left(\tfrac{n-\tau}{a}\right)\mathrm{d}n = 1, \\ &iii) \quad 0 < C_\psi = \int_0^{\infty}\tfrac{|\Psi(\omega)|^2}{\omega}; \Psi(\omega) = \int_{-\infty}^{\infty}\psi\left(\tfrac{n-\tau}{a}\right)\mathrm{e}^{-\mathrm{j}\omega n}\mathrm{d}n \end{aligned} \tag{3.7}$$

where $\Psi(\omega)$ is the Fourier transform of $\psi(\omega)$. To satisfy assumptions for the time-frequency analysis, waves must be compact in time as well as in the frequency representation. There are several types of mother wavelets which can be used (e.g. Gaussian, Haar, Daubechies, Morlet etc.) (for details see [27]). One of the most

commonly used is the complex Morlet wavelet and can be defined as product of a complex sine wave and a Gaussian window[28]:

$$\psi(n) = \exp\left(\frac{-n^2}{2\sigma^2}\right)\exp(\mathrm{j}\omega_0 n) \tag{3.8}$$

where σ is Gaussian window width in time and ω_0 is the central frequency of the wavelet. The complex Morlet wavelet is a substantially complex exponential modulated by a Gaussian envelope.

In order to recalculate the pseudo-frequency, corresponding to the scale a, the following equation can be used [25].

$$\omega(\tau) = \frac{\omega_0}{a(\tau)} \tag{3.9}$$

where $\omega = 2\pi f$ and τ is the time shift.

For an orthogonal wavelet and under certain admissibility conditions (see (3.7)), we can define its inverse form, i.e. inverse continuous wavelet transform (ICWT), as

$$s(n) = ICWT\{S_{\mathrm{CWT}}(a,\tau)\} = \tfrac{1}{C_\psi}\int_\infty^\infty\int_{-\infty}^\infty S_{\mathrm{CWT}}(a,\tau)\tfrac{1}{\sqrt{a}}\psi^*\left(\tfrac{n-\tau}{a}\right)\mathrm{d}\tau\tfrac{da}{a^2} \tag{3.10}$$

where $C_\psi < \infty$ comes from the admissibility condition [29].

3.1.3 Autoregressive (AR) Process

A different approach in comparison with the methods mentioned above is the TFAR process. This method uses a parametric approach and creates a model generating an input signal. The analyzed signal $s(n)$ is then regarded as the output of a linear filter influenced by white noise ϵ with variance σ_{w}^2. The autoregressive process AR(p) of the order p for an input signal $s(n)$ can be described by model given by the equation

$$s(n) = c + \sum_{i=1}^{p} a_i s(n-i) + \epsilon(n) \tag{3.11}$$

where $a_i,\ i = 1,\dots p$ are the parameters of autoregressive model of the order p, c is the constant and $w(n)$ is the white noise.

The output spectrum can be described as [10]

$$S_{\mathrm{AR}}(f) = |H(f)|^2\,\sigma_{\mathrm{w}}^2 \tag{3.12}$$

where $H(f)$ is a linear time variant filter. Several methods for estimating AR model parameters can be used. The most common are the Burg method (BU), Yule-Walker method (YW), unconstrained least-squares method (LS) [10]:

$$S_{\mathrm{AR}}^{\mathrm{BU}}(f) = \frac{\widehat{E}_{\mathrm{p}}}{\left|1 + \sum_{k=1}^{p} \widehat{a_{\mathrm{p}}}(k)\mathrm{e}^{-\mathrm{j}2\pi fk}\right|^2}$$

$$S_{\mathrm{AR}}^{\mathrm{YW}}(f) = \frac{\widehat{\sigma}_{\mathrm{wp}}^2}{\left|1 + \sum_{k=1}^{p} \widehat{a_{\mathrm{p}}}(k)\mathrm{e}^{-\mathrm{j}2\pi fk}\right|^2} \tag{3.13}$$

$$S_{\mathrm{AR}}^{\mathrm{LS}}(f) = \frac{\widehat{E}_{\mathrm{p}}^{\mathrm{LS}}}{\left|1 + \sum_{k=1}^{p} \widehat{a_{\mathrm{p}}}(k)\mathrm{e}^{-\mathrm{j}2\pi fk}\right|^2}$$

where $\widehat{a_{\mathrm{p}}}(k)$ are estimates of the AR parameters, p is the lag order, $\widehat{E}_{\mathrm{p}}$ is the total least-squares error, $\widehat{\sigma}_{\mathrm{wp}}^2$ is the estimate minimum mean-square value for the pth order predictor and $\widehat{E}_{\mathrm{p}}^{\mathrm{LS}}$ is the residual least squares error.

For the AR process, appropriate selection of lag order p is an important aspect. Selecting a low level order leads to excessive smoothing of the spectrum. Furthermore, if the level of p is selected too high, a non-significant spectral coefficient can be arises as high peak. Several information criteria can be used to ensure optimal selection. Most commonly used are Akaike information criterion (AIC), Minimum description length (MDL), Hannan-Quinn information criterion (HQC) and Bayesian information criterion (BIC) [10], [30]:

$$AIC(p) = \ln \widehat{\sigma}_{\mathrm{wp}}^2 + \frac{2p}{N}, \tag{3.14}$$

$$MDL(p) = N \ln \widehat{\sigma}_{\mathrm{wp}}^2 + p \ln N, \tag{3.15}$$

$$BIC(p) = \ln \widehat{\sigma}_{\mathrm{wp}}^2 + \frac{p}{N} \ln N, \tag{3.16}$$

$$HQC(p) = \ln \widehat{\sigma}_{\mathrm{wp}}^2 + 2\frac{p}{N} \ln(\ln N) \tag{3.17}$$

where $\widehat{\sigma}_{\mathrm{wp}}^2$ is estimated variance of linear prediction error, p is the order and N is length of the signal. The order is selected as optimal when information criterion reaches minimum.

The AR process as described above provides the result only in the frequency domain. To obtain results in time-frequency domain (spectrogram) similar approach to STFT can be used, i.e. sliding observation window. Window functions commonly used in STFT can also be applied. We will denote this method as TFAR.

3.2 Co-Movement Measures

To evaluate similarity between two signals several co-movement measures are commonly used. The widely used methods in the TF domain are cross-spectrum and coherence [31]. For the two time series $x(n)$ and $y(n)$, $n = 1, \ldots, N$, the cross-spectrum coefficients measure the local covariance of these two variables in respective time and frequency/scale and can be defined as

$$CS_{xy}(f,n) = S_x(f,n)S_y(f,n)^* \tag{3.18}$$

where $S_x(f,n), S_y(f,n)$ are the time-frequency representations of signal x resp y. The symbol "*" denotes complex conjugation [20, 23]. If these representations are in complex numbers (depends on used method) the resulting cross-spectrum is also complex. Then we can define its square absolute value, i.e. power cross-spectrum (PCS) [23], as

$$|PCS_{xy}(f,n)|^2 = |S_x(f,n)|^2|S_y(f,n)^*|^2. \tag{3.19}$$

The phase part (quadrature spectrum) is then defined as [17]

$$\Phi_{xy} = \tan^{-1}\left[\Im(CS_{xy}(f,n))/\Re(CS_{xy}(f,n))\right] \tag{3.20}$$

where symbols $\Im$ and $\Re$ denote imaginary part and real part of cross-spectrum $CS_{xy}(f,n)$. In case of two identical signal the cross-spectrum is equal to one. If the input signals differs greatly, their cross-spectrum is approaching to zero.

Another possibility for measuring co-movement in the TF domain is coherence $COH_{xy}(f,n)$ of two time series. It can be defined as the PCS normalized to the squared module of the time-frequency representations of the analyzed signals [17].

$$COH_{xy}(f,n) = \frac{CS_{xy}(f,n)}{\sqrt{|S_x(f,n)|^2}\sqrt{|S_y(f,n)|^2}}. \tag{3.21}$$

This function is modified as follows when WT representations of signals are used as input: [17, 31]

$$COH_{xy}(f,n) = \frac{CS_{xy}(f,n)Q}{\sqrt{Q|S_x(f,n)|^2}\sqrt{Q|S_y(f,n)|^2}} \tag{3.22}$$

where Q is a smoothing operator defined as

$$Q(S) = Q_{\text{scale}}(Q_{\text{time}}(S(f,n))) \tag{3.23}$$

where Q_{scale} denotes smoothing along the wavelet scale axis and Q_{time} smoothing in time.

In an analogous manner, as for CWT or STFT (equations (3.10),(3.3)), we can calculate an inverse transform for co-movement measure, such as PCS or coherence.

3.3 Kernel Smoothing

Kernel analysis have been established as effective techniques and unified framework for pattern discovery. They are used to solve a non-linear problem by using a linear classifier. Let (x_i, Y_i), $i = 1, \ldots, N$ be a sequence of observations (x, Y), where in general $x \in R, x_i \in [0, 1]$ is a design point and Y is an observation. In our case Y can represent frequency value of significant spectral coefficient obtained from spectrogram and x is transformed equidistantly in $[0; 1]$ in correspondence to the time of values in Y. The dependency of value Y on x for fixed design regression model (x are not randomly chosen) can be described in following form

$$Y_i = m(x_i) + \epsilon_i, \quad i = 1, \cdots, N, \tag{3.24}$$

where m is an unknown regression function, and let conditions $E(\epsilon_i) = 0 \quad i = 1, \ldots, N,\ D(\epsilon_i) = \sigma^2 > 0 \quad i = 1, \ldots, N$ hold [32].

Kernel function K of smoothness μ, order (ν, k) and $K \in \mathcal{S}^{\mu}_{\nu,k}$ must satisfy following conditions [33, 34]:

$$\begin{aligned} &i) \quad K^{(-j)} = K^{(j)} = 0, && j = 0, \ldots, \mu - 1 \\ &ii) \quad \textstyle\int_{-1}^{1} x^j K(x)dx = \begin{cases} 0 & 0 \le j < k, j \ne \nu \\ (-1)^{\nu}\nu! & j = \nu \\ \beta_k \ne 0 & j = k \end{cases} \end{aligned} \tag{3.25}$$

where ν,k are non-negative integers and $0 \le \nu < k,\ \mu \ge 1$.

The general formula for a kernel estimator can be expressed by the form

$$\hat{m}(x) = \sum_{i=1}^{n} W_i(x, h) Y_i, \tag{3.26}$$

where $W_i(x, h)$ are weight functions depending on h, i, x and K. Bandwidth $h = h(n)$ is a positive constant and K is a kernel. Denote $K_h(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right)$. According to the weighted function we can distinguish several estimators. One of the most commonly used is Nadaraya-Watson (NW) estimator [32]:

$$\hat{m}(x, h) = \frac{\sum_{i=1}^{n} K_h(x_i - x) Y_i}{\sum_{i=1}^{n} K_h(x_i - x)}. \tag{3.27}$$

where $\sum_{i=1}^{n} K_h(x_i - x) \ne 0$. If $\sum_{i=1}^{n} K_h(x_i - x) = 0$, than we can define $\hat{m}(x, h) = 0$.

3.4 Singular Value Decomposition

The singular value decomposition (SVD) is useful instrument in solving least-square problems via the process of orthogonal decomposition of rectangular matrix. The method follow two steps, the decomposition and the reconstruction [10], [35].

The first step of SVD is to make trajectory matrix from the input time series $x(n), n = 1, \dots N$ without any missing values. The trajectory matrix T with $K \times L$ dimension and takes the form

$$T = \begin{pmatrix} x_1 & x_2 & \dots & x_L \\ x_2 & x_3 & \dots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \dots & x_N \end{pmatrix} \tag{3.28}$$

The window L such that $2 < L < N/2$ to embedded initial time series $\mathbf{x} = (x_1, x_2, \dots, x_N)$ is defined by user. In this way we map time series $\mathbf{s}$ into the lagged sub-vectors $T_i = (x_i, x_{i+1}, \dots, x_{i+L-1})', i = 1, 2, \dots, K$, where $K = N - L + 1$. Consequently we apply on the trajectory matrix T singular value decomposition to obtain trajectory matrices $T_i, i = 1, \dots, L$. From an eigenvalue analysis of TT' we collect the eigenvalues $\lambda_1 \geq \dots \geq \lambda_T$ where $r = \text{rank}(TT')$ and the corresponding left and right singular vectors, respectively denoted as U_i and V_i such that $T_{K\times L} = U_{K\times K} D_{K\times L} V'_{L\times L}$, where U is orthonormal matrix, D is the diagonal matrix of the order L and V is a square orthonormal matrix [35]. We can write

$$\begin{aligned} T &= T_1 + T_2 + \dots + T_r \\ &= U_1\sqrt{\lambda_1}V'_1 + \dots U_r\sqrt{\lambda_r}V'_r \\ &= \textstyle\sum_{i=1}^{r} U_i\sqrt{\lambda_1}V'_i. \end{aligned} \tag{3.29}$$

In the second step of SVD we are focused on reconstruction. The first task is selection of M leading components of decomposed time series and exclude the reminding $(r - M)$ associated to the noise [35]. That is

$$T = \sum_{i=1}^{M} U_i\sqrt{\lambda_1}V'_i + \epsilon \tag{3.30}$$

where the first summands corresponds to the time series and ϵ denotes an error term.

4 Optimization of AR Parameters

As we already mentioned in Chapter 3.1.3, the results of AR process considerably depends on model parameters. To obtain the best possible spectral representation it is crucial to select optimal method for estimation of AR coefficients with regard to the nature of input time signal. In this chapter we apply simulations to evaluate precision of spectral estimation of selected methods for AR coefficients estimation. Therefore, on the basis of simulation, we analyze the behavior of Yule-Walker, Burg and least-squares method. Furthermore, in the case of AR process, next key factor is the choice of an optimal lag order on which the accuracy strongly depends. Therefore, it is good to investigate various optimization criteria for its optimal selection [10]. We examine the advantages and disadvantages of these estimation methods and selection of appropriate lag order and formulate recommendation for its usage.

4.1 Experiments and Results

To evaluate performances of AR methods, the autoregressive process of order 20 was created. Selection of the order was done with regard to desired complexity of signal. The requirements were several spectral peaks with diverse amplitude and diverse frequency spacing. Therefore we chose level of 20 which provides this sufficient complexity with reasonable computational requirements. AR coefficients were arbitrary selected to create signal with sufficient complexity and are presented in Table 4.1. Ideal spectral representation of the signal (using (3.13)) is shown in Fig.4.1.

a_0	1	a_1	–1.04	a_2	0.78	a_3	–0.88	a_4	0.45
a_5	0.35	a_6	0.34	a_7	–0.04	a_8	–0.40	a_9	–0.10
a_{10}	0.04	a_{11}	0.36	a_{12}	0.38	a_{13}	–0.32	a_{14}	–0.35
a_{15}	–0.11	a_{16}	–0.08	a_{17}	0.26	a_{18}	0.15	a_{20}	–0.09
a_{21}	–0.31								

Tab. 4.1: AR coefficients.

In next step we created signal in time domain using these coefficients. Signals of two length were produced to assess the impact of signal lengths. Based on AR order of 20, we chose length of short signal to be 50 samples and length of long signal to be 500 samples. For each signal length multiple realization in time domain were created using equation 3.11 with $\sigma_{\mathrm{w}}^2 = 1$. Number of realizations was set to $N = 10, 100, 1000$. In each simulation least-squares, Burg and Yule-Walker method

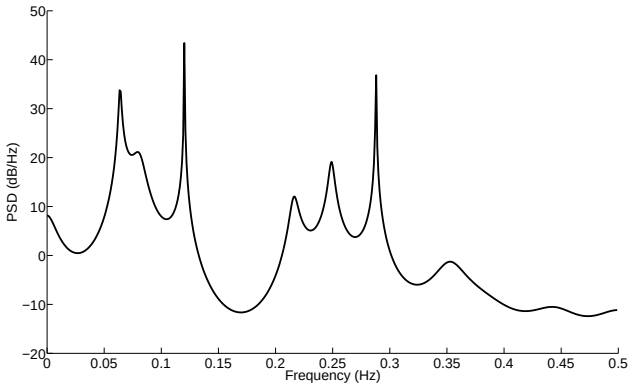


Fig. 4.1: Ideal spectral representation of simulated signal.

was applied to obtain estimated coefficients. Using equation (3.13) we obtained corresponding spectral estimation. To asses performance of both methods in the case of selecting non optimal lag order, the computation was done on selected range of order. The range was 1 to 30 in the case of short signal and 1 to 300 in the case of long signal.

To objectively measure quality of each method the mean square error (MSE) was calculated using ideal and estimated spectrum representation for each process. This was done according following formula [10].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y_i}\right)^2 \tag{4.1}$$

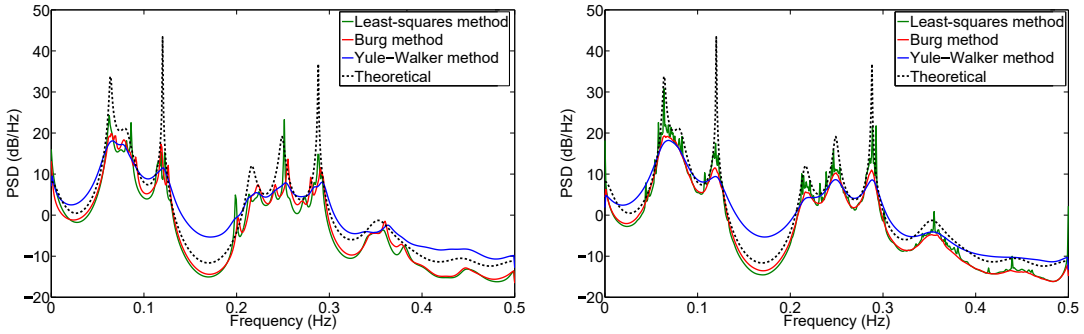
where Y_i is the value of theoretical spectral coefficient and $\widehat{Y_i}$ is its estimated counterpart. The lower the resulting number is the more accurate is the estimation.

Table 4.2 shows values of MSE in dB for selected methods. Lag order was set to optimum, i.e. 20 (this order was chosen while creating the original signal). The results suggest that with increasing number of simulations precision of spectral estimation also increases. In general, results of all methods are significantly better for the signal of longer length. The MSE is smaller for Yule-Walker method in all considered cases. Least-squares and Burg method show similar values in the case of the long signal. For the short signal Burg method provides better results.

Figure 4.2 (a) and (b) shows theoretical and estimated spectrum of short signal for 10 and 1000 simulations. Lag order was set to optimal value of 20. The frequency of x-axis is relative to sampling frequency (value of 0.5 in figures corresponds to 0.5 sampling frequency). We can see that the larger number of simulation leads to more precise estimation. Neither of examined methods was able to adequately describe rapid changes of spectral peaks for measuring on 10 simulations. In case of 1000 simulations least-squares method tends to be more accurate in the case of describing

	10 simulations		**100 simulations**		**1000 simulations**	
signal length	**short**	**long**	**short**	**long**	**short**	**long**
Yule-Walker	27.35	7.60	23.79	5.91	22.95	5.77
Burg	45.89	13.76	42.41	12.03	39.32	11.73
Least-squares	50.20	13.54	48.68	11.91	45.15	11.65

Tab. 4.2: Values of MSE in dB.



(a) 10 simulations

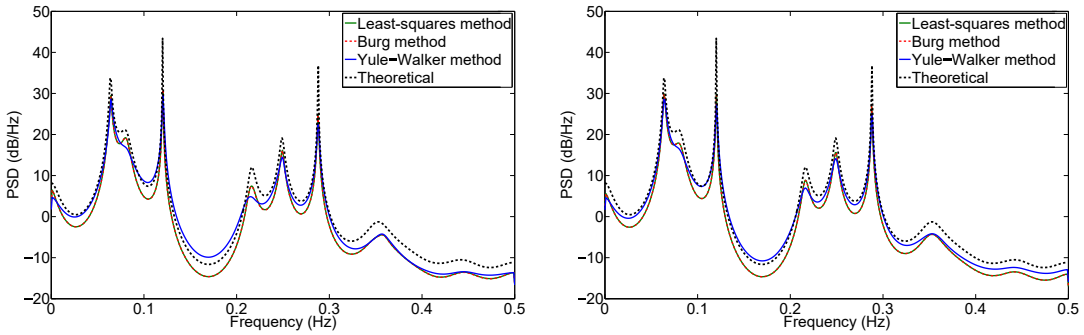
(b) 1000 simulations

Fig. 4.2: Theoretical and estimated spectrum of short signal, optimal lag.

amplitude of spectral peaks. However other two methods provide spectral line with greater smoothness. We can also see that the Burg method tends to be more accurate in capturing the shape of the graph than the Yule-Walker method.

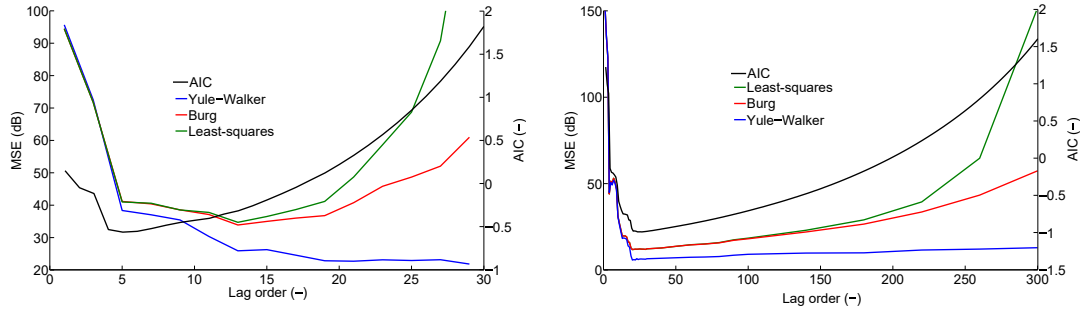
In Figure 4.3 (a) and (b) theoretical and estimated spectrum of long signal for 10 and 1000 simulations and optimal lag order is shown. Larger number of simulations also leads to more precise estimation and smoothness of the graph but in is not as significant as in the case above. Least-squares and Burg method show almost the same results and tend to be more precise in the capturing the shape of the graph than Yule-Walker method. However all methods were able to capture all significant spectral peaks.

In Figure 4.4 (a) and (b) we can see influence of the lag order on the precision of spectrum estimation. For signal of the length 1000, the minimum of AIC and of all methods (lowest error) corresponds to the order of original AR process, i.e. 20. We can see that the trend decrease rapidly to optimal lag order, after that the error tends to increase. In case of the short signal, the minimum lays on lower value that the optimal order. The trend also decreases rapidly to the lag order of 5. In case of Yule-Walker method the decrease continues, however error of Burg and least-squares method increase exponentially.



(a) 10 simulations (b) 1000 simulations

Fig. 4.3: Theoretical and estimated spectrum of long signal.



(a) short signal (b) long signal

Fig. 4.4: MSE and AIC of estimated spectrum for 1000 simulations, optimal lag.

4.2 Chapter Conclusion

To sum up our findings, we can state that Yule-Walker method proved to be more efficient when applied on a signal of the longer duration. On the other hand, Burg and least-squares method offers better peak detection for the short signal. Furthermore, the results showed that the performance of used methods also strongly depends on used lag order. Yule-Walker method performs best in scenarios, when higher lag orders are allowed. Burg and least-squares method is more suited in the cases, when we aim for lower lag orders.

5 Impacts of Input Data Character

In this chapter, the performance of individual TFA methods compared using different types of input data is proposed. Data from different disciplines show different characteristics, and therefore, the appropriate setting must be investigated. We take into account economic, technical, and simulated data. The parameters of the technical data are extensively described in the literature, this is due, among other things, to the fact that most of these data are known for their physical nature and accurate description of the background noise. The subsequent analysis is then simpler, which is reflected in the literature. In technical data, background noise (GWN, red noise, etc.) can be considered the rest after the removal of periodic components, whereas, in the case of economic data, the situation is not the same.

Economic data can be influenced by factors that are often unpredictable and may affect or change the character of the data. This can manifest as structural changes in the data, thus the question arises of how to analyze such data to obtain relevant results. Moreover in the case of economic data, their structure is more complicated. Usually, it contains structural trend-breaks, outliers, cyclical components of close frequencies which can occur or diminish in different time sub-periods (not during the whole time), or nested cycles with different frequency limited in time [36–39]. Moreover, the nature of economic indicators play an important role and can influence the character of the frequency structure, e.g. business cycles, financial cycles etc. Then, it is quite difficult to simulate the universal behavior of the economic series and its noising with a generalized artificial signal. Based on these differences, we offer instructions on how to approach selected TFA methods, taking into account the character of the data.

5.1 Simulated Data

As a representative of data composed of several cyclical components and background noise can be clearly described, we have created an artificial signal. The length was 2000 samples. In the time domain, this signal consists of four sine waves of four different frequencies. The following equation describes the created signal.

$$s(t) = \sum_{i=1}^{4} b_i(t) \sin(2\pi f_i t) \tag{5.1}$$

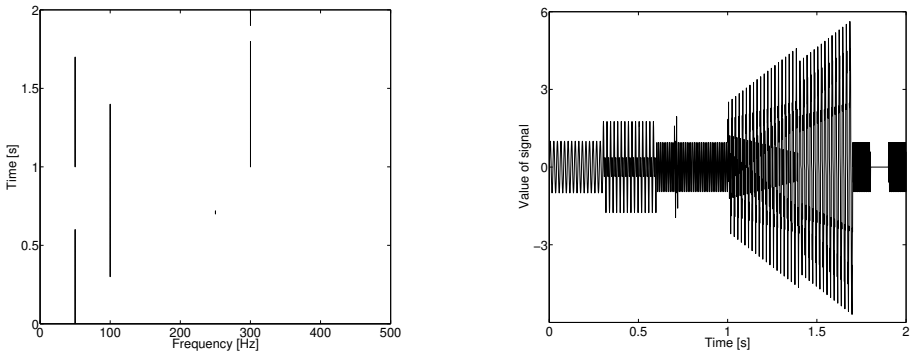
where f_i is the frequency of an individual wave (namely 50, 100, 250 and 300 Hz), t is the time in the range 0–2 seconds, b_i is its amplitude and it is defined as

$$
\begin{aligned}
b_1 &= \begin{cases} 1 & t \in \langle 0, 0.6 \rangle \,\mathrm{s}, \\ \frac{40}{7}t - \frac{33}{7} & t \in \langle 1.0, 1.7 \rangle \,\mathrm{s}, \\ 0 & \text{elsewhere}, \end{cases} \\
b_2 &= \begin{cases} 1 & t \in \langle 0.3, 1.4 \rangle \,\mathrm{s}, \\ 0 & \text{elsewhere}, \end{cases} \\
b_3 &= \begin{cases} 1 & t \in \langle 0.70, 0.72 \rangle \,\mathrm{s}, \\ 0 & \text{elsewhere}, \end{cases} \\
b_4 &= \begin{cases} 1 & t \in \langle 1.0, 1.8 \rangle \cup \langle 1.9, 2.0 \rangle \,\mathrm{s}, \\ 0 & \text{elsewhere}. \end{cases}
\end{aligned} \tag{5.2}
$$

As shown in Fig. 5.1 the resulting signal consists of several simple ones. We chose this approach to test ability of selected methods to identify following moments:

- Very short signal (on frequency 250 Hz)
- Short break of signal (on frequency 300 Hz)
- Longer break of signal (on frequency 50 Hz)
- Two close signals (on frequency 50 and 100 Hz)

Ideal spectral representation along with time representation of the resulting signal is shown in Fig. 5.1(note that for illustration, the signal is shown without noise). For better simulation of actual conditions, white Gaussian noise was added. The signal to noise ratio (SNR) was set to 15dB. Several methods were used to estimate the spectrogram of the simulated data, namely the estimation via AR and STFT and CWT. The simulated signal was divided into equidistant blocks of a length of 96 values.



(a) Ideal spectrum (b) Time representation

Fig. 5.1: Simulated signal.

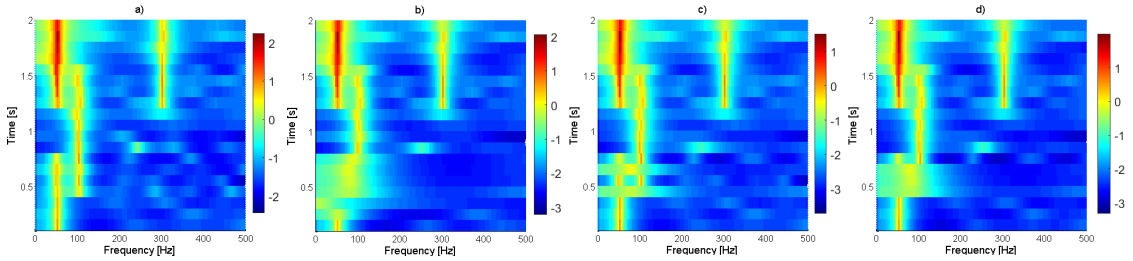


Fig. 5.2: Comparison of information criteria a) AIC, b) MDL, c) HQC, d) BIC.

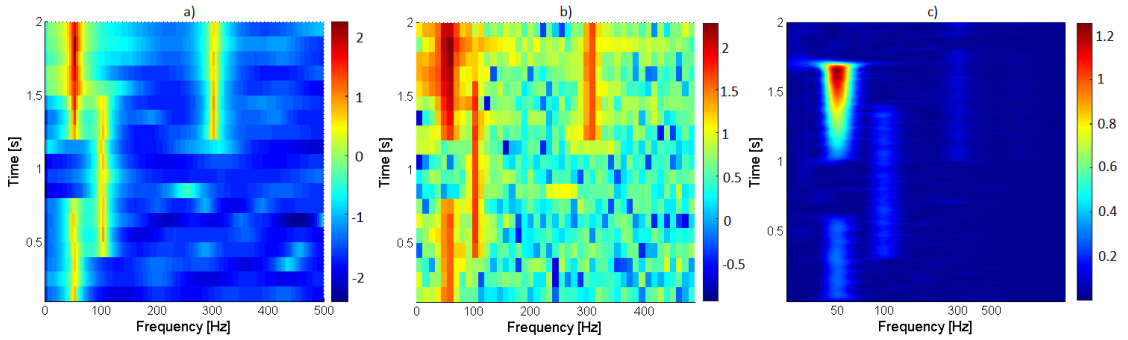


Fig. 5.3: TFA using a) TFAR, b) STFT, c) CWT.

5.1.1 Application

For estimating AR coefficients in TFAR method, due to a large number of data points, the Yule-Walker method was chosen [10]. In order to optimize AR lag order, several information criteria were considered. We chose four different information criteria, namely AIC, MDL, HQC and BIC (see eq. (3.14)–(3.16)). The optimal lag length corresponding to the chosen criteria was estimated separately for each data block; therefore, the resulting lag length varies. The best results were achieved for AIC, as shown in Fig. 5.2.

In case of CWT, we used a complex Morlet wavelet as the mother wavelet using a scale parameter with a maximum value $a = 64$. The graphical representation of the wavelet spectrogram was modified to be comparable with the results of the other methods. The frequency axis was recalculated to a logarithmic scale. The resulting estimates are presented in Fig. 5.3.

5.1.2 Summary

When comparing the ideal spectrum (Fig. 5.1a) with the results of the spectrogram using the TFAR (Fig.5.3a), we can identify all long term components. Identifying the short term component b_3 using formula (5.2) is possible but not so clear. The

estimate of the spectrogram via STFT (Fig. 5.3b) provides similar results as with AR in all simulated components. All these three methods mentioned above were unable to recognize rapid time changes in signal components denoted as b_1 and b_4 in formula (5.2). The results of the wavelet spectrogram (Fig. 5.3c) show its ability to identify all components, including rapid time changes; however, the precision of identification depends on the optimization of the scale parameter, the choice of wavelet type, and graphical 3D representation.

To sum up, the achieved results show a good ability to identify simulated components for the AR process and wavelet analysis. In comparison to all other methods, i.e. CWT and STFT, the AR spectrogram shows smooth results with the best frequency resolution but cannot capture rapid time changes. The wavelet analysis does not provide such good frequency resolution, but only this method was able to identify rapid time changes.

5.2 Economic Data

Economic data or economic statistics are data (quantitative measures) describing an actual economy, past or present. These are typically found in the time-series form, covering more than one time period or in cross-sectional data in one time period. For time-series data, reported measurements can be hourly (e.g. for stock markets), daily, monthly, quarterly, or annually. Estimates such as averages are often subjected to seasonal adjustment to remove weekly or seasonal-periodicity elements, for example, holiday-period sales and seasonal unemployment.

In this chapter, we use a gross domestic product (GDP) to represent economic data with not clearly descriptive background noise. The motivation was to show the influence of economic data on the TF estimate, and GDP was chosen as a suitable significant macroeconomic indicator. As it is common in econometric analyses of such data, we use seasonally adjusted quarterly data of GDP, volume index in OECD reference year 2005 [40] of the United States (US) and United Kingdom (UK) in 1956/01-2014/04, Korea in 1970/02-2014/02. All variables are in first-order difference of natural logarithms (FODLOG).

5.2.1 Application

After transformation (into FODLOG values), the data was analyzed using CWT. We investigated the impact of scale value and type of wave. In the first step, scale setting was established on 64, 128, and 256 values which are commonly used. Concerning sample size, data structure, and anticipated events, we took the optimal scale of 128. We used three mother wavelets from the group of waves, namely Morlet,

complex Morlet, and Daubechies. Morlet wavelet was chosen as it is commonly used for this type of application. Complex Morlet wavelet is based on standard Morlet with the advantage of providing complex results, making it possible to obtain phase part (quadrature) of the spectrum. Evaluation of this part of the spectrum will be subjected to further research. Daubechies wavelet of order five was chosen as it is the second commonly used and can assess lower frequencies and business cycles. The results of time-frequency wavelet representations of cospectra for USA, UK, and Korea are given in Fig 5.4.

In the case of cospectra estimate via TFAR process, we used Burg approach for coefficient estimates on 20 samples with 50% overlay and Hann window. The optimal value of the lag order was based on AIC criteria. For better illustration, frequencies were recalculated to show business cycles in years using $f = \frac{1}{T}$. This recalculation was done with respect to sampling frequency (4 samples per year). Minimal observable business cycles correspond to 0.5 of f_s and is 0.5 year.

5.2.2 Summary

Comparison of results can be done in the following direction. The first is a comparison among chosen mother wavelet; the second is among data. Focusing on cospectra for complex Morlet waves, we can identify the co-moving areas across a range of periods (i.e. 1/frequencies) from very long cycles (till 8 years), across business cycles (8 to 1.5 years) to short cycles (1.5 years and less). In some cases (USA-UK, USA-Korea), the most significant co-movement can be found around 5 years long period, while in another case (UK-Korea), the most significant co-movement can be found for very long periods, see Fig. 5.4. In the case of the Morlet mother wave, the results of USA-UK show the most significant co-movement for a very long period (more than 20 years). We can also find some co-movement among 2.5 and 9 years which is not so significant in this case. The co-movement for UK-Korea is significant among 2.5-20 years and in case of USA-Korea there is a range of significant periods starting at 2.5 years and including very long periods (more than 20 years). Application of Daubechies mother wave produces similar results to Morlet mother wave, but there is no such significant co-movement for shorter periods, i.e. for periods range 2.5 to 20. The complex Morlet wave gives a possibility to study phase shift, which is impossible with the usage of Morlet or Daubechies wave. Applying Morlet and Daubeschie to our data revealed rather longer cycles. Therefore, the authors suggest removing such components via advanced filtering method using Baxter-King or Christiano-Fitzgerald band-pass filter [41] to get better visibility of other co-movement periodicities. A specification of range can easily establish the frequency range that we want to remove before using a proper band-pass filter.

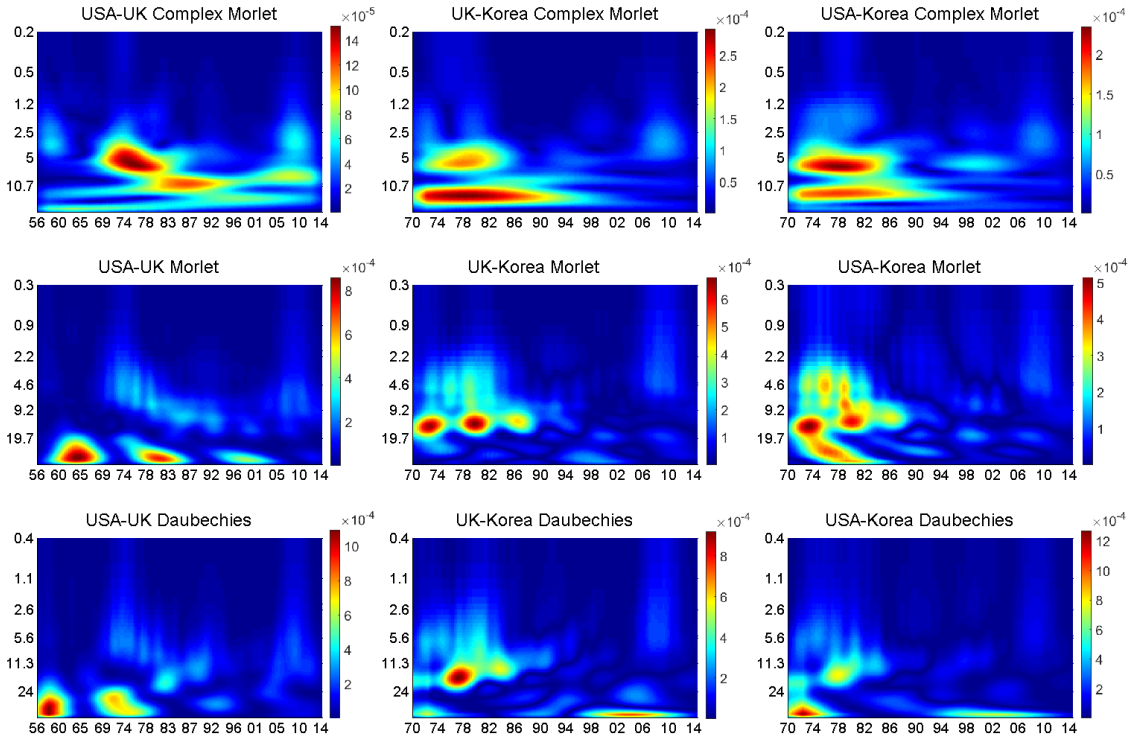


Fig. 5.4: CWT cospectrum (x-axis represent time in years and y-axis represent time cycles in years).

Figure 5.5 shows results for co-movement measuring via cospectrum for AR process. For all countries, we can identify strong significant co-movement in a wide range of frequencies, predominately in long periods. Opposite to wavelet analysis, we can also see significant co-movement in rapid changing periods (shorter than 2.5 years). Unfortunately, with respect to the sample size, the better time resolution is impossible. This confirms the fact given in [42] which recommends for AR approach rather than monthly data or higher frequency data (weekly, daily etc.). We can admit that in such case the AR method will confirm results from wavelet analysis much better than in quarterly data.

The wavelet approach to co-movement measures shows a good ability to capture

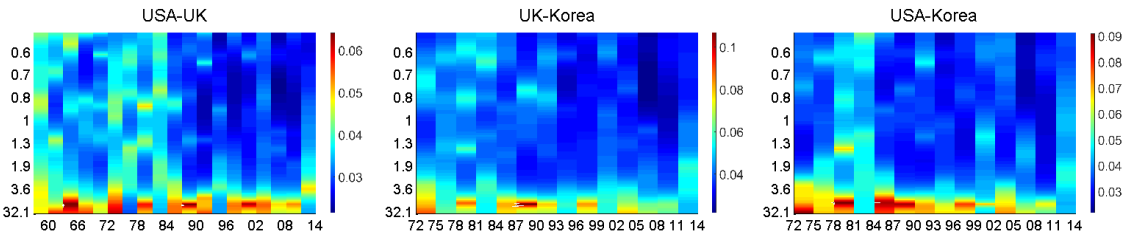


Fig. 5.5: AR cospectrum (x-axis represent time in years and y-axis represent time cycles in years).

rapid changes in time. However, its ability to identify the length of business cycles strongly depends on mother wavelet selection. Our results suggest that complex Morlet wavelet was able to capture short business cycles better and for this, we recommend its usage for data where long cycles are not present. The second possibility is to remove the permanent component via a band-pass filter with respect to the aim of the consequent analysis. Another property of this wavelet is its ability to provide complex output, making phase spectrum evaluation possible. Daubechies wavelet provided good results for long business cycles. However, short business cycles can be shown as less significant. In the case of co-movement measured via the AR process, our results confirm better frequency resolution, but in the case of a small sample size (which our quarterly data are), the time resolution is not as good as for higher frequency data. Therefore, we suggest using these methods for time series where monthly data are available and when frequency/business cycles resolution is more critical than time resolution.

5.3 Engineering Data

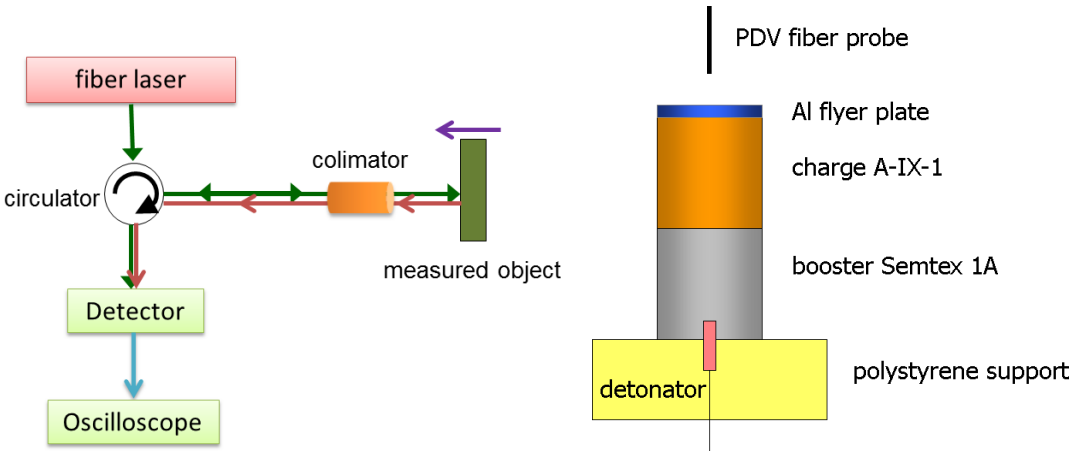
Another selected data type was Photonic Doppler Velocimetry Data (PDV) data. Their parameters are based on physical nature, their structure thus differs from economic data and the background noise is descriptive. For applying the selected methods of TFA, we used recorded data of an aluminum metal plate acceleration by detonation products of brisant high explosive. The data was obtained using PDV [43], [44]. Used setup is shown in Fig. 5.6(a). A fiber laser with wavelength λ_0 of 1550 nm was used to feed a 3-port circulator. Light from the circulator travels towards the probe (collimator or bare fiber end), where it partially exits and partially reflects back. The light that exits the probe reflects from the surface of a measured object and reenters the probe. If the object is in motion, the frequency of reflected light is changed by Doppler shift f_d. This frequency is then combined with the non-Doppler-shifted frequency of the laser source f_0. The resulting signal with frequency f_b equal to the Doppler-shift is captured by the detector and visualized by the oscilloscope. For a more detailed description of PDV, see [45] [46]. In the case of our measurement, no amplifier was used between the detector and the oscilloscope. Detailed schematic of measured object can be seen in Fig. 5.6(b) [43]. To recalculate Doppler-shift to the velocity of the moving target v following formula (5.3) can be used. Light from the circulator travels towards the probe (collimator or bare fiber end), where it partially exits and partially reflects back. The portion of the light that exits the probe travels towards the measured surface, reflects back, and enters the probe. If the target moves, the light has a Doppler-shifted frequency). The back-reflected part is non-Doppler-shifted and has a frequency of the laser source

f_0. The two different frequencies combine and create beats with frequency f_b equal to the difference between the shifted and non-Doppler-shifted one. The beat signal is captured by the detector and visualized by the oscilloscope.

$$v = \left(\frac{\lambda_0}{2}\right) f_\mathrm{b} \tag{5.3}$$

where λ_0 is the wavelength of the non-Doppler-shifted light (the laser wavelength), f_b is the Doppler-shift and v is the resulting velocity.

Two signals obtained using PDV were selected. Both with a sampling rate of 25Gs/s and length of 20µs giving us 500k samples. The maximum available frequency that our data contained was limited by the bandwidth of the used oscilloscope, which was 4GHz. Time representation of both signals is shown in Fig.5.7. The red arrow denotes when the Al plate accelerated by the detonation destroyed PDV fiber probe (level of signal rapidly falls). After this moment, the optic fiber is compromised, and the measured signal contains basically unusable noise.



(a) PDV schematic (b) Test setup schematic

Fig. 5.6: Scheme for data acquisition (Source:[43]).

5.3.1 Application

To make the process of description of the trend of significant spectral coefficient more manageable we divided the algorithm into several steps. After the data acquisition the procedure was done in following steps:

- Preliminary analysis of data, selection of region of interest using STFT
- Application of STFT, TFAR and wavelet transform

The first step was a selection of the part of the signal containing the required information. The range of considered signals was defined as follows. The beginning

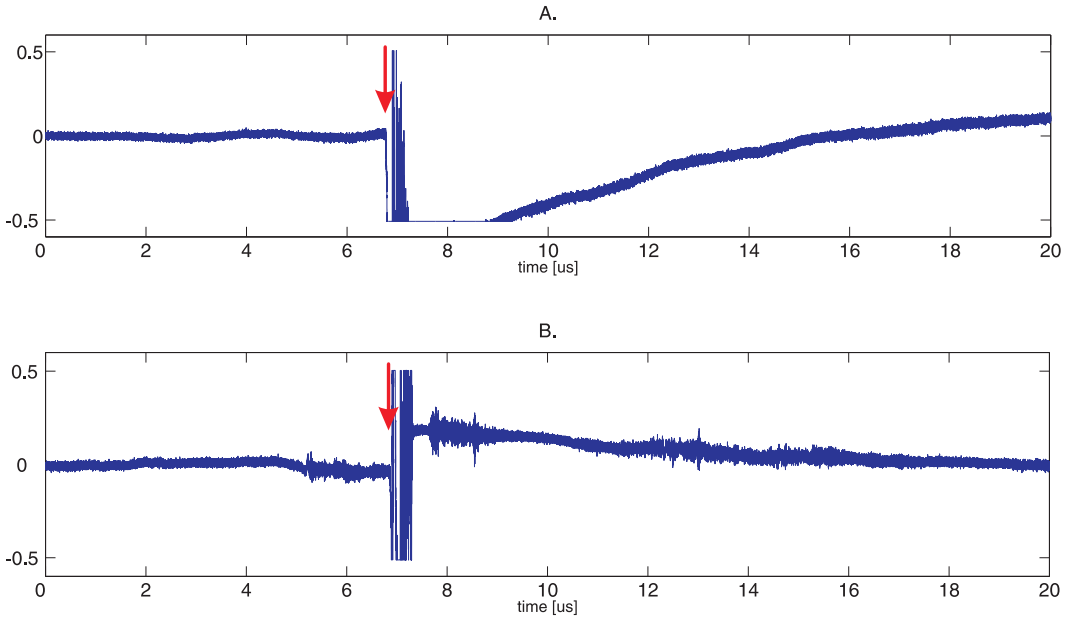


Fig. 5.7: Time representation of signal A and B.

of this time range was taken as the start of the shock wave in the Al plate. The end was done by the destruction of optic fiber (rapid fall of signal). The exact time of destruction is easily visible even in simple time representation. Unfortunately, the exact location of the start of the shock wave is difficult to identify/determine based only on the time representation precisely. However, we can perform a preliminary TFA of the signal. Therefore based on this TFA, we can select the appropriate sample range. For this aim, the STFT method is sufficient (low computational requirements, easy implementation), Hann window with length of 5000 samples with an overlap of 2000 samples was used.

We can see the demonstration of preliminary STFT analysis results in Fig. 5.8 where a circle highlights the area of interest. The results look quite similar for both signals (A and B); therefore, we present only for signal A. Based on the detailed analysis, we select signal A samples in the range of 5.00 to 6.80 μs, which corresponds to 45k samples, for signal B in the range of 5.00 to 6.88 μs which corresponds to 47k samples. To have access to a low-frequency slope, we avoid any filtering of data.

5.3.2 Summary

We applied STFT, but this application was made on the preselected range described above. The establishment of parameters was motivated by the intention to keep a

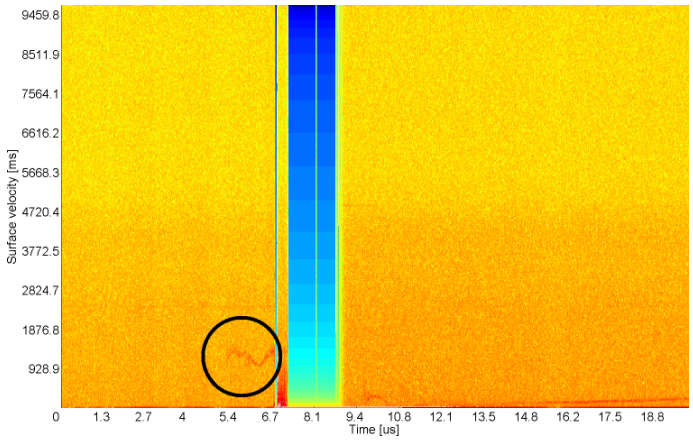
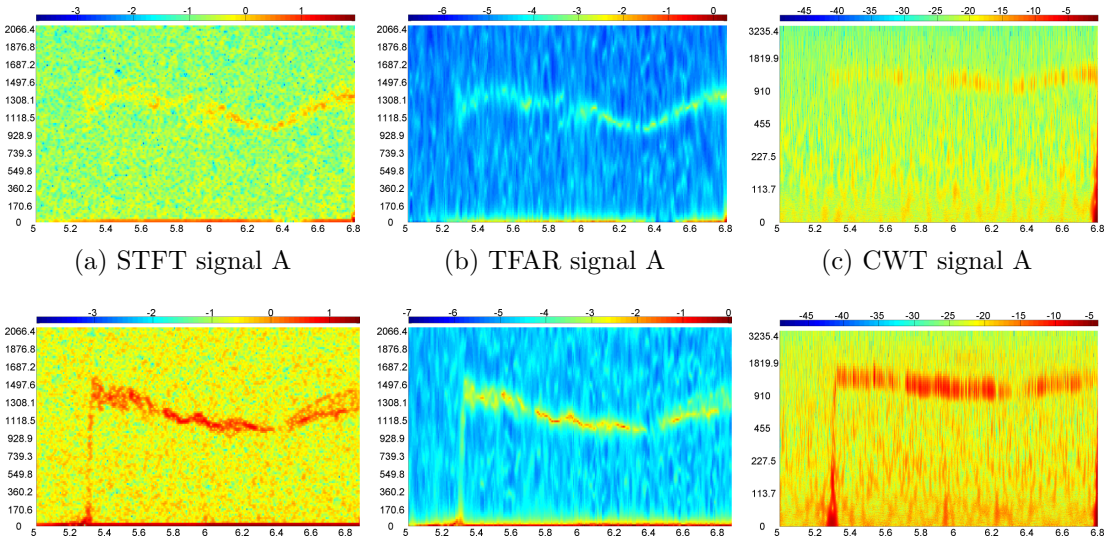


Fig. 5.8: Preliminary spectral estimation of signal A.

balance between time and frequency resolution. Therefore we used a window length of 1024 samples with an overlap of 900 samples. Due to wide overlap, the Hann filtering window was used to suppress the influence of earlier signal samples. Results are presented on the Fig. 5.9 a) and b).

In the next step, we obtained time-frequency representation using the AR process. Application of this method was also done on the preselected range described above. Motivation for parameter selection was the same as for STFT; therefore, we choose a window length of 1024 samples with an overlap of 900 samples and a Hann filtering window. Considering the sample size of signal and size of the window, we choose the Yule-Walker method for AR coefficients to estimate which is more suitable for long signals [47]. Selection of lag order was done separately for each window (i.e. 1024 samples). To determine optimal lag order, we used the Akaike information criterion, which provides good results in the case of a similar signal. Results are presented on the Fig. 5.9 c) and d).

The lastly applied method was wavelet analysis. The aim was to get more information via better time resolution. Therefore, we used a complex Morlet wavelet and scale 1:512. Results are presented on the Fig. 5.9 e) and f). As shown in Fig. 5.9 STFT and AR provide fair results. In the case of CWT, significant spectral components are less visible, making their identification slightly more difficult. Nevertheless, one advantage is a more detailed resolution of the time axis. This allows us to be more precise in determining the time of the beginning of the shock wave in the AL plate.



(a) STFT signal A (b) TFAR signal A (c) CWT signal A

(d) STFT signal B (e) TFAR signal B (f) CWT signal B

Fig. 5.9: Spectrogram estimation of both signals ($x-$axis: time, $y-$axis: frequencies, $z-$axis: spectral values).

5.4 Chapter Conclusion

Based on obtained results, we can see that in the case o economic data with a specific structure with several cyclical components (not necessarily based on sinus function), wavelets can perform better than STFT. This can be traced to the fact that the main advantages of wavelets are: the applicability on stationary and non-stationary time series; the flexibility of choice of mother wavelet with respect to the character of inputs; the ability to uncover unique, complicated patterns over time and a good time resolution. However, when using technical data (where the basis function is primarily a sine), the STFT provides sufficiently good results, and there is no need to choose a suitable mother wavelet and a suitable range of scales.

6 Enhanced TF Representation

To highlight important spectral components we propose combination of several TF methods. In each method background noise is depict with different characteristics. However significant spectral components should be captured in most cases. Based on such assumption we should be able to suppress the noise and highlight required components by using their combination. The procedure of this method is shown in Fig. 6.1. This procedure can be perceived as an alternative to significance tests, which we will discuss in Chapter 7. As mentioned in Chapter 5 the characteristics of input data must be taken into account. Therefore, we will show the application on different types of data; i.e. economics and engineering.

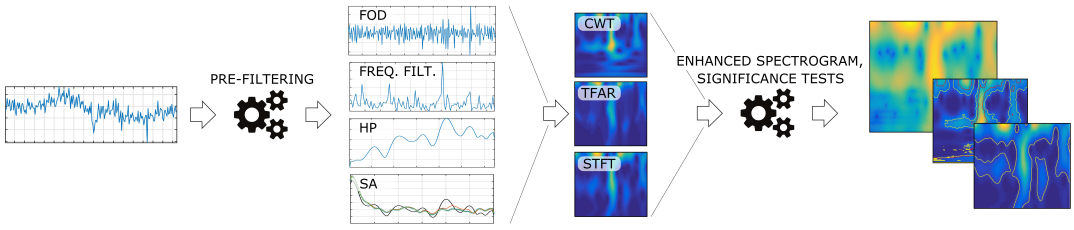


Fig. 6.1: Enhanced modelling of TF spectrograms.

6.1 Combination of TF methods

To obtain the best possible TF representation we combined results from the CWT, TFAR and STFT approach. Since the main focus was on the amplitude part of the spectra we have omitted phase part of complex spectra S_{CWT} and S_{STFT}. In case of focus on amplitude and phase components whole signal can be used for subsequent processing.

Firstly we align time axis (time resolution) of spectral representations S_{CWT}, S_{TFAR} and S_{STFT} so each spectrum would correspond to one another. All three vectors representing resolution in time have linearly increasing trend so for the time axis alignment the only requirement was to adjust starting and ending point for each method. We omitted first and last 15 columns of S_{CWT}, we denoted this remaining matrix as S'_{CWT}. By doing this we ensured corresponding time axis for all three methods.

Secondly we needed to align the frequency/scale axis of S'_{CWT}, S_{TFAR} and S_{STFT}. The frequency range of S_{TFAR} and S_{STFT} was cropped to correspond the range of S'_{CWT} which was 1 year to 10 years cycles. Resulting frequency/business cycles vectors $\overline{f_{\mathrm{TFAR}}}$ and $\overline{f_{\mathrm{STFT}}}$ had a linearly increasing trend however trend of $\overline{f_{\mathrm{CWT}}}$ was

non linear. To obtain corresponding vectors we matched each point of $\overline{f_{\mathrm{CWT}}}$ with one value of $\overline{f_{\mathrm{TFAR}}}/\overline{f_{\mathrm{STFT}}}$ with 1.4% tolerance:

$$\begin{aligned} |f_{\mathrm{CWT}} - f_{\mathrm{STFT}}| &\leq 0.014 \left|\max(\overline{f_{\mathrm{CWT}}}; \overline{f_{\mathrm{STFT}}})\right|, \\ |f_{\mathrm{CWT}} - f_{\mathrm{TFAR}}| &\leq 0.014 \left|\max(\overline{f_{\mathrm{CWT}}}; \overline{f_{\mathrm{TFAR}}})\right|. \end{aligned} \tag{6.1}$$

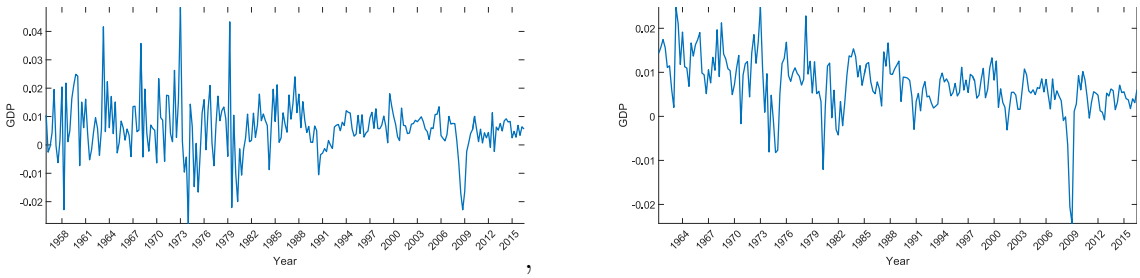
With this step we have gained adjusted TF matrices S'_{TFAR} and S'_{STFT} making all three methods aligned. For the methods combination we selected simple multiplication. We used combination of CWT and TFAR ($S_{\mathrm{CWT,TFAR}}$) and combination of CWT, TFAR and STFT ($S_{\mathrm{CWT,AR,STFT}}$):

$$\begin{aligned} S_{\mathrm{CWT,TFAR}} &= S'_{\mathrm{CWT}} S'_{\mathrm{TFAR}}, \\ S_{\mathrm{CWT,TFAR,STFT}} &= S'_{\mathrm{CWT}} S'_{\mathrm{TFAR}} S'_{\mathrm{STFT}}. \end{aligned} \tag{6.2}$$

6.2 Application on Economic Data

6.2.1 Data Describtion

As a representative of economic data with not clearly descriptive background noise, we use seasonally adjusted quarterly data of GDP. We selected volume index in OECD reference year 2010 [40] of the United Kingdom (UK) in 1956/01-2016/03 and Group of 7 (G7) in 1961/02-2016/03. All variables are in FODLOG (Fig. 6.2). G7 countries are: Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States.



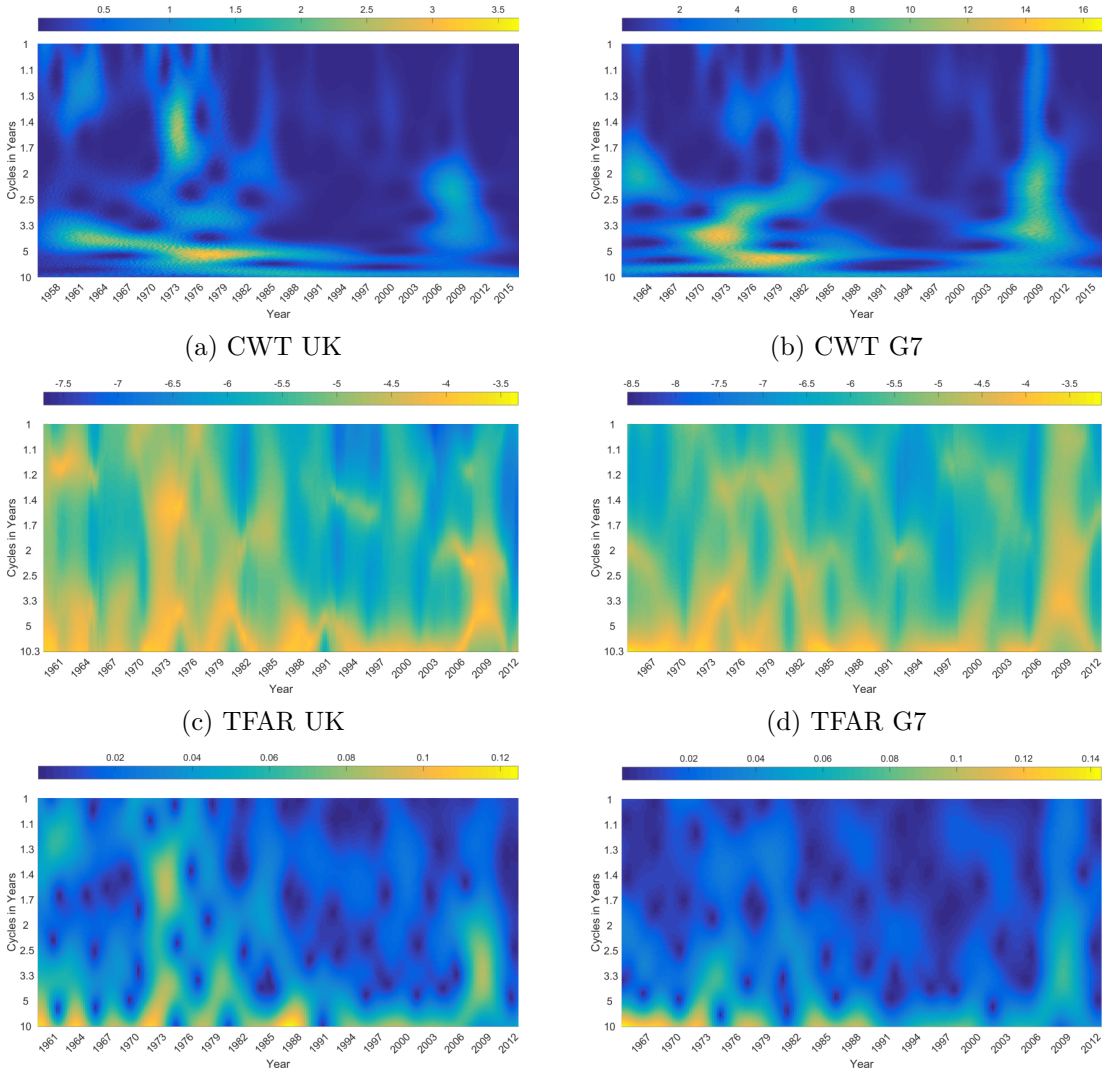
(a) UK (b) G7

Fig. 6.2: GDP of UK and G7 in time domain.

Our analyses consists from several steps. In the first, we analyse data using CWT. We set scales to correspond range of 1 year to 10 years, with 257 individual scales. As mother wavelet we selected complex Morlet with center frequency $f_{\mathrm{b}} = 1.5$. The complex Morlet wavelet is based on standard Morlet with the advantage

of providing complex results making it possible to obtain phase part (quadrature) of spectrum. In case of TF estimation via TFAR process we used Burg approach for coefficient estimates on 30 samples with 29 samples overlay and Hann window. Optimal value of lag order was based on AIC criteria. Parameters of STFT were set to correspond TFAR settings (30 samples, 29 samples overlay, Hann window) to simplify the process of combination of methods.

The data and results for UK and G7 are presented graphically in Fig. 6.2a-b, in Fig. 6.3a-f and Fig. 6.4a-d. There are four types of figures. Namely time representation of GDP for UK and G7 (Fig. 6.2a-b), TF transform via CWT (Fig. 6.3a-b), TF transform via AR (Fig. 6.3c-d), transformation via STFT (Fig. 6.3e-f) and adjustment of CWT picture with the help of AR (Fig. 6.4a-b) and with the help of TFAR+STFT (Fig. 6.4c-d).

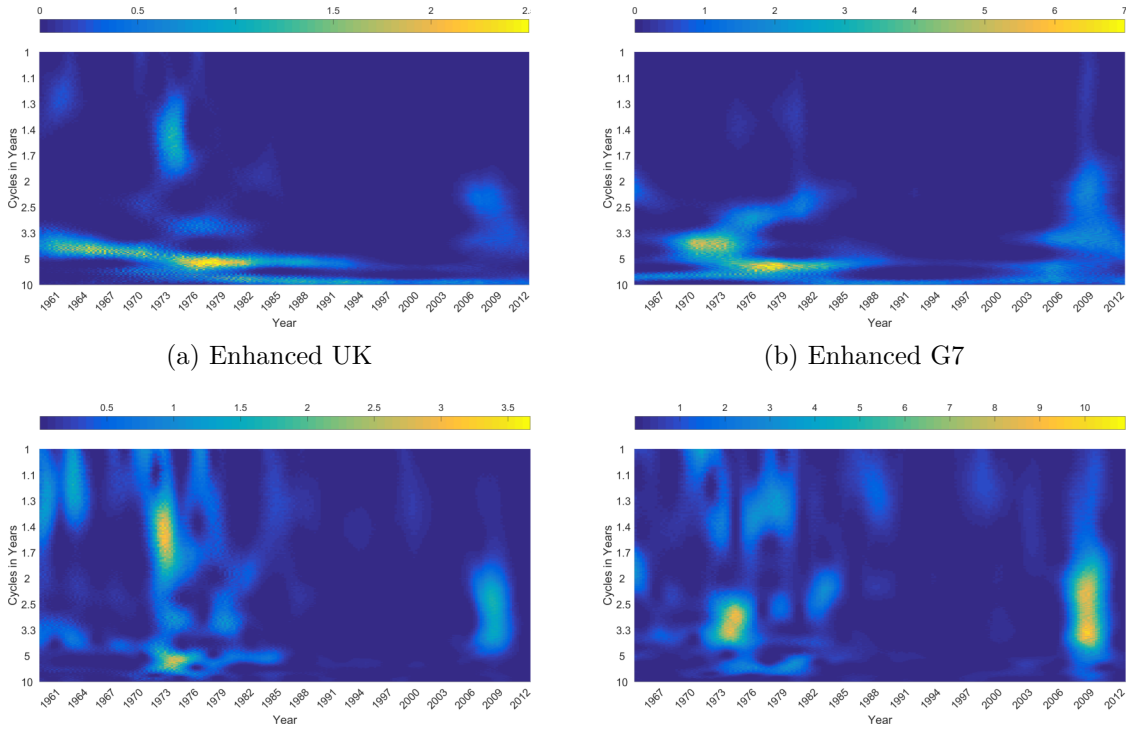


(a) CWT UK (b) CWT G7

(c) TFAR UK (d) TFAR G7

(e) STFT UK (f) STFT G7

Fig. 6.3: Spectrum of GDP of UK and G7.



(a) Enhanced UK (b) Enhanced G7

(c) Enhanced UK (d) Enhanced G7

Fig. 6.4: Adjustment of TF methods.

6.2.2 Results

Focusing on time representation given in Fig. 6.2a-b we can conclude following. In the time representation of the United Kingdom data we can see two sub-periods with different volatility. Between 1956-1988 and 1989-2015. There are also visible several moments with higher/lower level of the data, i.e. structural breaks (1958, 1964, 1968, 1973, 1979 and 2008) given by events in UK economy such as oil crises, financial crisis. In case of G7 there is similar problem with volatility, but it is not such visible as in UK. In contrast with UK, the G7 data has slowing decreasing trend with higher volatility between 1961-1988. Thereafter (1989-2015) the data character looks similar to the UK case. In G7 we can see similar structural breaks (1973, 1979 and 2008).

After a short analyses of time representation of the data we apply TF approaches. Firstly we modelled CWT (Fig. 6.3a–b), consequently TFAR (Fig. 6.3c–d) and STFT (Fig. 6.3e–f). As we expected CWT provides results with very good time resolution. We can see several important areas across time and frequency. Focusing on TFAR representation the results are not so clear from time perspective as CWT, but they give us better information from frequency perspective similarly as STFT. Therefore, we decided to do adjustment of CWT picture with the help of TFAR and

TFAR+STFT according to the calculation (eq. 6.1,6.2) described in Combination of TF methods.

The results of adjustments can be seen in the Fig. 6.4a–d. Focusing on UK situation and in comparison with enhanced figure (Fig. 6.4a) with CWT figure (Fig. 6.3a) we see sharper picture with suppressed noise. Therefore, we can easily identify the most important events in UK data from time and frequency perspectives. We can found three most important even in UK. The first is between 1960–1988 and can be distinguished into two subperiods; 1960–1972 and 1973–1988. These results corresponds with time domain description. In addition to the time domain we find that such event have reaction in the approximately 4 years (in the first subperiod 1960–1972) and around 5 years (in the second subperiod 1972–1988) cycles. The second important area arise between years 1973–1976 and has the impact on economic reaction in short cycles about length proximately 1.5 year showing quick reaction. The last important area is between 2007-2010 covering business cycle frequencies (from 4 to 2 year frequencies) and compare to the previous one it seems that it has not such impact in UK as previous events. To be sure with such conclusion and for cross validation we add additional adjustment of three TF approaches leading into the Fig. 6.4c. The result confirm results from the adjustment of CWT and TFAR and that events between 1970–1976 have stronger impact on UK economy.

When we focus on the results for G7, we can find some similarities as well as dissimilarities. Again, we can see (Fig. 6.4h) that the most important area is between 1970–1974 in 4 years cycles and 1974-1981 in 5 years cycles consisting one period 1970–1982. The second important area is between 2007–2010 (financial crisis) which cover the range of business cycles, i.e. 1.5-5 years cycles. After adding the second adjustment for cross validation of obtained results presented in the Fig. 6.4d we see conformity in important area identification. Also in this case for sureness and cross validation we added adjustment of TF approaches leading into the Fig. 6.4d. The result confirm results from the adjustment of CWT and TFAR and that events between 1970-1976 have stronger impact on G7 economy than financial crisis in 2007–2009.

Comparing results from economy point of view, we can see, that in UK oil crisis has bigger impact than financial crisis, while from the perspectives of G7 countries impact of financial crisis was stronger. Obtained results can be used for consequent macro or micro-econometric analysis to search for dependencies or relations with other economic aspects. Or it can motivates researcher in next steps which could be decomposition analysis on specific component of corresponding frequency which can be used in analysis of bilateral causalities.

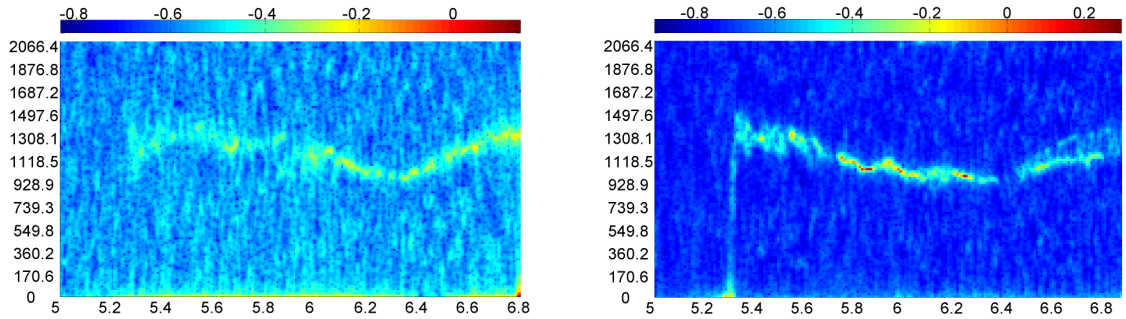
6.3 Application on Engineering Data

6.3.1 Data Describtion

Another selected data type was Photonic Doppler Velocimetry Data (PDV) data. Their parameters are based on physical nature and their structure is thus different from economic data. As input data we took STFT, TFAR and wavelet spectral representation of the data from Fig. 5.9. Equations (6.1) and (6.2) were then used. Because the main focus was on the amplitude part of the spectra we used only the amplitude part of STFT and wavelets.

6.3.2 Results

Resulting modified spectrogram is in Fig. 6.5. We can see that the scatter of background noise is smoothed and the data signal is more clearly visible. Even on this type of input data, it was confirmed that the method can highlight the required components in the spectrogram and therefore provides required advantages for further processing.



(a) signal A

(b) signal B

Fig. 6.5: Enhanced PDV data ($x-$axis: time, $y-$axis: frequencies).

6.3.3 Post-processing

The next step was focused on kernel estimate of the trend of significant spectra peaks. Identification of such trend allows better determination of materials properties at very rapid load than simple maxims connection. We firstly created vectors containing the position of detected peaks. With respect to the frequency value we took the highest spectrogram value $\pm\Delta$ to specify close surrounding area. According to empirical results we established $\pm\Delta = 7.5\%$. Using this thresholding we specified a mask which contained the region of our interest (region of required signal). By implementing this step we gained a vector of peak positions suitable for kernel analysis[32].

To support decision beneficial to kernel estimate we investigated two parametric and one non-parametric model. Namely the deterministic polynomial model (Fig. 6.6c,d), stochastic AR model (Fig. 6.6a,b) and non-parametric NW kernel estimate (Fig. 6.6e,f).

Polynomial model for the dependent value y and an qth degree is defined as:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \cdots + \beta_q x^q \tag{6.3}$$

where $\beta_0, \beta_1, \ldots$ are unknown polynomial coefficients and x is an independent variable. We investigated degree of polynomial $q = 2-30$ and Vandermonde matrix was used to obtain polynomial coefficients [48]. AR model was calculated using (3.11). We investigated lag order $p = 1 - 20$; optimal orders were selected using AIC criterion. NW kernel estimate was done using (3.24)–(3.26) for a kernel of the order $\nu = 0; k = 2$, smoothness $\mu = 2$ and bandwidth $h = 0.02$ (see Chapter 3.3) [32, 49]. Values were optimized via the generalized cross-validation method [50].

In Table 6.1 you can see the evaluation of model fit via MSE (4.1) and coefficient of determination R^2 for parameters with best fit. R^2 was is computed as: [51]

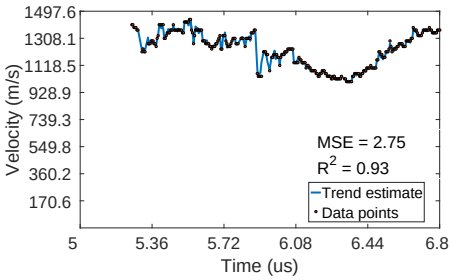
$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \widehat{Y_i}\right)^2}{\sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)^2} \tag{6.4}$$

where Y_i is the value of observed spectral coefficient, $\widehat{Y_i}$ is its estimated counterpart and $\bar{Y}$ is the mean of the observed data. The higher the resulting number is the more accurate is the estimation.

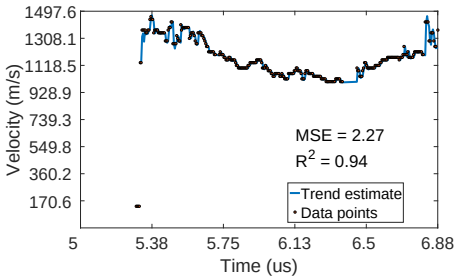
In both signals, Fig. 6.6, we can identify the same shape of the curve. Comparing them, we can see a rising edge of the curve at the start of signal B. This shape of curve path corresponds with expectations of material behavior during explosion. It is missing in signal A and is probably caused and influenced by the explosive event. Therefore, the start of signal A was identified after this part. In both signal cases we can identify a three-level decrease with osculation followed by its increase. The general tendency has a similar parabolic shape in both signals.

A detailed analysis of Fig. 6.6 reveals the following facts. Signal A: the graphical representation reveals a short time decrease followed by a short time increase and stagnation of the signal at a higher level. The second and third part copies this tendency, but both parts are moved to the lower level having a stair shape. The last part takes a parabolic shape with the second half of the signals. In the case of signal B we can identify three similar parts of decrease. Generally, the dynamics of the curve is smaller. The final part of the curve for signal B, taking a parabolic shape, is similar to the curve for signal A.

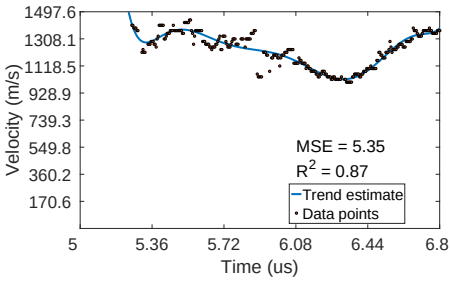
The comparison of estimated curve shapes (Fig. 6.6) and measured quality values (Tab. 6.1) reveal, that for signal A the best approach is the NW estimate while for



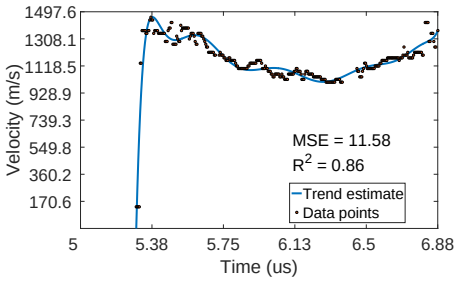
(a) Signal A: AR estimate $p = 6$



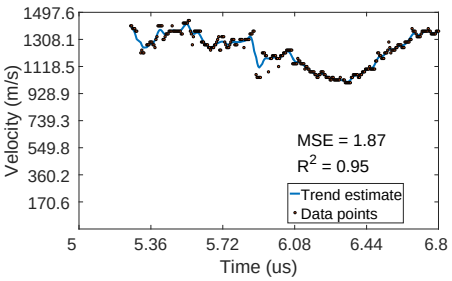
(b) Signal B: AR estimate $p = 3$



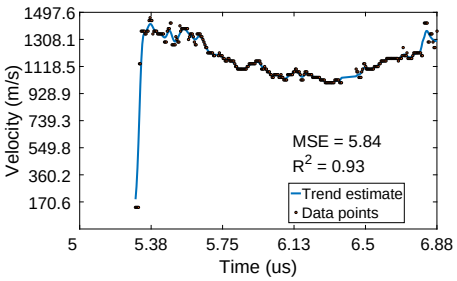
(c) Signal A: Polynomial est. $q = 8$



(d) Signal B: Polynomial est. $q = 11$



(e) Signal A: NW estimate



(f) Signal B: NW estimate

Fig. 6.6: Modelling of the curve fit.

Model	**Parameters**	**MSE**	R^2
	Signal A		
Polynomial	$q = 8$	5.35	0.87
AR(p)	$p = 6$	2.75	0.93
NW	$\nu = 0, k = 2, \mu = 2, h = 0.02$	1.87	0.95
	Signal B		
Polynomial	$q = 11$	11.58	0.86
AR(p)	$p = 3$	2.27	0.94
NW	$\nu = 0, k = 2, \mu = 2, h = 0.02$	5.84	0.93

Tab. 6.1: Evaluation of the model fit via MSE and R^2.

signal B it seems to be the AR process. However, AR appears problematic when focusing on capturing the rising edge because in such a case it was not captured. This is caused by a loss of data in the initial part of the signal - area of omission depends on lag order of the AR process corresponding to the signal character. Another problem with AR is in the case of missing points (non equidistant); the approximation of the resultant curve looks under-smoothed. For this reason, we prefer the NW estimate (Fig. 6.6e and f as the best even when it has a little bit worse measured quality values for signal B, because the rising edge was clearly captured. The worst curve fit estimation was achieved via the polynomial model (Fig. 6.6c and d). As the results show, in such an approach, the estimated fit is over-smoothed. Also, its ability to describe a rising edge is worse compared to the NW estimate. To sum up, the kernel estimate provides the best curve shape estimate for such data.

6.4 Chapter Conclusion

If we review results, by combining several TF approaches we were successful in background noise suppression. Consequently, events of interest became more visible and their identification in time, as well as in frequency was easier. An example of the possible use of this identification is the trend detection in a spectrogram. This approach can also be taken as a supplement to the significance testing with the investigation of background noise description, which will be described in Chapter 7.

7 Standard Significance Tests

Time-frequency transform can give reasonable results of both perspectives, time and frequency, in one moment. In some branches such as engineering, the physical nature of inputs is obvious and gives valuable information. We can assume the existence of several harmonic components corresponding to the specific frequency during all times of given input. Unfortunately, in other scientific disciplines, such as economy or sociology it may not be so simple. Applications of TF analyses have been so far limited by the fact that it was impossible to draw any implications on the statistical significance. Thus, significance testing of obtained results is welcomed. The original contribution in the spectrogram testing was provided by Torrence and Compo [23], followed by Ge [24,52]. Both provided a framework for testing individual spectrograms as well as testing of co-movement representation. We denote them both as standard testing approach (STA).

7.1 STA on Individual Spectrograms

The basic work discussing significance testing TF representation is by Torrence and Compo [23]. Their paper aimed to answer the question how one can distinguish statistically significant results from those due to pure randomness. Authors present comparison of wavelets to the windowed Fourier transform and propose tests for wavelet power spectra developed by deriving theoretical wavelet spectra for white and red noise processes. Ge [24], motivated by the Torrence and Compo (TC98) [23] work, derived the sampling distributions of the wavelet power and power spectrum of a (GWN) in a rigorous statistical framework. He proved that the results given by [23] are numerically accurate when adjusted by a factor of the sampling period. Similar approach to TC98 can be found in the work of Schulte et al. [53] or James and Fleming [22]. As an alternative approach we can use combination of several TF methods to suppress the noise and highlight required components of TF transform as presented in Chapter 6.

Thus, according to TC98 [23] and Ge [25] we can do following. The distribution for the Fourier power spectrum is

$$\frac{N|S_{\mathrm{STFT}}(m,f)|^2}{2\sigma^2} \sim \frac{1}{2} P_k \chi^2(2) \tag{7.1}$$

where N is the number of points, $S_{\mathrm{STFT}}(m,f)$ is given by (3.4), σ^2 is the variance of time series $s(n)$. Such testing statistic is distributed as chi-squared distribution χ^2 with two degree of freedom $\chi^2(2)$. The corresponding distribution for the local

wavelet power spectrum with usage Morlet wavelet is

$$\frac{|S_{\mathrm{CWT}}(a,\tau)|^2}{\sigma^2} \sim \frac{1}{2} P_k \chi^2(2) \tag{7.2}$$

at each time n and scales a, $S_{\mathrm{CWT}}(a,\tau)$ is given by (3.6) The 1/2 factor removes degree of freedom factor from the χ^2 distribution. If the wavelets are real, the distribution on the right-hand side would be chi-squared with one degree of freedom $\chi^2(1)$. The value of P_k is the mean spectrum at the Fourier frequency k that corresponds to the wavelet scales a. In case of gaussian white noise background spectrum this is equal to one [23, 25].

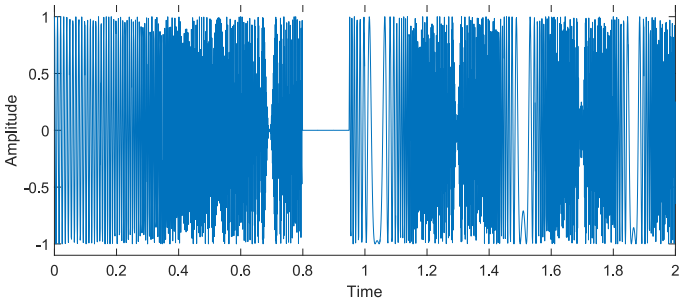
After finding the appropriate background spectrum and choosing a particular confidence for χ^2 for 95% ($\alpha = 5\%$ risk) we can construct confidence contour line for each scale. Thus, if the peak in the wavelet power spectrum is significantly above the background spectrum (in our case GWN), than we can assume the peak is true feature with a certain percentage confidence.

7.1.1 STA on Individual Spectrograms Application

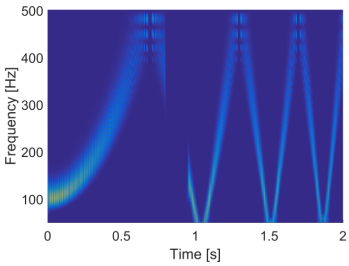
To determine the characteristics of STA and select optimal parameters of the chosen method we have created a deterministic signal (Fig. 7.1a). By doing this we obtained a signal with the known properties allowing us to asses the quality of method performance. The signal was constructed to include frequency sweep: exponential at the beginning, linear afterwards and containing a drop to zero at the end. In contrast to test signal used in [24], we aimed to create a signal of a higher complexity compared to the simple signal with just one harmonic component. Time representation of the signal can be seen in Fig. 7.1a and its spectral estimation via CWT and STFT in Fig. 7.1b,c.

Before applying the methods described in the methodology part we firstly corrupted the synthetic signal Fig. 7.1a. We added the GWN to the signal; the SNR was set to the moderate level of 3dB. The type of noise and its level was selected to correspond with Ge [24]. Consequently we estimated time- frequency representation via CWT and STFT. For estimating STFT, we used Hann window [54], for CWT estimation we used Morlet wavelet with central wavelet frequency $f_0 = 0.8125$ (this selection was motivated by wide use of this wavelet in economic applications). The resulting figures for noised signal can be seen in Fig. 7.2a,b.

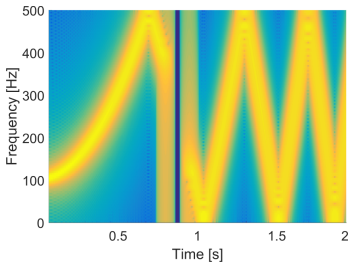
In the following step, we identify significant components of estimated spectra. To determine significant components we proceed according to (7.1) and (7.2). Due to the character of the noise, which was additive GWN, we were able to use the simplification of $P_k = 1$ (for details see [25]). The level of risk α was set to the two



(a) Time representation

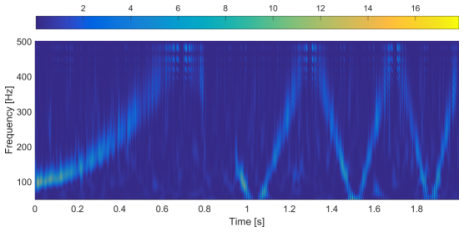


(b) Estimation using CWT

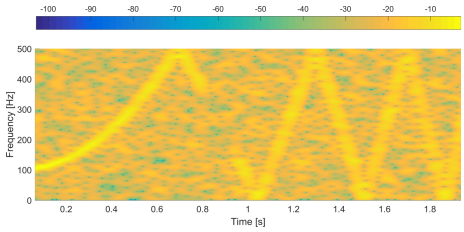


(c) Estimation using STFT

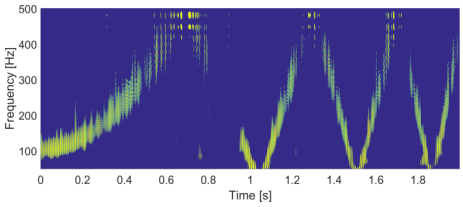
Fig. 7.1: Signal without GWN.



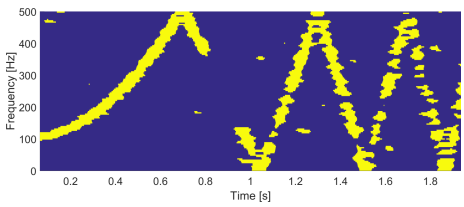
(a) CWT spectrum of noised signal



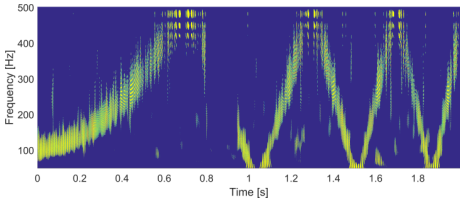
(b) STFT spectrum of noised signal



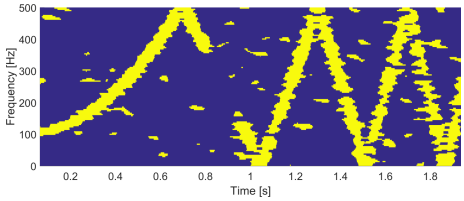
(c) Significant components of CWT for $\alpha = 0.05$



(d) Significant components of STFT for $\alpha = 0.05$

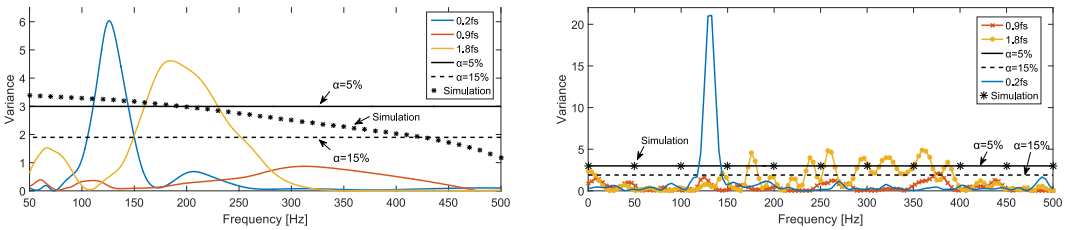


(e) Significant components of CWT for $\alpha = 0.15$



(f) Significant components of STFT for $\alpha = 0.15$

Fig. 7.2: Spectral estimation of noised signal and its significance testing.



(a) CWT power spectrum

(b) STFT power spectrum

Fig. 7.3: Normalized power spectrum in three time slices points and corresponding 5% and 15% signigicance level.

levels $\alpha = 5$ and 15%. As Ge [24] wrote, in such cases the significance level can empirically be relaxed to 15% or even higher depending on the particular problem.

The resulting figures for estimated spectra for both risk levels are given in Fig. 7.2c–f. The figures show only the identified significant component without the real value.

To demonstrate the method, we took three time instants $t = 0.2, 0.9$ and 1.8 s. Fig 7.3 shows the CWT and STFT spectra in these time points together with the corresponding significance levels of 5% and 15%. The frequencies of values that are placed above the line for the selected α are identified as significant.

7.1.2 Simulations of Background Noise Levels

To verify the accuracy and suitability of Torrence and Compo tests, we used simulations based on the background spectrum construction. These simulations are used to identify our own critical values on the basis of repeated random sampling to identify empirical distribution. We generated 50000 GWN repetitions with the length of 2000 points with variance level corresponding to the signal variance. For each sequence, the CWT and STFT spectrum was calculated. After the STFT calculation, we took the values for each frequency from all iterations, gaining vectors of 50000 values. After that we separately calculated 95% quantile of empirical distribution for each vector (shown as asterisk in Fig7.3 b). We can see that the obtained values approximately match Torrence and Compo tests. Not to complicate things for CWT with cone of influence [25] we selected a time point in middle of CWT spectrogram ($t = 1$s). The reason is that the character of GWN does not affect the accuracy of the result. For the time point $t = 1$s , we took the values for all scales (frequencies). By doing this for all 50000 iterations, we gained vectors containing 50000 values for each scale (frequency). After that we separately calculated 95% quantile of empirical distribution for each vector (shown as asterisks in Fig7.3 a). We can see that this

significance level is not constant and that its level decreases with frequency, which in return enables for higher spectral components to be identified as significant.

Using a synthetic signal, we analyzed the behavior of testing procedure for the estimated spectrogram with respect to the GWN background. We employed Torrence and Compo significance tests based on χ^2 distribution and verified their accuracy and suitability for more complex signal using simulations of background noise levels. The results suggest that in the case of STFT it is sufficient to use Torrence and Compo tests even for a more complex signal. In the case of CWT simulations of background noise levels, the results differ from Torrence and Compo. Therefore, we recommend using the Torrence and Compo method for less complex signals.

7.2 STA on Co-movement

As mentioned in Chapter 3.2, co-movement measures are used to evaluate similarity between two inputs (signals or time series). In case of cross-spectrum analysis the output data is three-dimensional and resembles spectrogram. Therefore, it is advisable to use some form of testing as described above. The general formula for the significance testing of the power wavelet cross-spectrum (PWCS) was firstly proposed by TC98 [23]. They derived white- and red- noise wavelet power spectra and used them further to establish a null hypothesis for the significance test of a peak in the wavelet power spectrum and cross-spectrum. This work was consequently improved by Ge [25, 55] who also proposed formulas derived specifically for the GWN series.

As TC98 [23] and Ge [25, 55] wrote, for two independent GWN series $x(n)$ and $y(n)$ with variances σ_x^2, σ_y^2 and wavelet spectrogram $S_x(a,\tau), S_y(a,\tau)$, we can calculate the PWCS $|S_{xy}(a,\tau)|^2$, which is the product of two χ^2- distributed random variables. As pointed out by Wells et al. [56] and TC98, the distribution of the square absolute value of the normally distributed variable is $\chi^2(2)$. In the case when both power wavelet spectra are based on two independent GWNs, the χ^2-distribution are both with 2 degrees of freedom and the non-centrality parameters are zero [25,55,56]. That is, as proved Ge [25, 55] since

$$\frac{|S_x(a,\tau)|^2}{\sigma_x^2/2} \sim \chi^2(2), \quad \frac{|S_y(a,\tau)|^2}{\sigma_y^2/2} \sim \chi^2(2), \tag{7.3}$$

their product is

$$\frac{|S_x(a,\tau)|^2}{\sigma_x^2/2}\frac{|S_y(a,\tau)|^2}{\sigma_y^2/2} \sim W_2. \tag{7.4}$$

In (7.4), W_2 (based on Wells et al. [56]) denotes the probability distribution with

the probability density function

$$f(z) = 0.5K_0(z^{1/2}), \tag{7.5}$$

where $K_0(z^{1/2})$ is a modified Bessel function of order zero for the complex wavelets [56]. The significance level $Z(1-\alpha)$ for the risk α can be deduced from $1-\alpha$ percentile of the W_2 distribution [23, 25].

After rearranging the terms we obtain

$$\frac{|S_{xy}(a,\tau)|^2}{\sigma_x^2\sigma_y^2} \sim \frac{1}{4}W_2. \tag{7.6}$$

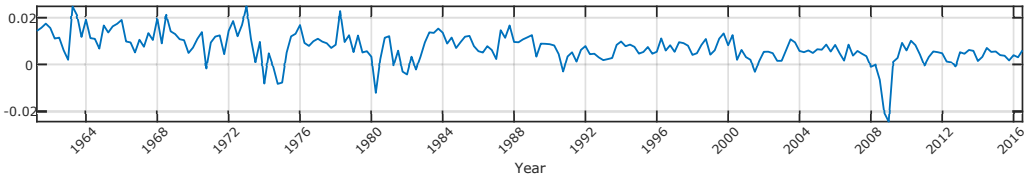
7.2.1 STA on Co-movement Application

To determine the characteristics of STA for co-movement we selected real economic data rather than their simulation, because such real data contain structural breaks and reactions on economic even. In case of its simulation some of unpredictable factor can be suppressed, which is not welcomed. Our approach consists from several steps. In the first step we estimate WCS and wavelet power cross spectrum (using (3.19)). In the second step we identify significant regions according to [23] approach (7.6). In the third step we used simulations of background noise levels to identify empirical critical values. The simulations were done according the Chapter 7.1.2 with 1000 iterations. At the end we compare both results (according to STA and simulations of background noise levels) and formulate recommendations.

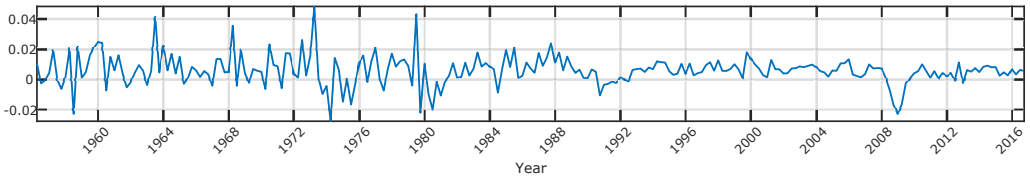
For the demonstration of discussed approach we use data described in Chapter 5.2 (Fig. 7.4a–c). The motivation for data selection were: i) sufficient data range (it is desirable to have detailed time resolution); ii) in case of UK and G7 we expect co-movement, because UK is a member of G7 and thus support validation of purposed method; iii) in case of Korea we expect lower level of co-movement with G7, because Korea is not the member of G7. Additionally, with respect to the Brexit we were interested into analysis before such even having character of structural break which will affect the data.

For CWT estimation we used complex Morlet wavelet with central wavelet frequency $f_0 = 1.5$ as the mother wavelet. We set scales corresponding to the range from 2 years to 40 years cycles, with 334 individual scales. The power WCS were calculated according to the formula (3.19). The resulting figures for the data can be seen in Fig. 7.5a-c. In case of simulations of background noise levels we firstly generate two independent GWN series with σ^2 corresponding to the background noise. Then, we estimate its CWTs and consequently power WCSs with 334 individual scales (corresponding to the data setting). Consequently, we generate 1.000

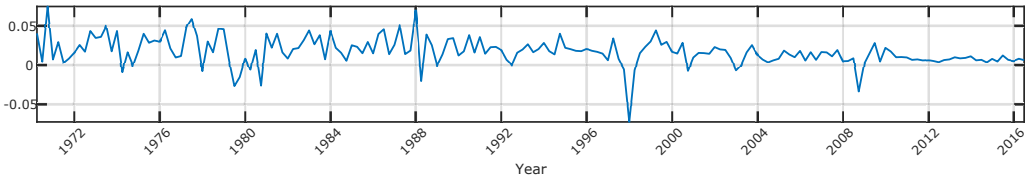
repetitions with the length of 1.000 points. To avoid complications with cone of influence in CWT we selected a time point in middle of WCS spectrogram. For each frequency (scale) from all iterations we save the obtained values. After that, we calculated 95% quantile of empirical distribution separately for each vector (scale). By this simulation we were able to obtain critical values for each corresponding scale.



(a) G7



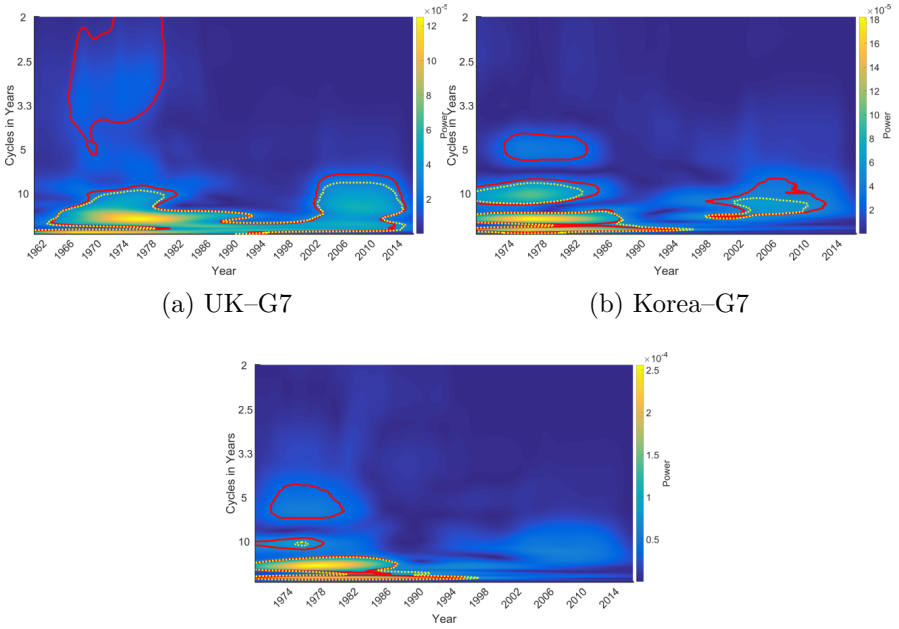
(b) UK



(c) Korea

Fig. 7.4: Time series representation.

The results of co-movements are presented in Fig. 7.5a–c below. The y-axis was re-calculated from frequency to the cycles per year for the better interpretation of an economic inputs. Thus, given the quarterly character of the data, we denote the sampling frequency f_s to be 4 samples per year. The red curve in the figures defines significant area identify via the simulations of background noise levels, the yellow dashed curve defines significant area according to [23]. Focusing on the results of co-movement (Fig. 7.5a–c) with respect to the identification of significant area via [23] we can state following. All figures show existence of co-movement between countries which matches our assumptions. That is, the UK-G7 (Fig. 7.5a) has higher level of co-movement caused by being UK in G7. The Korea and G7 (Fig. 7.5b) has less areas with co-movement, but generally both UK-G7 and Korea-G7 has co-movement in the time of oil crisis 1970-1982 and in the time of financial crisis 2001–2010. Such strong co-movement of Korea with G7 is also visible in Fig. 7.5c. All these significant co-movement (dashed curves in the Fig. 7.5a–c) were identified for long cycles, i.e. 10 years length cycles and longer. These results were confirmed by simulations of background noise levels. Focusing on same figures (Fig. 7.5a–c),



(a) UK–G7

(b) Korea–G7

(c) Korea–UK

Fig. 7.5: Co-movement between countries, x-axis: the time, y-axis: the frequency measured in the cycles per year and z-axis: the power WCS.

but on significant areas identified by the simulations of background noise levels, we can see an additional areas in all measured co-movements. That is, the additional co-movement between UK-G7 (Fig. 7.5a) in the time of oil crisis 1968-1982, but for shorter cycles, i.e. for the cycles of the approximately 2–4 years. Such additional co-moved area reflects the faster response in the country economy to the oil crisis, such as small and medium-sized enterprises, loans, investments etc. In the case of the co-movement between Korea-G7 (Fig. 7.5b) we can see also additional co-movement in the period 1974-1982 but for cycles of the length 4–6 years. Similarly, between Korea-UK (Fig. 7.5c) there is also the additional co-movement in the period 1972–1980 for cycles of the length 4–6 years. Here, the results again confirm our expectation of direct interconnection of UK and G7 and the weaker interconnection of Korea and G7. In all the cases results confirm the globalization of the economies and the stronger impact of the oil crisis on the world economy then the financial crisis.

The comparison and the verification of the accuracy and suitability of STA approach via the simulations of background noise levels is presented on Fig. 7.7. For the illustration of the significance level we firstly preselect the three time slices of the WCS (for example of UK-G7 (Fig. 7.6)). Selected time slices are denoted by the numbers 1–3 and corresponds to the time 1976, 1998 and 2006. The figure (Fig. 7.7) demonstrates power WCS curves and the significance level identified by

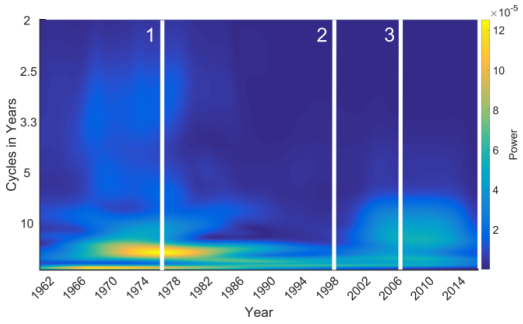


Fig. 7.6: Selection of slices in co-movement (1: 1976, 2: 1998, 3: 2006).

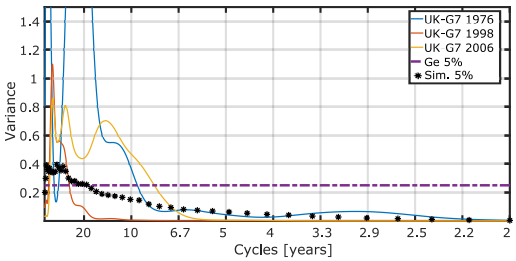


Fig. 7.7: Comparison of STA and simulations of background noise significance level.

STA (dashed-dotted line) and by the simulations of background noise levels (stars). We can see that the simulations of background noise significance level is not a constant, resp. it is decreasing with the frequency. Thus, the simulations of background noise levels enables the additional power WCS components to be significant. The difference between significant level of STA and the simulations of background noise levels is most likely caused by the heteroscedastic character of an input data, when in the time period of commoved series are the sub-periods with the different variances (you can see Fig. 7.5). An additional source of the difference between the significance level could be the scale range selection.

Using real data we analysed the behaviour of testing approach for the estimated wavelet power cross spectrum with respect to the GWN background. We use two approaches, namely the statistical significance according STA, and the identification of critical value using simulations of background noise levels. While STA uses comparison with constant critical value identified via the modified Bessel function of order zero corresponding to the risk, the simulations of background noise levels identify critical value on the basis of repeated random sampling to identify empirical distribution.

7.3 Investigation of Background Noise

In this sub chapter we focus on evaluation of background noise for determination of significance level for spectrum/spectrogram estimate. We investigate input signal which is viewed as composition of several cyclical components occurring in different time sub-period (not in whole time). Moreover, we admit that after removing estimated periodic components the rest of such signal does not have to be GWN. Therefore, we investigate identification of significant level on the basis of empirical distribution and noise analysis and its comparison with the results for GWN background.

We consider several types of input data. The first type is engineering data such as measurement of some noisy signal or deterministic technical data. The basic characteristics of such group of signals are composite of various harmonics corresponding to the different frequency which can be visible during all time of the signal. The second fact is, that after removing such periodic component the resultant noise is in expected form, for example white noise, red noise etc. Thus, simulation of such kind of signal can bring research in the case of changing the noise variance or changing frequency in signal during time. But the background spectrum will be still in pre-define form, in our case GWN.

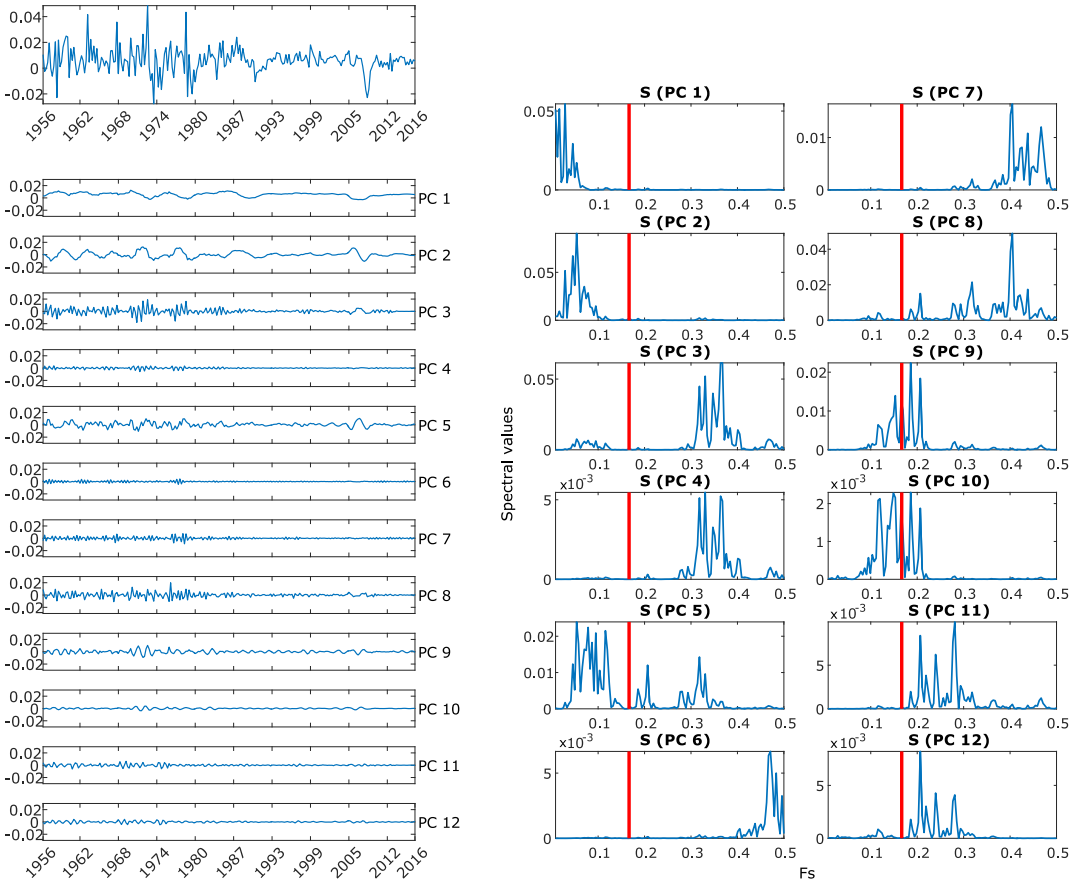
In practice, across different disciplines, there are signals or time series for which exact definition of its character is not so clear as in technical signals. As an example we can take economic time series, which have structural breaks. Such data can be viewed as a composition of several cyclical components which can occur in different time sub-period (not in whole time). The nature of an economic indicator play an important role and can influence the character of nested cycles. In such way the background noise is usually taken as a weakly stationary series and is obtained in dependence on analytical approaches (filtering, regression, decomposition etc.). Then, it can not be excluded that noise will absorb other components and therefore may no longer have the character of GWN. In such situation it is appropriate to verify the nature of the noise component, even if such data. And further, to compare whether identification of critical values using noise analysis is at the same level as in the case of GWN.

From the reason written above we use data described in Chapter 5.2. Given the quarterly character of the data, we can denote sampling frequency $f_s = 4$ samples per year.

We proceed in following steps. Firstly, we decompose input time series via singular vector decomposition (SVD) [35] into 12 component denoted as PC 1–12 and estimate corresponding frequency. Consequently, we divide components into two groups: i) components corresponding to the trend and cyclical component, ii) components corresponding to the noise.

Secondly, we consist noise by summation of components from second group (see [10]) and we estimate histogram. Thereafter, we fit it by predefine group of distributions, namely Gaussian, Student, Gamma and Rician distribution. We select the best fitting distribution.

Thirdly, we generated 10.000 replications of noise of fitted distribution to obtain 95% quantile and we establish confidence level.



(a) Time (b) Spectrum

Fig. 7.8: SVD decomposition (full data sample).

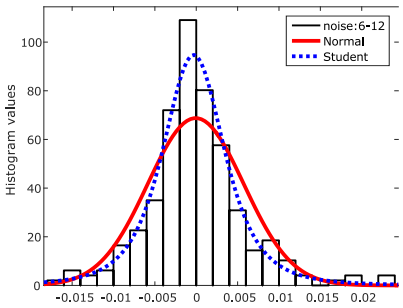
7.3.1 Full Data Sample

In the first step of analysis we performed SVD (details in Chapter 3.4). We decomposed full data sample into 12 principle components (PC 1–12). We used the FODLOG GDP data as input for eq. (3.28). We used the resultant trajectory matrix to obtain the orthogonal matrix, the diagonal matrix and the square orthonormal matrix. Using these as input for eq. (3.29) we got individual principle components. Results can be seen in Fig. 7.8. The choice of 12 components was motivated by [57] based on the cyclical components of the input signal whose recurring movements range from 6 to 32 quarters. To support the decision which components will be added as the noise component during reconstruction step, we used estimates of spectra. Thus, following knowledge from [57] we found, that PC 1–5 have the spectral peaks in the business cycle frequency while the others have not.

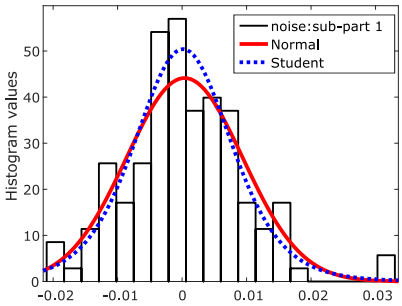
Spectra of each principle component can be seen in the Fig. 7.8, components with frequency greater than $0.17 f_s$ (vertical red lines in spectrograms) represents business cycles bellow 6 quarters and can be considered as a noise. With respect to

the heteroscedastic character of input data and estimated spectra for PC 9 and PC 10 we can not exclude that the noise will not be corrupted by some of the component. This fact, however, leads us to verify the nature of the noise component, i.e. to examine Gaussian distribution of the noise.

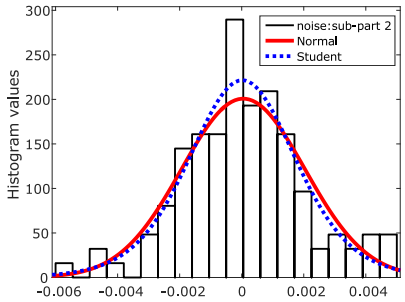
In the second step, for the noise component obtained using reconstruction of PC 6-12 (using eq. (3.30)) denoted as "noise:6-12" we estimates histogram. We also tests whether the"noise:6-12"has Gaussian distributed or not by goodness of fit test and by the Jarque-Berra test [51]. Both tests showed that the data are not Gaussian distributed. In such case we can not use GWN as the background spectrum for TF significance testing. Therefore, we are going to identify own critical values with the help of simulations of background noise levels. Thus, estimated histogram of"noise:6-12"was fitted by pre-define distributions, namely Student, Gamma and Rician distribution. As we can see in the Fig. 7.9a the best fit was achieved for Student distribution with 3 degree of freedom. The consensus with Student distribution was confirmed by goodness of fit and by the Jarque-Berra test.



(a) Full data sample



(b) Sub-part 1

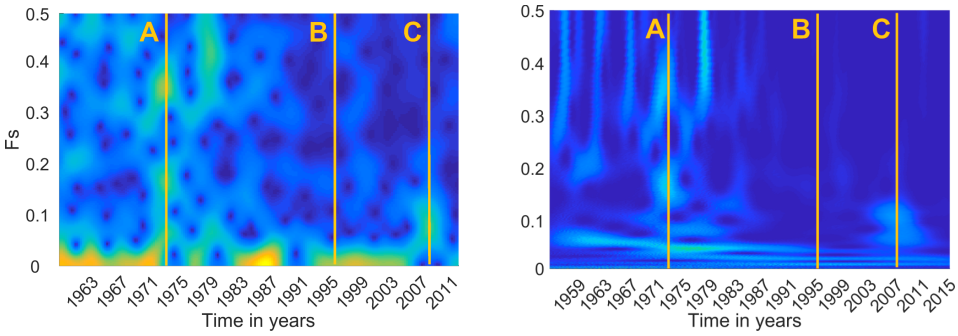


(c) Sub-part 2

Fig. 7.9: Histogram of the noise with fitted distributions.

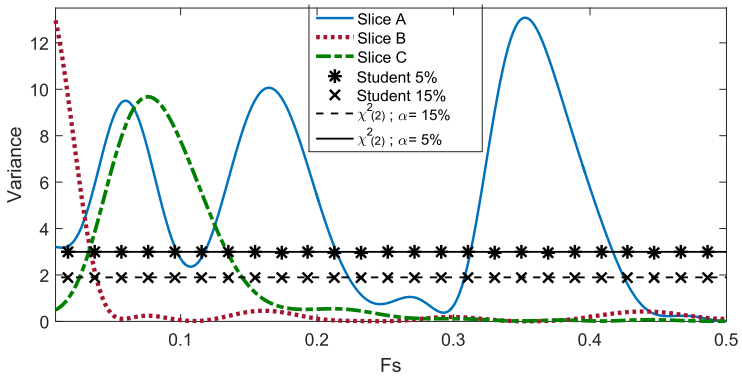
In the last step, we generated 10.000 replications of noise with Student's t-distribution with 3 degrees of freedom to identify own critical values and compare it with the $\chi^2(2)$ critical value [24]. This was based on the finding that the noise components have character of Student distribution with 3 degree of freedom. The result of simulations of background noise levels in comparison to $\chi^2(2)$ for STFT is

given in the Fig. 7.11a and for CWT in Fig. 7.11b. To give example of simulations of background noise levels we constructed spectral estimation of the time series using CWT and STFT approach. In case of STFT Hann window of 30 samples with 29 samples overlap was used. In case of CWT complex Morlet wavelet with central frequency of 1.5 was used. Resulting spectrograms are shown in Fig. 7.10. Three selected time slices are shown as lines A, B and C. Using equations (7.1) and (7.2) we recalculate all frequency values in each STFT and CWT slice (Fig. 7.11a,b).

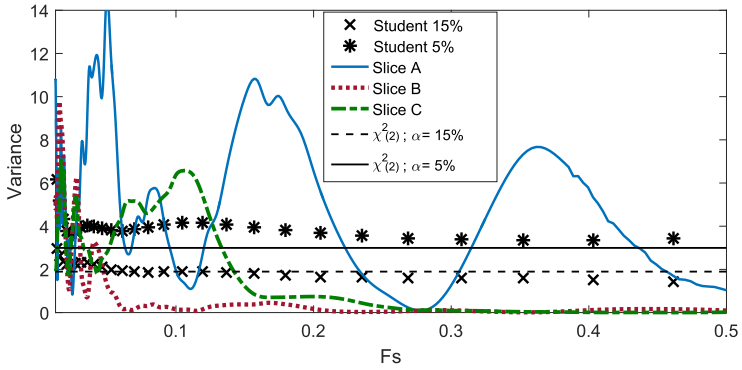


(a) STFT (b) CWT

Fig. 7.10: Spectrograms with slices identification.



(a) STFT



(b) CWT

Fig. 7.11: Noise simulations for ”noise:6-12” for full sample size.

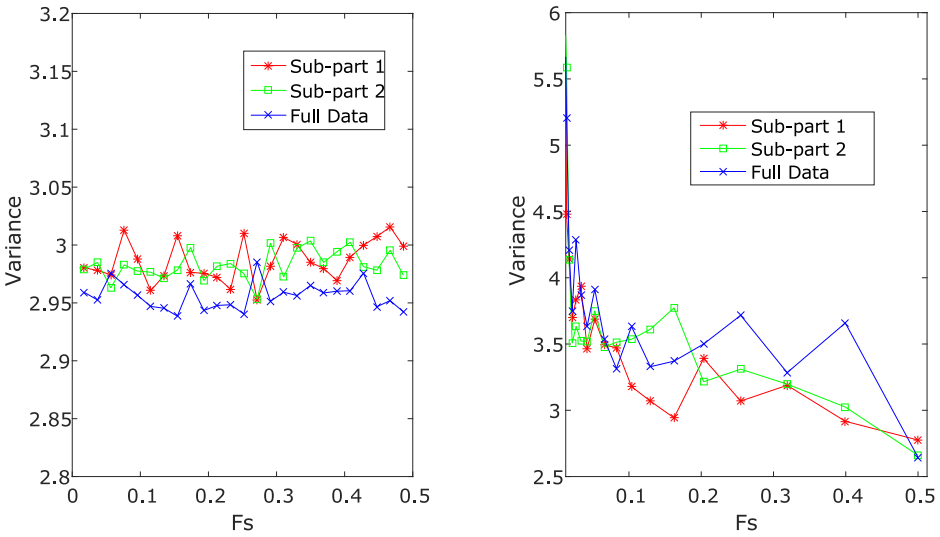
We can see, that significance level for 5% as well as 15% for STFT are moreover the same. For CWT the results of noise simulations are a little bit higher at low frequency range and slowly decrease to the $\chi^2(2)$ levels. In both figures significant components can be found above critical levels in dependence to the α. Such results (significant component) can be used for consequent analysis based on data character. In our case consequent economic analysis of significant components.

7.3.2 Sub-Periods

To be sure with our findings above and with respect to heteroscedastic character of the data (the variance is not the same during observing period) we made the same analysis on the data divided into two sub-parts. We can see two sub-parts; the first is in 1956/1-1988/2 denoted as sub-part 1 with variance $\sigma_1^2 = 1.46 \cdot 10^{-4}$. The second is 1988/3-2016/3 denoted as sub-part 2 with $\sigma_2^2 = 3.53 \cdot 10^{-5}$. The total variance of the full data sample is $\sigma^2 = 9.54 \cdot 10^{-5}$.

Similarly to analysis on full sample size we made SVD on both parts. Obtained noises (noise:sub-part 1, noise:sub-part 2) were tested whether they are Gaussian or not. While the goodness of fit test confirmed the Gaussian character for both noises, the Jarque-Berra indicated Gaussian distribution only the noise for the second sub-sample. Therefore, we fitted histograms (see Fig. 7.9b and c) similarly as the full data sample. Again, the Student distribution with 7 degree of freedom fitted the best. The consensus was confirmed by goodness of fit and by the Jarque-Berra test. Consequently, on the basis of noise simulations we identified own critical values and compared them with $\chi^2(2)$ critical value [24]. Out analysis confirmed that significance level for 5% as well as 15% for STFT are moreover same. The noise critical values for both sub-parts were a slightly higher than full data set (Fig. 7.12a). For CWT the noise simulation results follow slowly decreasing trend (Fig. 7.12b).

We investigated identification of significant level for empirical distribution using noise simulations and its comparison with the values for GWN background. Based on analyses above we suggest the following. Firstly, if the data are heteroscedastic, we recommend their division into sub-parts according to their variances. After identification of important components using SVD, we recommend testing of the reconstructed noise character. If we do not have information about the type of background noise or if we may expect that it is not GWN, we recommend fit noise distribution. Consequently we recommend identification of own critical values on noise simulation basis. By this way we can identify own critical value and perform relevant testing.



(a) STFT (b) CWT

Fig. 7.12: Noise simulations for noise:sub-parts 1 and 2 of data.

7.4 Chapter Conclusion

Based on obtained results we see that a problem can arise if the data contains segments with different volatility (i.e, heteroscedastic data). An experiment conducted on economic data (expressly GDP), which shows such a character due to structural shocks, confirmed that this can be a problem. Therefore, this fact should be taken into account during subsequent testing. As a result of disregarding this, the interpretation of the obtained results could be inaccurate or misleading (thus introducing errors in subsequent analyzes). In the next chapters, we offer a solution to this problem by proposing an adaptive method of testing.

8 Segmentation Based Testing

As stated in Chapter 1 in some data types, e.g. economic, heteroscedastic behavior occurs, i.e. that data variance is changing in time. We can expect that the STA can be sensitive to this fact, because of the assumption of constant variance over time. Therefore, based on STA, we propose the significance test of the TF co-movement on the segmentation bases. After the segmentation of the background noise according to the levels of variance in both input signals, we use the STA approach in each segment with the corresponding variance. As the results show, the segmentation allows us to identify significant co-movement with respect to the local variance, which can reveal additional significant co-movement areas. In the following text, we denote this method as segmentation adaptive based testing (SAB).

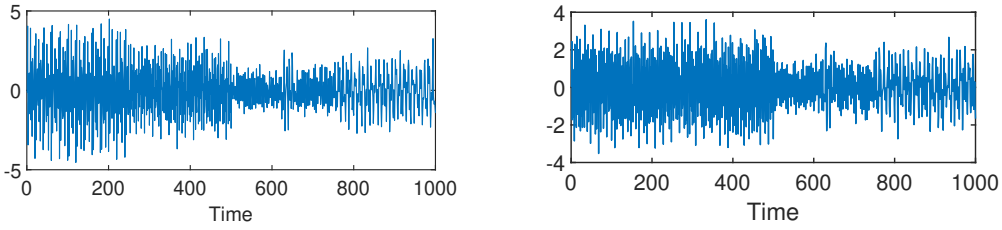
8.1 SAB Methodology

The idea of a segmentation adaptive based testing is based on the fact, that time series x or/and y may be heteroscedastic. That is, the variance of the time series is not fix, but changing during the time period. One of the cases may be a quick/step change of variance. Then, we can split the time range of the time series into the sub-periods according to its variance level. The case when the variance change is smooth is presented in the Chapter 9

Assume that both time series x, y are heteroscedastic. Then, we are able to identify the moments (variance breaks) after which the variance arise or decrease. It can be done by expert estimate or by statistical testing [51]. Comparing sub-periods of both series we can establish the segments (SG) of the time reflecting heteroscedasticity in x and y. Consequently, we can identify critical value for significance testing in each segment by STA [25, 55]. An algorithm for segmentation adaptive based testing follow these steps:

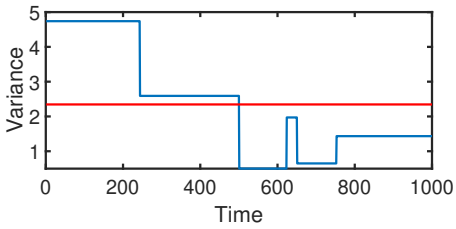
1. Identification of sub-periods in the time series x, y according to their variance levels
2. Splitting the whole time period into the segments for all levels of variance
3. Estimation of PWCS
4. Calculation of the significant level of PWCS with respect to the variance in the segment

Very important aspect of PWCS testing is an interpretation of the results which can be influenced by the methodology of testing. That is, the determination of the reference time period in relation to the results are evaluated. This can often be seen in the economics, where the events have the leads/lags effect. Thus, we have

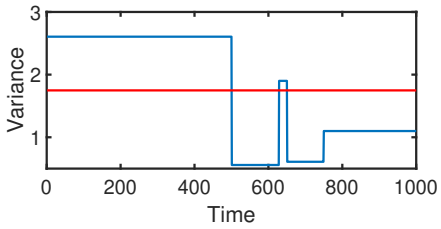


(a) Signal A

(b) Signal B



(c) Variance A: 6 sub-periods



(d) Variance B: 6 sub-periods

Fig. 8.1: Signal A and B and their variance. The red line denotes variance of full dataset.

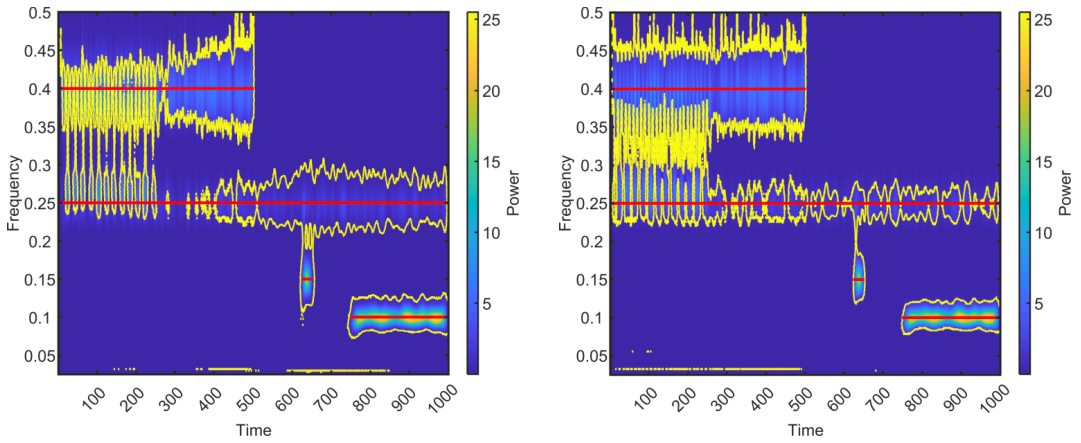
to state, if we are interested in evaluation of economic even with respect to the whole time period or with respect to its lead/lag influence.As already stated, this is a typical problem of economic data and therefore they will be used to demonstrate the SAB method.

8.2 SAB Application

8.2.1 Simulated Data

For the testing of the segmentation algorithm we created two artificial signals of the length of 1000 samples and sampling frequency $f_s = 1000$ Hz. The time domain representation of the signal consist of 4 and 5 sine waves respectively each of different frequency. Thus the total number of segments for both series is 6. Ideal spectral representation along with time representation of resulting signal is shown in Fig. 8.1. For better simulation of real conditions, GWN was added. The signal to noise ratio was set to 12 dB.

For CWT estimation of simulated signals we used complex Morlet wavelet with the central wavelet frequency $f_0 = 1.5$ as the mother wavelet. We set the scales corresponding to the range from $0.01 f_s$ to $0.5 f_s$, with 490 individual scales. The PWCS is calculated according to eq. (3.19).



(a) Standard testing

(b) Segmentation adaptive based testing

Fig. 8.2: PWCS of simulated signals. The red line denotes ideal PWCS.

Firstly, we test estimated PWCS (eq. (3.19)) via STA (eq. (7.6)) with fix variances for all data set, i.e. $\sigma_x^2 = 2.35, \sigma_y^2 = 1.73$ (Fig. 8.2a). With knowledge of the number of segments, we do SAB testing on six segments (SG-6), as shown in Fig. 8.1. For preliminary analysis of the case when the number of segments needs to be determined, Matlab function *findchangepts* was used. Changes in standard deviation with minimum residual error improvement of 10 was used as criteria for this function. Minimal length of segment was set to 15 samples. A more detailed analysis of selection of optimal segments number can be found in the following subchapter.

Consequently, in each segment, we provide STA with corresponding variance in segment (Fig. 8.2).

Comparing the test results for estimated PWCS via STA and SAB testing (Fig. 8.2) with ideal cross-spectrum (Fig. 8.2b) we can see that SAB approach propose more precise results. For the frequency $f_c = 0.4 f_s$ and $f_c = 0.25 f_s$ the SAB testing captured better the simulated significant co-movement compare to the STA testing. We also investigated case when the number of segments was 10. The results of PWCS show same significant areas as in case of 4 segments.

8.2.2 Real Data

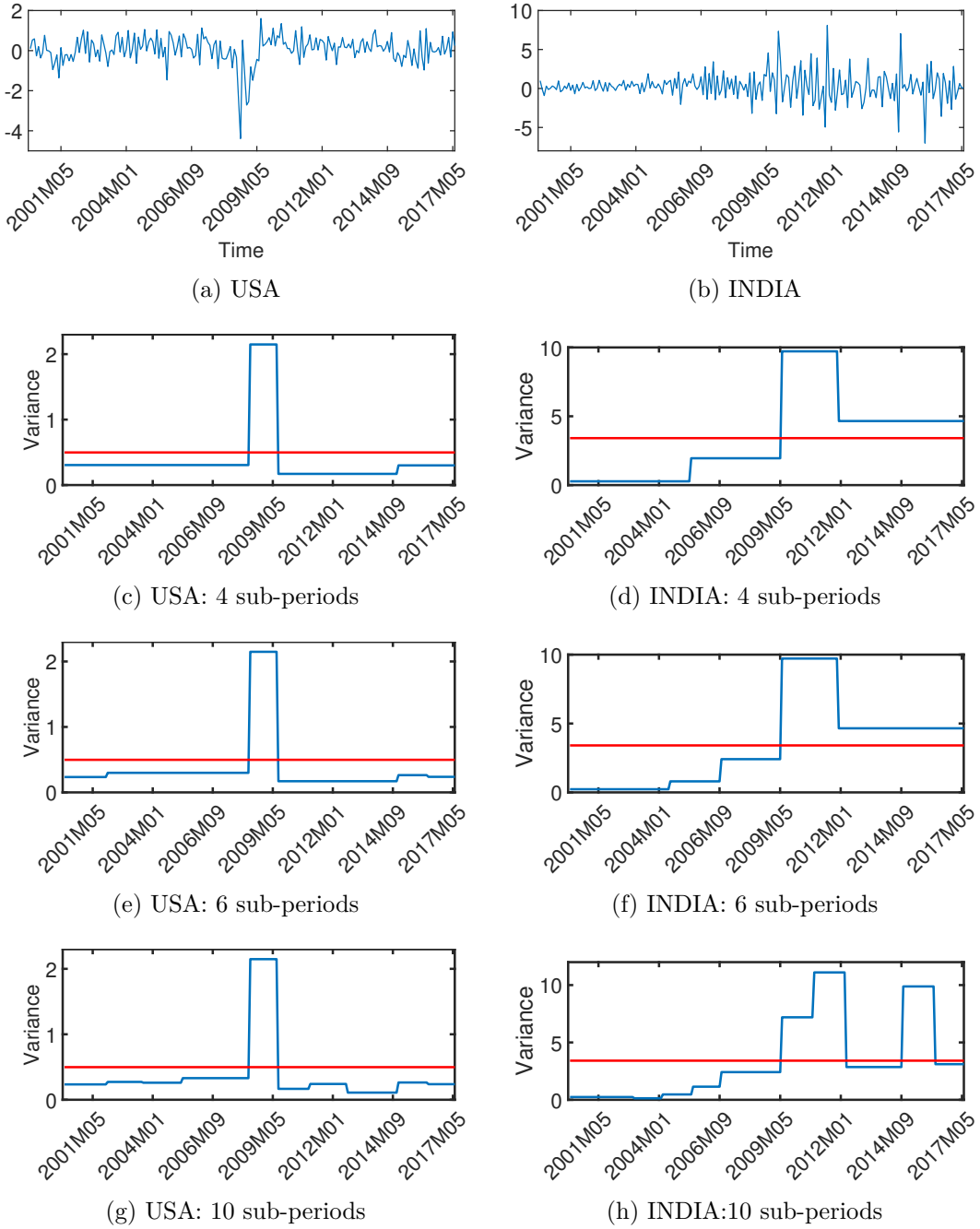
For the real application we use seasonally adjusted monthly data of industrial production index (IPI) of USA and India in 2001/M1-2017/M6 as the selected Asia Pacific countries. The IPI as a monthly indicator is widely used for assessing both the current state and the short-term outlook for GDP. One of the main reasons

why the IPI is considered to be a good proxy for GDP is that the value added by industrial production represented a substantial share of GDP and therefor display strong co-movements with GDP [58]. The motivation for IPI data selection are: i) sufficient sampling frequency (monthly) leading to sufficient sample size (with respect to the application); ii) heteroscedasticity of the data series; iii) expectation of co-movement between series (financial crisis in 2008). The GDP data used in the previous chapters do not always meet these conditions and therefore the IPI is used. All variables were transformed into FODLOG.

Focusing on the real data we firstly identify sub-periods according to the time series volatility. We assume that if the data are not heteroscedastic (i.e. are homoscedastic), then the STA and SAB testing will propose the same results. The number of segments was also checked via Matlab function, as for simulated data. We split both time series into 4, 6 and 10 segments (SG-4,6,10). The input time series and determination of sub-periods are presented in Fig. 8.3. We can see that data of India are heteroscedastic; the data of USA has problem with significantly higher volatility between time 2008–2010. When we compare identified sub-periods with the time of economic events in each country, we can take six segments (SG-6) as an optimum number of segments. In spite of this, we also make the calculations for four (SG-4) and ten (SG-10) segments.

For CWT estimation we used complex Morlet wavelet with central wavelet frequency $f_0 = 1.5$. We set scales corresponding to the range from 0.5 years to 5 years cycles, with 381 individual scales. The PWCS were calculated according to (3.19).

An estimation of the PWCS in Fig. 8.4a) was tested by STA testing and compared with results after SAB testing (Fig.8.4b–d). Comparing this results we can see several differences. Via SAB testing: i) we identify additional significant area in very short cycles 0.5-0.63 years in 2002-2003; ii) the co-movement in very short cycles 0.6-0.7 years has longer duration (2013–2015) then in STA testing (2014); iii) the co-movement in cycles of the length approximatelly 2 yeras is not significant in 2008–2009, because of the dominance of longer cycles (approx 3-5 years). Therefore, from interpretation point of view we can state following. The STA testing reveal, that the financial crisis in 2008–2009 in USA has strong impact on both economies. By measuring the co-movement of both series (USA and India) in the time period 2001-2017 this event was the most important influencing factor which predominantly influenced mutual behaviour of both countries in long and middle-term cycles. Focusing on reaction of the country and interpreting the co-movement with respect to this reaction, i.e. the IPI volatility of the country as the consequence of economic event, we can found also significant co-movement in very short cycles in two subperiods. This result is confirmed by proxy of PWCS segmentation adaptive based testing for the different number of segments.



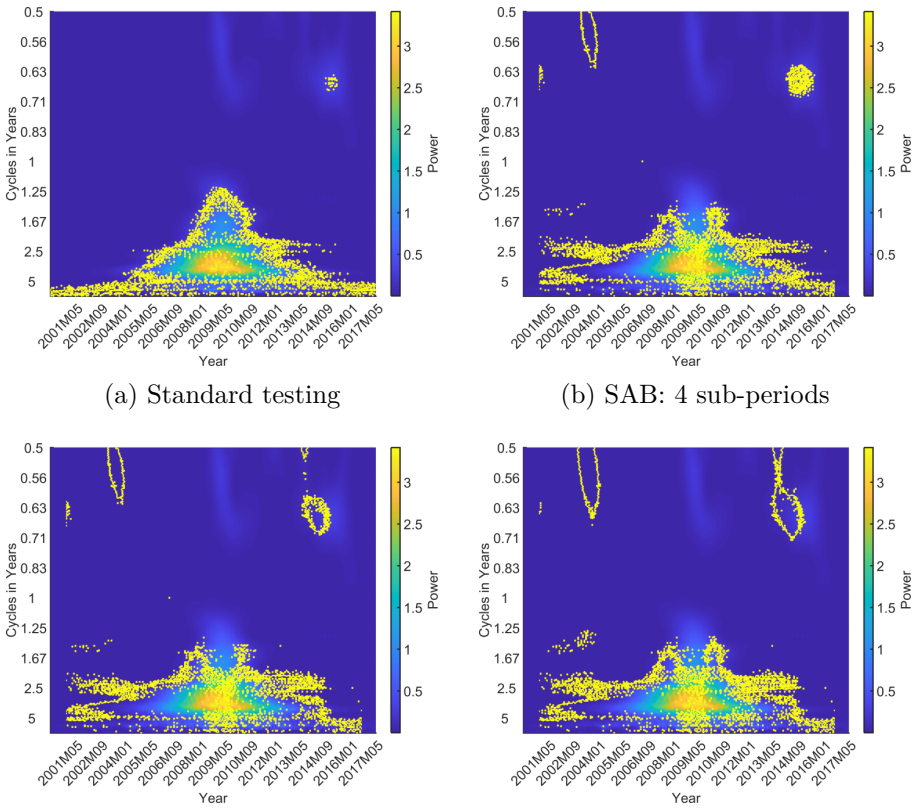
(a) USA (b) INDIA

(c) USA: 4 sub-periods (d) INDIA: 4 sub-periods

(e) USA: 6 sub-periods (f) INDIA: 6 sub-periods

(g) USA: 10 sub-periods (h) INDIA:10 sub-periods

Fig. 8.3: IPI signals and their variance. The red line denotes variance of full dataset.



(a) Standard testing

(b) SAB: 4 sub-periods

(c) SAB: 6 sub-periods

(d) SAB: 10 sub-periods

Fig. 8.4: PWCS of the IPI: STA testing and SAB. The red line denotes ideal PWCS.

8.3 SAB Segments Optimization

The previous chapter shows that determining the number of segments with respect to local variation is essential for the correct use of the SAB method. Therefore in this section, we present the theoretical background for the identification of an optimum number of segments for the SAB approach of significance testing of the co-movement. The number of segments is set with the help of the heteroscedasticity test and the test for comparing variances in the time series segments.

8.3.1 SAB Optimization Methodology

Let us suppose that both time series $x(t)$ and $y(t)$ are heteroscedastic in the time range $t = 1, \ldots, T$. Thus, for each series, we are able to identify moments when the data volatility increases or decreases. Consequently, we can determine the corresponding time segments via the identified moments. In some cases it is easy to identify the moments by an expert estimate. In the cases when the assessment of these structural moments in the graphical form of the data presentation is not clear, we can confirm the expected moments via the use of heteroscedasticity test [59].

Let us focus on each series separately. Assume that we applied test of the equality of variances on the time series $x(t)$ and we identified n moments of volatility changes at 1% significance level, or $n + k$ at 5% significance level, or $n + k + l$ at 10% significance level. Now, we split the time range into $n+k+l$ segments according to all identified time moments for 1% significance level. In each segment we calculate the variance and we compare the identified variances on the basis of standard F-test (test of equality of variances) in j and $j + 1$ segment, $j = 1, \ldots, n + k + l - 1$. Let the null hypothesis hold for the variances in the j and $j + 1$ segment

- $H_0 : \sigma_j^2 = \sigma_{j+1}^2$ (joining of segments), if $S_j^2/S_{j+1}^2 \in CI$

where S_j^2 is the sample variance in j-th segment, $CI = (F_{\alpha/2}(\nu_1-1, \nu_2-1), F_{1-\alpha/2}(\nu_1-1, \nu_2-1)$ is the confidence interval determined by the quantiles of F-distribution with ν_1 degree of freedom in segment j, ν_2 degree of freedom in segment $j+1$ and the risk α is in the pre-defined difference of approx. max. 1%. Then, the variances are taken as equal and we connect both segments into one. The minimum measurements in each segment are set to 40 values. As a result we obtain $SG_{\text{opt},x}$ number of segments, $n \leq SG_{\text{opt},x} \leq n + k + l$. After applying the same algorithm on the time series $y(t)$ we obtain $SG_{\text{opt},y}$ number of segments.

In the second step, for both time series, we sort the ascending identified moments by time and we focus on the comparison of identified volatility levels between both time series with respect to the identified sub-segments. The resultant SG_{opt} number of segments split the time range into the SG_{opt} sub-segments. If the time series are not heteroscedastic, then we use the whole time range for each series, i.e. $SG_{\text{opt}} = 1$.

8.3.2 Real Data

The data set for an empirical analysis consists of the seasonally adjusted monthly data of (IPI) from the OECD database for G8 countries (i.e. Canada, Franc, Germany, Italy, Japan, the United Kingdom (UK) and the United States (US)) in the range January 1993 – December 2017 [60]. We examine co-movements between the growth cycles of the US and all others G8 countries. The data have been transformed to FODLOG values (Fig. 8.5 right column).

8.3.3 Settings of the Methods

For the PWCS estimation, we used the complex Morlet wavelet with the central wavelet frequency $f_c = 1.5$ as the mother wavelet. We set scales corresponding to the range from 0.5 year to 10-year cycles, with 257 individual wavelet scales. Table 8.1 presents the selection of optimum number of segments ($SG_{\text{opt},x}$) for each country according to the approach described in the methodology and the total number of segments (SG_{opt}) for SAB testing for the particular pairs (the US and a country).

Segments for each country $SG_{\text{opt},x}$		**Total number of segments** SG_{opt}	
Canada	6	Canada&US	7
France	6	France&US	9
Germany	4	Germany&US	7
Italy	4	Italy&US	7
Japan	4	Japan&US	7
Russia	4	Russia&US	7
UK	6	UK&US	9
US	4		

Tab. 8.1: Optimal number of segments.

8.3.4 Results

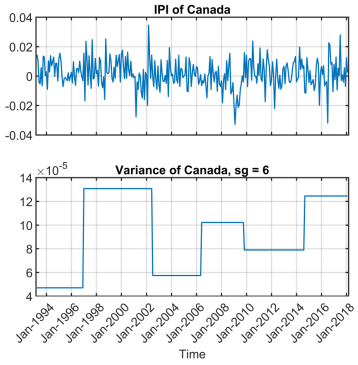
The empirical analysis consists of several steps. In the first step we apply the heteroscedasticity test on the time series. After the confirmation of heteroscedasticity we proceed with the identification of optimum number of segments and corresponding time moments for each time series. Then, we split the time range into segments according to the optimum number of segments for the pairs of the time series, i.e. for the US and a G8 country (see Sec. 8.3.3). In the second step, we estimate PWCS and apply SAB testing for the particular pairs. In the third step, as a supplement, we identify significant co-movement via STA approach.

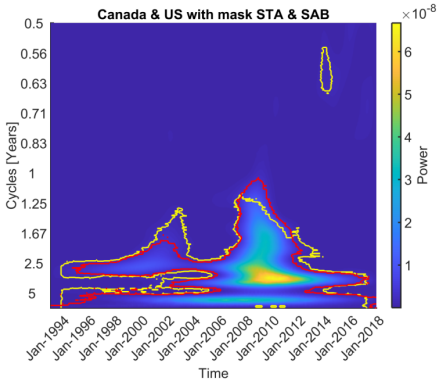
The results of co-movements are presented in Fig. 8.5 below. We use the following description in all the figures: x-axis represents the time, y-axis represents frequency measured in the cycles per year and z-axis represents the PWCS. The figures show a two-dimensional projection of three-dimensional charts. The intensity of each contour represents the relative importance of the different periodicities and time. The y-axis was re-calculated from frequency to the cycles per year to enable better interpretation of economic inputs. Thus, given the monthly character of the data, we denote the sampling frequency f_{s} to be 12 samples per year. The red curve in the figures defines the significant area identified via STA approach, the yellow dashed curve defines the significant area according to SAB testing. In the case of Canada & US (Fig. 8.5), Italy & US and Russia & US (Fig. 8.5) both testing approaches (STA and SAB) identify similar significant areas of co-movement. In the case of Canada and Italy we can say that the 2008 crisis was not the only important structural change during the analyzed time contrary to Germany or Japan. In the

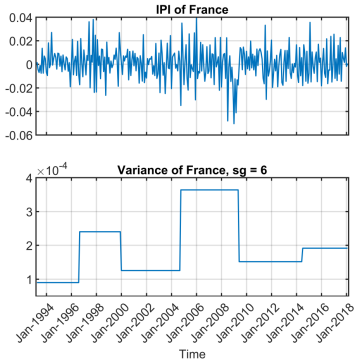
case of Russia we can see that the 2008 crisis was also an important structural break. Further, in the period 1994–2000 there was a different structural change causing a higher volatility level in this period leading to the extension (1997–2004) of the co-movement area with the US in the cycle of the length 1.5–5 years. In the case of France & US (Fig. 8.5) and UK & US (Fig. 8.5) the SAB approach reveals an additional significant area before the 2008 crisis in the cycles of the length of approx. 2.5 years. Specifically, in the case of France in 1994–2004 and in the case of UK in 1995–2001.

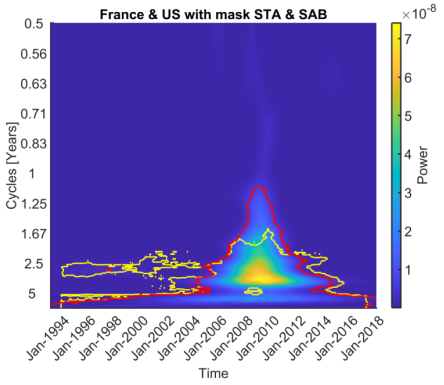
In the case of Germany & US (Fig. 8.5) and Japan & US (Fig. 8.5) we can see that the significant area, identified via SAB testing compared to the STA, does not cover the cycles of a shorter length during the 2008 crisis. That is, in the case of Germany and Japan, the significant co-movement with the US occurs around the 2008 crisis in the cycles of the length 1.5–5 years. In both these countries we can state that the 2008 crisis was the most important structural break during the analyzed time.

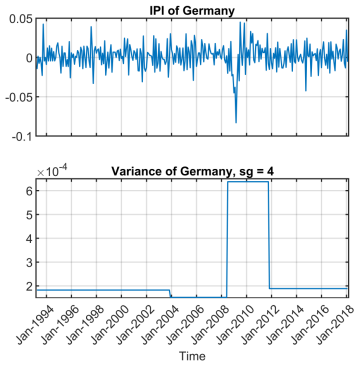
Using real data we assess the co-movements between the US and G8 countries with regard to the impact of the structural change of the financial crisis in 2008. The financial crisis caused a structural break in many economics, which led to the change in volatility level in the economic indicators. We found at least three levels of volatility during the given time range in all analyzed time series. Therefore, in the consequent co-movement analysis via PWCS we propose the SAB testing which takes into account the local volatility of the time series. We set the optimum number of segments for co-moved time series and identify significant co-movement areas. The results confirm the globalization of the economies and the impact of the 2008 financial crisis on the world economy. In some countries we reveal additional co-movement areas (France, the UK) or the extension of co-movement areas into a longer time period or into an additional cycle length (Japan, Russia). In other countries (Germany, Japan and Italy) we also find that the co-movement in the 2008 crisis was only in the business cycle frequencies, i.e. between the cycles of the length of 1.5–5 years. The most energy of the PWCS, i.e. the most important co-movement between US and G8 countries, was identified in the business cycle frequencies. The achieved results can be used for further analyses, e.g. to look for interconnections between countries co-movement and bilateral trade, in order to reveal additional information about the globalization and interconnections of economies.

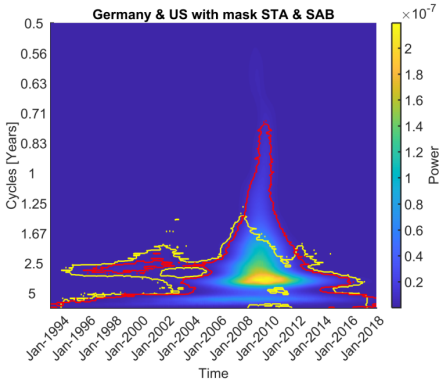

IPI of Canada
Variance of Canada, sg = 6
Time

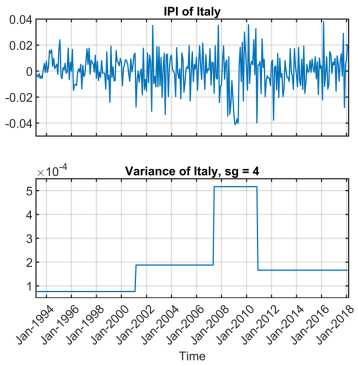

Canada & US with mask STA & SAB
Cycles [Years]
Power
Time

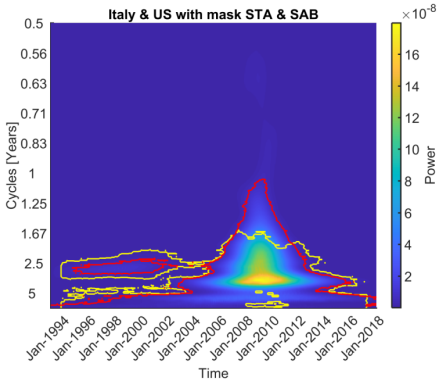

IPI of France
Variance of France, sg = 6
Time


France & US with mask STA & SAB
Cycles [Years]
Power
Time


IPI of Germany
Variance of Germany, sg = 4
Time


Germany & US with mask STA & SAB
Cycles [Years]
Power
Time


IPI of Italy
Variance of Italy, sg = 4
Time


Italy & US with mask STA & SAB
Cycles [Years]
Power
Time

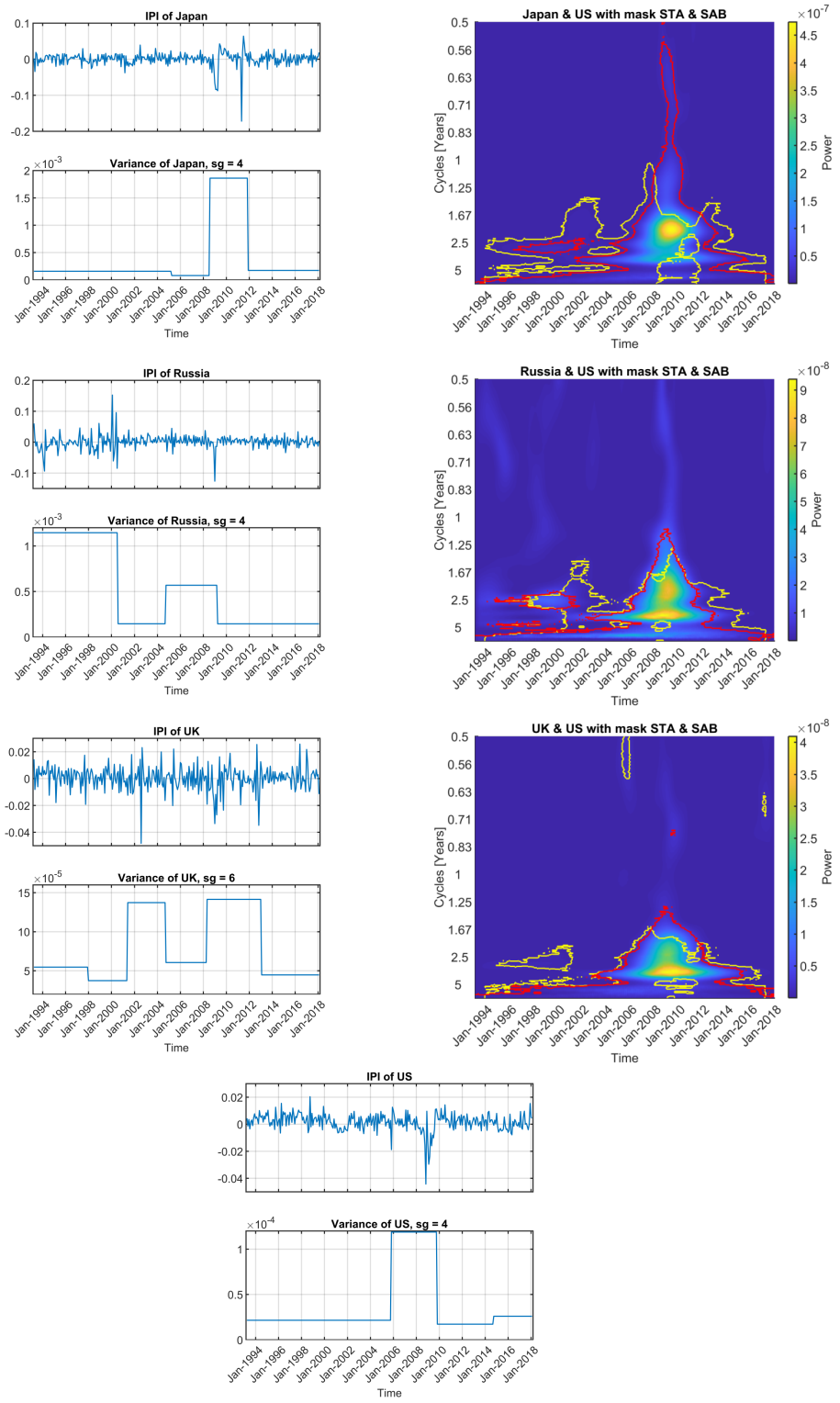


Fig. 8.5: Co-movement between countries, IPI of each country and variance of their segments.

8.4 Chapter Conclusion

This chapter was focused on the significance testing of the TF co-movement on the segmentation bases. We investigate the case when the input data are heteroscedastic. Firstly we propose segmentation of the data according to its levels of variance and we offer a method for optimal segmentation. Secondly, we propose testing of significance of power wavelet cross-spectrum with respect to the corresponding variance in each segment via STA test. The demonstration was done on simulated and economic data (i.e. IPI). The comparison of the results shows that segmentation can reveal additional significant co-movement areas in heteroscedastic data. Moreover, it confirms, that in the case of different volatility levels of inputs, the STA testing needs more careful interpretation and SAB can provide more complex results. Therefore, we recommend using the SAB method in case of in the case of significant or visible volatility in the data, which allowed segmentation accordingly.

9 A Local-Adaptive-Based Testing

In this chapter, we focus on the case where data volatility changes continuously (not quick/step change as in the previous chapter). For this purpose, real economic data that shows just this behavior were selected.

This chapter is based on following published journal paper:

POMĚNKOVÁ, J.; KLEJMOVÁ, E.; KUČEROVÁ, Z. "Cyclicality in Lending Activity of Euro Area in pre- and post- 2008 Crisis: A Local-Adaptive-Based Testing of Wavelets". **Baltic Journal of Economics**, 2019, vol. 19, no. 1, p. 155-175. DOI: 10.1080/1406099X.2019.1596466

Abstract The paper deals with the identification of time-frequency regions describing cyclicality of bank loans before, during and after the 2008 crisis via wavelets. We bring new methods and findings about the short and medium cycles of loans provided to corporates and households in the Euro Area in 2000–2017 using seasonally unadjusted monthly data. We have recognized an impact of the crisis on data volatility which further influences the type of significance testing of wavelet spectrograms. To avoid this influence we propose: (1) an adaptive spectrogram testing based on Torrence and Compo approach and (2) robustness analysis via enhanced spectrogram modelling tested by simulations of background noise levels. Both cross-checked approaches prove the sensitivity of standard wavelet tests on data volatility. The results confirm the usability of the new approaches and show that the crisis in 2008 influenced the cyclical behavior of both categories of economic sectors, but in a different way.

9.1 Introduction

The banking sector plays a special role in the monetary transmission mechanism and produces waves of cyclical behavior with a strong propagation in the real economic sector. In [61] the idea that economic booms improve the borrowers' balance sheets and net worth and support the lending activities of banks and thus spending, investment and as such the output of the growing economy is discussed. On the other hand, recessions bring the opposite transmission with a negative impact on the economy in distress. Discussions about fluctuations in lending activity and factors causing these fluctuations have been quite frequent since the financial crisis of 2007 and 2008 during which the world economy faced a drastic drop in the lending activity, particularly in the case of large loans (most of which are syndicated

loans). The drop was caused by a worsened access of banks to deposit financing, as documented in [62].

It is a well-known fact that banks tend to behave pro-cyclically and reinforce the credit and economic cycle, i.e. they lose their credit underwriting practices and massively provide loans in the period of economic growth and severely tie the practices and limit their lending activities just before and in crisis times. The extreme case when banks limit providing loans is called a credit crunch. Also, in times of globally integrated financial markets, a negative economic shock can be very quickly transmitted to other countries and may produce negative spill-over effects and a financial crisis in the world economy. Therefore, the analysis, identification and recognition of cyclical behavior and changes in lending activity in the economy are in the forefront of economic research and policy makers can thus adjust policy measures not only to economic cycles but also to credit cycles. This can reveal how quick and flexible short-, medium- and long-term reactions of banks and economic sectors to a shock in the economy can be. In our paper, we do not study factors influencing the behavior of banks (i.e. supply and demand factors). We focus on the development of methods which could give us an information about the time and the frequency behavior of the selected economic time series representing the volume of loans provided to basic economic sectors.

In this context, such fluctuations in lending activity can be generally defined as credit cycles, or more generally as financial cycles. However, financial cycles have not been satisfactorily defined and identified in empirical research so far. In [63] is stated that it is possible to describe the financial cycle using quantity (the volume of credit) and price (residential property prices, equity prices, risk premia etc.) variables. According to [64], the character of a financial cycle has changed since the early 1980s and was caused particularly by financial liberalization and less strict monetary policy after leaving Keynesian stop-go macroeconomic policy. While financial cycles have an impact on economic cycles ([65]), traditional macroeconomic policy is not able to address them.

As pointed out in [66], 'economic time series are an aggregation of components operating on different frequencies'. In [67] is proved that as for the financial cycle, 'there is not consensus in the literature on which variables to include in the analysis'. Further, as documented in [68] wavelets allow to 'distinguish the case that a series is the sum of several cycles at different frequencies from the case that the series is characterized by structural changes'. The TF modelling allows the investigation of the spectral character of time series with respect to time. In this way, we can analyze how various cyclical components, i.e. long, medium and short cycles, as well as seasonal component (a very short cycles), of a particular time series evolve through the time [20].

The early methodology of financial cycle analyses contains filtering and decomposition methods from the simplest to more advanced, turning point analysis, and the combination of both. Consequent approaches include frequency domain methods for an identification of hidden cyclical components. One group of researchers who used the early methods states that financial cycles are longer than business cycles ([64, 65, 69]). Talks about 5–20 years with a cross-country median around 15 years. Contrary to this group of researchers, another group proved that in a certain small group of countries (Germany, the Czech Republic, Hungary or the Netherlands), there are shorter financial cycles which are close to the business cycle frequencies ([67, 70, 71]). The third group of researchers make their findings more general ([67, 71, 72]) as they see the variation of the financial cycles' length at the country level which reflects heterogeneity between countries. In [72] is pointed out that financial cycles tend to differ from the business cycle counterparts, and that the identified length of financial cycles differs according to the definition of financial cycles which is given by the used methodology.

The last current group of scientists provide more complex results related to cyclical properties of time series by investigating the issues by early methods as well as by the use of wavelets. In [73] the financial cycle length is measured using quarterly data via wavelets in the case of developed and emerging economies. Their results show that in developed countries financial cycles are longer than business cycles. In [74] the power wavelet spectrum (PWCS) is used to estimate three types of time series, i.e. three types of quarterly data, and identifies several cyclical regions across all frequency range. Similarly, in [71] the PWCS is used on several type of quarterly data for European countries and also identify a wide range of time-frequency regions differing across countries and indicators. Their findings, using wavelets, confirm the statements of [66, 68], and the conclusion of [72].

Applications of TF analyses, where wavelets belong, have been so far limited by the fact that it was impossible to draw any implications on the statistical significance. The original contribution in the spectrogram testing (STA) was provided by Torrence and Compo [23], followed by Ge [24,52]. The STA assumes the fixed variance during all-time range of data. We use a modified form of this test considering that the variance in time series may vary for a certain sub-period, even for a short duration. Then strong events, such as the 2008 crisis, may cause a change in the data volatility. This may suppress the significance of other events which have a lower level of the data volatility and thus may suppress the importance of other cyclicality behavior in a specified time range. If there are no changes in the data variance, both (standard and modified) forms of the test produce the same results.

The paper deals with the identification of the time-frequency regions describing cyclical behavior of the bank loans with a special attention to the pre-, post- and

2008 crisis. The paper focuses on evaluating how the specific shock, i.e. the financial crisis in 2007 and 2008, could affects the cyclical behavior of given indicators. We found an important impact of the crisis on data volatility which may further influence the significance of wavelet spectrograms estimates. Therefore, keeping in mind this volatility influence, we propose: (i) testing of wavelet spectrogram via standard STA test and its robustness check via simulations of background noise levels; (ii) testing of wavelet spectrogram via modified form of STA called local-adaptive-based testing; and (iii) robustness analysis of wavelet testing via enhanced spectrogram modelling also tested by the simulations of background noise levels. We show that Continuous Wavelet Transform (CWT), i.e. wavelet spectrogram, compared to the Short-Term Fourier Transform (STFT), i.e. STFT spectrogram, is much more influenced by the data volatility during the standard STA testing. To demonstrate the newly proposed method, we use the monthly data of bank loans provided to corporates and households in the Euro Area in 2000–2017.

Presented paper investigates an application of proposed methods only on unadjusted monthly data. Therefore, the results obtained via presented methodology is interpreted from the cyclical point of view where the seasonal component (i.e. very short cycles) is taken as a part of cyclical behavior.

While many authors focus on the medium and long cycles of selected price- and volume-based measures of financial or credit cycles and use quarterly data, we bring new methods and findings about the short and medium cycles of loans provided to corporates and households in the Euro Area in 2000–2017. By using seasonally unadjusted monthly data, we were able to identify time-frequency regions for higher frequencies. Moreover, we distinguish between the sector of corporates and households as the lending activity is motivated by different factors. The call for new modelling strategies and adequate reactions of macroeconomic policies to the changes of financial cycles is described in [63]. In [75] the role of empirical research for the purpose of describing the features of financial cycles and designing macroprudential policies is emphasized. In this way, we propose a method of identifying mostly short and medium lending cycles of corporates and households, which is quite important for prompt reactions of policy makers and proper implementation of economic policies.

We aim to answer, by applying the cross-checked approaches (i–iii) described above, the following economic questions: Did the shock, represented by the financial crisis of 2007 and 2008, influence the cyclical behavior of lending activity in the Euro Area and if so in which lengths of the cycles (i.e. in which frequencies) were the reactions the strongest? Was the character of the cyclical behavior different before, during and after the crisis? Are there any differences in the character of cyclicality in the two analyzed sectors? Does our approach bring new possibilities for modelling

strategies of policy makers? The results confirm the usability of the newly proposed approaches in our case (the research was conducted only on seasonally unadjusted data), particularly in the short cycles and show that the crisis in 2008 had an important impact on the cyclical behavior of corporates and households, but in a different way. Moreover, the most remarkable influence of the crisis on the cyclicality was identified in the case of households.

9.2 Literature Review

Evolution of methodological approaches analysing credit and financial cycles and its characteristics reflects evolution of business cycle analysis. As many authors agree [71, 72, 74], the body of the literature for financial cycle analysis remains nascent. The methodological approaches started with i) time domain analysis of turning points identification [65] and was followed by ii) detrending via frequency-based filtering [64, 69]. Several authors use more sophisticated models such as unobserved component models [71] or structural models [67, 70]. The next step was iii) application of frequency domain methods which allows identification of spectral components, i.e. periodicities hidden in the data. Currently, iv) the time-frequency methods especially wavelets [68, 71, 74], which combine both time and frequency point of view and allow describe the cyclical behavior of data with respect to the time, are at the forefront of methodological approaches.

The approach of turning points identification, mostly based on Bry-Boschan or similar algorithm, describe the cyclical character of the data via calculation the distance between two peaks (local maxima) of an unobserved time series suffering from the fact that the data may consist of several cyclical components, i.e. hidden periodic component. This insufficiency was partially solved by applying detrending methods such as frequency-based filters (high-pass or band-pass filters) [64] which allow selection of pre-defined frequency range from the data, or multivariate model-based filters [67, 71]. Some authors [64, 67, 72] combine both filtering and dating approaches to improve achieved results and bring more robust conclusions. In [70] the alternative methods are used, i.e. combination of Bayesian model and singular decomposition followed by Fourier spectral analysis. An extension of methodology about the frequency domain methods allows identification of spectral components, i.e. periodicities hidden in the data. Both approaches (detrending/filtering and dating) fight the problems with moreover statistical character based on assumptions for dating methods application (sensitivity for trend extraction), or expected frequency range for the financial cycles filtering (no consensus for frequency range of financial cycles), or application of filtering methods themselves (ideal filter approximation, edge effect problem [64, 67, 70–72]. And even the frequency techniques highlight the

cyclical behavior of data, they unfortunately were not able to describe the temporal character of identified cyclical behavior.

An alternative approach is the time-frequency methods especially wavelets, which become current state method for the financial cycles analysis. The well-known TF methods include STFT [14], or TFAR [76]. While turning point approaches require pre-specified rules or mathematical apparatus to identify local extrema of time series, and even the frequency domain techniques have no prior assumptions for the financial cycles frequency range, their combination does not bring information of the time localization of the frequency which is easily proposed by the wavelets [74]. As written in [66], contrary to the time representation of time series, the wavelets map the original time series as a function of time and frequency revealing, how each periodic component of the time series changes over time. As we can study in [77–79], the wavelet analysis allows decomposition of even non-stationary economic time series into the different frequencies which after summation constitute the original series. Via this approach we can assess the relative importance of a different frequency component trough time and how such relationship changes over time, which make from wavelets a very useful tool for analyzing financial cycles. As many researchers agree, the main advantages of wavelets are applicability on non-stationary time series, flexible settings of parameters reflecting data character, ability to uncover unique complicated patterns over time and good time resolution [20, 36, 39, 66, 78, 80, 81]. Therefore, it is worth to investigate their use in financial cycle analysis.

9.3 Methodology

We use CWT, TFAR, STFT [10] for the TF modelling of input time series. Since these techniques are well known, we will not provide their description. For the significance testing of the TF transform, we use the standard test (STA) according to [23, 24, 52]. Additionally, we propose the local-adaptive-based testing (LAB) for the cases when the variance in the time domain may vary over the time. For the robustness check of CWT results we use simulations of background noise levels (as described in Chapter 7.1.2). For the cross-check of CWT tested by STA and by simulations of background noise levels, and CWT tested by adaptive LAB approach we propose robustness enhanced spectrogram modelling also tested by the simulations of background noise levels.

9.3.1 Local-Adaptive-Based (LAB) Testing

Both testing statistics presented in (7.1) and (7.2) are formed as the power value of the spectrogram of a noise signal normalized by the signal variance in the time

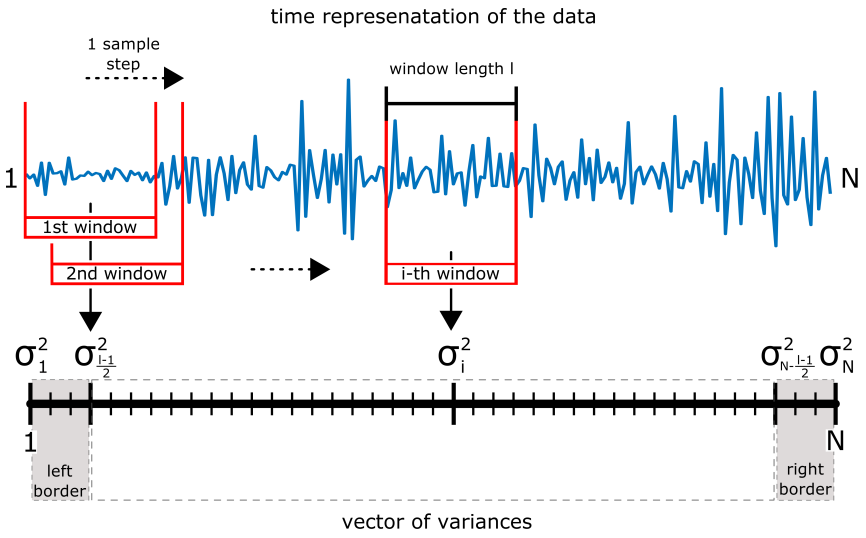


Fig. 9.1: LAB testing diagram.

domain. In the case of an input signal with strongly localized fluctuations of the signal strength, the total variance may not sufficiently describe the character of the data. It is, therefore, not surprising that events, such as the 2008 crisis, may have a strong impact causing a suppression of other events. To avoid this problem, we propose an adaptive form of STA testing named a local-adaptive-based testing (LAB). In the case when the data does not have such problem, the STA and LAB testing produce same results.

The LAB testing is based on the evaluation of significance via STA in each time n with respect to the sliding window l. Let us have the time series $s(n), n = 1, \ldots, N$ and set up the time window length l, l is an odd number. The vector of local variances $\sigma_n^2 = 1, \ldots, N$ of the same length as the time series $s(n)$ is calculated on the sliding window l with the sliding one step ahead (Fig. 9.1). In the border regions, edge effect may occur, because there is a limited number of observations available for the calculations. In these cases we use the first and the last l observations of the $s(n)$ for border variances σ_n^2 calculation. In the middle region we use l observations of the $s(n)$. Then we use (7.1) and (7.2) for thresholds calculation.

The localization allows us to assess the spectral components with reference to its surrounding events. The range/scope of these events is selected by the length l of the sliding window. To set an appropriate time window length, we must take into account the requirement for at least 35 points (approx. 3 years). To maintain sufficient adaptability to rapid signal fluctuations, we set the window length to $l = 49$ (i.e. approx. 4 years window).

9.3.2 An Enhanced Spectrogram Modelling

For the significance testing on the basis of STA, we have to know the background noise character. When analyzing economic indicators, we can assume that the background noise is GWN. In some cases (application of pre-filtering, heteroscedasticity in the data etc.) this assumption need not to be satisfied. To avoid such a case, we alternatively suggest the combination of several TF approaches and in the following we call the resultant TF transform "enhanced". The enhanced transform is the display of the CWT, TFAR and STFT spectrograms of one time series in one chart to obtain the best possible TF representation. We investigate mainly the amplitude part of the spectra. The phase part of complex spectra S_{CWT} and S_{STFT} is not investigated.

Firstly, we align the time axis of all obtained spectra S_{CWT}, S_{TFAR} and S_{STFT} to match each other. Since the trend of all three time vectors is linearly increasing, it is sufficient to adjust the starting and ending point for each method. We omit the first and last 15 columns of S_{CWT}, we denote the remaining matrix as S'_{CWT}. Hereby we ensure the correspondence of the time axes for all three methods. Secondly, we align the frequency/scale axis of S'_{CWT}, S_{TFAR} and S_{STFT}. The frequency range of S_{TFAR} and S_{STFT} is cropped to correspond with the range of S'_{CWT}, which was 6–192 months (0.5–16 years) cycles. The resulting frequency cycles vectors $\overline{f_{\mathrm{TFAR}}}$ and $\overline{f_{\mathrm{STFT}}}$ have a linearly increasing trend, however, the trend of $\overline{f_{\mathrm{CWT}}}$ is non-linear. To obtain the corresponding vectors we match each point of $\overline{f_{\mathrm{CWT}}}$ with one value of $\overline{f_{\mathrm{TFAR}}}$ and $\overline{f_{\mathrm{STFT}}}$ with 1.4% tolerance:

$$\begin{aligned} |\overline{f_{\mathrm{CWT}}} - \overline{f_{\mathrm{TFAR}}}| &\leq 0.014 \max(\overline{f_{\mathrm{CWT}}}; \overline{f_{\mathrm{TFAR}}}), \\ |\overline{f_{\mathrm{CWT}}} - \overline{f_{\mathrm{STFT}}}| &\leq 0.014 \max(\overline{f_{\mathrm{CWT}}}; \overline{f_{\mathrm{STFT}}}). \end{aligned} \tag{9.1}$$

With this step, we obtain the adjusted TF matrices S'_{TFAR} and S'_{STFT} by making all three methods aligned. The combination of methods is done by a simple multiplication and is called the "enhanced TF picture"

$$S_{\mathrm{TF}} \leq S'_{\mathrm{CWT}} S'_{\mathrm{TFAR}} S'_{\mathrm{STFT}}. \tag{9.2}$$

9.4 Application

9.4.1 Data

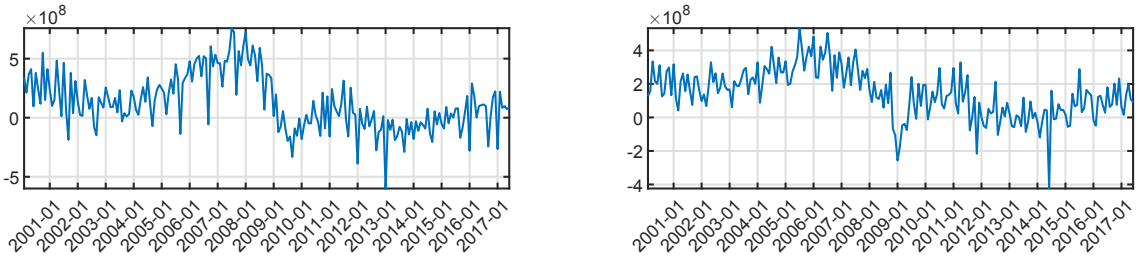
In order to identify the cyclical behavior of financial data, we use the seasonally unadjusted real monthly data of bank loans provided to corporates (Corporates) and

households (Households) in the Euro Area in 2000/M1–2017/M05 [82]. All variables are in first-order difference of natural logarithms (FODLOG). To be more precise, we focus on the credit cycle as we use the level of provided credit as one possible approach of how to measure the financial cycle, i.e. we use a quantity-based measure and not a price-based measure of the cycle. As proved by relevant empirical studies, the chosen methodology is applicable to our date range (see [20,36,67,68,73,80,81]. From the first overview of an input time series in Fig. 9.2a–b, we can see that the time series of Households and Corporates contain a long-term trend which goes through visible expansion and recession phases. In both cases, we can see several structural breaks.

In our analyses we use seasonally unadjusted data, because the aim of the paper is focused on the cyclical behavior of financial data (lending activities) where the seasonal component is taken as part of cyclical behavior. The information about cyclical character containing seasonal behavior is valuable, because the analysis of unadjusted data: (i) better reflects the real behavior of subjects (households and corporates) which can be influenced by seasonality; (ii) it can bring more valuable information to policy makers than adjusted data: they can react better to prevent disruptions of the economic cycle because of seasonality or they can reduce its possible negative effect. The use of adjusted time series may lead to losing some information, which could reduce the efficiency of monetary policy and limit the achievement of the objectives.

There are also methodological aspects [83] to use unadjusted data: (i) the seasonal component is not independent from the cyclical component and can change in time; (ii) 'the evaluation of the seasonal component provided by an adjustment method is hampered by the fact that the true seasonal component remains a theoretical and imprecise concept, never liable to direct observation'; (iii) 'the objectives of seasonal adjustment appear multiple and implicit. Is it to obtain the best estimate of the trend-cycle component, the best estimate of the seasonal component itself? Each objective will generate its own quality criteria'; (iv) 'the expected content of a quality report usually differs according to the user'. That is, the different adjusting method can produce different adjusted time series which can adjust more or less than a seasonal component and that seasonal and other cyclical components can interact.

Generally, in the case of monthly data we can expect that the seasonal component will range in frequencies up to 12 months. With respect to the ability of wavelets to model the time-frequency behavior of the time series, we can consider that the wavelet spectrogram in the range of 2–12 month length cycles (in our paper short-run cycles) can contains in the case of seasonally unadjusted time series seasonal and some cyclical component. Our complementary analysis confirm an interaction



(a) Corporates (b) Households

Fig. 9.2: Input data for Euro area.

of both cyclical components in the frequency range 6–12 month in our figures and that the seasonal component changes during time. This fact will be kept on mind during interpretation of the results.

Therefore, because we aim to identify the time-frequency regions describing mainly short- and medium-term cyclicality of bank loans before, during and after the financial crisis in 2008 and with respect to the ability of wavelets to model time-frequency character of the data it is worth to leave from economic as well as methodology point of view seasonal effect in the data.

9.4.2 Settings of TF Methods

In the case of the TF estimation via the TFAR, we use the Burg approach for coefficient estimates on 40 samples with 39 samples overlay, and the Hann window. The optimal value of the lag order is based on AIC criteria. The parameters of the STFT are set to correspond to the TFAR settings (40 samples, 39 samples overlay, Hann window) to simplify the process of the methods combination.

For the CWT transform calculation, we set the scales corresponding to the range of half a year to 16 years, with 388 individual scales. We select the complex Morlet wavelet with the center frequency $f_\mathrm{b} = 1.5$. That is, for the time vector with $N = 209$ samples $t = 2000/\mathrm{M1} - 2017/\mathrm{M5}$, we set the vector of the period T to be equidistantly distributed between maximal (T_max) and minimal (T_min) length of the period $T_\mathrm{max} = 16$ years (192 months), $T_\mathrm{min} = 0.5$ years (6 months) corresponding to the vector of frequency f with minimum and maximum of:

$$f_\mathrm{min} = 1/T_\mathrm{max} = 0.0625\,\mathrm{year}^{-1}, \quad f_\mathrm{max} = 1/T_\mathrm{min} = 2\,\mathrm{year}^{-1}. \tag{9.3}$$

For the number of scales 388, we can set the vector of scales s

$$s = \frac{f_\mathrm{b}}{f\delta_t}; \quad s_\mathrm{min} = \frac{f_\mathrm{b}}{f_\mathrm{max}\delta_t} = 9, \quad s_\mathrm{max} = \frac{f_\mathrm{b}}{f_\mathrm{min}\delta_t} = 288, \tag{9.4}$$

for $f_b = 1.5$ and $\delta_t = 1/f_s = 1/12$ (for monthly data).

In the case of the TF estimation via the TFAR, we use the Burg approach for coefficient estimates on 40 samples with 39 samples overlay, and the Hann window. The optimal value of the lag order is based on AIC criteria. The parameters of the STFT are set to correspond to the TFAR settings (40 samples, 39 samples overlay, Hann window) to simplify the process of the combined methods.

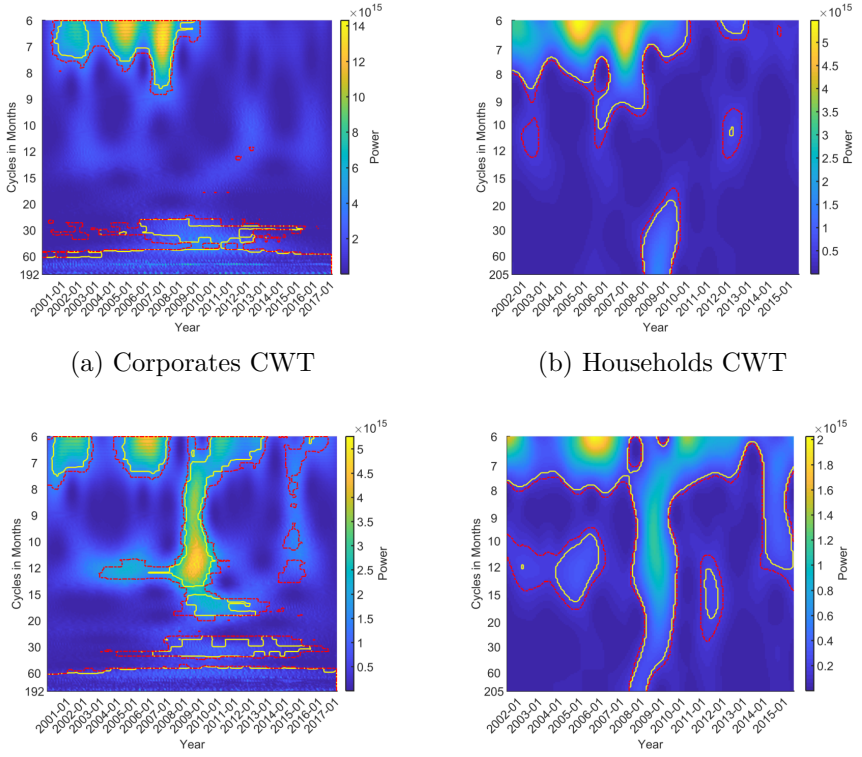
9.5 Empirical Results

After the preliminary analysis, we follow these steps: (i) we perform the CWT and STFT modelling of each series to obtain the spectrograms; we use the significance testing via STA. For the robustness check of the results, we apply simulations of background noise levels, according to the Chapter 7.1.2, with 1000 iterations; (ii) we perform adaptive LAB testing of the wavelet spectrograms; (iii) we do the cross-check of the (i)–(ii) results via enhanced spectrogram transform tested for its robustness by simulations of background noise levels; (iv) we compare and discuss the achieved results.

Although the TF method, especially the CWT, allows the modelling of a non-stationary time series, we decide to transform all input time series via the first order difference (FOD), because it can easily remove the long-term trend. Moreover, it is not possible to do the standard logarithmic transform before FODLOG for scattering reduction of the data, i.e. we can expect the persistence of a long-term component in CWT as an edge effect. Additionally, we check the results for both series, with and without a long-term trend. In accordance with the graphical processing and to insure a better visibility of detected areas, we decide to use detrended data via FOD. The long-term trend (i.e. cycles from 48 to 192 months) in TF transform of all indicators was present during all time. To make the orientation in the description of the results easier, we are going to divide the cyclical behavior into three basic regions: the short-run cycles (SR–C) of duration <12 months, the short cycles (S–C) of duration 12–20 months and the medium cycles (M–C) of duration 20–48 months. Denote that the wavelet spectrogram in the short-run cycles in our paper can contains seasonal and some cyclical component.

9.5.1 CWT and STFT Spectrograms Tested by STA

In our empirical analysis, we focus on the CWT modelling mainly due to its wide range of economic applications and its popularity among economists. Further, the CWT has a better time resolution compared to the STFT and the TFAR [20,81,84]



(a) Corporates CWT

(b) Households CWT

(c) Corporates STFT

(d) Households STFT

Fig. 9.3: CWT and STFT spectrogram tested by STA. The yellow/red (dotted line) curves in all figures indicate a significant area found by STA/simulations of background noise levels.

which is important for economic applications in general. To confirm the CWT results, we use the STFT method which estimates the spectrogram in the moving time window, therefore, it can differ in low frequencies (due to the window used). Both CTW and STFT transforms are tested by standard STA test. For the robustness of the results both transforms (CWT and STFT) are also tested by simulations of background noise levels.

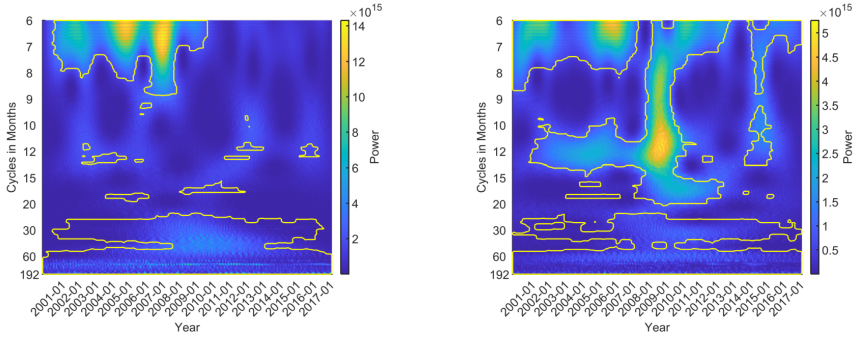
Comparing the CWT (Fig. 9.3a, c) and the STFT spectrograms (Fig. 9.3b, d), we can find two differences in the significance areas: (i) using the CWT transform, we identify a long-term trend component covering cycles of the approximate length of 48 months during the time range, while in the case of the STFT it is not present; (ii) the significant areas for the CWT and the STFT are different. While for STFT, STA and simulations of background noise levels identify similar significant regions, in case of CWT, simulations of background noise levels show a wider and also an additional area of significance.

The first difference mentioned above can have several reasons. It can be caused be the existence of the edge effect of the CWT transform called the cone of influence (COI). As written in [23], the COI is a usual problem for finite-length time series and may occur at the beginning and at the end of the spectrogram or the PWS representation. The second reason is the nature of the Fourier transform [81]. The persistence of a long-term trend component could be also expected with respect to the FODLOG transform, as we mentioned at the beginning of the Results section. Such an assumption is partially confirmed by the results of the STFT which is not so sensitive to the long-term trend. In other words, if the data contain the long-term trend component represented by the long cycles (low-frequency component) occurring in the time period shorter than the moving window part (which is 40 samples in our case, i.e. approximately 36 months), the STFT will not identify it. Despite the fact that STFT has this limitation, we can identify the existence of the long-term trend component in the sub-period of Households and Corporates. Then, we can admit the existence of cycles of 30–48 months duration despite the existence of the CWT edge effects.

In the case of the second difference (except the long-term trend component explained above) we assume that it is caused by STA testing which evaluates the significance with respect to the fixed variance calculated in all-time range and does not consider the data character, i.e. the volatility of the values. Then, an event (such as the 2008 crisis) may suppress the significance of other events. It is also important to note that for the STFT results, the resultant figures have a shorter time axis caused by the STFT methodology. Thus, CWT transform tested by classical STA approach can lead to misleading results. Therefore, the STA test seems to be insufficiently adaptable to the volatility changes.

9.5.2 LAB significance testing

As the next step to adapt to the volatility changes in CWT testing, we proceed with an adaptive form of STA testing, i.e. local-adaptive-based testing. Comparing Fig. 9.3a,c and Fig. 9.4 and taking into account the adaptive nature of LAB testing, we can conclude that the LAB testing of PWS spectrograms generally confirms the CWT spectrogram tested by simulations of background noise levels. The differences in the significant areas identified via STA and the LAB testing in CWT figures is caused by the changing volatility in the data. This fact should imply that STA approach may not be able to find all significant regions. That is, the denominator of the testing statistic given in (7.2) is a constant value and it is the total variance for the time series (i.e. one fixed number). Therefore, in this case we prefer the results of the LAB testing of CWT due to its adaptability.



(a) Corporates (b) Households

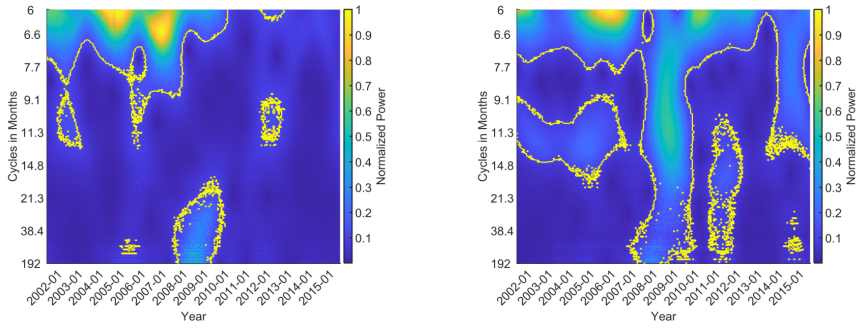
Fig. 9.4: Wavelet spectrograms tested by LAB test.

9.6 Robustness Analysis

9.6.1 Enhanced Spectrogram Modelling

As an alternative approach which considers data volatility we propose the enhancement spectrogram modelling (defined in Chapter 6) tested via simulations of background noise levels. The enhanced transform is the display of the CWT, TFAR and STFT spectrograms of one-time series in one chart for obtaining the best possible TF representation. Since the CWT has a better time resolution compared to the STFT and TFAR (they are better in the frequency resolution) [20, 81], we decide to use the combination of the TF results via simple multiplication. This method is based on a simple idea that important components in the same positions (time and frequency) will be strengthened and methodical residues should be suppressed. The resultant spectrograms according to (9.1) have been called 'enhanced' and are presented in Fig. 9.4a–b. Due to the STFT and TFAR limitations we expect a worse ability to capture the long-term trend component. Since the application of STFT and TFAR causes the shortening of the sample size, all figures below (Fig. 9.4) are in the shorter time range 2001/9–2015/9 compared to the CWT (Fig. 9.3a–d). Further, the comparison of CWT and enhanced spectrograms is evaluated in this shortened time range. All enhanced spectrograms are tested via simulations of background noise levels.

Comparing Fig. 9.3a, c, Fig. 9.4 and 9.5, and taking into account the methodology nature of individual approaches, we can conclude that the enhanced approach generally confirms the influence of data volatility on the CWT spectrogram testing. That is, we confirm that it is important to take into account the data character during the TF significance testing.



(a) Corporates (b) Households

Fig. 9.5: Enhanced TF pictures.

9.6.2 Comparison of Achieved Results

Since the results show that the enhanced spectrograms and LAB testing of CWT spectrograms generally provide the same results, we will demonstrate the difference between STA testing and the LAB testing of CWT spectrograms. The summary of the significant regions is shown in Tab. 9.1.

For Households and Corporates, both testing approaches for the short-run cycles have, in general, similar results till 2013. In 2013–2017 there is the difference in case of Household, where the LAB approach identify additional cyclicality (i.e. seasonal or cyclical component). In the short cycles, the LAB testing identified additional significant cyclical behavior compared to STA. In the case of medium cycles, the LAB testing identified a wider time-region with a significant cyclicality. Such similarities are possible because, analyzing Fig. 9.2, we can see that the Households and Corporates time series do not show a serious problem with the data volatility. The bigger differences occur when comparing both testing approaches. STA was not able to reveal any cyclicality (i.e. seasonal or cyclical component) in the short-run cycles and any cyclicality in the short cycles in Households after 2013. In case of Corporates STA was not able to reveal any cyclicality in short term cycles. In the case of the medium cycles, the identified areas via STA are smaller. Such a shortcoming of STA testing was expected also from the comparison with the STFT results.

As we stated in the sections above, the significance via STA does not consider variation in the data variability and takes a fixed variance for all data samples. Therefore, this testing indicates the significant areas of the cyclicality with respect to the all-time range. This leads to the suppression of other possible significant areas by the biggest shocks (the financial crisis). Moreover, the identified areas are smaller or are not even identified. On the other hand, using the LAB testing uncovers and confirms the existence of previously omitted areas. In this way, we are

	Test	Short-run cycles (<12 months)	Short cycles (12–20 months)	Medium cycles (20–48 months)
Corporates	**STA**	2001/Q1–2002/Q3 2003/Q3–2009/Q1	–	2004/Q3–2015/Q1
	LAB	2000/Q4–2010/Q1	2002/Q1–2004/Q3 2008/Q2–2011/Q1 2012/Q1–2013/Q2 2015/Q2–2016/Q3	2004/Q3–2015/Q1
Households	**STA**	2000/Q4–2002/Q2 2004/Q4–2007/Q1 2008/Q2–2013/Q1 2014/Q2–2016/Q1	2007/Q1–2012/Q2	2006/Q1–2016/Q3
	LAB	2000/Q1–2007/Q3 2008/Q2–2013/Q1 2014/Q2–2017/Q1	2001/Q2–2012/Q2 2014/Q1–2015/Q2	2000/Q1–2017/Q2

Tab. 9.1: Significant area of wavelet spectrogram over frequency intervals.

able to evaluate the cyclical behavior of lending activities in the time window with a higher precision.

Based on the performed analyses and after a detailed examination of the results, the methodological findings can be summarized in the following recommendations. In the case of constant volatility in the data, STA approach is plausible. We can also recommend and use this test if we want to evaluate any event in the time series with respect to the all-time range. In cases of increasing or decreasing volatility in the data, the overall variance can be affected. In this case, we recommend LAB testing, due to its adaptability, or enhanced spectrogram modelling.

9.7 Discussion

In this chapter, we have presented and compared the results of three approaches (CWT and STFT tested by STA and simulations of background noise levels, LAB testing of CWT, enhanced spectrogram modelling) applied to the cyclicality description of bank loan activities. We recognized that the volatility in the data influences the type of testing and we, therefore, recommend using the LAB testing of CWT or enhanced approach. In this section, we discuss interpretations of achieved results in the economic context.

First, the general results given by CWT with the STA testing (Fig. 9.3a–d, Tab. 9.1) show that the financial crisis was reflected in both economic sectors, but

in a different way. We found the most extensive reaction to the crisis in the behavior of Households. The reactions were strong particularly in the years around the crisis across all frequencies (from short-run to long cycles), and in the years just before and after the crisis in the quick movements represented by the short-run cycles (i.e. in seasonal and some cyclical components). In the case of Corporates, the financial crisis did not cause any significant reactions during and after the crisis. We can find only one reaction, i.e. the disappearance of short-run cyclicality.

Next, due to new approaches, i.e. adaptive LAB testing or enhanced spectrogram modelling, we can evaluate the above-mentioned general results with a better precision. In the case of Corporates, there is an important area in the short-run cycles which covers seasonal and cyclical components and during the crisis. After the crisis, there is no significant area, probably as a result of the credit crunch. The LAB testing for Corporates (Fig. 9.5a, Tab. 9.1) reveals the important pre-crisis and crisis period 2002–2010 in the short-run frequencies (up to 12 months; i.e. seasonal and some cyclical components) and the unique short cyclicality (12–20 months). Thus, we can see that the reaction of Corporates to the crisis was very limited without any strong impact on the post-crisis time.

At the same time, the situation for Households is different; the financial crisis of 2008 can be taken as an important factor having a strong impact on cyclicality. The new approaches (Fig. 9.4b and 5b, Tab. 9.1) show several important areas: (i) the first area can be identified in the time 2000–2016 in the short-run cycles (up to 12 months; i.e. in the seasonal and some cyclical components); (ii) the second in the time 2001–2012, 2014–2015 in the short cycles (12–20 months); (iii) the third in the medium cycles in the time 2006–2012. We can see a very strong reaction to the crisis, which was reflected across all frequencies around and after the year 2008. Before 2008, a significant area lies (similarly as for Corporates) in the short-run cycles which covers seasonal and some cyclical components. After the crisis, we can find an important area in the short-run and the medium cycles. Therefore, we can confirm our previous results that the cyclical movement of loans to Households during the financial crisis was the most significant. In the case of both sectors, we can also see the medium cycles (12–20 months), but due to the COI (discussed in the previous sub-section), we can admit its existence in 2007–2011, i.e. shortly before, during and after the crisis.

Overall, our results can be summarized as follows. The financial crisis of 2007 and 2008 had a significant impact on the cyclical behavior of both categories of economic sectors analyzed in our paper, but in a different way and with a different intensity. Moreover, the character of the cyclical behavior was different before and after the crisis. In the case of Corporates, we do not see any significant cyclical behavior after the crisis. This fact could be caused by a stronger position of large

firms which have more financing possibilities than small firms and both banks and nonbanks provide short-term financing to large firms rather than smaller firms [85]. Small firms are also connected to small banks [86] and this may influence the lending activity of small banks as they are more influenced by economic distress and information asymmetries compared to large banks [87]. However, the ongoing banking consolidation and a reconstruction of internal organization of banks in recent years have caused the fact that the bank size is not as important for small business lending as before [88]. As such, this sector could be presented as a relatively stable sector with the least volatile lending activities. On the contrary, in the case of Households sector, the crisis was a very important event causing changes in the volume of lending activity and thus economic or financial distress may cause huge fluctuations in the spending of households which are also substantial for the economic growth of a country. This fact is documented in [85] who argue that banks tend to limit lending to households while they may rise loans to firms at the same time. However, in [89] is stated that it is lending to firms, and not lending to households, that has a positive impact on economic growth and limits income inequality through the financial development, better capital allocation and economic transformation. Therefore, the lending activities in the sector of Households showed the most extensive cyclical behavior and this sector could be characterized as the sector which was influenced by the financial crisis most significantly. In this context, it would be advisable to stabilize these fluctuations in lending activities of households using various economic policy measures. Many studies confirmed the existence of the cyclical behavior of the lending activities (see [61, 90–92] and others). However, we bring new findings about the short- and medium-term cyclicality in these two economic sectors via the TF transform.

9.8 Chapter Conclusion

This chapter deals with the identification of the time-frequency regions describing the cyclicality of bank loans activity before, during and after the financial crisis in 2008. We proposed the local-adaptive-based testing and the so-called enhanced time-frequency spectrogram modelling and compare them with standard classical testing of CWT and STFT spectrogram. The demonstration of the methods was proposed only on seasonally unadjusted monthly financial data of bank loans provided to two categories of economic sectors in the Euro Area in 2000–2017. We identified areas of cyclicality in the lending activities and found an important impact of the crisis on data volatility which further influenced the type of significance testing of wavelet spectrograms. The results confirmed the usability of the newly proposed approaches, i.e. LAB testing of CWT spectrogram and enhanced spectrogram modelling, in those

case the data have a changing volatility. In the case of constant volatility, the STA testing approach of TF transform gave the same results.

The proposed approaches extend the methodology related to studying short and medium cycles as they can be applied also in the case of volatile data. Applying the methods, which were conducted only on seasonally unadjusted monthly data, enables us to detect a faster reaction to a specific change, which further enables a faster response from the economic policy makers.

From the economic point of view, the results show that the crisis in 2008 had an impact on the cyclical behavior of corporates and households, but in a different way. The most remarkable influence of the crisis was identified in the households. Here, the reaction was relatively strong during and after the crisis. Therefore, we can identify Households as a more sensible sector reacting to changes, such as the 2008 crisis. At the same time, the less dynamics is apparent in Corporates, with severe fluctuations mainly in the period before the financial crisis. Therefore, Corporates seems to be less affected by the crisis compared to Households and this finding may be of high importance for policy makers when formulating and implementing their economic policy measures in a period of crisis.

On the other hand, the cyclicality of Corporates was mainly dominated by the short-run cycles (which includes seasonal components) and the medium cycles. Therefore, Eurozone policy makers should take this information into account. We need to point out that the presented research did not distinguish between the supply and the demand factors of lending activities, i.e. whether the movement was caused by banks or by economic agents. In this context, we bring interesting findings showing the cyclicality in economic sectors of corporates and households.

10 Co-movement (Sub-)Indicator

This chapter is based on following published journal paper:

POMĚNKOVÁ, J.; KLEJMOVÁ, E.; MALACH, T. "Co-movement (Sub-) Indicator as the Measurement of the Synchrony of EA and Visegrad Group Countries". **Journal of Economics**, 2020, vol. 68, no. 3, p. 231-251.

Abstract The paper deals with the construction of a co-movement indicator suitable for assessing the synchrony between countries. The indicator is represented as a time series and its construction is based on a reconstruction of a co-spectrum measure from the time-frequency to the time domain. We use the statistically significant part of the power wavelet co-spectrum for pairs of countries. An advantage of the newly proposed co-movement indicator is no loss of observations (such as in the case of correlation) and a possibility to construct sub-indicators which correspond to the predefined frequency range, e.g. business cycle frequencies. In such a way we can obtain a decomposition of the (total) co-movement indicator (which covers the full range of frequencies) into, for example, short-run cycles, medium and long business cycles and long-run cycles. The proposed methodology is demonstrated on the US and EA monthly data of industrial production index in 1991-2018. A further application is performed on the EA and Visegrád Group Countries with the same data type and time range. The results confirm the globalization of the world economy and a transfer of the 2008 crisis into European countries economies. Further, it is shown that the synchrony of the EA and Visegrád Group of Countries is the most striking in medium business cycle frequencies covering the cycles of the length 1.5-4 years. In the case of Hungary, the synchrony is also visible in the short cycles of the length <1.5 years. The obtained co-movement (sub-)indicators can be further used in regression analyses for researching the relation with economic indicators, such as bilateral trade or investment to reveal additional information about interconnections of the countries.

10.1 Introduction

The globalization of economies has been reflected in many areas of the particular country economy, and thus is still at the forefront of economic analysts' interest. It can be studied in many contexts, i.e. business cycle synchronization between countries, financial globalization and its effect on monetary policy, emerging markets, financial or goods markets, the influence of globalization on social life, etc. The

transfer of shocks from the economy of one country to other countries can significantly influence the evolution of regional as well as world economy. The crisis in 2008 which started in the USA, or oil crisis in 1970s are one of the most visible examples. In case of the European area we can mention another example which is the adoption of Euro currency leading to the Optimum Currency Area (OCA). Specifically, the OCA started a huge amount of analyses of business cycle (BC) synchrony in various forms.

The analysis and measuring of globalization was also reflected in the development of methodological tools. Standard techniques, such as regression analysis, correlation, co-integration, vector autoregressive modelling, all in its basic or modified form are still in usage and can give valuable results [36, 93]. However, these methods are not flexible enough to provide an in-depth view into the cyclical character of economic data and do not capture the cyclical structure as a function of time. In the last decade wavelets, especially co-movement wavelet analysis, have become a popular instrument. This methodology has been known in engineering for a long time, but it is quite young in economics. An increase in its came with the analysis of BC synchrony, OCA and the reflection of structural shocks in economies. Wavelets can be applied on non-stationary data, they have a very good time-frequency localization, and it is possible to reconstruct the transformed values back into the time domain [29].

The globalization between China and G7 countries via wavelets was assessed by [80]. They proved that the co-movement between the countries differs with respect to the frequency in which the co-movement is being evaluated. In [36] the synchrony between China, Japan, the US and other Asia Pacific countries is analyzed via wavelet power spectrum and found out that the strength of BC synchronization fluctuates across frequencies and over time. Another study in this field [20] uses cross wavelet power spectra and wavelet coherence to show that a closer BC synchronization can lead to a stable and effective monetary union. In [39] different approach is used: the rolling wavelet correlation between the stock markets' index returns of the PIIGS countries with the UK and Germany at different frequencies. Their results show that in low-frequencies the PIIGS stock markets are more correlated with Germany, while in high-frequencies with the UK. The interconnection of financial globalization and monetary policy effectiveness was studied [94], where is empirically quantified that the financial globalization has a net impact on monetary policy effectiveness. The relation between the Korean Republic and the European Union, which signed the trade agreement in 2010 (EU-Korea FTA), motivated [95] to research the globalization effect between Visegrád countries (V4: the Czech Republic, Hungary, Poland and the Slovak Republic) and the Republic of Korea. The author identifies the impact of South Korean direct investments on trade. He states

that the EU-Korea FTA agreement is an important step in the process of globalization of economies serving the interests of the most competitive economies. Further he claims, that V4 countries are passive players and will continue to remain so as long as they are highly dependent on the effects brought about by foreign direct investments.

This chapter deals with the description of the synchrony between two time series via the construction of a co-movement indicator. The indicator is constructed on the basis of time-frequency co-movement measure, i.e. coherence between two time series. After the identification of the significant part of the coherence via statistical testing, the significant co-moved part of the time series is reconstructed into the time domain leading to the time representation of the co-movement indicator. An advantage of this indicator is its time domain representation, which contains information about the significant frequency of the co-movement regions. Further, the methodology of its construction allows a selection of frequency regions according to the researcher's focus, e.g. BC frequencies, or the construction of the corresponding sub-indicators. The proposed indicator is demonstrated on the industrial production index of the US, EA and Visegrád Group of countries between years 1991-2018. The indicator thus enables us to asses how the co-movement between the US and EA, EA and V4 countries evolves.

10.2 Methodology

10.2.1 An Algorithms for the Co-movement Indicator

In the following, we propose the diagram (10.1) and the description of the algorithm for the co-movement indicator construction. Let us two time series $x(n)$ and $y(n)$, $n = 1, \dots, N$. The construction follows these steps:

Step 1: The estimation of the co-movement measure, i.e. PWCS or coherence, for both time series.

Step 2: The partition of the co-movement measure into regions with the significant and insignificant co-movement, i.e. the construction of the mask

$$M(a,\tau) = \begin{cases} 1, & \dfrac{|S_{xy}(a,\tau)|^2}{(\sigma_x^2\sigma_y^2)} \geq 0.25Z(1-\alpha), \\ 0, & \dfrac{|S_{xy}(a,\tau)|^2}{(\sigma_x^2\sigma_y^2)} < 0.25Z(1-\alpha) \end{cases} \tag{10.1}$$

where the threshold $thr = 0.25Z(1-\alpha)$ is given by TC98 testing (7.5).

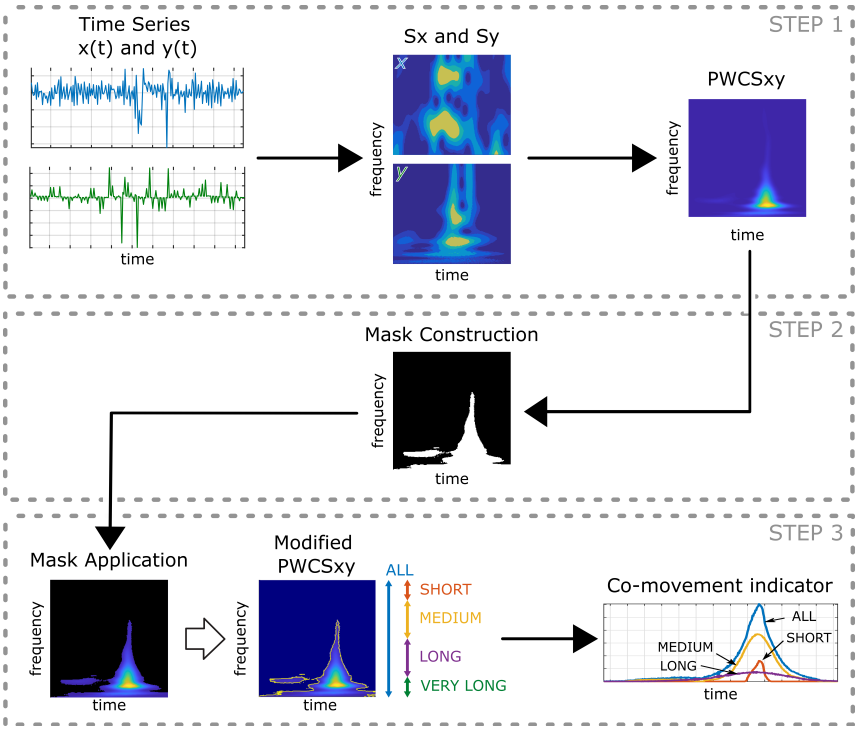


Fig. 10.1: Co-movement indicator construction.

Step 3: Creation of modified PWCS (MPWCS) which contains only the significant part of PWCS by applying the mask:

$$MPWCS(a,\tau) = |S_{xy}(a,\tau)|^2 \, M(a,\tau). \tag{10.2}$$

Consequently the ICWT is used to transform this product into the time domain, i.e. the construction of the time representation of the co-movement indicator. Or, ICWT of the pre-defined frequency region (e.g. BC frequencies) of this product into the time domain to construct the time representation of the co-movement sub-indicator corresponding to the pre-define frequency region.

10.3 Data Description

The data set consists of the seasonally adjusted monthly data of IPI from OECD database [40] for the the US, EA19 (EA) and Visegrád Group of countries (V4), i.e. the Czech Republic, Hungary, Poland and the Slovak Republic, between years 1991/M1-2018/M5.

As we examine TF selective filtering based on co-movement between the growth cycles of selected countries, we transform the data to FODLOG values. For the time-series co-movement via PWCS calculation we set the scales corresponding to

the range of 1 year to 10 years, with 257 individual scales. We selected the complex Morlet with the center frequency $f_b = 1.5$ as the mother wavelet [29].

The co-movement indicator was constructed for all frequencies (all scales) leading to the time representation of the total co-movement indicator. We also construct the time representation of the co-movement sub-indicators corresponding to the frequency sub-ranges. The division of the frequencies re-calculated to the length of cycles is the following: the short-run cycles (SC) of duration <1.5 years, the medium BC of duration 1.5–4 years, the long BC of duration 4-8 years and the long-run cycles (LC) of duration >8 years.

10.4 Results

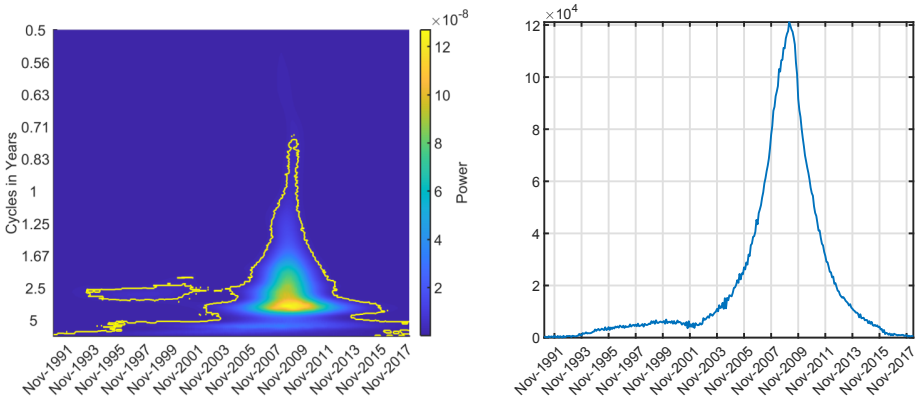
The application was done on the data described in the previous part according to the algorithm described in the methodology (Section 2.4). First we demonstrate the proposed algorithm on the case of co-movement between the US and the EA. Secondly, we apply this algorithm on the co-movement of the EA and V4 countries and present the resultant time domain figures for the total co-movement indicator as well as for the sub-indicators corresponding to the pre-defined frequency regions.

10.4.1 Demonstration of the Co-movement Indicator

The proposed algorithm demonstrated on the case of the US and EA co-movement follows the steps given in Section 2.4. Firstly, we estimate PWCS (Figure 10.2a)) of both countries. Secondly, we construct the mask on the basis of TC98 significance testing of PWCS, i.e. we identify significant and insignificant parts of PWCS according to the formula (10.1). Then, we multiply the PWCS (normed by variances of the US and EA time series) by the mask to obtain time-frequency co-movement representation. Thirdly, we calculate the inverse transform of the normed PWCS which is an analogy as for CWT (eq. (3.10)), and we obtain the time representation of the total co-movement indicator in the time domain (Figure 10.2b).

Further, for a detailed look into the co-movement indicator behavior, we split the masked normed PWCS into the frequency sub-regions according to the description given in Section 3.1. Similarly, as in the case of total co-movement indicator, we calculate the inverse transform of a frequency sub-region of the masked normed PWCS, which leads to the time domain representation of the sub-indicators and corresponds to the pre-defined frequency regions (Figure 10.3).

The first overview of the PWCS (Figure 10.2a) suggests that the most significant event in both countries is the global economic crisis in 2008. Therefore, the shape of the total co-movement indicator, i.e. the full range of frequencies, is not surprising.



(a) PWCS (b) Total co-movement indicator

Fig. 10.2: Co-movement of the US and EA.

We can see an increase of the synchrony in 2002-2008, with a peak in 2008, followed by a decrease of the synchrony in 2008-2014. The inflexion points determine the period 2006-2011 of the concave indicator shape. In this interval we can find the most energy of the PWCS and thus the strongest co-movement of both countries.

When we focus on the sub-indicators in Figure 10.3 in the pre-defined frequency sub-ranges, we can see a similarity of the medium BC co-movement indicator shape with the total co-movement indicator shape (i.e. the curve corresponding to all frequencies in Fig. 10.3). Based on this we conclude that in the BC frequencies of the duration 1.5-4 years the economies are the most synchronised. The other sub-indicators, i.e. SC, long BC and LC frequencies, contribute to the overall synchronisation less, because in the PWCS there are no other significant areas in the frequency and the time (see PWCS, Figure 10.2a)). The long BC sub-indicator summed with the medium BC indicator cause the total indicator enlargement, while the SC sub-indicator causes the augmentation of the peak in 2008. The LC indicator has the lowest contribution to the total co-movement indicator. In both countries, the EA and the US, the global crisis in 2008 significantly influenced developments in the countries and demonstrated the interdependence, and thus globalization of economies.

10.4.2 Application on the EA and V4 Countries

Similarly to the demonstrated case, in the case of the EA and Visegrád countries we firstly look at the PWCS figures (Figure 10.4a-d). In all four figures we can identify the most significant co-movement during the global financial crisis, more or less extended across the frequencies, depending on the country. In the case of

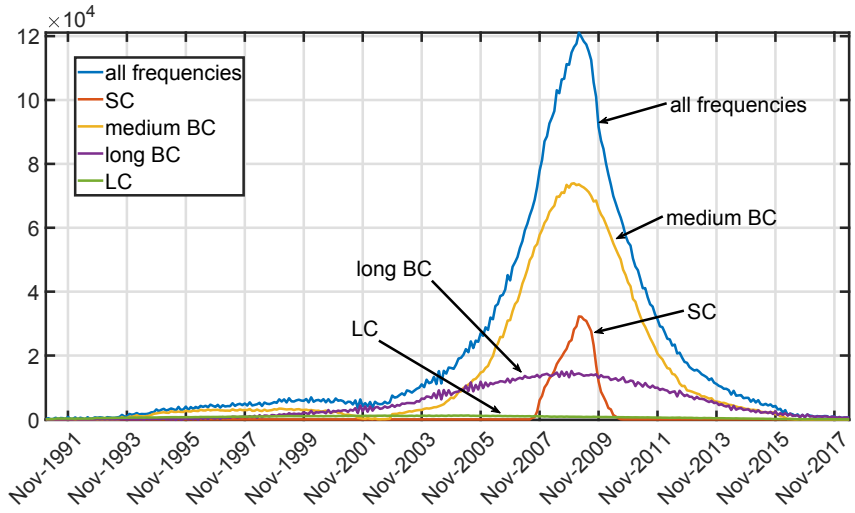
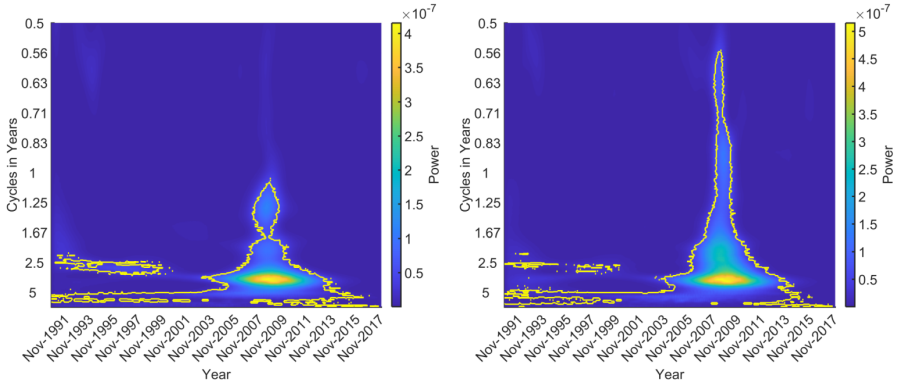
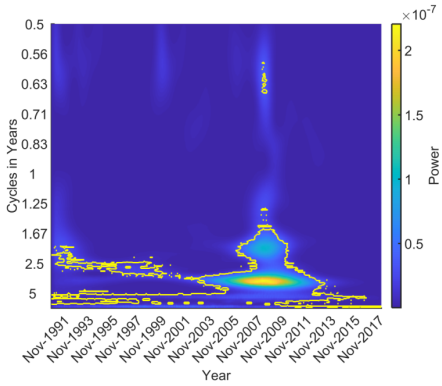


Fig. 10.3: Co-movement indicator for the US and EA in the frequency regions.

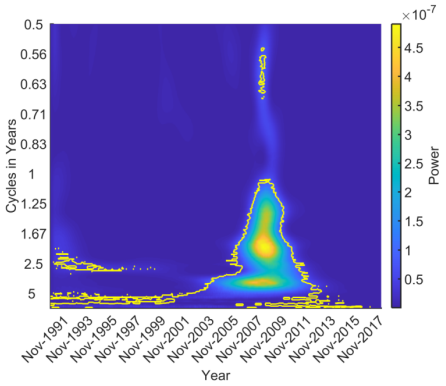


(a) CZ & EA

(b) HU & EA



(c) PL & EA



(d) SK & EA

Fig. 10.4: PWCS of EA and V4 countries. The yellow curves in all figures indicate a significant area of PWCS.

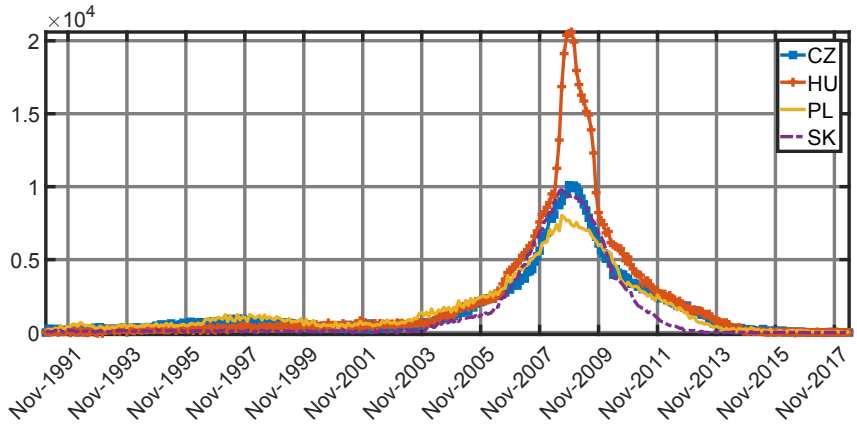


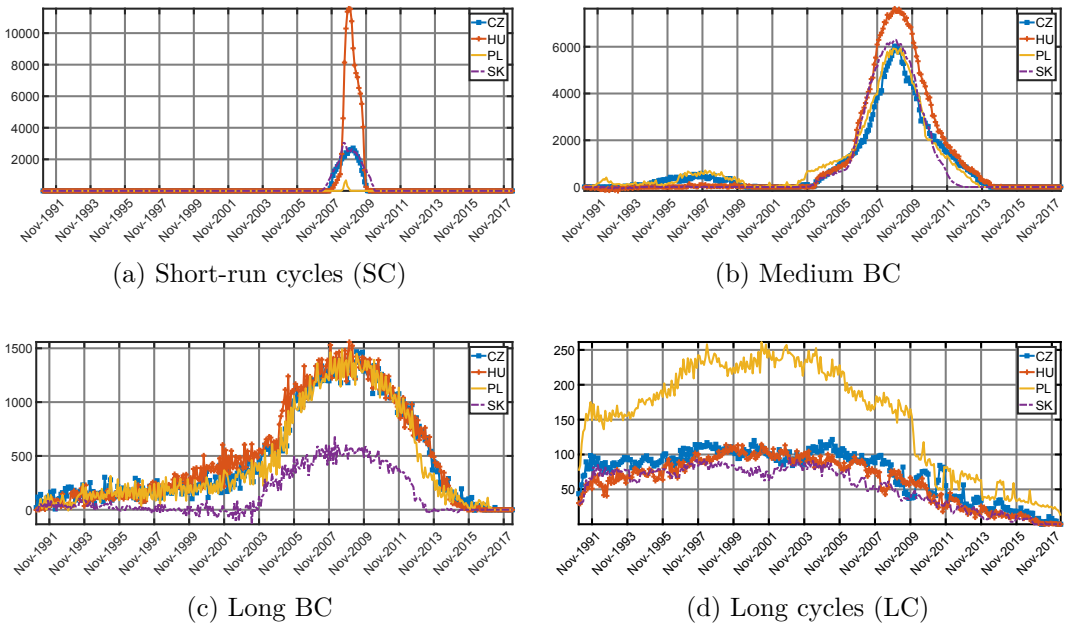
Fig. 10.5: The total co-movement indicators for V4 and the EA (all frequencies).

the EA and the Czech Republic, the Slovak Republic and Poland the significant co-movement prevails in the LC, long BC and medium BC. In the case of the EA and Hungary the significant co-movement is visible in all frequencies in the 2008 crisis. All these facts are reflected in the co-movement indicator as well as in the sub-indicators.

The total co-movement indicators of the EA and all V4 countries for all frequencies is shown in Figure 10.5. We can see a proximity of the co-movement of the Czech Republic, Poland and the Slovak Republic with the EA during the time range; in the case of Poland the indicator achieved a little bit lower value of the peak in the 2008 crisis. In the case of Hungary, the indicator has a similar shape, except for the shape in the time 2007-2009 when a significantly higher level of peak was achieved. This fact is explained by the SC co-movement sub-indicator.

Further, we focus on the sub-indicators. In the case of SC co-movement sub-indicator (Figure 10.6a) we can see a proximity of the Czech Republic and the Slovak Republic in the shape of the curve. In the case of Poland, the SC sub-indicator does not contribute much to the overall country co-movement, i.e. the EA and Poland do not have co-movement in these frequencies. The opposite situation has occurred for Hungary where the co-movement in the SC plays an important role. As we can see in the PWCS figure (Figure 10.4), the co-movement with the EA covers all frequencies during the crisis time, which is reflected in the SC sub-indicator.

In the medium BC frequencies, similarly to the US and EA case, the Visegrád countries and the EA are the most synchronized and the medium BC co-movement sub-indicator (Figure 10.6b) contributes most to the total co-movement indicator (with respect to the amplitude of the indicator). Therefore, we can expect a similar reaction of countries to the situation in the EA which primarily reflected the situation in the US. In the case of Hungary we can see a higher position of the peak



(a) Short-run cycles (SC)

(b) Medium BC

(c) Long BC

(d) Long cycles (LC)

Fig. 10.6: Co-movement indicators for the US and EA in the frequency regions.

in 2008. The inflexion points of all medium BC co-movement sub-indicators are determined by the period 2006-2010. In this interval there is the highest amplitude (i.e. the most energy) of the PWCSs.

The long BC co-movement sub-indicators (Figure 10.6c) between the EA and V4 countries have a similar shape for the Czech Republic, Poland and Hungary. The co-movement sub-indicator for cycles longer than 8 years (Figure 10.6d), i.e. LC sub-indicators, contributes less to the total co-movement indicator in all countries, and thus to the co-movement of the country and the EA. In the case of the Czech Republic (with respect to the indicator amplitude) we can state that the long BC sub-indicator has the same contribution to the total indicator as the SC sub-indicator.

10.5 Chapter Conclusion

This chapter is focused on the description of the synchrony between countries via a co-movement (sub-)indicator. The indicator is constructed as the reconstruction of the significant power wavelet co-spectrum from the time-frequency to the time domain. There are two advantages of the indicator: i) contrary to the correlation coefficient, concordance index or other metrics, the co-movement indicator is a time series represented in the time domain without any loss of observations; ii) the methodology of co-movement indicator allows the construction of the total

co-movement indicator as well as the construction of the sub-indicator which corresponds to the predefined frequency range (e.g. BC frequencies). According to the aim of interest we can divide PWCS into frequency sub-regions and then construct sub-indicators.

The proposed methodology is demonstrated on the case of the US and EA. We found out that both subjects are synchronized mainly in medium business cycle frequencies covering cycles of the length 1.5–4 years. Further, we could see an increase of the countries co-movement in the area of ± 6 years around the global economic crisis 2008. A consequent application is presented on measuring the synchrony between the EA and Visegrád Group of Countries. Similarly, as in the case between the US and EA, the synchrony of EA and Visegrád Group of Countries is the most important in medium business cycle frequencies. Only in the case of Hungary we additionally identified the same importance of short cycles of the length less then are 1.5 years for the synchrony. That is, the crisis in Hungary was reflected in a wider range of frequencies then are in the Czech Republic, Poland and the Slovak Republic. The results confirm the fact of globalization of economies and the transfer of important shocks, such as the 2008 global economic crisis into European countries.

The obtained co-movement (sub-)indicators can be further used in regression analyses, as presented by [80] or [96] for the research of the relation with economic indicators, such as bilateral trade or investment to reveal more information about economic interconnections and influencing factors.

11 Co-movement Selective Detection Filter

This chapter is based on following published journal paper:

POMĚNKOVÁ, J.; KLEJMOVÁ, E. "Co-movement Selective Detection Filter to identify time series co-movement indicator or to filter out symmetric economic shocks". **Digital Signal Processing**, 2021, vol. 114, p. 1–14.
DOI: 10.1016/j.dsp.2021.103033

Abstract The paper deals with designing a mask suitable for a selective filtering of data. The design of the mask is performed in the time-frequency domain and the selection is based on the co-movement measure of time series. We propose two approaches for the mask construction: i) hard thresholding based on χ^2 testing; ii) adaptive based thresholding. The proposed mask can be used for time series filtering in which we obtain either the adjusted time series or the construction of the time series containing only the co-moved parts. Further, after computing an inverse transform we can obtain time series with/without the co-moved area applicable for consequent econometric analyses. The paper provides recommendations concerning the selection of a particular approach in a given situation. The proposed methodology is demonstrated on the adjustment of industrial production index of Euro Area and selected G8 countries about co-movement with the US.

11.1 Introduction

A large number of econometric analysts use filtering when processing data. They are usually interested in a decomposition into the long-term trend and oscillations. After that, the filtered time series is taken as an input to further econometric analyses. Filtering techniques are viewed not only from the perspective of removing the trend component, but also from the perspective of identifying trend-breaks, outliers, or its removing ability [97, 98].

For a long time, the filtering of time series prevailed in the time-domain represented by deterministic [99, 100] or stochastic methods [101], or their combinations [102]. The time series processing in the time-domain is simple, but is not able to remove a specific frequency range and, in some cases, is inflexible in the long-term trend modelling. Unfortunately, time domain approaches are weak in capturing a cyclical character of the time series and need parameter optimization. Therefore, analysts began to use methods in the frequency-domain, i.e. low-, high- or band-pass filters [31,103,104], or a windowed filter as proposed by [105]. Such approaches are more flexible and are widely used in the economic area for the business cycle

(BC) analyses [106]. An alternative point of view is filtering via eigenvalue-based decomposition [57, 107, 108], which allows a decomposition of the time series into the number of components. Then, some of the components can be removed and the others can be reconstructed into the time domain. The drawback is an arbitrariness of the decision what will be removed. This deficiency can be solved by the TF selective filtering [109].

The TF methods have come to the forefront of the interest in many scientific disciplines; engineering [26, 110], geoscience [111, 112], biomedical [113]; social sciences, or economics [38,114]. The popular TF techniques are the Short Term Fourier transform, the time-frequency autoregressive method or the wavelet transform. As Wang et. al [115] wrote, wavelet transform has a good TF localization which enables the focus on any detailed time and frequency domains of input time series. The methods can be used for analysing time series, either individual or in the co-movement. The latter is very popular among economists [20, 116, 117]. As many experts agreed, the main advantages of wavelets are: the use for non-stationary data (i.e. we do not have to solve the problem of stationarity and possible autocorrelation in data), the ability to uncover unique complicated patterns over time and a good time resolution [20, 36, 39, 81, 110, 118]. The wavelets give a possibility to estimate and describe the cyclical structure as a function of time, and they show how the different components of a particular time series evolve [20]. We may say that the increase of wavelet use among economists was brought by the research in the optimum currency area (OCA), by the analysis of the business cycle synchronization and by the the globalization of economies [37, 93, 118, 119]. TF methods can also be used for thresholding or shrinkage, which are popular as instruments for the image and signal de-noising [115, 120]. A different approach common in engineering uses neural networks [121] or a combination of methods to improve the performance of signals [122, 123].

The use of hard, soft and adaptive threshoding applications in engineering great popularity of wavelets among economists motivates us to improve our earlier study [84]. This previous research employed the hard threshold for the co-movement-selective filter which was applied to filtering out symmetric macroeconomic shocks from individual time series. There, the hard thresholding was based on the analyst's experience, while the current study proposes a more sophisticated approach based on statistical testing. We apply the approach according to Torrence and Compo (TC98) [23] who were the first to propose algorithms for significance testing of power wavelet spectrum, the cross-spectrum and the linear coherence. An improvement of their work was provided by Ge [25, 55]. The algorithm presents several assumptions to test the significance of the power wavelet cross-spectrum. That is, inputs are two independent GWN and thus the power wavelet co-spectrum is the product of two

χ^2-distributed random variables. Further, using the Bessel function, we can test whether the power wavelet cross-spectrum coefficients are significant with respect to the variance of each time series. Notice that the authors of the mentioned publications work with a constant variance of input time series. An alternative approach for the testing is using the simulations of background noise levels as proposed in [17,26].

11.1.1 Chapter Contribution and Organization

We have determined that the literature does not deal with the following: firstly, how the TF transform of the time series with a transient character and changing volatility should be tested for an objective identification of its significant parts; secondly, how we can adjust time series about transient parts; thirdly, how we can identify the time domain representation (without a loss of data) of time series co-movement (the loss of data is a typical problem of e.g. correlation).

On the basis of the TF analysis and standard testing according to TC98 [23] (for the detail description of this approach, see Chapter 7) we propose a novel approach useful for applications:

• a method for mask construction based on the significance testing for co-movement selective detection filtering. Considering the character of data, we investigate two approaches for the mask construction: i) via hard thresholding based on the χ^2 testing according to TC98 [23]; ii) via adaptive thresholding considering the data character.

• an approach for identification of two-types of time series: either the time series adjusted about the co-moved part, or the time series containing just the co-moved part. Both series can be obtained after the co-movement selective filtering and the inverse transform in the time domain.

• an investigation of proposed approaches on simulated and real data.

• a recommendation for the use of proposed approaches.

To demonstrate the proposed methodology, we use an industrial production index of the the Euro Area (EA) and selected G8 countries, i.e. US, EA, UK, Japan and Russia. Based on the existence of globalization of economies in the world, we expect the existence of the co-movement. All these countries somehow reacted to the US sub-prime mortgage crisis in the country cyclical movements of economic activity, thus we can observe fluctuations in the country BC, in investment activities or consumer spending etc. Therefore, the proposed methodology is useful for economists, especially policy makers, to analyse cyclical movements and the synchronization as well as dissimilarities of macroeconomic indicators. The contribution of the method is illustrated by removing the symmetric shock from the macroeconomic time series in the TF domain with the US as the reference time series.

11.2 The Wavelet Transform, Co-movement Measures and its Testing

The algorithm for the co-movement selective filtering presented here is designed for the WT because of its advantages, especially very good time resolution and usability for non-stationary time series [20, 110, 124], and for the power wavelet co-spectrum, as the co-movement measure [23, 31]. It can be modified for the Short Term Fourier transform and for different co-movement measures such as coherence [23, 25, 55]. As these methods are well known, we do not describe them in this paper. Instead, we focus on the description of testing approaches for the co-movement measure.

11.3 An Algorithm for the Co-movement Selective Detection Filter

An algorithm for the co-movement selective detection filter is based on two processes, i.e. transformation plus analysis and reconstruction. The transformation process consists of the TF modelling (WS and PWCS analyses) and masking. The reconstruction process inversely transforms time series from the TF to the time domain. Figure 11.1 proposes a block diagram describing this algorithm.

11.3.1 An Algorithm for the Co-movement Selective Detection Filter

The proposed method for the identification of the time series co-movement indicator or for filtering out a symmetric behaviour is designed as follows:

1. **Time-frequency transform**
   Transform time series $x(t)$ and $y(t)$ using CWT resulting in wavelet spectrograms $S_x(a,\tau), S_y(a,\tau)$ (eq. (3.6)) and wavelet cross-spectrum $S_{xy}(a,\tau)$ (eq. (3.18)) respectively. Alternatively, another co-movement measure as coherence could be used.
2. **Decision about type of thresholding**
   Decide the method for identifying the threshold of wavelet cross-spectrum coefficients WCS and find the threshold thr, i.e. decide for the method of the mask M design via standard or adaptive thresholding (see Sec. 11.4).
3. **Mask design**
   Divide the wavelet cross-spectrum in power form, i.e. PWCS $|S_{xy}(a,\tau)|^2$, into regions with significant and insignificant co-movement of $x(t)$ and $y(t)$ based

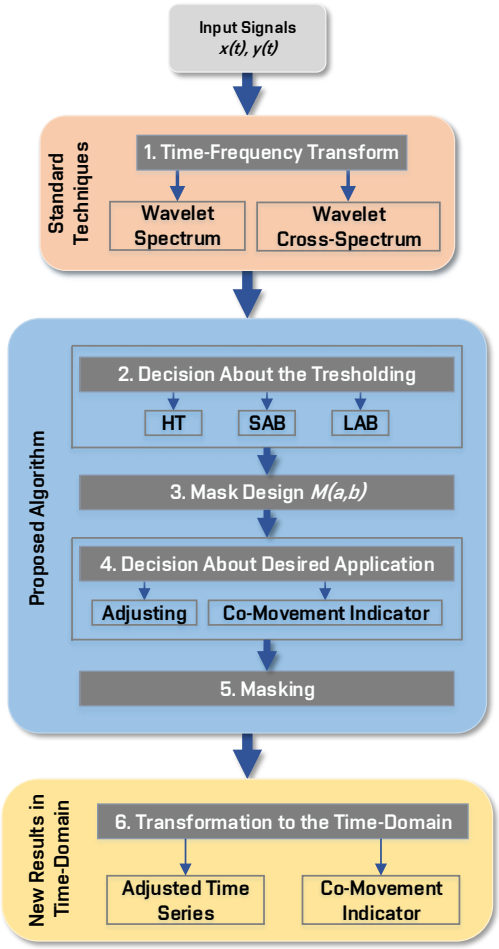


Fig. 11.1: Block diagram of co-movement detection filter algorithm.We use the following abbreviations: WS denotes the wavelet spectrogram coeffcient (Sec. 3.6), PWCS denotes the power wavelet cross-spectrum (Sec. 3.18), HT denotes Hard-Threshold Masking (Sec. 11.4.2), SAB denotes Segmentation-Adaptive-Based Masking (Sec. 11.4.3) and LAB denotes Local-Adaptive-Based Masking (Sec. 11.4.3).

on the threshold thr. That is, we construct the mask M:

$$M(a,\tau) = \begin{cases} 1 \text{ (significant co-movement)}, & |S_{xy}(a,\tau)|^2 \geq thr \\ 0 \text{ (insignificant co-movement)}, & |S_{xy}(a,\tau)|^2 < thr \end{cases} \tag{11.1}$$

The term "significant co-movement" denotes statistically significant PWCS values which will be identified via statistical testing. See Sec. 11.4 for a detailed description of this testing.

The PWCS coefficients are used due to the complex valued cross-spectrum for the majority of practically used mother wavelet functions.

4. **Decision about desired application**

   Decide what the desired application is. We investigate two cases: i) if the

application is to remove the co-moved part from the time series, or ii) the inverse task, i.e. keeping the co-moved part of the time series and remove the part with distinct cyclical behaviour from the time series.

5. **Masking**
   i) Create a modified wavelet spectrogram (MWS) by masking i.e. co-movement selective detection filtering (adjusted time-frequency transform)

$$MWS_x(a,\tau) = (1 - M(a,\tau)) * S_x(a,\tau).$$

   Analogously for the time series $y(t)$.
   ii) Create a modified PWCS (MPWCS) by masking, i.e. selecting the distinct cyclical behaviour leading to the time-frequency transform of a co-movement indicator

$$MPWCS(a,\tau) = |S_{xy}(a,\tau)|^2 * M(a,\tau).$$

6. **Transforming to the time domain**
   Transform inversely product of masking from the previous step via inverse continuous wavelet transform (ICWT, eq. (3.10)). We can obtain: i) adjusted time series in the time domain)

$$\widetilde{x}_{\text{adj}}(t) = ICWT\{MWS_x(a,\tau)\}.$$

   Analogously for the time series $y(t)$, we can get $\widetilde{y}_{\text{adj}}(t)$.
   ii) the co-movement indicator represented in the time domain

$$\widetilde{x}_c(t) = ICWT\{MPWCS(a,\tau)\}.$$

   That is, we can construct the time representation of the co-movement indicator. Or, we inversely transform the pre-defined frequency region (e.g. a sub-part of MPWCS in BC frequencies) of this product into the time domain to construct the time representation of the co-movement sub-indicator corresponding to the pre-defined frequency region.

11.4 Mask Design

In this section we firstly discuss possible approaches for thresholding and and suggest when the method is suitable for use. Furthermore we focus on the description of the mask corresponding to the thresholding method, i.e. Step 5 in the algorithm described in Chap. 11.3.

11.4.1 Mask Design

Let us consider the PWCS coefficients for the time series $x(t)$ and $y(t)$. Based on the TC98 significance testing of PWCS we are going to design the mask. We follow two basic approaches. One is based on the hard threshold given by the testing via STA leading to the so called hard-threshold (HT) masking. The other is based on the adaptive threshold identification leading to two possibilities, i.e. local-adaptive-based (LAB) threshold and segmentation-adaptive-based (SAB) threshold. This adaptivity is in time. In both adaptive cases, we propose an improvement of STA in the adaptive form.

If an analyst focuses on the time series adjustment about the co-moved part, or on the construction of the co-movement indicator with respect to the full time range in order to identify the most important events in the time series, we recommend the use of HT masking, i.e. STA.

The idea of SAB and LAB testing considers the situation when the variance of the time series x or/and y in the TD may vary for some sub-period, even for a short duration. In this case, the adaptive masking may be more suitable, because there may exist events (such as the financial crisis in 2008) having a higher level of amplitude in the co-spectrum, which may suppress the significance of other events. These events can be usually visible in the time representation of the data (structural breaks, outlier or cause changes in the volatility of the data).

11.4.2 Hard-Threshold (HT) Masking

The HT masking is based on STA significant testing which is follows the work of Torrence and Compo ([23], pp. 69 and 76) for the special case if the background spectra is Gaussian white noise and PWCS is the S_2 distribution (see also Ge [25, 55, 56]). As described in Chapter 7, the significance level $Z(1-\alpha)$ for the risk α can be deduced from $1-\alpha$ percentile of the S_2 distribution [23, 25].

Thus, the mask M (see eq. (11.1)) for the HT approach is given by

$$M(a,b) = \begin{cases} 1 & |S_{xy}(a,\tau)|^2 \geq thr \\ \\ 0 & |S_{xy}(a,\tau)|^2 < thr \end{cases} \tag{11.2}$$

where the threshold thr (given by STA according to TC98)

$$thr = \frac{1}{4}\sigma_x^2\sigma_y^2 Z(1-\alpha)$$

is a fix scalar number for the risk α for all PWCS coefficients. The value of $Z(1-\alpha)$ is calculated by STA (eq. (7.6), (7.5)).

11.4.3 Adaptive-Based Masking

Segmentation-Adaptive-Based (SAB) Masking

The SAB testing is suitable if we are able to identify the sub-segments with different volatility in the data. Firstly, for each time series $x(t)$ and $y(t)$, we identify the moments of the change of the time series variance. It can be done by expert estimate (usually in the case if the data are filtered by the long-term component and take the form of fluctuation around x-axis) or by statistical testing [125]. Secondly, we arrange all moments for both time series in the ascending time-order and we split the time range into the segments (SG) reflecting volatility changes in $x(t)$ and $y(t)$. Consequently, we identify the critical value for the significance testing in each segment by STA.

The proposed SAB masking method is designed as follows:

1. Identify the moments of the variance change of the time series $x(t), y(t)$ via expert estimate or statistical testing.
2. Arrange all identified moments for both time series in the ascending order and split the time range $t = 1, \ldots, n$ into the segments $SG_j, j = 1, \ldots J$ reflecting volatility changes in $x(t)$ and $y(t)$.
3. Estimate the PWCS and split it into the segments $PWCS_j, j = 1, \ldots J$ according to segments SG_j.
4. Construct the segments of the mask $M_j(a, \tau)$ corresponding to the segment SG_j, i.e. calculate $M_j(a, \tau)$ in each segment SG_j according to eq. (11.2) with respect to the variances of the time series in $j-$th segment. That is,

$$M_j(a,\tau) = \begin{cases} 1 & |S_{xy,j}(a,\tau)|^2 \geq thr_j \\ \\ 0 & |S_{xy,j}(a,\tau)|^2 < thr_j \end{cases} \tag{11.3}$$

   Here, in the SG_j segment, the threshold

$$thr_j = \sigma^2_{x,j}\sigma^2_{y,j} 0.25 Z(1-\alpha)$$

   is a fixed scalar number, α is a risk, $\sigma^2_{x,j}\sigma^2_{y,j}$ are variances for time series x, y in the time segment SG_j and $|S_{xy,j}(a,\tau)|^2$ is the corresponding part of PWCS. The value of $Z(1-\alpha)$ is calculated by STA test, i.e. eq. (7.6) and (7.5). Compared to the HT masking, in the case of SAB masking the threshold is the vector $thr = (thr_1, \ldots thr_J)$ adaptively changing with respect to the variances in segment.
5. Construct the mask $M(a, b)$ as the composition of the $M_j(a, \tau), j = 1, \ldots, J$ mask segments, i.e.

$$M(a,\tau) = (M_1(a,\tau), \ldots M_J(a,\tau)) \tag{11.4}$$

where the $j-$th $(j = 1, \ldots, J)$ segment of the mask $M(a, \tau)$ corresponding to the time segment SG_j as described in eq. (11.3).

Local-Adaptive-Based (LAB) Masking

The LAB testing is suitable if the variability of the data slowly increases or/and decreases, once or several times during the time range of the series x or/and y. Before starting the LAB algorithm, we have to set the value of l - the length of a sliding window. Consequently, we can identify the critical value for a significance testing in each segment by STA.

The proposed SAB masking method is designed as follows:

1. Select the length l of the sliding window as an odd number.
2. Estimate the PWCS for the time series $x(t), y(t)$.
3. Calculate the mask $M_t(a, \tau)$ in each time $t = 1, \ldots, n$ according to (11.5) with the variances as described in (11.6), i.e.:

$$M_t(a,\tau) = \begin{cases} 1 & |S_{xy,t}(a,\tau)|^2 \geq thr_t \\ \\ 0 & |S_{xy,t}(a,\tau)|^2 < thr_t \end{cases} \tag{11.5}$$

and the threshold

$$thr_t = \sigma^2_{x,t}\sigma^2_{y,t}0.25Z(1-\alpha)$$

is the fixed scalar number in the $t-$th sliding window, α is a risk. The value of $Z(1-\alpha)$ is calculated by STA test, i.e. (7.6), (7.5). The variance $\sigma^2_{x,t}$ is calculated as follows

$$\sigma^2_{x,t} = \begin{cases} \frac{1}{l-1}\sum_{i=1}^{l}\left(x(i)-\bar{x}\right)^2 & t \in 1, \ldots (l-1)/2 \\ \\ & t \in n-(l-1)/2+1, \ldots, n \\ \frac{1}{l-1}\sum_{i=t-(l-1)/2}^{t+(l-1)/2}\left(x(i)-\bar{x}\right)^2 & t \in (l-1)/2+1, \ldots, n-(l-1)/2 \end{cases} \tag{11.6}$$

where l is the odd number representing the sliding window length, $\bar{x}$ is the mean value of the time series x in the sliding window. The variance $\sigma^2_{y,t}$ is calculated accordingly. Compared to the HT masking, in the case of LAB masking the threshold is the vector $thr = (thr_1, \ldots thr_n)$ adaptively changing with respect to the variances in the sliding window.

4. Construct the mask $M(a, \tau)$ as the composition of the $M_t(a, \tau), t = 1, \ldots, n$, i.e.

$$M(a,\tau) = (M_1(a,\tau), \ldots, M_n(a,\tau)) \tag{11.7}$$

where the $t-$th $(t = 1, \ldots, n)$ part of the mask $M(a, \tau)$ is calculated according to eq. (11.6).

11.5 Experimental Results

To demonstrate the methods described above we present them both on simulated signals and on real economic data. In practice, across various non-technical disciplines, there are signals or time series for which the exact description of its character is not as clear as in technical signals [110]. While in the case of engineering, signals can be simulated as the simplification of the composition of several harmonic component, in the case of economic data, their structure is more complicated. Usually, it contains structural trend-breaks, outliers, cyclical components of close frequencies which can occur or diminish in different time sub-periods (not during the whole time), or nested cycles with different frequency limited in time [36–39]. Moreover, the nature of economic indicators play an important role and can influence the character of the frequency structure, e.g. business cycles, financial cycles etc. Then, it is quite difficult to simulate the universal behaviour of the economic series and its noising with a generalized artificial signal.

Therefore, we decided to model an artificial signal as a simplification of basic features in economic data. That is, a composition of signals which have co-movement in time-limited long-term trend (low frequency component), short time-limited middle-term co-movement with the high amplitude, middle-term co-movement during the whole time period (i.e. cyclical fluctuations in BC frequencies) and short-term trend (cyclical fluctuations of high frequency, such as seasonality) in the first half of time-period.

The quality of the identification of significant co-movement part in co-spectra (i.e. TF components) is evaluated via two metrics [126]. The first one evaluates how many TF components were significant and were not identified as significant by the test, i.e. relevant parts were not identified:

$$\mathrm{M1} = \frac{FN}{TP}. \tag{11.8}$$

The second metrics evaluates how many TF components were not significant and were identified as significant by test, i.e. irrelevant parts were identified:

$$\mathrm{M2} = \frac{FP}{TN}. \tag{11.9}$$

Here, TP (True Positives) is the number of correctly identified TF components in co-spectra; FN is the number of TF components in co-spectra which were significant but were not identified as significant; FP (False Positives) is the number of TF components in co-spectra which were insignificant but were identified as significant; and TN is the number of TF components in co-spectra which were insignificant and were identified as insignificant.

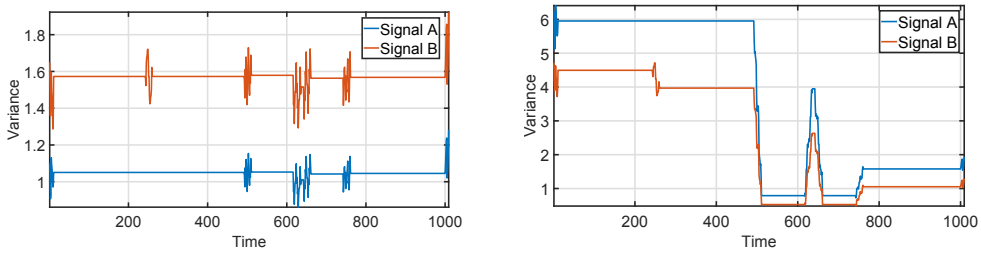
11.5.1 Simulated Data Description

For testing purposes, we have created four artificial signals, each of the length 1000 samples. Figure 11.2a) illustrates their time domain representation with constant variance, Figure 11.2b) with segmented variance (changing volatility), each for two signals. All signals were noised with signal-to-noise (SNR) ratio $SNR = 10, 3.16, 2$. In the simulation, we tried to approach the behavior of economic time series in the field of business cycle and synchrony analysis, therefore we selected $SNR = 10$ and 3.16 as stated above.

To be in correspondence with the real data analyses, we use the following settings during the analyses of artificial signals. For the PWCS calculated according to the eq. (3.19), we set the scales in the range 1–10 years divided into 257 individual scales. Further, we use the complex Morlet with the center frequency $f_{\mathrm{b}} = 1.5$ as a mother wavelet. The LAB testing is done according to Sec. 11.4.3. The sliding window length used in the eq. (11.6) is the same for both signals (A and B) and is set to $l = 36$ samples, which corresponds to 3 years. The SAB testing is done according to Sec. 11.4.3. As the scope of this paper is not to investigate the optimal method for variance segmentation, the number of SG_j segments is set to match the number of segments in the artificial signals. Thus, $j = 5$ for the signal A and $j = 6$ for the signal B.

To be able to quantify the accuracy of the proposed SAB and LAB methods, we have created the so called benchmark figure of ideal PWCS $\pm$ *spread* for co-spectral components (see Fig. 11.3a. This was done to include energy spread in frequency for each individual wavelet. The size of the frequency spread is set as $\pm 15\%$ from the maximum in the center frequency of each individual wavelet. As the result, we can see a wider co-spectral components represented as yellow blocks in the figures (e.g. Fig. 11.3a). In the next step, we calculate the metrics M1 and M2 using the benchmark representation and the masked PWCS. Further, we noise all signals as mentioned in the first paragraph of this subsection and then estimate PWCS of signals with constant variance and PWCS of signals with segmented variance.

In the following figures (Figs. 11.3–11.6), the x-axis represents time, the y-axis represents specific periods (cycles in years) and the z-axis represents the values of spectrogram. The figures show a two-dimensional projection of three-dimensional charts. The intensity of each contour represents the relative importance of the different periodicities and time, i.e. from dark blue (low amplitude) to yellow (high amplitude) colour. The yellow curve denotes the mask in all figures.



(a) Signals with constant variance.

(b) Signals with changing (segmented) variance.

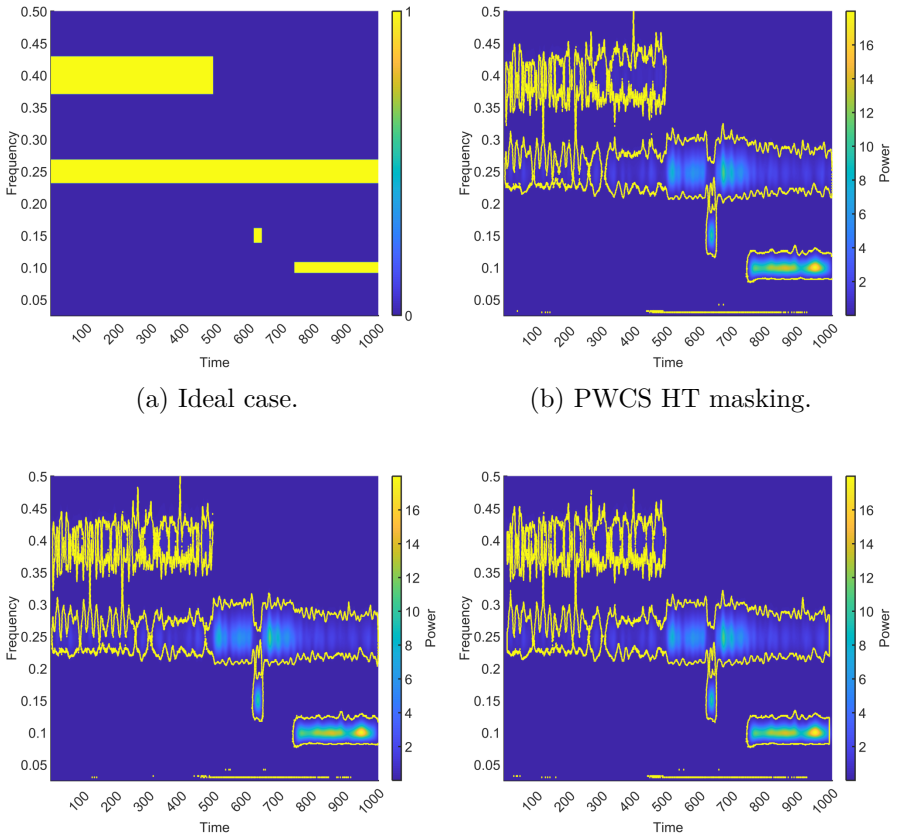
Fig. 11.2: Behaviour of a variance of simulated signals in the time.

SNR = 10						
	FN	**M1** [%]	**Δ M1** [%]	**FP**	**M2** [%]	**Δ M2** [%]
HT	$6.44 \cdot 10^3$	12.62	–	$5.83 \cdot 10^4$	13.58	–
SAB	$6.74 \cdot 10^3$	13.31	−0.60	$5.76 \cdot 10^4$	13.42	0.16
LAB	$7.12 \cdot 10^3$	13.95	−1.34	$5.46 \cdot 10^4$	12.72	0.86
SNR = 3.16						
	FN	**M1** [%]	**Δ M1** [%]	**FP**	**M2** [%]	**Δ M2** [%]
HT	$1.25 \cdot 10^4$	32.66	–	$5.25 \cdot 10^4$	13.94	–
SAB	$1.27 \cdot 10^4$	33.10	−0.43	$5.21 \cdot 10^4$	13.80	0.14
LAB	$1.28 \cdot 10^4$	33.60	−0.94	$5.16 \cdot 10^4$	13.66	0.28
SNR = 2						
	FN	**M1** [%]	**Δ M1** [%]	**FP**	**M2** [%]	**Δ M2** [%]
HT	$1.89 \cdot 10^4$	59.13	–	$4.65 \cdot 10^4$	12.15	–
SAB	$1.89 \cdot 10^4$	58.81	0.32	$4.65 \cdot 10^4$	12.14	0.01
LAB	$1.89 \cdot 10^4$	58.93	0.19	$4.61 \cdot 10^4$	12.02	0.13

Tab. 11.1: Metrics for constant variance- averages of 1000 simulations.

11.5.2 Results for Constant Variance

Next, we identify the significant co-movement of PWCS via HT, LAB, SAB masking for $SNR = 10$ (Figs. 11.3b,c,d) and for $SNR = 3.16$ (Figs. 11.4b,c,d). Then we calculate the metrics $FN, FP, M1, M2, \Delta M1, \Delta M2$ (Tab. 11.1). The metrics $\Delta M1, \Delta M2$ describe how the metrics changed for SAB, LAB with respect to HT. As we can see in Figs. 11.3, 11.4 and Tab. 11.1, there are no big differences between the results for HT, SAB and LAB masking when the variance of signals is constant. That is, the metrics $\Delta M1, \Delta M2$ are mostly lower than 1%; in the case of LAB ($SNR = 10$) the metric $\Delta M1 = -1.34$ is a little higher than 1%.



(a) Ideal case.

(b) PWCS HT masking.

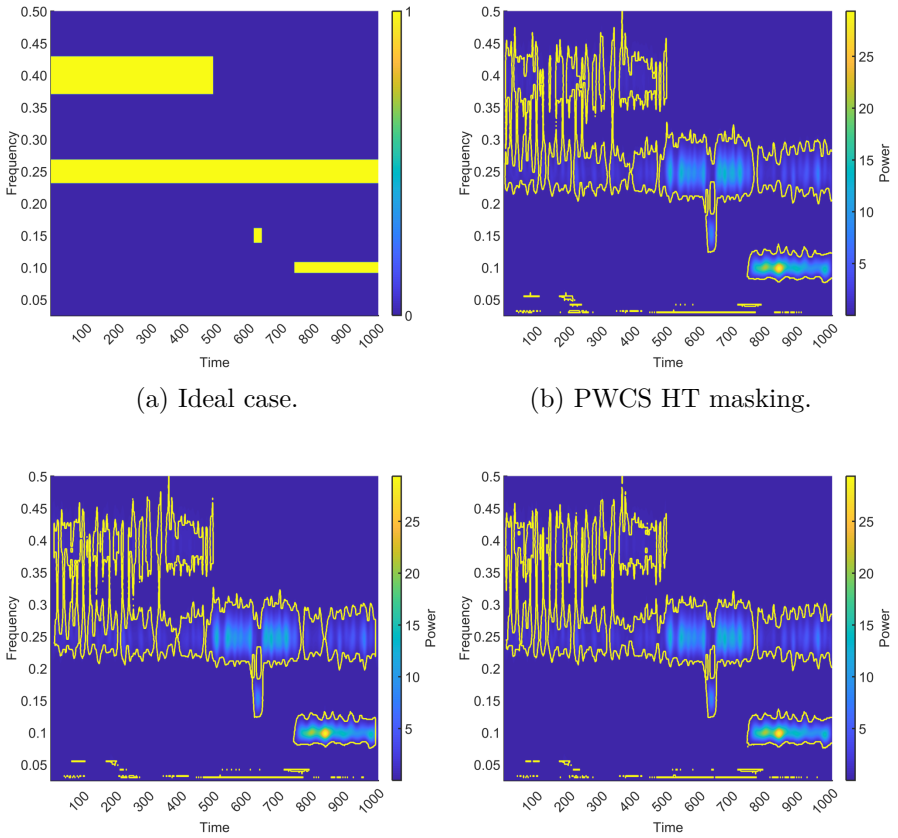
(c) PWCS SAB masking.

(d) PWCS LAB masking.

Fig. 11.3: PWCS and its estimate for constant variance and $SNR = 10$.

11.5.3 Results for Segmented Variance

As for the signals with segmented variance, we also identify the significant co-movement of estimated PWCS via HT, LAB, SAB masking for $SNR = 10$ (Figs. 11.5b,c,d) and for $SNR = 3.16$ (Figs. 11.6b,c,d). Then we calculate the metrics $FN, FP, M1, M2, \Delta M1, \Delta M2$ (Tab. 11.2). Comparing Figs. 11.5b–d with Fig. 11.5a we can see that HT masking of PWCS was not able to identify well the frequency component corresponding to frequency 0.25 in the second half of the time. Moreover, in the case of 0.40, the frequency component PWCS HT masking covered a wider range of surrounding components than SAB, LAB masking. This fact is also documented in Tab. 11.2. The metrics $\Delta M1, \Delta M2$ describe how the metrics changed for SAB, LAB with respect to the HT. As we can see, there are differences between the results for HT, SAB and LAB masking for segmented variance compared to the constant variance. That is, for the $SNR = 10$, the metric $\Delta M1$ is mostly 11.5% higher; the metric $\Delta M2$ is mostly 3% higher. For the $SNR = 3.16$,



(a) Ideal case.

(b) PWCS HT masking.

(c) PWCS SAB masking.

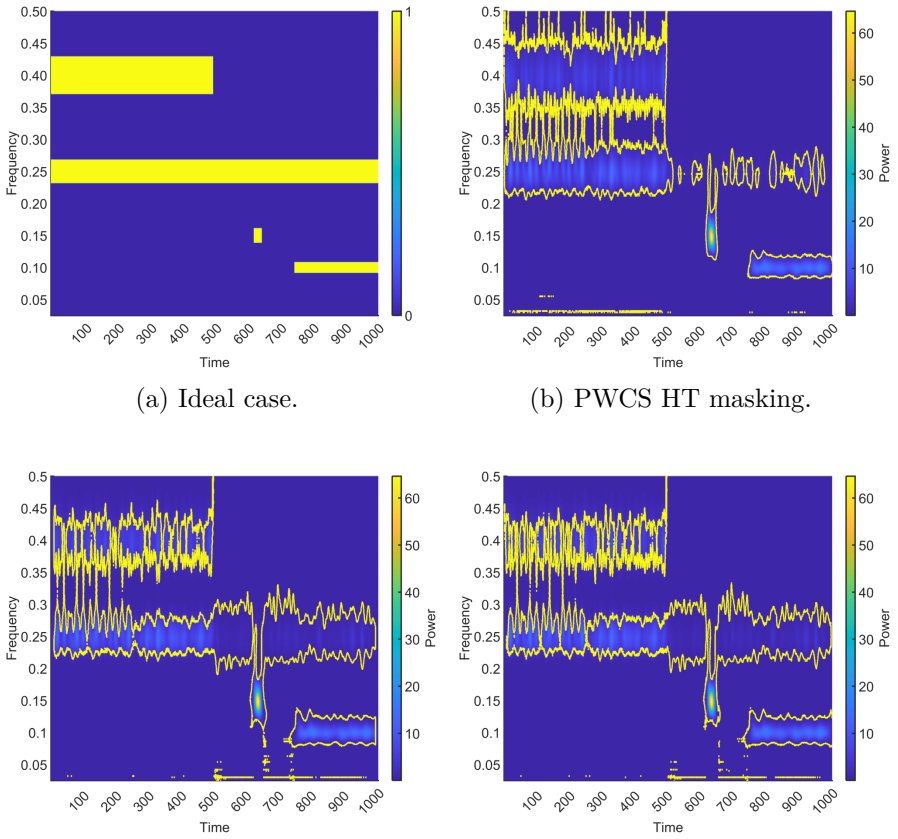
(d) PWCS LAB masking.

Fig. 11.4: PWCS and its estimate for constant variance and $SNR = 3.16$.

the metric $\Delta M1$ is mostly 0.5-1% higher; the metric $\Delta M2$ is again mostly 3% higher. We can see that the metric $\Delta M1$ is more sensitive to the noise level. An additional simulation for SNR=2 confirmed the sensitivity of $\Delta M1$ and the slow decrease of $\Delta M2$. That is, the growing noise level causes the increase of false negative components and thus the decrease of the level of improvements measured by $\Delta M1$ in HT, SAB and LAB. The level of $\Delta M2$ keeps a roughly the same level, which means that the level of spurious significance given by HT is corrected. Thus we can conclude that in the case of segmented variance, HT masking can produce worse results and should be replaced by adaptive masking (SAB, LAB). The graphical comparison is visualized in Figs. 11.5 and 11.6.

11.5.4 Summary of the Results of Simulations

To summarize the results from the simulation, we can give the following recommendation. Before the significance testing of co-movement measure, an analyst should



(a) Ideal case.

(b) PWCS HT masking.

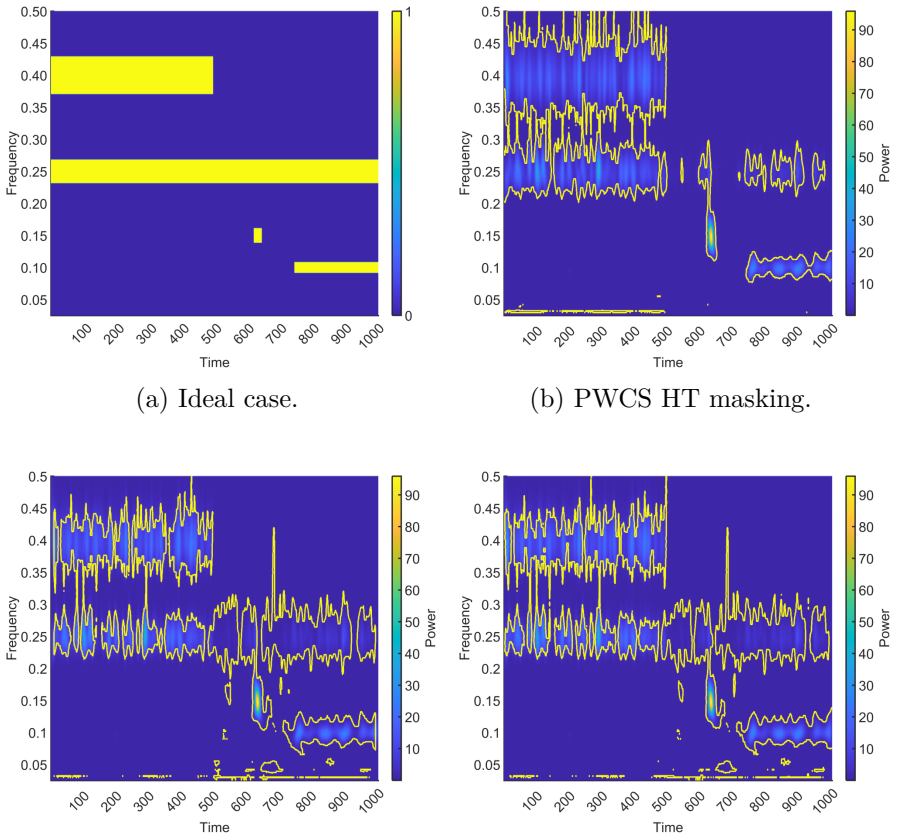
(c) PWCS SAB masking.

(d) PWCS LAB masking.

Fig. 11.5: PWCS and its estimate for changing variance and $SNR = 10$.

SNR = 10						
	FN	**M1** [%]	**Δ M1** [%]	**FP**	**M2** [%]	**Δ M2** [%]
HT	$9.01 \cdot 10^3$	21.75	–	$6.49 \cdot 10^4$	17.80	–
SAB	$4.73 \cdot 10^3$	10.24	11.50	$5.53 \cdot 10^4$	14.79	3.00
LAB	$4.70 \cdot 10^3$	10.15	11.60	$5.39 \cdot 10^4$	14.36	3.44
SNR = 3.16						
	FN	**M1** [%]	**Δ M1** [%]	**FP**	**M2** [%]	**Δ M2** [%]
HT	$1.07 \cdot 10^4$	26.60	–	$5.77 \cdot 10^4$	15.54	–
SAB	$1.58 \cdot 10^4$	26.23	0.37	$4.86 \cdot 10^4$	12.77	2.77
LAB	$1.50 \cdot 10^4$	25.85	0.75	$4.71 \cdot 10^4$	12.35	3.19
SNR = 2						
	FN	**M1** [%]	**Δ M1** [%]	**FP**	**M2** [%]	**Δ M2** [%]
HT	$1.49 \cdot 10^4$	41.50	–	$4.85 \cdot 10^4$	12.75	–
SAB	$1.85 \cdot 10^4$	57.09	−15.58	$4.33 \cdot 10^4$	11.23	1.52
LAB	$1.85 \cdot 10^4$	56.87	−15.37	$4.20 \cdot 10^4$	10.84	1.91

Tab. 11.2: Metrics for changing variance - averages of 1000 simulations.



(a) Ideal case.

(b) PWCS HT masking.

(c) PWCS SAB masking.

(d) PWCS LAB masking.

Fig. 11.6: PWCS and its estimate for changing variance and $SNR = 3.16$.

identify the behaviour of the time series volatility. If the time series have a constant variance, then HT masking is a plausible instrument. If one time series, or both, indicates a changing variance during the time (heteroscedasticity), then adaptive masking is a proper way how to obtain relevant information. This recommendation may be particularly useful for time series for which heteroscedasticity is expected, as in the case of economic time series.

11.6 Application of the Results

11.6.1 Real Data Description

To demonstrate the proposed methodology we use the seasonally adjusted monthly data of IPI from the OECD [40] database which are commonly used among economists for business cycle modelling as the macroeconomic indicator of country economy. With respect to the globalisation of economies we focused on the EA and selected

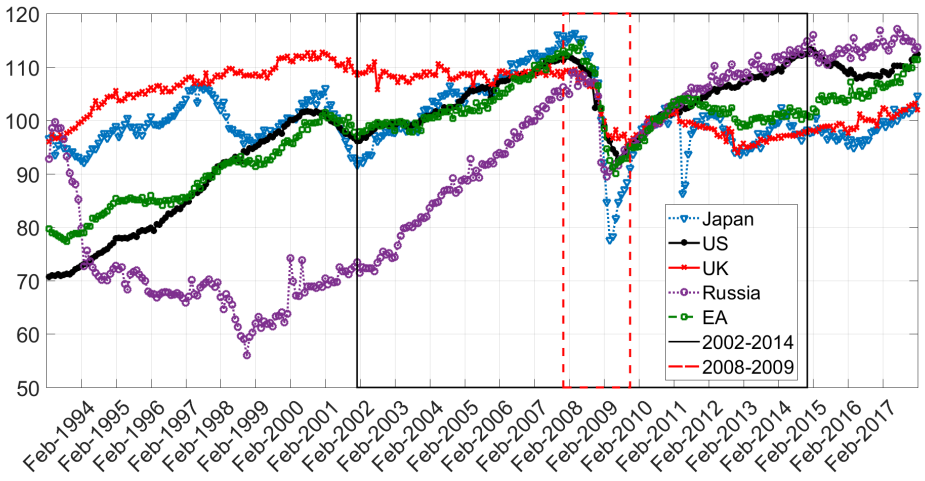


Fig. 11.7: Industrial production index of selected countries (in the levels).

G8 countries [40]: the US, Japan, Russia, and the UK. The sample period starts with July 1975 and ends in December 2017 for all countries except Russia. In the case of Russia, the available data are in the range from January 1993 to December 2017. This selection was motivated by the following facts: the US was, for a long time, the leading world economy causing the crisis in 2008; the EA is taken as a representative of 19 European economies; The UK was preselected because of Brexit; Japan is an Asia Pacific representative economy; and Russian is taken as an East European Asia Country. We examine TF selective filtering based on co-movement between the growth cycles of the US and the selected countries. The data were transformed to FODLOG values which represent the growth business cycles [127] of selected countries.

Figure 11.7 displayes these business cycles (data in levels) in the time domain. Further, the data are transformed into the growth business cycles (i.e. fluctuation around a potential product) and are used for the synchrony analysis via wavelets as usual by economists [20, 36, 39, 81, 118].

As a preliminary analysis, assuming the existence of synchrony between the US and selected countries, we calculate the correlation coefficients of business cycles of selected countries. The synchronization among countries during the economic crisis in 2008 is also illustrated in Fig. 11.7. Here we can see the tendency of the curves to converge especially arround 2008–2009 time window, which is caused by a structural symmetric economic shock, i.e. the global financial crisis. Comparing correlation coefficients in the overview given in Tab. 11.3, we can see the influence of the sample size on levels of the data and on the FOD transform (growth business cycle). We focused on the difference between the time period around the crisis, i.e. 2002–2014, and the available sample size. As we can see from the FOD transform,

US	Levels			FOD transform		
	1978-2017	1993-2017	2002-2014	1993-2017	2002-2014	2008-2009
Japan	0.7546***	0.3459***	0.6255***	0.1825***	0.1904**	0.3599**
UK	0.7605***	−0.0365	0.0607	0.1390**	0.1540*	0.2131
EA19	0.9721***	0.9312***	0.7268***	0.2963***	0.3732***	0.4090**
Russia		0.6483***	0.6669***	0.0781	0.1621**	0.

Note: statistically significant at: ***1%, **5%, *10%

Tab. 11.3: Correlation coefficients between US and selected countries.

the selected countries represent situations with a generally stable correlation during the crisis time in 12-year window (Japan, EA), with an increase (Russia) as well as a decrease (UK) in the correlation significance. The table also briefly compares the correlations for the data without any transform (in the levels).

11.6.2 Settings for Implementation

During the analyses, we use the following settings. For the PWCS calculation, we set wavelet scales corresponding to the range of 1 year to 10 years divided into 257 individual scales. We select the complex Morlet wavelet with the center frequency $f_b = 1.5$ as mother wavelet. For LAB testing we set a sliding window for 3 years, i.e. 36 samples, with 1 sample step ahead. The choice of Morlet wavelet was motivated by the fact that it is a widely used wavelet for the co-movement analysis by economists. For the SAB testing, we will describe the number of segments during the presentation of particular results. After masking the country wavelet spectrogram, we use its inverse transform to obtain filtered time series. In such a way, the inverse transform of the whole wavelet spectrogram works as a band-pass filter with respect to the scales setting.

11.6.3 Demonstration of the Proposed Mask Construction

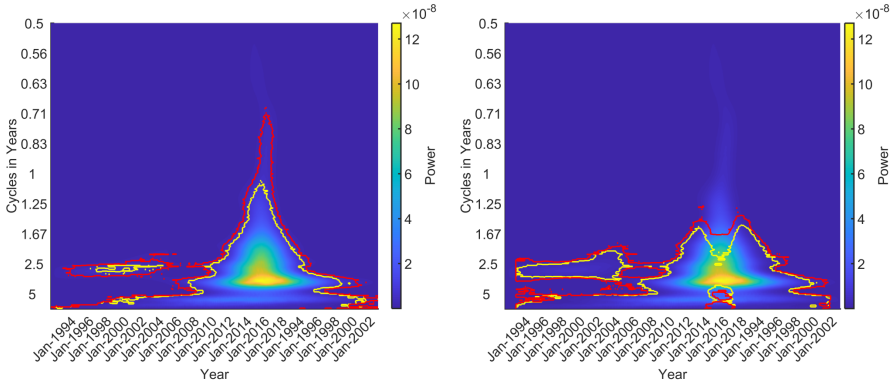
For the demonstration of the proposed approaches, we choose the US and EA countries. We concentrate on the removal of all co-movements with the US from the EA data. The TF representation re-calculated into the time-scale form is given in the PWCS figures (Fig. 11.8(a),(b),(c)). After the testing of the obtained PWCS via STA (Fig. 11.8a)), via SAB (Fig. 11.8b)) and LAB testing (Fig. 11.8c)) we construct the masks. Then, we partition the TF plane into two regions, with and without significant co-movement. In all figures, the border is highlighted as a yellow solid-line curve. After the EA spectrogram masking, we inversely transform wavelet coefficients which correspond to the part without a co-movement. The obtained time

series, for all three masks, are depicted in Fig. 11.9. The dotted black line is the original time series $\widetilde{x}$ (i.e. growth business cycle) obtained via eq. (3.10), the green line represents the adjustment via HT, the red line represents the adjustment via LAB, the blue line represents the adjustment via SAB. The same denotation is used in Figs. 11.10-11.12. To ensure a better visibility of the detected areas, we zoomed the figure to a shorter time range i.e. 2006–2012 (the shape of the curves are the same before 2006 and after 2012). In this way we obtain time series adjusted for significant co-movement parts with the reference (the US) country.

As presented in Fig. 11.8, we identify the mask via three approaches leading to three adjusted time series. The HT masking produce the mask covering cycles of the range 5–0.7 years. The most energy of co-spectrum is concentrated in the cycles 5–1.5 years; thus, the adjustment via the HT masking removes mainly long and business cycles, as well as part of short and very short cycles, from the original EA data. As a result, the EA's time series will reduce the fluctuations in the time around the crisis. In other words, due to the fact that the mutual movement of the countries manifested itself in many different periods then the removal of these components from the time series results in its greater smoothness with respect to the temporal localization of the co-movement. Next, the HT masking is constructed as a selection of a co-movement of countries relative to the full time range. Thus, if a significant event occurs, such as the 2008 crisis (which is reflected by a significantly higher amplitude in the spectral and co-spectral component) and if we evaluate co-movement in the whole time range, the significance of other spectral components will be significantly lower.

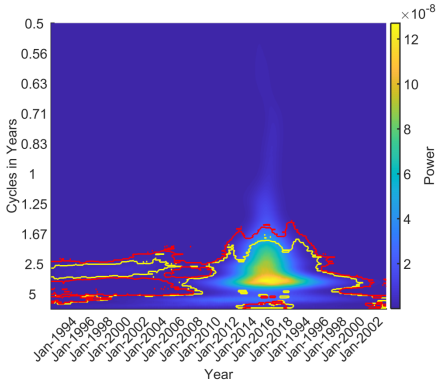
If we use the LAB or SAB masking, we evaluate the significance in the window that adaptively calculates the variance. Thus, it marks the components that contribute to the overall variance as significant. Because of this, some spectral components in other areas are not suppressed, and thus gain in significance. In the case of the US and EA, the LAB and SAB masking concentrates most of the PWCS energy into long cycles and shows that the most important co-movement is in long and business cycles, while the HT masking shows also very short cycles. As the result of the adaptability, the high-frequency components are not removed from the EA, which results in a lower volatility reduction in the crisis period.

To validate the results of SAB, LAB masking, we provide simulations of background noise levels (the red line in Fig. 11.8). We can see the proximity of SAB, LAB masking with simulations of background noise levels for presented volatile data. Thus, the difference in the significant region identified via STA testing (i.e. HT masking) and simulations of background noise levels confirmed the influence of the volatility on the testing. The detailed description of simulations of background noise levels can be found in Chapter 7.1.2 and [17, 26].



(a) PWCS of EA: HT masking.

(b) PWCS of EA: SAB masking.



(c) PWCS of EA: LAB masking.

Fig. 11.8: Power wavelet co-spectra. The yellow curve denotes the respective mask and the red curve denotes simulations of background noise levels.

If we assess the EA data, adjusted by the HT masking about the co-movement with the US, then we can see that the fluctuation in industrial production without a linking to the US is smaller. Conversely, the local effect of interdependence during the crisis period (using adaptive masking) results in a greater fall in the index value than in the long-term time horizon (using HT masking). Furthermore, the adjustment of the local co-movement in relation to the unadjusted data points to a larger drop in values, i.e. a larger structural break. Thus, the interdependence with the US economy led to a deepening of the structural breakdown of the crisis (a high correlation has given a reason to believe that there will be a significant reaction and deepening).

11.6.4 Application to Other Countries

In this section we present the results for Japan, the UK and Russia. The goal is the same as in the previous part, i.e. to filter out the common cycles of the country with respect to the US. We are going to present only the adjusted time series of the country indicator in the TD (the WCS and PWCS figures with three types of masking are available upon request).

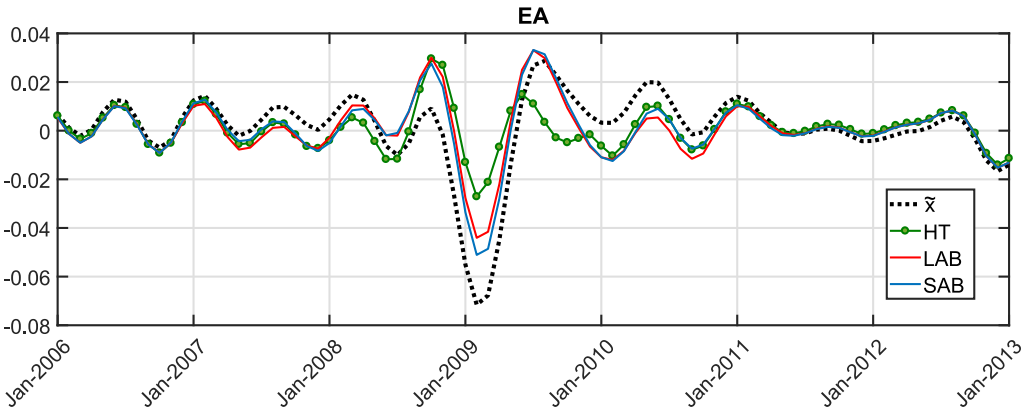


Fig. 11.9: IPI of EA adjusted of co-movement with US.

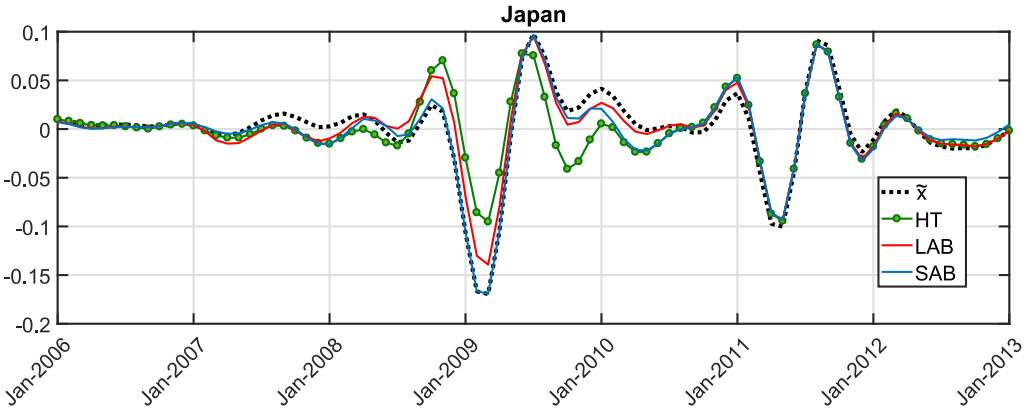


Fig. 11.10: IPI of Japan adjusted of co-movement with the US.

As for **Japan**, the adjustment of industrial production index Fig. 11.10 via the HT masking leads to filtering out the cycles of 1–5 years. As a result, fluctuations during the 2008 crisis will diminish in the time series. In contrast to the adaptive masking, we remove a larger spectrum of cyclical components (long, short and very short cycles), which leads to a greater smoothness of adjusted time series around the crisis time. As can be seen from the time series graphs (Fig. 11.10), there is another fall in mid-2010 after the structural break in 2008. However, the adjusted and original values are at the same level. Japan's adjustment with the US, in this structural break in terms of BC synchronism, seems to have no effect. The adaptive masking in the Japan case shows the most significant co-movement with the US to be in the BC (1.7–5 years), while the HT masking shows, in addition, co-movement in the area of very short cycles (1.5–0.7 years). As a result of adaptability, the high frequency components are not removed from Japan, and further, BC are adjusted in the pre-crisis period. This results in less volatility during the crisis.

If we assess the adjusted Japan data via the HT masking about the co-movement with the US, we can see that the fluctuation in industrial production without linking

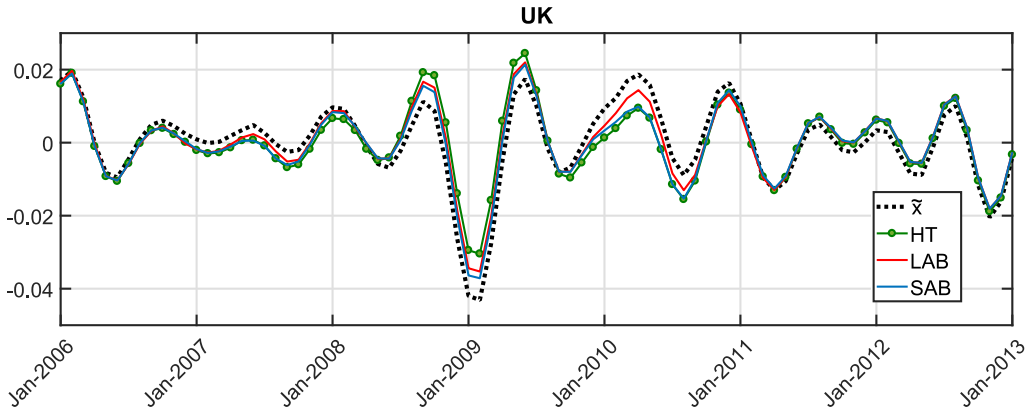


Fig. 11.11: IPI of UK adjusted of co-movemen with US.

to the US is smaller. On the contrary, the local effect of interdependence during the crisis period results in a greater fall in the index value compared to the HT masking. Furthermore, the adjusting of local co-movement in relation to the un-adjusted data leads to a larger structural break. Thus, the interdependence with the US economy leads to the deepening of the structural break. In the case of Japan, the amplitude of the drop in the crisis time is larger than in the case of the EA.

The adjustment of the **United Kingdom** industrial production index Fig.11.11 via the HT masking leads to filtering out the cycles of 2–5 years. As a result, the UK's time series will reduce fluctuations during the 2008 crisis. If we compare the adaptive masking with HT masking (Fig. 11.11), we can see a very similar mask. The PWCS testing shows that the most significant co-movement is in the BC in the range of 2–5 years. Contrary to Japan and the EA, in the UK PWCS there are not very short cycles. Therefore, the adjusted time series via all three approaches are very similar, with primary long- and medium-term cycles adjustments. We conclude that the co-movement of these economies is rather low, as evidenced by the small difference in the adjusted values of the time series. Thus, lower interdependence with the US economy did not cause such a deepening of the structural break as in the case of the EA and Japan.

The adjustment of the **Russia** index about the co-movement with the US (Fig. 11.12) via the HT masking caused filtering out the cycles of 1.5–5 years, mostly in the period around the crisis, i.e. 2006–2010. As a consequence, the fluctuation in the crisis time is reduced. In the case of adaptive masking, we can see the most significant co-movement in the BC, but spread over a wider time period than just around the crisis. Contrary to the HT masking, we see a larger amount of PWCS energy in the pre- (2001–2007) and the post- (2009–2014) crisis time. The crisis of 2008 does not occur in the case of adaptive masking for SAB; as for LAB, it can be seen only in 2–4 year cycles. As a result of the adaptability, a very small number of

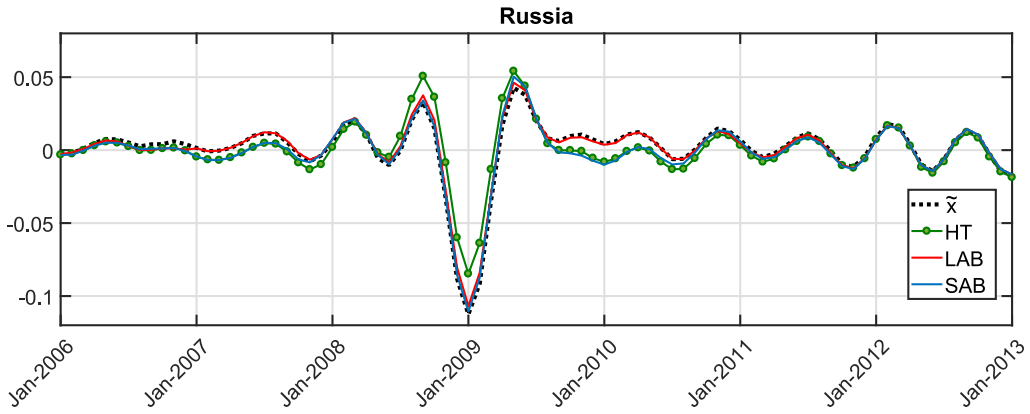


Fig. 11.12: IPI of Russia adjusted of co-movemen with US.

cyclical components are removed from the Russia data. Therefore, the adjusted and unadjusted values in 2008 are close to each other. On the other hand, in the LAB masking before and after the crisis period, we can see a smaller drop in values than the SAB and HT masking. Notice that even with the HT masking, the differences between the adjusted and unadjusted values are the smallest of all considered countries. Similarly, as in the case of the UK, we can see lower interdependence with the US economy, which probably did not cause such a deepening of the structural break as in the case of the EA and Japan.

In our work, we selected global economies which are synchronised primarily in global features and therefore indirect economic consequences (i.e. idiosyncratic shock) specific for the EA were not removed, because they were not synchronised with the US. From the co-movement analysis, we selected the financial crisis around 2008 as it is the strongest and most visible global shock which was transferred to the whole world. The adjustment showed that economies such es Euro Area and Japan are more synchronised with the US than with the United Kingdom and Russia. Furthermore, larger economies have larger idiosyncratic shocks which can be an important cause of macroeconomic volatility. To explain the above mentioned non-removal of certain symmetric economic shocks, our experience reveals ([84]) that the reason can be found in economic consequences of global shocks which can be transferred into various economic areas and aspects (foreign trade, investigation, banking sector, saving etc). Consequently, it can be reflected into a wide range of time series frequencies. As we found out in [84], for example in the case of Visegrad Countries, small open economies depend mainly on developments in Europe, i.e. there is a strong connection to some bigger economy (e.g. Germany), and thus the effect of removing co-movement part is most evident between larger economies such as the Euro Area countries and the US.

The information about the removal or non-removal of symmetric economic shocks is valuable for further economic research, so this question is worth investigating in a detailed economic context.

11.7 Chapter Conclusion

This chapter presents time-frequency selective filtering for the time series adjustment based on the time series co-movement measure. We propose a mask which can be used for selective filtering (adjustment) on statistical basis. The adjustment means removing common components from time series with respect to the reference time series. We investigate two approaches, via hard thresholding based on the χ^2 testing and via adaptive thresholding considering the data character. As the result of the co-movement selective filtering (which includes masking and inverse transform) we obtain two time series, i.e. the time series adjusted about the co-moved part and the time series containing just the co-moved part. The adjusted time series can be then used for consequent econometric analyses. The validation of the proposed method is done in MATLAB on the application of symmetric shock removal from selected G8 countries with the US as the reference country.

Considering the type of mask construction, our research leads to the following recommendations. If an analyst focuses on the time series adjustment about the co-moved part with respect to the full time range in order to identify the most important events in the time series, we recommend the use of hard-threshold masking, i.e. STA. If an analyst is interested in the adjustment of economic event with respect to its lead/lag influence, especially in the case when this event causes changes in the data volatility, then the adaptive masking (SAB or LAB) is a valuable instrument. The choice between the LAB and SAB approaches depends on the evolution of data volatility.

The presented approaches are able to provide an in-depth analysis of the time series. This can be done via adjustment for the significant symmetric shocks measured in the TF domain. In this way, we can investigate the global and regional country specific cyclical behaviour. It can be also done via investigating the time series representation of adjusted part as an inverse transform of significant co-movement regions leading to the construction of a co-movement indicator.

As can be seen in literature, a large number of economic studies uses TF domain, especially wavelets, for an individual time series analysis as well as for a co-movement analysis. Thus, the presented approach for the time-frequency selective filtering, or for the construction of the co-movement indicator, can be applicable and can reveal additional information about the investigated problem focusing on adjusted time series.

Conclusion

The doctoral thesis is focused on the current problems and shortcomings of time-frequency analysis and subsequent significance testing. The presented literature review shows current progress and gaps in this field. We found that the literature does not deal with the influence of the data character, i.e. data volatility, on the significance testing of the time-frequency methods. To include and encompass this required data character we selected three types of data. The technical data were chosen as a signal with known parameters, economic data as a signal where factors that are often unpredictable may affect or change the character of the data, and simulated data for verification. Using these types of data we focus on the issue of statistical testing of selected data in order to verify the standard methods and to propose methods for cases where the data volatility is changing over time.

Firstly, a brief methodological background of the most widely used TFA methods and co-movement measures is presented. Based on this, we analyze and assess selected methods and formulate a recommendation for working with such methods. We formulate recommendations for AR process optimization. We list the advantages and disadvantages of selected parametric and non-parametric TF methods taking into account data character.

Based on the knowledge mentioned above, the Objective I. is answered. We propose an approach for the enhancement of TF representation leading to background noise suppression. We denote this approach as "Enhanced TF representation". The core of this method is the combination/multiplication of several TF approaches. Thus, based on this, we can easily identify important areas in the TF representation. In specific cases, such as economic data the application of the designed methodology allows a more straightforward interpretation from time and frequency perspectives. Moreover, it can also be taken as a supplement to the significance testing or simulations of background noise levels.

Secondly, we propose the evaluation of standard significance tests on synthetic and real data. Here, we analyze the behavior of the testing procedure for the estimated spectrogram with respect to the GWN background noise. We use both the statistical significance with respect to the χ^2 distribution and the identification of critical values using simulations of background noise. We examine the advantages and disadvantages and formulate recommendations for its usage also with respect to the signal variance. By evaluating the influence of background noise on test accuracy, we find that the standard method may fail in some cases. In the case of an input signal with strong changes in volatility, such as selected economic data, the

total variance may not sufficiently describe the character of the data. Taking into account the gained knowledge, we further focus on economic data that have this character.

Findings mentioned above help to solve Objective II. We propose two modified methods of significance testing. We denote them as segmentation adaptive based (SAB) and local adaptive based (LAB) testing. Both of these methods take into account the possibility of changing the volatility of input data and adapting to it. The SAB method proposes segmentation of the data according to its levels of variance and thus providing better results when the changes in data variance have step character. The LAB method uses a sliding window and is, therefore, better when the variance change is gradual. We also confirm that in the case of different volatility levels in inputs, the significance testing needs a more careful interpretation of the results.

Thirdly the Objective III. is answered. We examine the possibility of using significance tests for subsequent data filtering. We use the statistically significant part of the power wavelet co-spectrum to construct a co-movement selective detection filter suitable for assessing the synchrony between two signals. We propose a mask construction that can be used for selective filtering, i.e. adjustment, on a statistical basis. The adjustment means removing common components from the time series with respect to the reference time series. We investigate approaches based on standard and newly proposed SAB and LAB testing. The advantage of the proposed co-movement selective detection filter is no loss of observations (such as correlation). Moreover, it is possible to construct sub-indicators that correspond to the predefined frequency range. In such a way, we can obtain a decomposition of the (total) co-movement indicator, which covers the full range of frequencies, into the required range.

Bibliography

[1] B. R. Gupta and V. Kumar, "Time-frequency analysis of asymmetric triaxial galaxy model including effect of spherical dark halo component," *International Journal of Astronomy and Astrophysics*, vol. 05, no. 02, pp. 106–115, 2015. [Online]. Available: https://doi.org/10.4236%2Fijaa.2015.52014

[2] K. Vida, K. Olah, and R. Szabo, "Looking for activity cycles in late-type kepler stars using time-frequency analysis," *Monthly Notices of the Royal Astronomical Society*, vol. 441, no. 3, pp. 2744–2753, may 2014. [Online]. Available: https://doi.org/10.1093%2Fmnras%2Fstu760

[3] P. Wang, J. Gao, and Z. Wang, "Time-frequency analysis of seismic data using synchrosqueezing transform," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 12, pp. 2042–2044, dec 2014. [Online]. Available: https://doi.org/10.1109%2Flgrs.2014.2317578

[4] H. Ding and B. F. Chao, "Detecting harmonic signals in a noisy time-series: the z-domain autoregressive (AR-z) spectrum," *Geophysical Journal International*, vol. 201, no. 3, pp. 1287–1296, apr 2015. [Online]. Available: https://doi.org/10.1093%2Fgji%2Fggv077

[5] S. Stanković, I. Orović, and V. Sucic, "Averaged multiple l-spectrogram for analysis of noisy nonstationary signals," *Signal Processing*, vol. 92, no. 12, pp. 3068 – 3074, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168412001892

[6] S. Liu, D. Wang, T. Li, G. Chen, Z. Li, and Q. Peng, "Analysis of photonic doppler velocimetry data based on the continuous wavelet transform," *Review of Scientific Instruments*, vol. 82, no. 2, p. 023103, feb 2011. [Online]. Available: https://doi.org/10.1063%2F1.3534011

[7] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, mar 2015. [Online]. Available: https://doi.org/10.1016%2Fj.seizure.2015.01.012

[8] M. R. Canal, "Comparison of wavelet and short time fourier transform methods in the analysis of EMG signals," *Journal of Medical Systems*, vol. 34, no. 1, pp. 91–94, oct 2010. [Online]. Available: https://doi.org/10.1007%2Fs10916-008-9219-8

[9] Y. Xu, S. Haykin, and R. Racine, "Multiple window time-frequency distribution and coherence of EEG using slepian sequences and hermite functions," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 7, pp. 861–866, jul 1999. [Online]. Available: https://doi.org/10.1109%2F10.771197

[10] J. G. Proakis, C. L. Nikias, C. M. Rader, F. Ling, M. Moonen, and I. K. Proudler, *Algorithms for Statistical Signal Processing*, 1st ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

[11] D. Walnut, *An Introduction to Wavelet Analysis*, ser. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, 2002. [Online]. Available: https://books.google.cz/books?id=atOStnsp9a0C

[12] H. Resnikoff and R. Wells, *Wavelet Analysis: The Scalable Structure of Information*. Springer New York, 2002. [Online]. Available: https://books.google.cz/books?id=poMzLbTn6JUC

[13] V. Sebesta, R. Marsalek, and J. Pomenkova, "The modified empirical mode decomposition method for analysing the cyclical behavior of time series," in *ECMS 2013 Proceedings edited by: Webjorn Rekdalsbakken, Robin T. Bye, Houxiang Zhang*. ECMS, may 2013. [Online]. Available: https://doi.org/10.7148%2F2013-0288

[14] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Birkhäuser Boston, 2001. [Online]. Available: https://doi.org/10.1007%2F978-1-4612-0003-1

[15] L. Huang, Q. Kemao, B. Pan, and A. K. Asundi, "Comparison of fourier transform, windowed fourier transform, and wavelet transform methods for phase extraction from a single fringe pattern in fringe projection profilometry," *Optics and Lasers in Engineering*, vol. 48, no. 2, pp. 141–148, feb 2010. [Online]. Available: https://doi.org/10.1016%2Fj.optlaseng.2009.04.003

[16] J. Zhong and Y. Huang, "Time-frequency representation based on an adaptive short-time fourier transform," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5118–5128, oct 2010. [Online]. Available: https://doi.org/10.1109%2Ftsp.2010.2053028

[17] X. Jiang and S. Mahadevan, "Wavelet spectrum analysis approach to model validation of dynamic systems," *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 575–590, feb 2011. [Online]. Available: https://doi.org/10.1016%2Fj.ymssp.2010.05.012

[18] D. Pollock, D. Pollock, R. Green, and T. Nguyen, *A Handbook of Time-series Analysis, Signal Processing and Dynamics*, ser. Signal processing and its applications. Academic, 1999. [Online]. Available: https://books.google.cz/books?id=sf5xEq1N8MwC

[19] F. Abramovich, Y. Benjamini, D. L. Donoho, and I. M. Johnstone, "Adapting to unknown sparsity by controlling the false discovery rate," *The Annals of Statistics*, vol. 34, no. 2, pp. 584–653, apr 2006. [Online]. Available: https://doi.org/10.1214%2F009053606000000074

[20] C. Aloui, B. Hkiri, and D. K. Nguyen, "Real growth co-movements and business cycle synchronization in the GCC countries: Evidence from time-frequency analysis," *Economic Modelling*, vol. 52, pp. 322–331, jan 2016. [Online]. Available: https://doi.org/10.1016%2Fj.econmod.2015.09.009

[21] J. A. Schulte, C. Duffy, and R. G. Najjar, "Geometric and topological approaches to significance testing in wavelet analysis," *Nonlinear Processes in Geophysics Discussions*, vol. 1, no. 2, pp. 1331–1363, aug 2014. [Online]. Available: https://doi.org/10.5194%2Fnpgd-1-1331-2014

[22] P. M. A. James, R. A. Fleming, and M.-J. Fortin, "Identifying significant scale-specific spatial boundaries using wavelets and null models: spruce budworm defoliation in ontario, canada as a case study," *Landscape Ecology*, vol. 25, no. 6, pp. 873–887, mar 2010. [Online]. Available: https://doi.org/10.1007%2Fs10980-010-9465-2

[23] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bulletin of the American Meteorological society*, vol. 79, no. 1, pp. 61–78, 1998.

[24] Z. Ge, "Significance tests for the wavelet power and the wavelet power spectrum," *Annales Geophysicae*, vol. 25, no. 11, pp. 2259–2269, nov 2007. [Online]. Available: https://doi.org/10.5194%2Fangeo-25-2259-2007

[25] Z. Ge, "Significance tests for the wavelet cross spectrum and wavelet linear coherence," *Annales Geophysicae*, vol. 26, no. 12, pp. 3819–3829, dec 2008. [Online]. Available: https://doi.org/10.5194%2Fangeo-26-3819-2008

[26] D. Wang, S. Sun, and P. W. Tse, "A general sequential monte carlo method based optimal wavelet filter: A bayesian approach for extracting bearing fault features," *Mechanical Systems and Signal Processing*, vol. 52-53, pp. 293–308, feb 2015. [Online]. Available: https://doi.org/10.1016%2Fj.ymssp.2014.07.005

[27] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. Elsevier Science, 2008. [Online]. Available: https://books.google.cz/books?id=5qzeLJljuLoC

[28] M. X. Cohen, "A better way to define and describe morlet wavelets for time-frequency analysis," *NeuroImage*, vol. 199, pp. 81–86, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811919304409

[29] *Time-Frequency Signal Analysis and Processing*. Elsevier, 2016. [Online]. Available: https://doi.org/10.1016%2Fc2012-0-00024-5

[30] H. Seddighi, *Introductory Econometrics: A Practical Approach*, ser. Economics, business, finance. Routledge, 2012. [Online]. Available: https://books.google.cz/books?id=tjJpuQAACAAJ

[31] C. Croux, M. Forni, and L. Reichlin, "A measure of comovement for economic variables: Theory and empirics," *Review of Economics and Statistics*, vol. 83, no. 2, pp. 232–241, may 2001. [Online]. Available: https://doi.org/10.1162%2F00346530151143770

[32] M. Wand and M. Jones, *Kernel Smoothing*. Chapman and Hall/CRC, dec 1994. [Online]. Available: https://doi.org/10.1201%2Fb14876

[33] I. Horová, "Some remarks on kernels," *Journal of Computational Analysis and Applications*, vol. 2, no. 3, pp. 253–263, 2000.

[34] I. Horová, P. Vieu, and J. Zelinka, "OPTIMAL CHOICE OF NONPARAMETRIC ESTIMATES OF a DENSITY AND OF ITS DERIVATIVES," *Statistics & Risk Modeling*, vol. 20, no. 1-4, jan 2002. [Online]. Available: https://doi.org/10.1524%2Fstrm.2002.20.14.355

[35] L. Trefethen and D. Bau, *Numerical Linear Algebra*, ser. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, 1997. [Online]. Available: https://books.google.cz/books?id=bj-Lu6zjWbEC

[36] A. N. Berdiev and C.-P. Chang, "Business cycle synchronization in asia-pacific: New evidence from wavelet analysis," *Journal of Asian Economics*, vol. 37, pp. 20–33, apr 2015. [Online]. Available: https://doi.org/10.1016%2Fj.asieco.2015.01.004

[37] Y. Funashima, "The fed-induced political business cycle: Empirical evidence from a time–frequency view," *Economic Modelling*, vol. 54, pp. 402–411, 2016.

[38] C. Su, X. Yin, R. Tao, O.-R. Lobonţ, and N.-C. Moldovan, "Are there significant linkages between two series of housing prices, money supply and short-term international capital?–evidence from china," *Digital Signal Processing*, vol. 83, pp. 148–156, 2018.

[39] A. K. Tiwari, M. I. Mutascu, and C. T. Albulescu, "Continuous wavelet transform and rolling correlation of european stock markets," *International Review of Economics & Finance*, vol. 42, pp. 237–256, mar 2016. [Online]. Available: https://doi.org/10.1016%2Fj.iref.2015.12.002

[40] "Organisation for economic co-operation and development (OECD)." [Online]. Available: http://stats.oecd.org/Index.aspx?DatasetCode=SNA_TABLE1

[41] J. Poměnková, "Business cycle identification," *Journal of Economics*, vol. 60, no. 9, pp. 899–971, 2012.

[42] J. Blumenstein, J. Pomenkova, and R. Marsalek, "Comparative study of time-frequency analysis approaches with application to economic indicators," in *ECMS 2012 Proceedings edited by: K. G. Troitzsch, M. Moehring, U. Lotzmann*. ECMS, may 2012. [Online]. Available: https://doi.org/10.7148%2F2012-0291-0297

[43] J. Pachman, M. Künzel, O. Němec, and S. Bland, "Characterization of al plate acceleration by low power photonic doppler velocimetry (pdv)," in *40th International Pyrotechnics Society Seminar, Colorado Springs, USA*, 2014, pp. 13–18.

[44] M. Kunzel, O. Němec, and J. Pachman, "Optimization of wall velocity measurements using photonic doppler velocimetry (pdv)," *Central European Journal of Energetic Materials*, vol. 12, no. 1, 2015.

[45] D. H. Dolan, "Accuracy and precision in photonic doppler velocimetry," *Review of Scientific Instruments*, vol. 81, no. 5, p. 053905, may 2010. [Online]. Available: https://doi.org/10.1063%2F1.3429257

[46] D. B. Holtkamp, "Survey of optical velocimetry experiments - applications of PDV, a heterodyne velocimeter," in *2006 IEEE International Conference on Magagauss Magnetic Field Generation and Related Topics*. IEEE, nov 2006. [Online]. Available: https://doi.org/10.1109%2Fmegaguss.2006.4530668

[47] P. Broersen, "The quality of lagged products and autoregressive yule–walker models as autocorrelation estimates," *IEEE Transactions on Instrumentation*

and Measurement, vol. 58, no. 11, pp. 3867–3873, nov 2009. [Online]. Available: https://doi.org/10.1109%2Ftim.2009.2021206

[48] D. Kalman, "The generalized vandermonde matrix," *Mathematics Magazine*, vol. 57, no. 1, pp. 15–21, 1984.

[49] J. Poměnková, "Remarks on optimum kernels and optimum boundary kernels," *Applications of Mathematics*, vol. 53, no. 4, pp. 305–317, jul 2008. [Online]. Available: https://doi.org/10.1007%2Fs10492-008-0028-7

[50] M. Jansen, "Generalized cross validation in variable selection with and without shrinkage," *Journal of Statistical Planning and Inference*, vol. 159, pp. 90–104, apr 2015. [Online]. Available: https://doi.org/10.1016%2Fj.jspi.2014.10.007

[51] W. Greene, *Econometric Analysis*. Prentice Hall, 2003. [Online]. Available: https://books.google.cz/books?id=JJkWAQAAMAAJ

[52] Z. Ge, "Corrigendum to" significance tests for the wavelet power and the wavelet power spectrum" published in ann. geophys., 25, 2259–2269, 2007," in *Annales Geophysicae*, vol. 31, no. 2. Copernicus GmbH, 2013, pp. 315–315.

[53] J. A. Schulte, C. Duffy, and R. G. Najjar, "Geometric and topological approaches to significance testing in wavelet analysis," *Nonlinear Processes in Geophysics*, vol. 22, no. 2, pp. 139–156, mar 2015. [Online]. Available: https://doi.org/10.5194%2Fnpg-22-139-2015

[54] J. Jan, *Číslicová filtrace, analýza a restaurace signálů*. Vutium, 2002.

[55] Z. Ge, "Corrigendum to "significance tests for the wavelet cross spectrum and wavelet linear coherence" published in ann. geophys., 26, 3819–3829, 2008," *Annales Geophysicae*, vol. 31, no. 2, pp. 317–317, feb 2013. [Online]. Available: https://doi.org/10.5194%2Fangeo-31-317-2013

[56] W. T. Wells, R. L. Anderson, and J. W. Cell, "The distribution of the product of two central or non-central chi-square variates," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1016–1020, sep 1962. [Online]. Available: https://doi.org/10.1214%2Faoms%2F1177704469

[57] M. de Carvalho, P. C. Rodrigues, and A. Rua, "Tracking the US business cycle with a singular spectrum analysis," *Economics Letters*, vol. 114, no. 1, pp. 32–35, jan 2012. [Online]. Available: https://doi.org/10.1016%2Fj.econlet.2011.09.007

[58] G. Fulop and G. Gyomai, "Transition of the OECD CLI system to a GDP-based business cycle target," OECD Composite Leading Indicators Background Note, 2012. [Online]. Available: https://www.oecd.org/sdd/leading-indicators/49985449.pdf

[59] D. Gujarati, *Basic Econometrics (Sie)*. McGraw-Hill Education, 2007. [Online]. Available: https://books.google.cz/books?id=RyiB71OVNswC

[60] Organisation for economic co-operation and development (oecd): National accounts. [Online]. Available: http://stats.oecd.org/index.aspx?DatasetCode=KEI

[61] M. Gertler and B. Bernanke, "Agency costs, net worth and business fluctuations," in *Business cycle theory*. Edward Elgar Publishing Ltd., 1989.

[62] V. Ivashina and D. Scharfstein, "Bank lending during the financial crisis of 2008," *Journal of Financial economics*, vol. 97, no. 3, pp. 319–338, 2010.

[63] C. Borio, "The financial cycle and macroeconomics: What have we learnt?" *Journal of Banking & Finance*, vol. 45, pp. 182–198, 2014.

[64] M. Drehmann, C. E. Borio, and K. Tsatsaronis, "Characterising the financial cycle: don't lose sight of the medium term!" 2012.

[65] S. Claessens, M. A. Kose, and M. E. Terrones, "How do business and financial cycles interact?" *Journal of International economics*, vol. 87, no. 1, pp. 178–190, 2012.

[66] L. Aguiar-Conraria and M. J. Soares, "The continuous wavelet transform: Moving beyond uni-and bivariate analysis," *Journal of Economic Surveys*, vol. 28, no. 2, pp. 344–375, 2014.

[67] G. Galati, I. Hindrayanto, S. J. Koopman, and M. Vlekke, "Measuring financial cycles in a model-based analysis: Empirical evidence for the united states and the euro area," *Economics Letters*, vol. 145, pp. 83–87, 2016.

[68] D. Kunovac, M. Mandler, and M. Scharnagl, "Financial cycles in euro area economies: a cross-country perspective," 2018.

[69] D. Aikman, A. G. Haldane, and B. D. Nelson, "Curbing the credit cycle," *The Economic Journal*, vol. 125, no. 585, pp. 1072–1109, 2015.

[70] R. B. Gonzalez, J. Lima, L. S. G. Marinho *et al.*, *Business and Financial Cycles: An Estimation of Cycles'; Length Focusing on Macroprudential Policy*. Banco Central do Brasil, 2015.

[71] G. Rünstler, H. Balfoussia, L. Burlon, G. Buss, M. Comunale, B. De Backer, H. Dewachter, P. Guarda, M. Haavio, I. Hindrayanto *et al.* (2018) Real and financial cycles in eu countries: Stylised facts and modelling implications. [Online]. Available: https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op205.en.pdf

[72] P. Hiebert, B. Klaus, T. A. Peltonen, Y. S. Schüler, P. Welz *et al.*, "Capturing the financial cycle in euro area countries," *Financial Stability Review*, vol. 2, 2014.

[73] M. Altăr, M. Kubinschi, and D. Barnea, "Measuring financial cycle length and assessing synchronization using wavelets," *ESPERA*, vol. 20, no. 3, 2017.

[74] F. Verona, "Time–frequency characterization of the us financial cycle," *Economics Letters*, vol. 144, pp. 75–79, 2016.

[75] C. Alcidi *et al.*, "Fiscal policy stabilisation and the financial cycle in the euro area," Directorate General Economic and Financial Affairs (DG ECFIN), European ..., Tech. Rep., 2017.

[76] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[77] M. Yogo, "Measuring business cycles: A wavelet analysis of economic time series," *Economics Letters*, vol. 100, no. 2, pp. 208–212, 2008.

[78] P. M. Crowley, "An intuitive guide to wavelets for economists," *Bank of Finland Research Discussion Paper*, no. 1, 2005.

[79] P. Conway, D. Frame *et al.*, "A spectral analysis of new zealand output gaps using fourier and wavelet techniques," Citeseer, Tech. Rep., 2000.

[80] J. Fidrmuc, I. Korhonen, and J. Poměnková, "Wavelet spectrum analysis of business cycles of china and g7 countries," *Applied Economics Letters*, vol. 21, no. 18, pp. 1309–1313, jul 2014. [Online]. Available: https://doi.org/10.1080%2F13504851.2014.920468

[81] Z. Ftiti, A. Tiwari, A. Belanès, and K. Guesmi, "Tests of financial market contagion: Evolutionary cospectral analysis versus wavelet analysis," *Computational Economics*, vol. 46, no. 4, pp. 575–611, sep 2014. [Online]. Available: https://doi.org/10.1007%2Fs10614-014-9461-8

[82] Euro area statistics: Banks balance sheet – loans [online database]. [Online]. Available: https://www.euro-area-statistics.org/banks-balance-sheet-loans?cr=eur

[83] G. L. Mazzi, D. Ladiray, and D. A. Rieser, "Handbook on seasonal adjustment," 2018.

[84] R. Marsalek, J. Pomenkova, and S. Kapounek, "A wavelet-based approach to filter out symmetric macroeconomic shocks," *Computational Economics*, vol. 44, no. 4, pp. 477–488, oct 2013. [Online]. Available: https://doi.org/10.1007%2Fs10614-013-9403-x

[85] M. Gertler and S. Gilchrist, "The role of credit market imperfections in the monetary transmission mechanism: arguments and evidence," *The Scandinavian Journal of Economics*, pp. 43–64, 1993.

[86] D. Holod and J. Peek, *The value to banks of small business lending.* Citeseer, 2007.

[87] A. K. Kashyap and J. C. Stein, "The impact of monetary policy on bank balance sheets," in *Carnegie-rochester conference series on public policy*, vol. 42. Elsevier, 1995, pp. 151–195.

[88] E. Takáts, "Banking consolidation and small business lending," *Available at SSRN 601027*, 2004.

[89] T. Beck, B. Büyükkarabacak, F. K. Rioja, and N. T. Valev, "Who gets the credit? and does it matter? household vs. firm lending across countries," *The BE Journal of Macroeconomics*, vol. 12, no. 1, 2012.

[90] A. N. Berger and G. F. Udell, "The institutional memory hypothesis and the procyclicality of bank lending behavior," *Journal of financial intermediation*, vol. 13, no. 4, pp. 458–495, 2004.

[91] V. Bouvatier and L. Lepetit, "Banks' procyclical behavior: Does provisioning matter?" *Journal of international financial markets, institutions and money*, vol. 18, no. 5, pp. 513–526, 2008.

[92] T. Helbling, R. Huidrom, M. A. Kose, and C. Otrok, "Do credit shocks matter? a global perspective," *European Economic Review*, vol. 55, no. 3, pp. 340–353, 2011.

[93] J. Fidrmuc, T. Ikeda, and K. Iwatsubo, "International transmission of business cycles: Evidence from dynamic correlations," *Economics Letters*, vol. 114, no. 3, pp. 252–255, 2012.

[94] G. Georgiadis and A. Mehl, "Financial globalisation and monetary policy effectiveness," *Journal of International Economics*, vol. 103, pp. 200–212, 2016.

[95] B. Michalski *et al.*, "Eu-korea fta and its impact on v4 economies. a comparative analysis of trade sophistication and intra-industry trade," *Comparative Economic Research. Central and Eastern Europe*, vol. 21, no. 1, pp. 7–23, 2018.

[96] P. B. Rana, T. Cheng, and W.-M. Chia, "Trade intensity and business cycle synchronization: East asia versus europe," *Journal of Asian Economics*, vol. 23, no. 6, pp. 701–706, 2012.

[97] L. P. Dutrieux, C. C. Jakovac, S. H. Latifah, and L. Kooistra, "Reconstructing land use history from landsat time-series: Case study of a swidden agriculture system in brazil," *International Journal of Applied Earth Observation and Geoinformation*, vol. 47, pp. 112–124, 2016.

[98] R. Metz, "Filter-design and model-based analysis of trends and cycles in the presence of outliers and structural breaks," *Cliometrica*, vol. 4, no. 1, pp. 51–73, 2010.

[99] I. Korhonen and A. Peresetsky, "What influences stock market behavior in russia and other emerging countries?" *Emerging Markets Finance and Trade*, vol. 52, no. 5, pp. 1210–1225, sep 2015. [Online]. Available: https://doi.org/10.1080%2F1540496x.2015.1037200

[100] L. Fredette, J. T. Dreyer, T. E. Rook, and R. Singh, "Harmonic amplitude dependent dynamic stiffness of hydraulic bushings: Alternate nonlinear models and experimental validation," *Mechanical Systems and Signal Processing*, vol. 75, pp. 589–606, 2016.

[101] T. Proietti, "The multistep beveridge–nelson decomposition," *Econometric Reviews*, vol. 35, no. 3, pp. 373–395, 2016.

[102] E. N. Salachas, N. T. Laopodis, and G. P. Kouretas, "The bank-lending channel and monetary policy during pre-and post-2007 crisis," *Journal of International Financial Markets, Institutions and Money*, vol. 47, pp. 176–187, 2017.

[103] J. Valle e Azevedo and A. Pereira, "Macroeconomic forecasting using low-frequency filters," *Oxford Bulletin of Economics and Statistics*, vol. 80, no. 1, pp. 39–64, 2018.

[104] D. S. G. Pollock, "Econometric filters," *Computational Economics*, vol. 48, no. 4, pp. 669–691, 2016.

[105] A. Iacobucci and A. Noullez, "A frequency selective filter for short-length time series," *Computational economics*, vol. 25, no. 1-2, pp. 75–102, 2005.

[106] A. F. Burns and W. C. Mitchell, *Measuring business cycles*. National bureau of economic research, 1946, no. burn46-1.

[107] K. Chowdhary and H. N. Najm, "Bayesian estimation of karhunen–loeve expansions; a random subspace approach," *Journal of Computational Physics*, vol. 319, pp. 280–293, 2016.

[108] M. C. Leles, J. P. H. Sansão, L. A. Mozelli, and H. N. Guimarães, "Improving reconstruction of time-series based in singular spectrum analysis: A segmentation approach," *Digital Signal Processing*, vol. 77, pp. 63–76, 2018.

[109] W.-H. Yang, S. H. Holan, C. K. Wikle *et al.*, "Bayesian lattice filters for time-varying autoregression and time–frequency analysis," *Bayesian Analysis*, vol. 11, no. 4, pp. 977–1003, 2016.

[110] A. Bhattacharyya, L. Singh, and R. B. Pachori, "Fourier–bessel series expansion based empirical wavelet transform for analysis of non-stationary signals," *Digital Signal Processing*, vol. 78, pp. 185–196, 2018.

[111] W. Liu, S. Cao, and Y. Chen, "Seismic time–frequency analysis via empirical wavelet transform," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 28–32, 2015.

[112] J. A. Schulte, "Statistical hypothesis testing in wavelet analysis: theoretical developments and applications to indian rainfall," *Nonlinear Processes in Geophysics*, vol. 26, no. 2, pp. 91–108, 2019.

[113] P. Karthick, D. M. Ghosh, and S. Ramakrishnan, "Surface electromyography based muscle fatigue detection using high-resolution time-frequency methods and machine learning algorithms," *Computer methods and programs in biomedicine*, vol. 154, pp. 45–56, 2018.

[114] Y. Jiang, H. Nie, and J. Y. Monginsidi, "Co-movement of asean stock markets: New evidence from wavelet and vmd-based copula tests," *Economic Modelling*, vol. 64, pp. 384–398, 2017.

[115] H.-Y. Wang, T. Jie, and M. Guo-Ying, "Study on the adaptive wavelet threshold denoising method for coal mine hoisting wire rope signals based on novel thresholding function," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 60, no. 2, pp. 99–103, 2018.

[116] J. C. Reboredo, M. A. Rivera-Castro, and A. Ugolini, "Wavelet-based test of co-movement and causality between oil and renewable energy stock prices," *Energy Economics*, vol. 61, pp. 241–252, 2017.

[117] M. Mandler and M. Scharnagl, "Financial cycles across g7 economies: A view from wavelet analysis," 2019.

[118] L. Aguiar-Conraria, P. Brinca, H. Gudjonsson, and J. Soares, "Optimal currency area and business cycle synchronization across us states." 2015.

[119] A. Rua, "Money growth and inflation in the euro area: A time-frequency view," *Oxford Bulletin of Economics and Statistics*, vol. 74, no. 6, pp. 875–885, 2012.

[120] S. Jangjit and M. Ketcham, "A new wavelet denoising method for noise threshold," *Engineering Journal*, vol. 21, no. 7, pp. 141–155, 2017.

[121] S. Liu, T. Liu, L. Gao, H. Li, Q. Hu, J. Zhao, and C. Wang, "Convolutional neural network and guided filtering for sar image denoising," *Remote Sensing*, vol. 11, no. 6, p. 702, 2019.

[122] Y. Li, K. Li, and Q. Lu, "Applying segmentation and classification to improve performance of smoothing," *Digital Signal Processing*, vol. 109, p. 102913, 2021.

[123] A. Morteza and M. Amirmazlaghani, "A novel statistical approach for multiplicative speckle removal using t-locations scale and non-sub sampled shearlet transform," *Digital Signal Processing*, vol. 107, p. 102857, 2020.

[124] D. F. Walnut, *An introduction to wavelet analysis*. Springer Science & Business Media, 2013.

[125] D. N. Gujarati, D. C. Porter, and S. Gunasekar, *Basic econometrics*. Tata McGraw-Hill Education, 2012.

[126] L. K. Westin. (2015) Receiver operating characteristic (roc) analysis. evaluating discriminance effects among decision support systems. [Online]. Available: http://nutkin.cs.umu.se/research/reports/2001/018/part1.pdf

[127] J. Bonenkamp, J. Jacobs, G. H. Kuper *et al.*, *Measuring business cycles in The Netherlands, 1815-1913: a comparison of business cycle dating methods*. Graduate School/Research Institute Systems, Organisation and Management, 2001.

List of Abbreviations

AIC	Akaike Information Criterion
AR	Autoregressive Process
ARMA	Autoregressive Moving Average
BC	Business Cycle
BIC	Bayesian Information Criterion
CI	Confidence Interval
COI	Cone of Influence
CWT	Continuous Wavelet Analysis
DFT	Discrete Fourier Transform
FODLOG	First-Order Difference of Natural Logarithms
FT	Fourier Transform
GDP	Gross Domestic Product
GWN	Gaussian White Noise
HT	Hard Threshold
HQC	Hannan-Quinn Information Criterion
IPI	Industrial Production Index
LAB	Local Adaptive Based Testing
MDL	Minimum Description Length
MSE	Mean Square Error
PDV	Photonic Doppler Velocimetry
PWCS	Power Wavelet Cross-Spectrum
SAB	Segmentation Adaptive Based Testing
SNR	Signal-to-Noise Ratio
STA	Standard Testing Approach

STFT	Short Time Fourier Transform
SVD	Singular Vector Decomposition
TF	Time-Frequency
TFA	Time-Frequency Analysis
TFAR	Time-Frequency Varying Autoregressive Process
WCS	Wavelet Cross-Spectrum