COMPARATIVE STUDY OF TIME-FREQUENCY ANALYSIS APPROACHES WITH APPLICATION TO ECONOMIC INDICATORS

Jiří Blumenstein
Department of Radio electronics
Brno University of Technology
Purkyňova 118
612 00 Brno, Czech Republic
Email: xblume00@phd.feec.vutbr.cz

Jitka Poměnková
Department of Radio electronics
Brno University of Technology
Purkyňova 118
612 00 Brno, Czech Republic
Email: pomenkaj@feec.vutbr.cz

Roman Maršílek
Department of Radio electronics
Brno University of Technology
Purkyňova 118
612 00 Brno, Czech Republic
Email: marsaler@feec.vutbr.cz

KEYWORDS

Time-frequency analysis, wavelets, multiple window method, AR process, periodogram

ABSTRACT

Presented paper deals with comparison of various methods for time-frequency representation of a signal with time-varying behavior. We choose methods such as wavelet analysis, multiple window method using Slepian sequences, time-frequency varying autoregressive process estimation and time-frequency Fourier transform representation (periodogram). We apply these methods first on the simple simulated artificial signal and we assess their performance. Then we proceed with application on the real data which is monthly data of the industry production index of European Union in the period 1990/M1-2011/M11. During the evaluation we focus on the results with respect to the time of global crisis. The results of the experiments are represented in the graphical form and briefly discussed.

INTRODUCTION

The description of time-frequency structure of signal has wide range of usage. Its application can be seen in many scientific areas such as engineering (Xu et al., 2011), medicine (Xu et al., 1999), economy and many others. In last several years these techniques are in the front of economic researchers which analyze comovement of economic indicators. In this sense the papers of (Rua, 2010), (Yogo, 2008) or (Hallett and Richter, 2007) and many others were written. Estimation of spectrogram or scalogram of input signal or time series depends on used methods and their parameters. We investigate in this article four basic methods such as wavelet analysis (Jan, 2002), multiple window method using the slepian sequences (MWM) (Xu et al., 1999), time-frequency varying autoregressive (AR) process (Proakis et al., 2002) spectrum estimation and time-frequency Fourier transform estimation (periodogram) (Jan, 2002). On the basis of simulations on the artificial well-known signal we analyze behavior of each method and search for their advantages, disadvantages and recommendations for their usage. Consequently we compare obtained results with the aim to give recommendation for methods application. In order to practically demonstrate and evaluate the performance of the chosen methods we apply them to the analysis of the real data which is the monthly data of the industry production index of the European Union countries in the period 1990/M1-2011/M11.

The paper is organized as follows: In the section Methodical Background we describe chosen methods of time-frequency analysis. Consequently, in the section Data, we briefly describe data used both for the simulation as well as for the practical application. After that, in the section Simulation, we show results of an application of chosen methods on simulated artificial data. The section Application presents results of real data analysis of the industry production index of the European Union and its brief economic interpretation. In both later sections results are graphically represented. The paper ends up with the conclusion and the list of used references.

METHODICAL BACKGROUND

Let us have a signal (a time series) \( y(n) \), \( n = 1, \ldots, N \). Under assumption that the time series contain a long-term trend, we can apply additive decomposition in the following form

\[
y(n) = g(n) + s(n) + c(n) + \varepsilon(n), \quad n = 1, \ldots, N, \tag{1}
\]

where \( g(n) \) denotes a long-term trend, \( s(n) \) is the seasonal component, \( c(n) \) is the cyclical component and \( \varepsilon(n) \) is the irregular component (a random noise). Focusing on analysis of cyclical movements around its long-term trend it is necessary to remove the long-term trend applying some filtering methods. When the seasonally adjusted data are not available (in other words the analyzed series contains the seasonal component), the seasonality should be removed by applying some corresponding method.

The spectrum of the signal (time series) \( y(n) \), \( n = 1, \ldots, N \) can be written as a Fourier sum (Hamilton, 1994)

\[
S_y(f) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-ijf}, \tag{2}
\]

where \( \gamma_j = \text{cov}(y(n)y(n+j)) \) is autocovariance between \( y(n) \) and \( y(n+j) \), \( i = \sqrt{-1}, f = 2\pi/N \).
Sample periodogram

Let $N$ denotes an odd integer, $M = (N - 1)/2$.

Let $f_j = 2\pi j/N$, $j = 1, \ldots, M$ and let

$$x_n = \begin{bmatrix} 1 & \cos(f_1(n - 1)) & \sin(f_1(n - 1)) & \ldots & \cos(f_M(n - 1)) & \sin(f_M(n - 1)) \end{bmatrix}^T.$$

Then

$$\sum_{n=1}^{N} x_n x_n' = \begin{bmatrix} N & 0 & 0 & \ldots & 0 & (N/2)I_{N-1} \end{bmatrix}.$$  Furthermore, let $y(n), n = 1, \ldots, N$ be any $N$ numbers. Then the following holds:

1. The value $y(n)$ can be expressed as

$$y(n) = \hat{\mu} + \sum_{j=1}^{M} \hat{\alpha}_j \cos(f_j n) + \hat{\delta}_j \sin(f_j n),$$

with $\hat{\mu} = N^{-1} \sum_{n=1}^{N} y(n)$ and for $j = 1, \ldots, M$

$$\hat{\alpha}_j = \left(2/N\right) \sum_{n=1}^{N} y(n) \cos(f_j(n - 1)), \quad \hat{\delta}_j = \left(2/N\right) \sum_{n=1}^{N} y(n) \sin(f_j(n - 1)).$$

2. The sample variance of $y(n)$ can be expressed as

$$\text{var}(y) = \frac{1}{N} \sum_{n=1}^{N} (y(n) - \bar{y})^2 = \frac{1}{2} \sum_{j=1}^{M} \left( \hat{\alpha}_j^2 + \hat{\delta}_j^2 \right),$$

end the portion of sample variance of $y$ that can be attributed to the cycles of frequency $f_j$ is given by

$$1/2 \sum_{j=1}^{M} \left( \hat{\alpha}_j^2 + \hat{\delta}_j^2 \right).$$

3. The portion of sample variance of $y$ that can be attributed to the cycles of frequency $f_j$ can equivalently be expressed as

$$1/2 \sum_{j=1}^{M} \left( \hat{\alpha}_j^2 + \hat{\delta}_j^2 \right) \approx \frac{4\pi}{N} \hat{S}_y(f_j),$$

where $\hat{S}_y(f_j)$ is the sample periodogram at frequency $f_j$.

Proof: (Hamilton, 1994).

Periodogram in time-frequency: Spectrogram

Practically, the periodogram is often estimated by the mean of a discrete Fourier transform

$$S_y(f) = \sum_{n=0}^{N-1} y(n)e^{-j2\pi fn},$$

In the assumption that the time-varying signal have a locally periodical behaviour, the time-frequency representation - spectrogram can be constructed by the mean of the Short Time Fourier Transform (STFT, here defined in the discrete form) at time $n$:

$$S_y(n, f) = \sum_{m=-\infty}^{\infty} y(m)\theta(n-m)e^{-j2\pi fm},$$

where $\theta()$ is the observation window with $N_w$ nonzero elements. According to the application and desired properties, different window functions can be used as the rectangular, the Hanning, the Hamming or the flat-top. A window is then slid over the analyzed signal and corresponding values of the STFT are computed. The (squared) module of the STFT results is then plotted in the 2D graph.

Multiple window method (MWM)

Similar approach for construction of the spectrogram has been proposed in (Xu et al., 1999). According the author the Slepian sequences are the eigenvectors of the equation

$$\sum_{m=0}^{N-1} \sin(2\pi W(n-m)) \nu_m(k) = \lambda(N, M) \nu_m(k) W,$$

where $N$ is the length of the eigenvectors (or data), and $W$ is the half-bandwidth that defines a small local frequency band centered around frequency $f$ : $|f - f'| \leq W$. Denote that Slepian sequences are orthogonal time-limited functions most concentrated in the frequency band $[-W, W]$.

The procedure how to compute the spectral estimate of $y(n)$ using the multiple window method written according to (Xu et al., 1999) consists from the following steps:

1. Specify $N$ and $W$, where $N$ is the number of data points, and $W$ depends on desired time-bandwidth product (or frequency resolution) $NW$.

2. Use (8) to compute the $\lambda_k$’s and $\nu_k$’s.

3. Apply $\nu_k$ to the entire length-$N$ data $y(n)$ and take the discrete Fourier transform

$$y_k(f) = \sum_{n=0}^{N-1} y(n)\nu_m(k)e^{-j2\pi fn},$$

where $y_k(f)$ is called $k$-th eigencoefficient and $|y_k(f)|^2$ the $k$-th eigenspectrum.

4. Average the $K$ eigenspectra (weighting by the reciprocal of the corresponding eigenvalues) to get an estimate of the spectrum

$$\hat{S}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \frac{1}{\lambda_k} |y_k(f)|^2.$$  (10)

Since the first few eigenvalues are very close to one, (10) can be simplified to

$$\hat{S}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \frac{1}{\lambda_k} |y_k(f)|^2.$$  (11)

The concept of the multiple window method can be easily extended into the time-frequency domain analysis, similarly to the spectrogram construction. In case of
MWM in TFA the idea is to apply a set of sliding windows and then takes the average (Xu et al., 1999)

\[ Y_{MW}(t, f) = \frac{1}{K} \sum_{k=0}^{K-1} |Y_k(t, f)|^2 \] (12)

where

\[ Y_k(t, f) = \int y(\tau) h_k(\tau - t) e^{-j2\pi f n} d\tau \] (13)

\( y(t) \) is the signal or time series to be analyzed. For the details you can see (Xu et al., 1999), (Frazer and Boashash, 1994). This equation stands for the time-frequency multivindow analysis in the continuous time domain. The modification to the discrete form, similarly to eq. 7 is straightforward.

**Time-frequency varying AR process**

As the alternative to the above mentioned non-parametric approaches, the time-frequency analysis can be performed in the parametric way. In such a case we assume (according to the well-known concept of the whitening filter), that analyzed time series \( y \) can be considered as the output of linear time invariant filter \( H(e^{j\omega}) \) driven by a with noise \( w \) with the variance \( \sigma_w^2 \). Note that \( \omega = 2\pi f \). The spectrum at the filter output is then \( S_y(f) = \sigma_w^2 |H(e^{j\omega})|^2 \). For this filter we can find so called whitening filter, on which output there is again a white noise. For the correctly designed whitening filter it holds \( H_w(e^{j\omega}) = 1/H(e^{j\omega}) \) (Jan, 2002).

The whitening filter can be implemented in the general form using the auto regressive moving average (ARMA) model. Its identification is not so simple, therefore moving average (MA) or auto regressive (AR) process are usually used. Using the AR process, the spectrum estimation can be done according to the formula (Proakis et al., 2002).

\[ \hat{S}_y(\omega) = \frac{\sigma_w^2}{1 - \sum_{i=1}^{\infty} a_i e^{-i\omega}} \] (14)

where \( a_i \) are the coefficients of the AR process of the order \( p \). An advantage of this approach is better frequency resolution. This depends on the order of the process which is smaller in comparison to the sample size \( N \). Several methods can be used for the AR process coefficient estimation like the Burg, Yule-Walker (Proakis et al. (2002)) or modified covariance method that we used in our previous paper (Sebesta (2011)).

In order to easily extend the AR process spectrum estimation to the time-frequency representation, a sliding window of the length \( N_w \) was moved across the analyzed signal (series) \( y(n) \). For each window shift, the AR process coefficients have been estimated and subsequently the spectrum has been computed and represented in the form of 2D plot.

**Wavelet transform**

Let us assume the time series is seasonally adjusted without the long-term trend. The continuous wavelet transform of the signal (the time series) \( y(n) \) with respect to the mother wavelet \( \psi_{a,\tau}(n) \) is defined as

\[ S_{CTW}(a, t) = \int_{-\infty}^{\infty} y(t) \frac{1}{\sqrt{a}} \psi\left(\frac{n - \tau}{a}\right) dt, \] (15)

\( a > 0, \tau \in R, \)

where the mother wavelet takes the form \( \psi_{a,\tau}(n) = \psi\left(\frac{n - \tau}{a}\right) \), \( \tau \) is the time shift, \( a \) is the parameter of dilatation (scale), which is related to the Fourier frequency. The numerator of the fraction \( \sqrt{a} \) ensures the conservation of energy (Jan, 2002).

To satisfy assumptions for the time-frequency analysis, wavelets must be compact in time as well as in the frequency representation. There exist a number of wavelets which can be used, such as Daubechie, Morlet, Haar or Gaussian wavelet (Gençay et al., 2002), (Adisson, 2002).

An inverse wavelet transformation is defined as

\[ y(n) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_{a,\tau}(t) S_{CTW}(a, t) \frac{da d\tau}{a^2}, \] (16)

where \( \psi_{a,\tau}(n) \) is the mother wavelet and \( S_{CTW}(a, n) \) is the continuous wavelet transform of time series \( y(n) \) defined in relation (15). For \( C_\psi \) hold \( 0 < C_\psi = \int_{0}^{\infty} |\Psi(\omega)|^2 d\omega < \infty \), \( \Psi(\omega) \) is the Fourier transform of \( \psi_{a,\tau}(n) \).

**DATA**

**Artificial data**

For the simulation purposes we used the following data. We assumed two sinusoidal waves with different periods which have been shifted and partially overlap in time. The resulting signal is given by following equation

\[ y = w_1 \sin(2\pi n/T_1) + w_2 \sin(2\pi n/T_2), \] (17)

where \( n = 1, \ldots, 263 \) is the time (in seconds, for the illustration),

\[ w_1 = \begin{cases} 1 & n = 10, \ldots, 180, \\ 0 & \text{elsewhere}, \end{cases} \]

and

\[ w_2 = \begin{cases} 1 & n = 100, \ldots, 249, \\ 0 & \text{elsewhere}. \end{cases} \]

are time windows of the two sinusoidal waves with period \( T_1 = 4s \) and \( T_2 = 16s \). The length of the artificial time series is chosen the same as the length of the real data (see next section). The time domain plot of the artificial signal is in Figure 1.
Real data
In order to demonstrate the performance of investigated methods in the real application, we used the data of the industry production index (IPI) of EU 27 in the period 1990/M1-2011/M11 (n = 263). Note that the length of this time series motivates our choice of the artificial signal length (see above). The monthly data are obtained from free database of Eurostat (Eurostat, 2011) and they are seasonally adjusted volume index of production, reference year 2005, mining and quarrying, manufacturing, electricity, gas, steam and air conditioning supply. Prior to the application of time-frequency analysis methods the input data of the industry production index is transformed by the natural logarithm and the long-term trend is removed with the use of Baxter-King filter (Guay, 2005).

SIMULATION
For simulation on the artificial data, the signals described in the preceding part in equation (17) were used. Let us remind that the simulated signal consists from two harmonic signals with periods of 4 and 16 seconds which partially overlap in time.

For the time frequency analysis (in figures denoted as TFA) - estimation of spectrogram we used time-frequency Fourier transform representation (periodogram) (Fig. 2), multiple window method (Fig. 3), time-frequency varying autoregressive process AR spectrum estimation with the lag order \( p = 6 \) (Fig. 4). For optimization of lag order we used the Akaike information criterion, for estimation of parameters of autoregressive process we used the Yulle Walker method (Green, 1997). We toke the sliding window of the length 64 moving through the time series with the one observation (sample) step. As the last method for time-frequency representation we used the continuous wavelet transform for scalogram (Fig. 5) computation. We took the Morlet wavelet as the mother wavelet and the maximum value of the scale parameter \( a = 30 \).

As the values of the spectrograms have been normalized, all figures has the same gray scale (Fig. 6). In the case of spectrogram plots (independently on the estimation method) we present the results in the whole normalized frequency range \((0, 1)\). Note that 1 here stands for the half of the sampling frequency and that there is a straightforward relationship between the normalized frequency and the cyclic component period \( T \):

\[
f_{\text{normalized}} = \frac{2}{T}.
\]

Normalized frequencies of 0.5 and 0.125 thus correspond to the periods of the artificial harmonic components \( T_1 = 4 \) and \( T_2 = 16 \), respectively. For better visibility of scalogram (wavelet analysis) results we limit the frequency range to \((0, 0.81)\) as no other components above the normalized frequency 0.81 were observable.

As we can see in the figures (2-5), we detect the simulated artificial components by all used methods. The highest frequency resolution is achieved by the AR process-based estimation, the lowest by the multiwindow and wavelet analysis. But the analysis results depend on the experiment setup and it is out of the scope of this paper to explore all the possible setups (e.g. the mother wavelet family) in order to generalize. In the case of wavelet analysis the results are by some reason slightly
moved towards the longer periods and are not well localized. Note that in contrast to other analysis methods, the frequency axis is in the logarithmic scale.

Influence of noise

In order to assess the performance of the TFA methods and particularly the influence of noise we setup a simple experiment. The additive white gaussian noise was added to the above mentioned data and the analysis with autoregressive process and wavelet transform were performed. The signal to noise ratio (SNR) of 0 dB was used. Thus the power of the useful signal is the same as the power of the noise. The results are shown on figures 7 and 8. It is possible to see that for such noised data, one of the signal components almost disappears in the plot of the wavelet analysis results while the AR estimation method works still well.

APPLICATION

Application was done on the real data described in the part "Data". Often there is no macroeconomic indicator with the long sample size, especially for small open economies. In most cases available data are quarterly data like in the case of gross domestic products (GDP), investment or consumption. The industry production index can be taken as an aggregated index containing the investments cycles, the household consumptions and production similar to GDP. It is available in monthly frequency. Therefore, with respect to the longer available sample size, we choose it for demonstration of discussed methods on the real data.

For time frequency analysis (TFA) of real data we used time-frequency Fourier transform representation (periodogram) (Fig. 9), multiple window method (Fig. 10), time-frequency changing autoregressive process $AR$ with the lag order $p = 6$ (Fig. 11). For optimization of lag order we used Akaike information criterion, for estimation of parameters of autoregressive process we used Yule Walker method. We toke the sliding window of the length 64 (Šebesta, 2011) moving through the time series about one observation. As the last method for the time-frequency representation we used the continuous wavelet transform for scalogram computation (Fig. 12). As in the previous case we took the Morlet mother wavelet and the maximum value of the scale $a = 30$.

As in the previous analysis of the artificial data, all figures have the same gray scale (Fig. 6). For better visibility, we limit the frequency range in all plots from the analysis of real data to the interval of $(0, 0.6)$ of normalized frequency. Remember again that 1 stands for the half of the sampling rate.
All presented figures identified two main exogenous shocks which affected not only short but also the long-term cycles in the economic activity of the Eurozone. The first, mortgage crisis began with collapse of price bubbles at the asset markets in the year 2007. The debt crisis followed two years after in the Eurozone. Both events were unique in its intensity in the last few decades. Obviously, the moving window approach (Fig. 10) identified only these two significant shocks and effects in long-term cyclical movements. Other previous changes in aggregate economic activity were obscured.

Figure 9 shows the contrast between the turbulent moving in low frequencies in the years 2007-2011 and cyclical movements during the nineties. The cyclical movements in the years 1997, 1998 and 1999 are sourced by low economic growth in Germany, Asian crisis and massive external imbalances at the capital accounts. The persistent effects of these imbalances were appeared at labor markets and investments (long-term cycles). Consumption changes were transmitted through decrease of credit money creation into the shorter waves (Kapounek, 2011). These two types of cyclical movements at different frequencies are evidently identified by wavelet analysis (Fig. 12).

Similar results, in comparison with wavelet analysis, provides time-frequency AR process, however there are not separated different types of cyclical movements between the years 2002 and 2006. As well, the mortgage and debt crisis consequences are slightly biased in the years 2008 and 2010.

CONCLUSION

In this paper we compared several methods for time-frequency analysis of given signal / time series. Chosen methods were time-frequency Fourier transform representation (periodogram), multiple window method using the slepian sequences, time-frequency varying AR process and the continuous wavelet transform. The methods have been first applied on the artificial data with the known structure, then we proceeded to the application on selected economic indicator. Note that the main limitation of the analysis is the short sample size of the analyzed data. At the final part of the paper, we tried to interpret the analysis results for the real data in the context of economic situation.

Estimation of time-frequency structure of simulated artificial data using all above discussed methods lead into similar results. In the case of wavelet analysis the identified spectrum makes impression to contain wider range of frequency components. A detailed look to the 3D representation of the chart would provide precise estimate of
the identified frequency components which is very close to the other simulated results and it will be presented during the conference.

To successfully apply the wavelet analysis, it is necessary to make a detailed analysis of mother wavelet choice. That will thus be the motivation for further research. In order to get the more realistic results of the time-frequency analysis of the economic indicators we therefore suggest to use two methods as minimum. We recommend to use the estimate of spectrogram (either periodogram, AR process based or multiwindow analysis) and to consult the results with the scalogram estimate obtained from the wavelet analysis.

In the case of multiwindow method the resultant time-frequency structure is estimated as the average over various used windows, so in this way we can cover area without any frequency by some average one and obtain spurious (or over-smoothed) result. Therefore we suggest for the work with real data to make a detailed look to the interconnection with field of the work (such in our case was economic situation) or to compare results with Fourier-transform based periodogram or wavelet analysis.

ACKNOWLEDGEMENT

The research described in the paper was supported by the Czech Science Foundation via grant n. P402/11/0570 with the title "Time-frequency approach for the Czech Republic business cycle dating" and by the project CZ.1.07/2.3.00/20.0007 WICOMT of the operational program Education for competitiveness. It was performed in the laboratories supported by the SIX project; the registration number CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation.

References


AUTHOR BIOGRAPHIES

**JIŘÍ BLUMENSTEIN** received his Master degree in electrical engineering from the Brno University of Technology in 2009. At present, he is a PhD student at the Department of Radio Electronics, Brno University of Technology. xblume00@phd.feec.vutbr.cz

**JITKA POMÉNKOVÁ** received the Ph.D. degree in applied mathematics at Ostrava University in 2005, the habilitation degree in Econometric and operational research at Mendelu in Brno in 2010. From 2011 she is the Senior researcher at Department of Radio electronics, Brno University of Technology. pomenkaj@feec.vutbr.cz

**ROMAN MARŠÁLEK** received his Master degree in control and measurements in 1999 from the Brno University of Technology and the Ph.D. in electronics and signal processing from Universite de Marne la Vallee, France in 2003. marsaler@feec.vutbr.cz