EXAMINATION OF RELATIONSHIPS BETWEEN TIME SERIES BY WAVELETS: ILLUSTRATION WITH CZECH STOCK AND INDUSTRIAL PRODUCTION DATA

MILAN BAŠTA
University of Economics, Prague, Faculty of Informatics and Statistics,
Department of Statistics and Probability,
W. Churchill Sq. 4, Prague, Czech Republic
e-mail: basm01@vse.cz

Abstract

Wavelet analysis is a relatively new approach to time series analysis which has also found its way to analysis of financial, economic and demographic time series. The paper provides examples from the literature how wavelets can be used for studying relationships between time series. Further, to give an illustration of practical implementation of wavelet analysis, the relationship between the Prague Stock Exchange PX Index and the Index of industrial production for the Czech Republic in the period from January 2000 to February 2016 is also analyzed, interesting results being obtained. Advantages of the wavelet approaches when compared to the traditional ones are discussed in the end.

Key words: *wavelets, time series, economy, comovement*

1. Introduction

There are several types of wavelet transforms such as the discrete wavelet transform (DWT), discrete wavelet packet transform (DWPT), maximal overlap discrete wavelet transform (MODWT), maximal overlap discrete wavelet packet transform (MODWPT), and the continuous wavelet transform (CWT). Each of the transforms has somewhat different properties, captures the characteristics of an input time series in a slightly different way, and can be more or less suitable for a given practical task at hand. However, all the transforms share the ability of providing time-scale information about an input time series or time-scale information about the relationship between (two or more) time series.

The goal of the paper is to illustrate the usefulness of wavelets for exploring relationships between time series by means of:

- providing an overview of examples in the literature studying relationships between economic time series by wavelets,
- analyzing the comovement of two time series for the Czech Republic, namely the Prague Stock Exchange PX index and the Index of industrial production.

The paper is organized as follows. Section 2 provides an introduction to the MODWT and to the notion of wavelet cross-correlation so that the idea of how studying relationships by wavelets works becomes clear to the reader. Section 3 gives practical examples of analyzing relationships by wavelets in the literature and emphasizes interesting results that have been obtained. Section 4 provides an analysis of the Czech Republic time series. Section 5 concludes.

2. MODWT-Based Wavelet Cross-Correlation

Let us assume two jointly stationary processes. The cross-covariance between the processes can be utilized as a traditional tool for exploring the strength of the linear relationship between the processes and for providing information about the lead/lag pattern. Other characteristics such as the cross-correlation, covariance or correlation can be derived from the cross-covariance and used for the examination of the relationship. Since the processes are jointly stationary, the cross-covariance does not vary over time.

If the frequency aspect of the relationship is required to be known between the jointly stationary processes, the cross-spectrum between the processes can be employed, being defined as the Fourier transform of the cross-covariance sequence. Since the cross-covariance sequence can be recovered from the cross-spectrum by the inverse Fourier transform, the cross-covariance sequence and the cross-spectrum carry the same information, but presented in a different way. Further measures, such as the amplitude and phase spectrum, can be derived from the cross-spectrum. A very useful measure is the squared coherency, which provides information about the strength of the linear relationship between the processes at different frequencies.

Methods based on wavelet transforms can be utilized for providing the scale aspect of the relationship between the stationary processes[1]. The MODWT-based wavelet cross-covariance and cross-correlation, or the MODWT-based wavelet covariance and correlation are favorite measures used for this purpose.

In order to illustrate the idea of how analyzing relationships by wavelets works, the MODWT-based wavelet cross-correlation will be introduced in Section 2.3. To introduce the MODWT-based wavelet correlation, the MODWT wavelet filters and the MODWT wavelet coefficients have to be introduced at first (Sections 2.1 and 2.2). Further, in Section 4.2 a simple generalization of the notion of wavelet cross-correlation is used which enables us to study not only the scale aspect of the relationship but also the temporal one.

2.1 MODWT Wavelet Filters

There are different families of wavelets (e.g. the Haar, D(4), D(6) etc. family). There is a set of the so-called MODWT wavelet filters associated with each family. Each MODWT wavelet filter in the set is associated with a specific level of wavelet analysis. The jth level (for $j = 1, 2, \ldots$) MODWT wavelet filter, denoted as $\{h_{j,l}: l = 0, \ldots, L_j - 1\}$, is a linear filter of length L_j. The filter fulfills several special restrictions (see e.g. Percival and Walden, 2006, Section 5.12, Table 202) and can be considered as an approximate bandpass filter for the range of frequencies $[1/2^{j+1}, 1/2^j]$.

2.2 MODWT wavelet Coefficients

Let us assume a stochastic process $\{X_t: t = \ldots, -1, 0, 1, \ldots\}$. The MODWT wavelet coefficients of level j ($j = 1, 2, \ldots$) for $\{X_t\}$, denoted as $\{W_{X,j,t}: t = \ldots, -1, 0, 1, \ldots\}$, are defined as (see e.g. Percival and Walden, 2006)

$$W_{X,j,t} \equiv \sum_{l=0}^{L_j-1} h_{j,l} X_{t-l}, \quad t = \ldots,-1,0,1,\ldots \qquad (1)$$

Here $\{W_{X,j,t}\}$ is associated with the dynamics of $\{X_t\}$ in the frequency range $[1/2^{j+1}, 1/2^j]$ and also with changes in $\{X_t\}$ on scale 2^{j-1} (Percival and Walden, 2006). Large levels are thus

[1] The notion of scale is closely related to the notion of frequency and will be defined in Section 2.2.

associated with long scales (long-run dynamics) and with low frequencies, whereas small levels are associated with short scales (short-run dynamics) and high frequencies.

If $\{X_t\}$ is stationary, $\{W_{X,j,t}\}$ is stationary too and has a zero mean. If $\{X_t\}$ is integrated of order $d \geq 1$, $\{W_{X,j,t}\}$ is stationary with a zero mean provided $L_1 > 2d$ (Percival and Walden, 2006, Section 8.2).

2.3 Wavelet Cross-Correlation

Let $\{W_{X,j,t}\}$ and $\{W_{Y,j,t}\}$ be the jth level MODWT wavelet coefficients for the processes $\{X_t\}$ and $\{Y_t\}$. Assuming joint stationarity of $\{W_{X,j,t}\}$ and $\{W_{Y,j,t}\}$ and zero means of $\{W_{X,j,t}\}$ and $\{W_{Y,j,t}\}$, the jth level wavelet cross-correlation (j = 1, 2, …) between $\{X_t\}$ and $\{Y_t\}$ at lag τ (τ =…, –1, 0, 1, …) is defined as (Whitcher et al., 2000)

$$\rho_{XY,j,\tau} \equiv \frac{\operatorname{cov}(W_{X,j,t}, W_{Y,j,t+\tau})}{\sqrt{\operatorname{var}(W_{X,j,t})\operatorname{var}(W_{Y,j,t+\tau})}} = \frac{E\left(W_{X,j,t}W_{Y,j,t+\tau}\right)}{\sqrt{E(W_{X,j,t}^2)E(W_{Y,j,t}^2)}}, \quad \tau = ...,-1,0,1,... \tag{2}$$

Since we assume stationarity, the subscript τ is not used in the denominator in the last part of Equation 2. Further, since $\{W_{X,j,t}\}$ and $\{W_{Y,j,t}\}$ are associated with changes in $\{X_t\}$ and $\{Y_t\}$ occurring on scale 2^{j-1}, the jth level wavelet cross-correlation provides information about the comovement and the lead/lag pattern between $\{X_t\}$ and $\{Y_t\}$ on scale 2^{j-1}.

2.4 Estimation of Wavelet Cross-Correlation

In real-life applications, only a realization of a finite length N of the stochastic process $\{X_t\}$ is available, denoted as $\{x_t: t = 0, …, N-1\}$. If no ad hoc assumptions on the value of x_t for $t < 0$ are imposed, the jth level MODWT wavelet coefficients for $\{x_t: t = 0, …, N-1\}$, denoted as $\{w_{x,j,t}\}$, can be calculated only for times $t = L_j - 1, …, N-1$ (we assume that $N \geq L_j$), i.e.

$$w_{x,j,t} \equiv \sum_{l=0}^{L_j-1} h_{j,l} x_{t-l}, \quad t = L_j - 1,..., N-1. \tag{3}$$

The jth level MODWT wavelet coefficients for $\{y_t: t = 0, …, N-1\}$, denoted as $\{w_{y,j,t}\}$, can be constructed in a similar manner.

Further, let $M_j \equiv N - L_j + 1 > 0$ be the length of $\{w_{x,j,t}: t = L_j - 1, …, N-1\}$ as well as of $\{w_{y,j,t}: t = L_j - 1, …, N-1\}$. Subsequently, $\rho_{XY,j,\tau}$ (j = 1, 2, …) can be estimated as

$$\hat{\rho}_{XY,j,\tau} \equiv \begin{cases} \dfrac{\frac{1}{M_j}\sum_{t=L_j-1}^{N-1-\tau} w_{x,j,t}w_{y,j,t+\tau}}{\left(\frac{1}{M_j}\sum_{t=L_j-1}^{N-1} w_{x,j,t}^2\right)^{1/2}\left(\frac{1}{M_j}\sum_{t=L_j-1}^{N-1} w_{y,j,t}^2\right)^{1/2}}, & \text{for } \tau = 0,..., M_j - 1, \\ \dfrac{\frac{1}{M_j}\sum_{t=L_j-\tau-1}^{N-1} w_{x,j,t}w_{y,j,t+\tau}}{\left(\frac{1}{M_j}\sum_{t=L_j-1}^{N-1} w_{x,j,t}^2\right)^{1/2}\left(\frac{1}{M_j}\sum_{t=L_j-1}^{N-1} w_{y,j,t}^2\right)^{1/2}}, & \text{for } \tau = -(M_j - 1),..,-1. \end{cases} \tag{4}$$

An approximate $(1-\alpha) \times 100\%$ confidence interval for $\rho_{XY,j,\tau}$ is given as (see e.g. Whitcher et al., 2000)

$$\left[\tanh\left\{\tanh^{-1}\left(\hat{\rho}_{XY,j,\tau}\right)-\frac{u_{1-\alpha/2}}{\sqrt{\frac{M_j}{2^j}-3}}\right\},\tanh\left\{\tanh^{-1}\left(\hat{\rho}_{XY,j,\tau}\right)+\frac{u_{1-\alpha/2}}{\sqrt{\frac{M_j}{2^j}-3}}\right\}\right], \tag{5}$$

where $u_{1-\alpha/2}$ denotes the $(1-\alpha/2)\times100$th percentile of the standard normal distribution.

3. Practical Examples from the Literature

In this section several practical examples from the literature are presented where wavelets have been used to analyze relationships between time series. Even though the wavelet techniques used in some of the literature papers may be different from the wavelet cross-correlation defined in Section 2, the idea of employing wavelets to explore the relationship between the time series is qualitatively similar to that presented Section 2. Specifically, the relationship is analyzed separately on different scales.

For example, Gallegati et al. (2014) explored the relationship between productivity and unemployment. They found a negative relationship in the long run, but a positive one in the short and medium run. Similarly, Gallegati et al. (2011) studied the relationship between wage inflation and unemployment rate on a scale-by-scale basis using quarterly US postwar data. They found a negative relationship at longer scales. They also studied the temporal stability of the results. Ramsey and Lampart (1998b) examined the relationship between consumption and income, and revealed a scale-specific dependence between the variables. Ramsey and Lampart (1998a) explored Granger causality between money and income on a scale-by-scale basis.

Understanding the nature of the Great moderation has also been of interest. Specifically, Crowley and Hughes Hallett (2014) explored the relationship between the components of GDP. They showed – both for US and UK quarterly data – that a decrease in volatility is especially observable at short scales, but not at long ones during the Great moderation, the moderation being more apparent in some components of GDP than in others. Gallegati and Gallegati (2007) carried out a study of (time-varying) wavelet variances of G7 countries' industrial production in order to understand the volatility moderation which occurred in the last decades. The authors reported a reduction in volatility which is uniform neither across scales nor across countries.

Business cycle synchronization has also been a popular research topic in wavelet analysis. Namely, Aguiar-Conraria and Soares (2011) used industrial production data in order to explore synchronization of business cycles for countries that have adopted, could have adopted or may adopt the Euro. The authors found that physical proximity is associated with stronger business cycle synchronization, and that France is leading the European cycle. Crowley and Mayes (2009) analyzed the relationship between the growth cycles of the three largest economies in the euro area (Germany, France and Italy), and reported interesting insights obtained from the wavelet analysis. Using real GDP, Crowley and Lee (2005) examined the comovement of business cycles across scales and over time for several European countries.

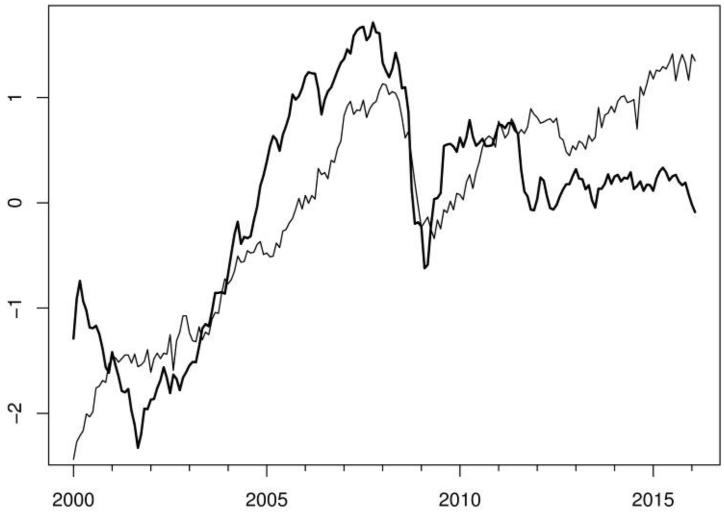
Other interesting papers have been published. Namely, Aguiar-Conraria et al. (2012) explored the relationship between the yield curve and the macroeconomy in the US in the period from 1961 to 2011. They estimated time-varying latent factors of the yield curve – namely, the level, slope and curvature – and explored time-varying relationships between these factors and macroeconomic variables. Jiang et al. (2015) applied wavelet analysis to study the relationship between money growth and inflation in China. Baubeau and Cazelles (2009) used wavelet analysis, and its ability to provide information on the time-scale content of a relationship, in order to assess the quality of two competing retrospective French GNP time series by comparing each time series with a non-retrospective benchmark.

4. Practical Illustration: PX Index and Industrial Production Index

In order to provide an illustration of the implementation of wavelets in analyzing relationships between time series, we study the comovement of the Prague Stock Exchange PX index[2] (PX, monthly time series) and the Index of industrial production for the Czech Republic[3] (IIP, monthly time series, basic index, seasonally adjusted) in the period from January 2000 to February 2016. Analyzing such a comovement can be interesting from the point of view of assessing whether stock markets can "predict" the economy or vice versa, and whether the relationship between the time series changes across scales and with time. Such issues are also relevant for constructing composite indicators such as those studied by Vraná (2014). We acknowledge the use of R software for performing the analysis.

The time series of the natural logarithm of PX will be denoted as $\{x_t\}$ and the time series of the natural logarithm of IIP will be denoted as $\{y_t\}$. Both $\{x_t\}$ and $\{y_t\}$ are presented in Figure 1. Scaling is performed by subtracting the mean of the time series and dividing the result by the standard deviation of the time series.

Figure 1: Scaled $\{x_t\}$ (in bold) and scaled $\{y_t\}$



Source: the author.

4.1 Wavelet Coefficients and Wavelet Cross-Correlation

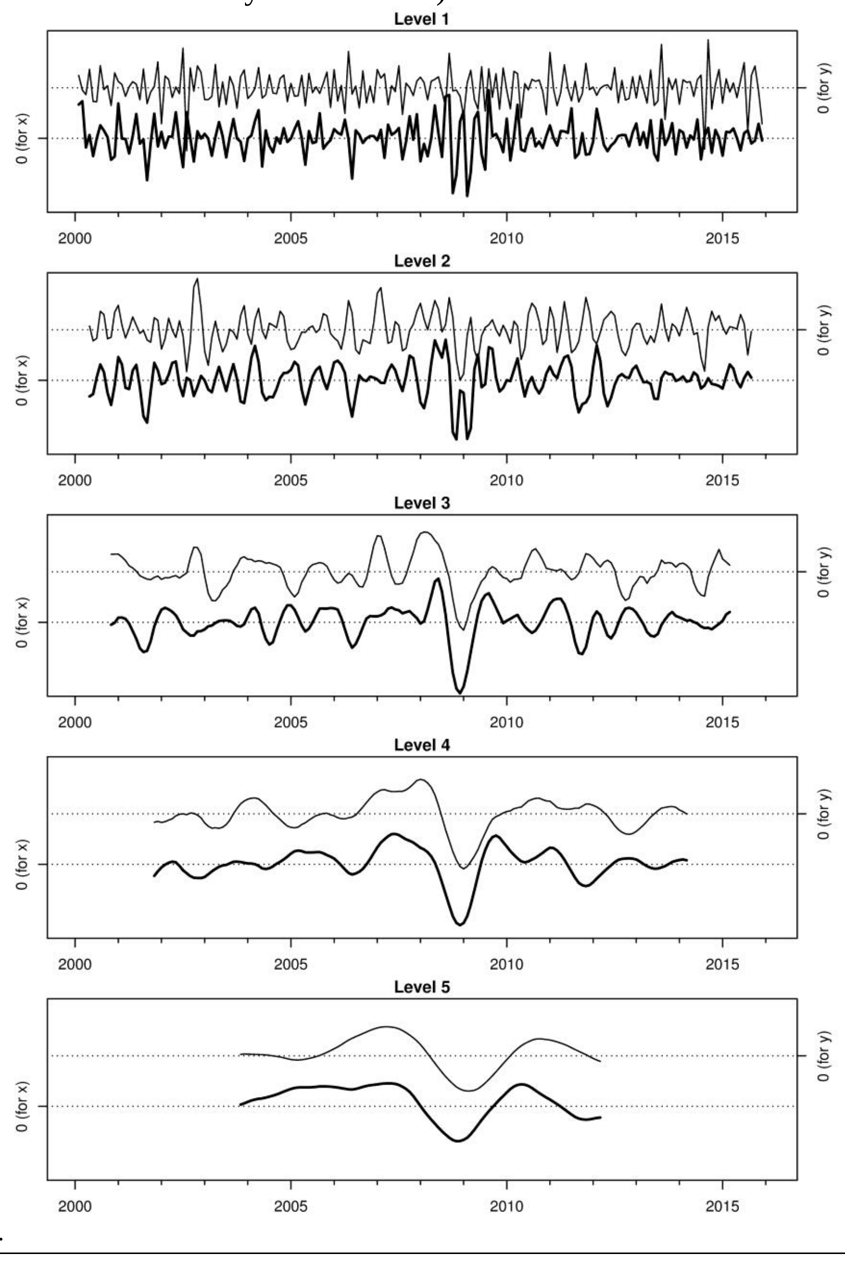
D(4) wavelet filters are used in the analysis[4]. MODWT wavelet coefficients of level 1 through level 5 (corresponding to scales 1, 2, 4, 8 and 16) both for $\{x_t\}$ and $\{y_t\}$ are presented in Figure 2. Since wavelet coefficients are obtained by linear filtering of the input time series with causal linear filters (see Equation 3), they would normally not be synchronized with the input time series. However, in Figure 2 synchronization is achieved by shifting the coefficients ahead by the amount given in Percival and Walden (2006, p. 118). Further, for a

[2] The data were retrieved from the Prague Stock Exchange web-site http://www.pse.cz in April 2016.
[3] The data were retrieved from the Czech Statistical Office web-site http://www.czso.cz in April 2016.
[4] $L_1 = 4$ for D(4) filters.

given level, the coefficients for $\{x_t\}$ and $\{y_t\}$ are plotted in one common subplot with two vertical axes, the left axis being associated with the coefficients for $\{x_t\}$, the right axis with those for $\{y_t\}$ (the positions of zeros on both the axes are denoted by the dotted lines). In the plots, the coefficients of each level are divided by their standard deviation.

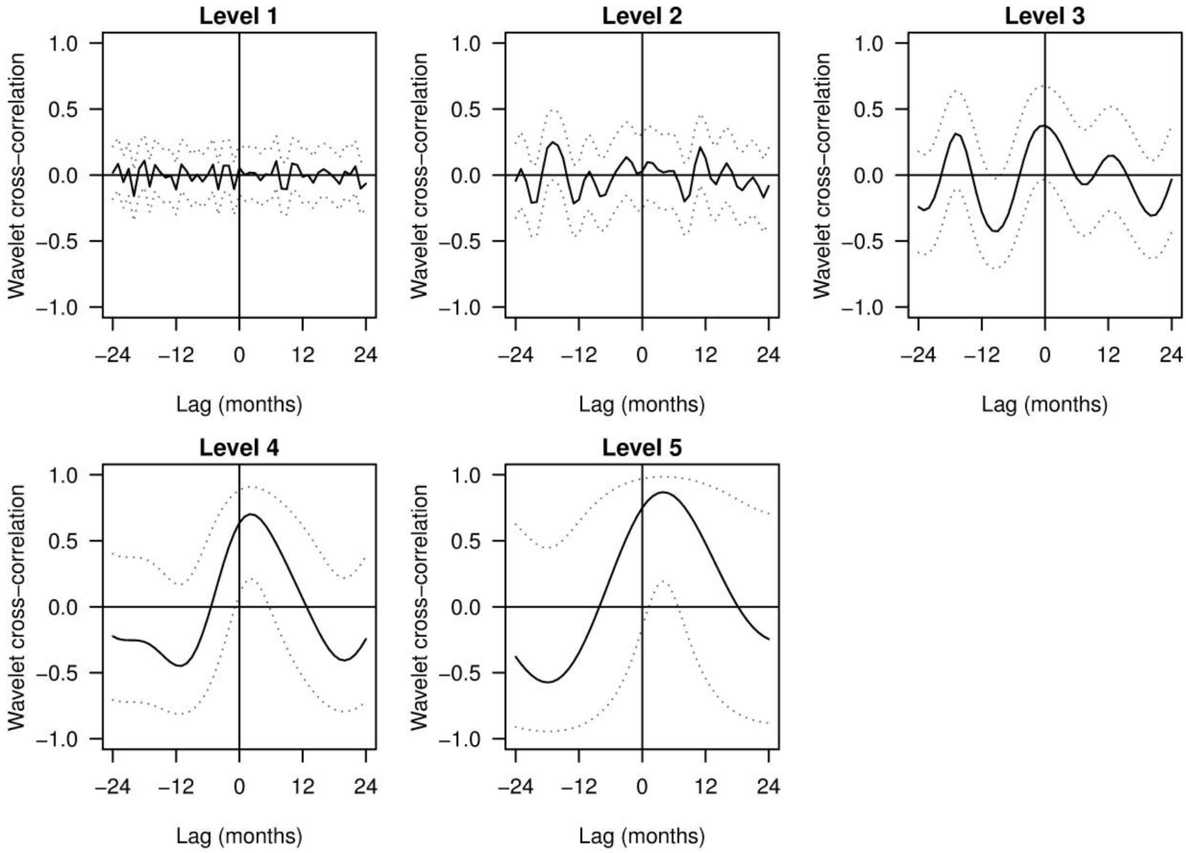
Figure 2 suggests that comovement between the time series $\{x_t\}$ and $\{y_t\}$ is scale-specific, stronger comovement being present at large levels (level 5) and no obvious comovement being present at small levels (level 1). These observations are also supported by the estimates of the wavelet cross-correlation presented in Figure 3 (the corresponding 95% confidence intervals are denoted by the dotted lines). In this figure, positive values of lag τ correspond to the situation where $\{y_t\}$ lags behind $\{x_t\}$. Consequently, Figure 3 suggests that there is comovement at level 4 and 5, $\{y_t\}$ lagging behind $\{x_t\}$ by several months at these levels. At level 3, the time series comove without any lag, while at level 1 and 2 no comovement is present.

Figure 2: D(4) MODWT wavelet coefficients for $\{x_t\}$ (in bold) and $\{y_t\}$ for level 1 through level 5 (after standardization and synchronization)



Source: the author.

Figure 3: The estimates of D(4) MODWT wavelet cross-correlations between $\{x_t\}$ and $\{y_t\}$ for level 1 through level 5



Source: the author.

4.2 Variation over Time

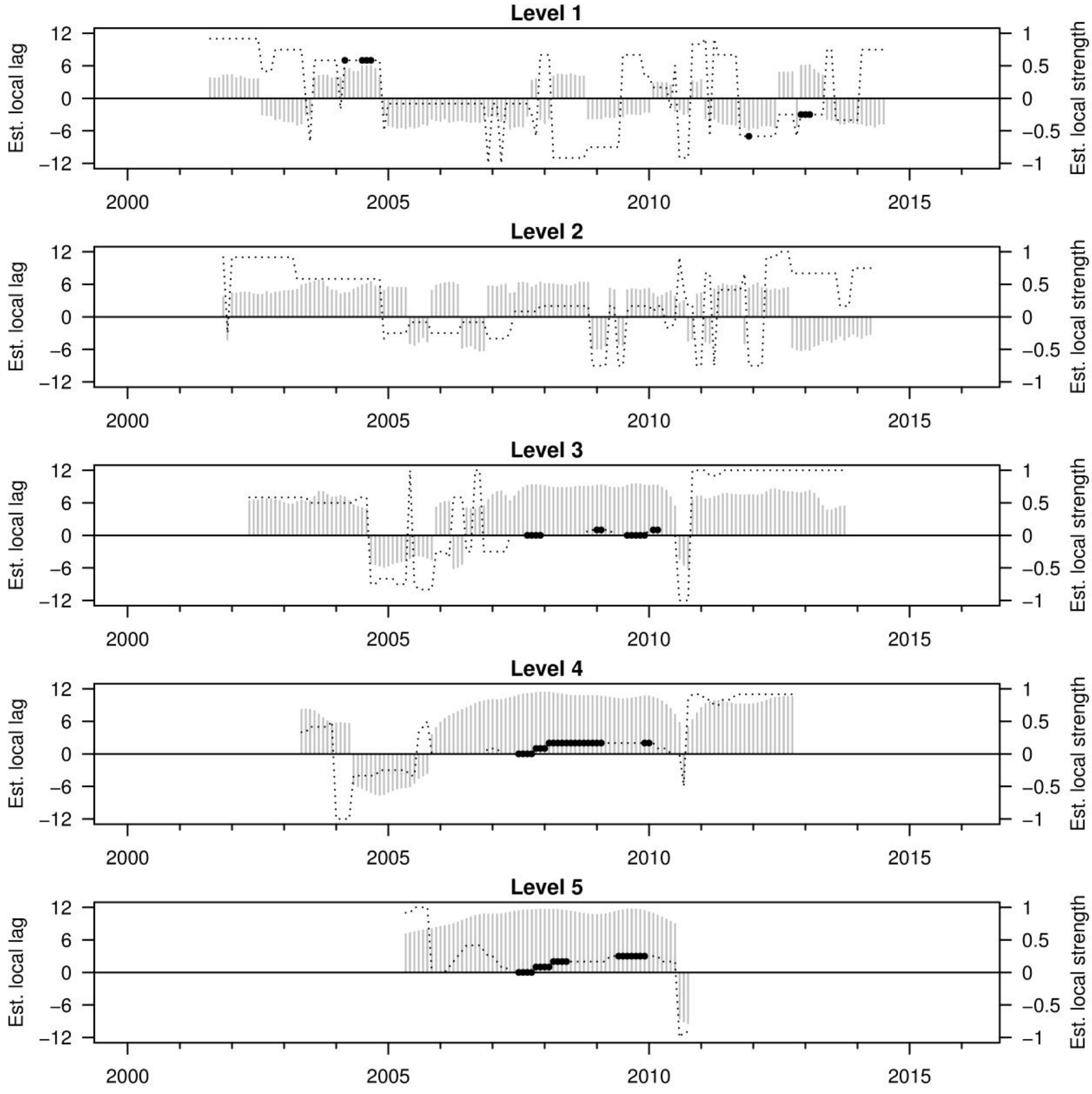
Upon a more careful examination of wavelet coefficients in Figure 2, we can see that the pattern of comovement between $\{x_t\}$ and $\{y_t\}$ is not persistent over time. To explore this pattern in more detail, we can make use of the ability of wavelets to provide not only the scale but also the temporal information about the relationship. A simple way to accomplish this is to assume a moving window approach.

We set the width of the window to 36 months (and move with the window from the start to the end of the time series). For a given level ($j = 1, \ldots, 5$) and for a given position of the moving window, we estimate $\rho_{XY,j,\tau}$ (for $\tau = -12, \ldots, 0, \ldots, 12$) and store the value of τ which provides the highest absolute value of the estimate of $\rho_{XY,j,\tau}$ (let us call this value of τ the "estimated local lag"). We also store the estimate of $\rho_{XY,j,\tau}$ at the "estimated local lag" (let us call this estimate of $\rho_{XY,j,\tau}$ the "estimated local strength of comovement"). For each level ($j = 1, \ldots, 5$), Figure 4 plots the "estimated local lag" (the left y-axis, dotted line) and the "estimated local strength of comovement" (the right y-axis, histogram-like vertical lines) against the time associated with the center of the corresponding moving window (x-axis).

Parametric bootstrap is used to assess the significance of the "estimated strength of comovement". Specifically, ARIMA models (assuming order of integration equal to 1) are estimated both for $\{x_t\}$ and $\{y_t\}$. Based on the estimated models, independent bootstrap realizations (one associated with $\{x_t\}$, the other one with $\{y_t\}$) of length N are generated. Subsequently, the wavelet cross-correlation combined with the moving window approach (described in the previous paragraph) is applied to analyze the comovement between the pair of the generated realizations, the dependence of the "estimated local strength of comovement" upon time (the center of the moving window) being stored. This procedure is repeated 5000 times.

Since the bootstrap distribution of the "estimated local strength of comovement" (obtained under the assumption of independence) does not vary with time, it is sufficient – for each level ($j = 1, \ldots, 5$) – to explore it at one specific time only and consider the 2.5% and 97.5% quantiles of the distribution as the critical values (assuming 5% significance) of the hypothesis test of no comovement at the given level (the null hypothesis). For each level, the hypothesis is tested at each single possible time. Those values on the curve of the "estimated local lag" in Figure 4 for which the corresponding hypothesis of no comovement is rejected are denoted by big black points.

Figure 4: The "estimated local lag" (left *y*-axis, dotted curve) and the "estimated local strength of comovement" (right *y*-axis, histogram-like vertical lines)



Source: the author.

We can see that the "estimated local lag" and the "estimated local strength of comovement" between $\{x_t\}$ and $\{y_t\}$ vary not only across levels but also with time, both being mostly (but not always) positive for levels 3, 4 and 5. Even though the "estimated local strength of comovement" is often rather large for these levels (3, 4, and 5), it is statistically significant only for several months in the period from 2007 to 2009. The corresponding "estimated local lag" is positive for these months and increases with the level (from level 3 to level 5), suggesting that $\{y_t\}$ lags behind $\{x_t\}$ at these levels.

For levels 1 and 2, the "estimated local lag" and the "estimated local strength of comovement" vary with time without any clear pattern. The "estimated local strength of comovement" is generally not too large (in the absolute value) and there are only a few months for which the "estimated local strength" is statistically significant.

5. Conclusion

We have given several illustrations how wavelet analysis can be useful for exploring the relationship between two time series as a function of scale and time. Two kinds of illustrations have been provided.

Firstly, several examples from the literature have been given where relationships between economic time series have been studied by wavelets and where interesting scale-specific relationships have been revealed. These illustrations support the idea that many economic phenomena exhibit scale-specific characteristics, and that traditional approaches to time series analysis may always not be suitable for uncovering these characteristics.

Secondly, we have explored the comovement between the Prague Stock Exchange PX index and the Index of industrial production for the Czech Republic in the period from the year 2000 to the year 2016. Using wavelets, we have shown that the relationship between the two time series varies across scales and with time, stronger comovement being present at longer scales. A particular strong and significant comovement at long scales occurred during the economic downturn in 2008 and 2009, PX index leading the Index of industrial production by a few months. No such strong comovement happened at short scales during this period or at any scales in other periods.

We have illustrated that the advantage of wavelet analysis (when compared to traditional methods of analysis such as cross-correlation analysis etc.) lies in the fact that it is capable of exploring the relationship as a function of scale and time, which leads to uncovering of those characteristics which remain unseen when explored by traditional methods.

It is interesting to note that the ability of wavelets to perform the analysis scale by scale and locally in time can also be employed in other fields of time series analysis such as time series forecasting or breakpoint and non-stationarity detection.

Acknowledgements

This paper was written with the support of the Czech Science Foundation (GA CR), grant no. P402/12/P507, Modelling of Financial and Economic Time Series - Application and Comparison of Wavelet and Traditional Methods.

References

[1] AGUIAR-CONRARIA, L. et al. 2012. The yield curve and the macro-economy across time and frequencies. In Journal of Economic Dynamics and Control, 2012, vol. 36, pp. 1950-1970.

[2] AGUIAR-CONRARIA, L., SOARES, M. J. 2011. Business cycle synchronization and the euro : A wavelet analysis. In Journal of Macroeconomics, 2011, vol. 33, pp. 477-489.

[3] BAUBEAU, P., CAZELLES, B. 2009. French economic cycles : A wavelet analysis of French retrospective GNP series. In Cliometrica, 2009, vol. 3, pp. 275-300.

[4] CROWLEY, P. M., HUGHES HALLETT, A. 2014. The great moderation under the microscope : Decomposition of macroeconomic cycles in US and UK aggregate demand. In GALLEGATI, M., SEMMLER, W. (eds.) Wavelet Applications in Economics and Finance. New York : Springer, 2014.

[5] CROWLEY, P. M., MAYES, D. G. 2009. How fused is the euro area core? An evaluation of growth cycle co-movement and synchronization using wavelet analysis. In OECD Journal: Journal of Business Cycle Measurement and Analysis, 2009, pp. 63-95.

[6] CROWLEY, P. M., LEE, J. 2005. Decomposing the co-movement of the business cycle: a time-frequency analysis of growth cycles in the euro area. In Bank of Finland Research Discussion Paper 12. Helsinki : Bank of Finland, 2005.

[7] GALLEGATI, M. et al. 2014. Does productivity affect unemployment? A time-frequency analysis for the US. In Gallegati, M., Semmler, W. (eds.) Wavelet Applications in Economics and Finance. New York : Springer, 2014.

[8] GALLEGATI, M. et al. 2011. The US wage Phillips curve across frequencies and over time. In Oxford Bulletin of Economics and Statistics, 2011, vol. 73, pp. 489-508.

[9] GALLEGATI, M., GALLEGATI, M. 2007. Wavelet variance analysis of output in G-7 countries. In Studies in Nonlinear Dynamics & Econometrics, 2007, vol. 11., iss. 3.

[10] JIANG, C. et al. 2015. Money growth and inflation in China : New evidence from a wavelet analysis. In International Review of Economics and Finance, 2015, vol. 35, pp. 249-261.

[11] RAMSEY, J. B., LAMPART, C. 1998b. The decomposition of economic relationships by time scale using wavelets : Expenditure and income. In Studies in Nonlinear Dynamics & Econometrics, 1998, vol. 3, pp. 23-42.

[12] RAMSEY, J. B., LAMPART, C. 1998a. The decomposition of economic relationships by timescale using wavelets. In Macroeconomic Dynamics, 1998, vol. 2, pp. 49-71.

[13] PERCIVAL, D. B., WALDEN, A. T. 2006. Wavelet methods for time series analysis. Cambridge : Cambridge University Press, 2006.

[14] VRANÁ, L. 2014. Czech business cycle chronology. In The 8th International Days of Statistics and Economics. Prague, 11.09.2014 – 13.09.2014. Slaný : Melandrium, 2014, pp. 1623-1632.

[15] WHITCHER, B. et al. 2000. Wavelet analysis of covariance with application to atmospheric time series. In Journal of Geophysical Research, 2000, vol. 105, pp. 14941-14962.