MIXED MODEL FOR SHORT-TERM PREDICTION OF TIME SERIES

Monika Hadaś-Dyduch

University of Economic of Katowice
Department of Statistical and Mathematical Method in Economics
40-287 Katowice, ul. 1 Maja 50
E-mail: monika.dyduch@ue.katowice.pl

Abstract: *The aim of the article is to present the original model for prediction of one-dimensional time series. The model described in the article is a mixed model based on a wavelet analysis. The model combines the creeping trend method with the wavelet analysis. The purpose of combining the above the method is a short-term prediction of time series - a dozen or so elements, although in principle the wavelet is applied to series of at least several dozen observations. The crawling trend method used has been modified in several basic stages. The proposed model was applied to the prediction of a time series presenting the social globalization index, understood as a component of the following indicators: international voice traffic, transfers, International tourism, migration, Patent applications by non residents filed through the Patent Cooperation, international students (sum of inbound and outbound number of tertiary students (% of population)), high technology exports (exports of products with high R&D intensity as share of total merchandise exports.), trade in cultural goods, trademark applications (applications to register a trademark with a national or regional Intellectual Property (IP) office by non residents in percent of all applications.), trade in personal services, McDonald's restaurant, IKEA stores.*

Key words: *wavelets, prediction, creeping trend, globalization, integration.*
JEL codes: *F10, C01, C02.*

1. Introduction

A forecast is a prediction of some future event or events. Forecasting is an important problem that spans many fields including business and industry, government, economics, environmental sciences, medicine, social science, politics, and finance. Forecasting problems are often classified as short-term, medium-term, and long-term. Short-term forecasting problems involve predicting events only a few time periods (days, weeks, and months) into the future. Medium-term forecasts extend from 1 to 2 years into the future, and long-term forecasting problems can extend beyond that by many years. Most forecasting problems involve the use of time series data. A Time series is a time-oriented or chronological sequence of observations on a variable of interest (Montgomery et al. , 2015).

The prediction can be made with different methods and models. These can be mechanical methods based on series analysis, trend models, cause-effect econometric models, analog methods, methods based on artificial neural networks, wavelets, etc. In the article, time wavelets and two adaptive models were used for prediction, i.e. the creeping trend method and the exponential equalization method. However, wavelets play a key role in the proposed original model. Important prediction methods include predictions based on artificial neural networks and wavelet analysis. Multifactorial predictors based on the use of artificial neural networks have been described extensively in (Freisleben, Ripper, 1997; Palit, Propovic 2005). In this way, "[...] implements models with a structure similar to ARMA, but with the use of implicit, non-linear transformations of explanatory variables, with parameters determined by teaching on historical data [...]" (Pełech-Pilichowski, Duda, 2008) . However, "[...] wavelet analysis is a kind of frequency analysis that allows to efficiently examine time-varying spectral characteristics of processes. Although it is not a prognostic technique per se, its distinguishing features, such as decomposition processes by frequency bands, good location properties over time, computational efficiency [...] (Bruzda, 2013) is "useful in predicting economic time series, especially those characterized by non-stationarity, manifesting short-term oscillations of varying amplitude, for which the links preceding in causal chains depend on the time scale (decision horizon) [...] "(Bruzda, 2013).

The aim of the article is to present the author's combination of the creeping trend method with wavelet analysis based on a discrete wavelet. A discrete wavelet was applied to the study.

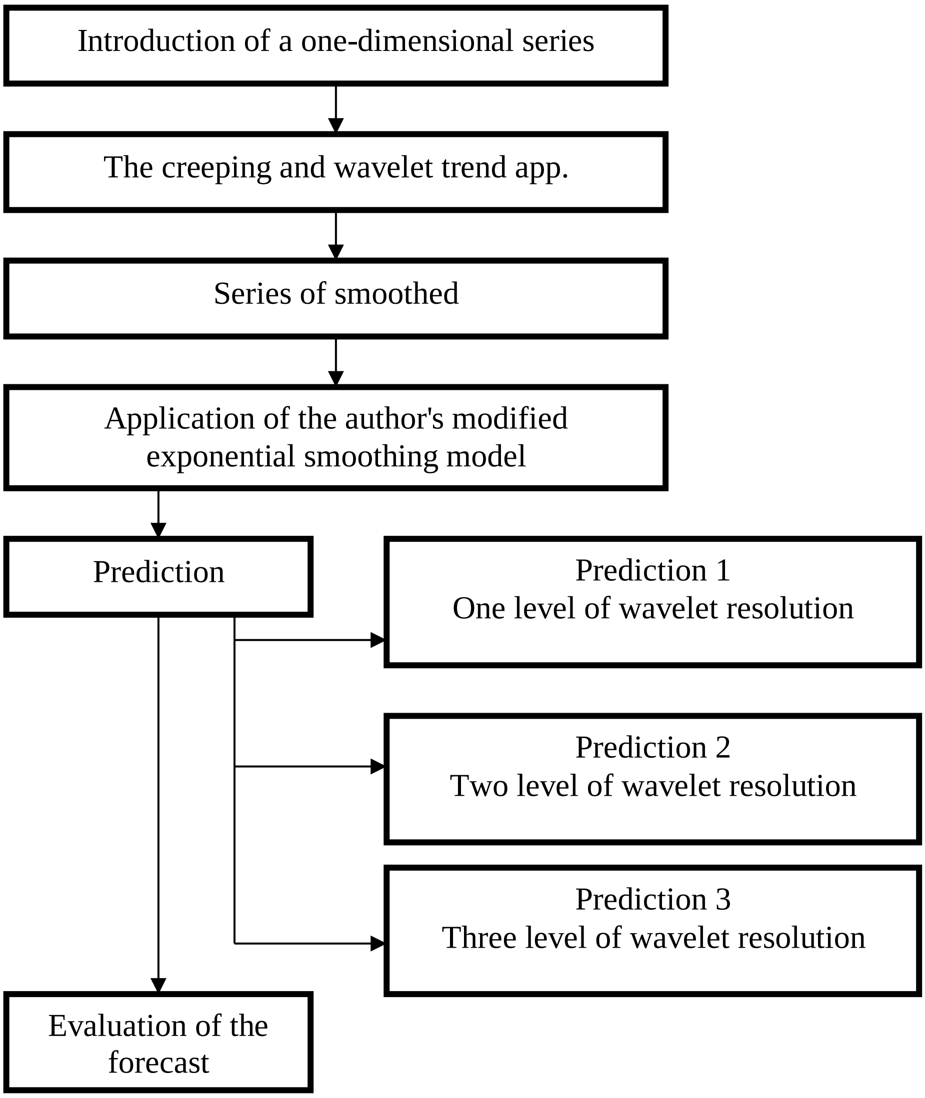
Mixed model, in the article is understood as a model integrating in one model various methods for analysis and prediction of time series. It is a model, understood as a model integrating different tools into one unit to obtain a model that is an effective prediction tool, in the sense of minimizing the error of post-prediction.

2. Methodology

The proposed algorithm for time series prediction using a discrete wavelet is an original algorithm. It is a model intended for short-term prediction based on a one-dimensional time series. It is a mixed model. Because it combines into one prognostic model: wavelet analysis and adaptive methods. The efficiency of such a model is much better than the efficiency of prediction of only individual adaptive models. The model presents a three-level decomposition of the time series. Of course, if the ranks contain a lot of observations, decomposition at higher levels of resolution is possible.

Fig. 1 presents a general scheme of the algorithm, and its detailed description is presented in the following figures.

Fig.1 The general scheme of the algorithm.



Source: Own elaboration.

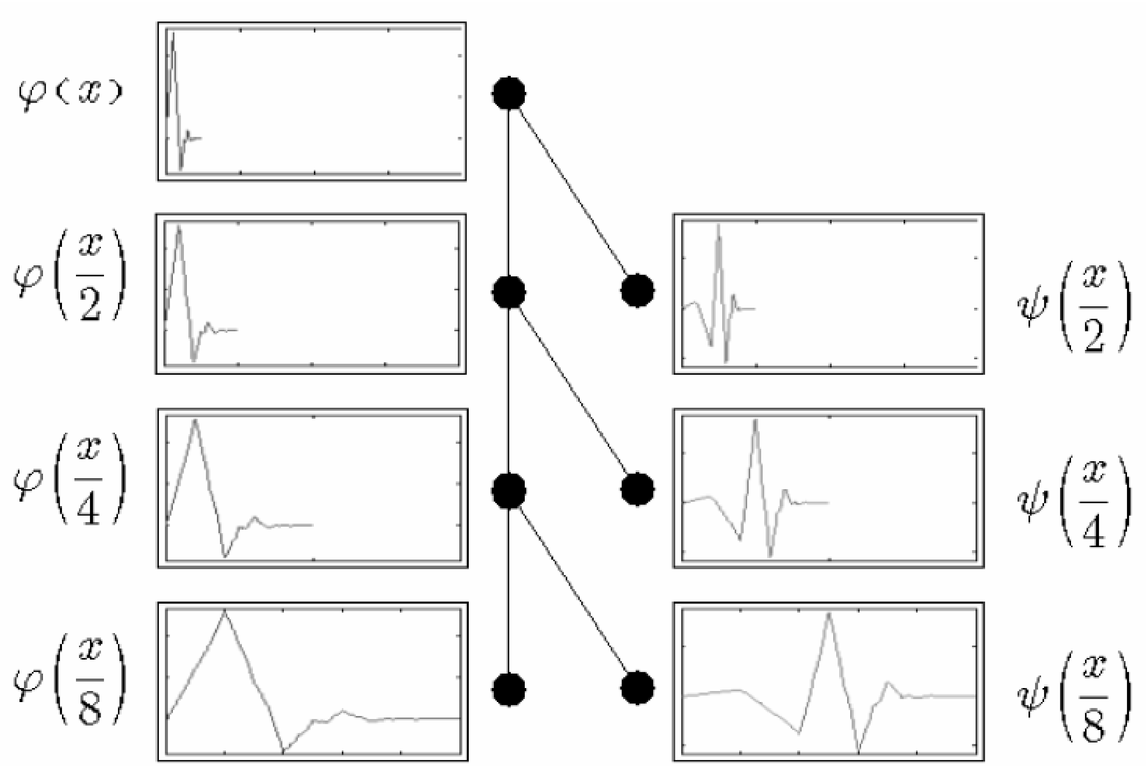
2.1. Wavelets

Wavelets are thus organized using two parameters:

- time k making it possible to translate the forms for a given level;
- scale 2^j making it possible to pass from a level j to the immediately lower

level in the underlying tree represented in Figure 4. In the first column of the figure we find the dyadic dilates (2 times, 4 times, 8 times, etc.) of the scaling function ϕ and in the second column, those of the wavelet ψ .

Fig. 2 Organization of wavelets



Source: Misiti, et al. 2013.

The simplest wavelet is the Haara wave, defined by A. Haar in 1910. A Daubechies wavelet was used in the study, introduced by Ingrid Daubechies in 1990 (see more in (Hadaś-Dyduch, 2015a, 2015b, 2016a, 2016b, 2016c)). The wavelet *db1* is simply the Haar wavelet. The main properties of the family of *dbn* wavelet are as follows (Misiti, et al. 2013):

- it is an orthogonal wavelet, associated with an MRA;
- it has compact support [0, 2n – 1] and the associated filters are of length 2n ;

- the number of vanishing moments is n and, in general, it is far from symmetric;
- the regularity is 0, 2n when n is sufficiently large.

The carriers of the basic scaling functions of Daubechies are the segments [0, 2M-1], and the carriers of the corresponding wavelets are sections [1-M, M]. The wavelets, in contrast to the scaling functions, do not contain any constants. For Daubechies scaling functions, the connection factors included in the matrix [Beylkin 1992, Ziółko 2000]:

$$F = \begin{pmatrix} 0 & 2 & 0 & 2 & 0 & 2 & 0 & \cdots \\ & 0 & 2 & 0 & 2 & 0 & 2 & \cdots \\ & & 0 & 2 & 0 & 2 & 0 & \cdots \\ & & & 0 & 2 & 0 & 2 & \cdots \\ & & & & 0 & 2 & 0 & \cdots \\ & & & & & 0 & 2 & \cdots \\ & & & & & & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

they meet the rule:

$$f_{m,n}^{j} = 2^{-2j} \int_{-\infty}^{\infty} \varphi\left(2^{-j}x - m\right) \frac{d\varphi\left(2^{-j}x - n\right)}{dx} dx = 2^{-j} f_{m-n}$$

where:

$$f_n = \begin{cases} 0 & gdy \quad n = 0 \text{ lub } n \geq 2M - 1 \\ \int_0^{\infty} \varphi(x - n) \frac{d\varphi}{dx} dx & gdy \quad 1 \leq n \leq 2M - 2 \end{cases}$$

$$\varphi_{m,n}(x) = 2^{\frac{m}{2}} \varphi\left(2^m x - n\right), m \in Z, \ n \in \aleph,$$

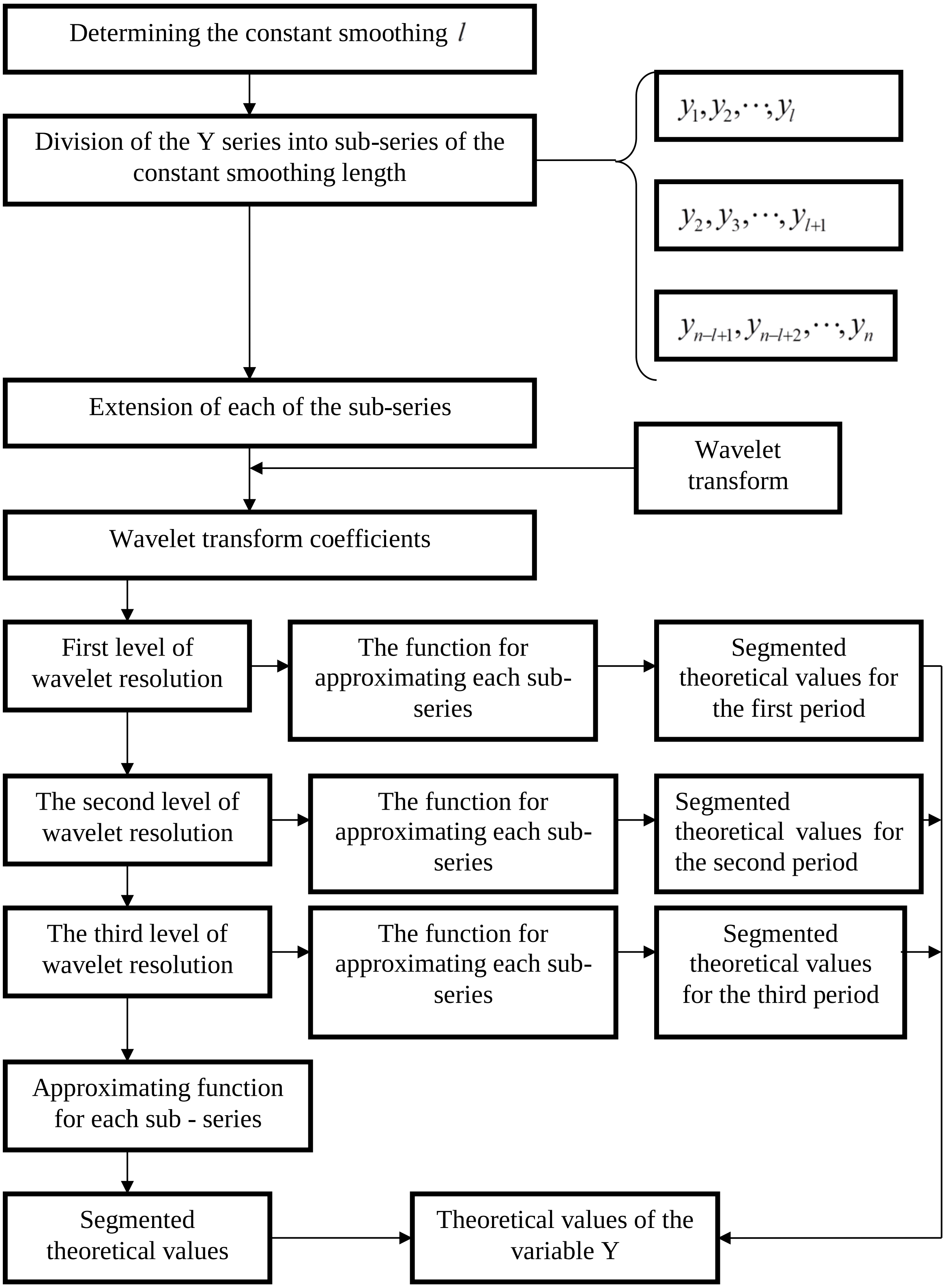
$$\aleph = \left\{n \in Z : \sup p\varphi_{m,n} \cap (0,1) \neq \varnothing\right\},$$

$$f_{-n} = -f_n$$

2.1. Model description

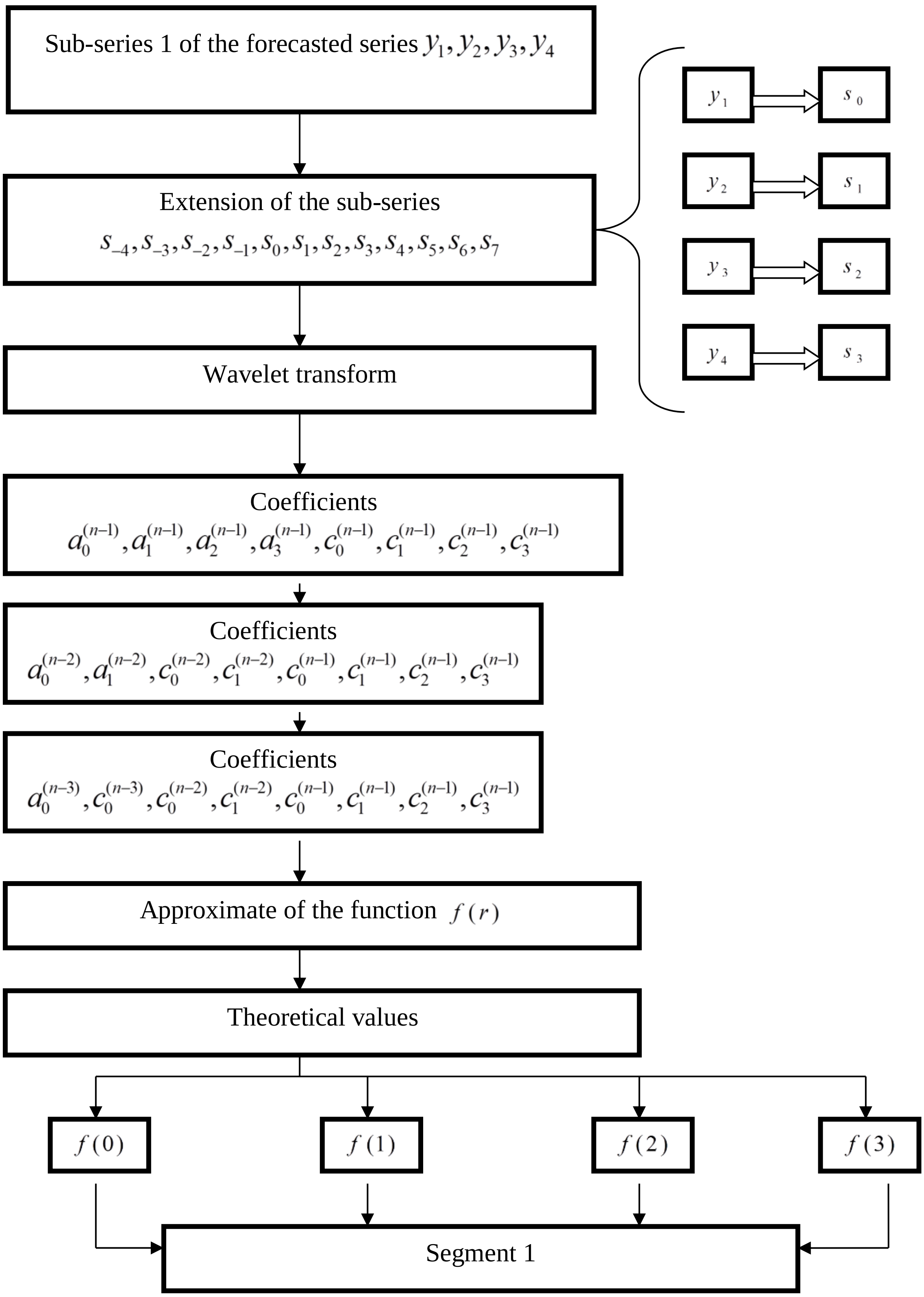
To the model on the input, we introduce a time series, which is a one-dimensional series, n-elements: $y_1,\ y_2,\ y_3,\ \ldots,\ y_n$. Introduced to the algorithm, the series is subjected to wavelet's decomposition, discrete wavelet. If necessary, the series is expanded, one of the commonly used methods or an original method. The decomposition of a series, or actually the level of decomposition, is subjective in this case. The longer the series, the higher the level of wavelet decomposition is possible. As it is known, the purpose of the time series decomposition, in general, is to: isolate the trend - illustrating the operation of the main causes; isolation of seasonal fluctuations - illustrating the effects of seasonal causes; consists in determining seasonal indicators, also known as periodicity indicators or seasonal fluctuation indicators; and the isolation of random fluctuations - illustrating the actions of accidental causes; the effect of their impact is expressed by the standard deviation of the residuals (calculated accordingly). The purpose of the decomposition of the series in the presented model are not only the abovementioned issues, but also saving an approximation function. In the model, we write an approximation function, appropriate for each level of resolution. With the approximating function, we follow the idea of the crawling trend method. As is known, in the classic approach on entering the crawling trend method we have a series of time observations for periods t. Then we determine the length of the segment of the sequence constituting k periods. For each string, we determine the linear trend using the classic method of least squares. In the next step, we calculate the theoretical values for each of the periods based on linear trend equations. For period 1, one calculation value will occur, for the second one two, because of the first and second sequence. However, for the third sequence, up to (n - k + 1) a string of three theoretical values for a given period. In order, we calculate the arithmetic means from the theoretical values for individual n periods. However, in the approach proposed in this model, theoretical values are determined based on the approximating function - wavelet function. The base is a wavelet approximation. In the subsequent stages, we follow the diagram shown in Figure 1 (Fig. 1 presents a general scheme of the algorithm, and its detailed description is presented in the following figures. Figure 2 details the modifications made in the adaptive prediction method of the series, which is the crawling trend metod). In the model, despite the predicted prediction method, the prediction based on the smoothed series can be performed by various methods.

Fig.3 Detailed description of the stage "Creep-waving trend application".



Source: Own elaboration.

Fig. 4 The general scheme of the algorithm.



Source: Own elaboration.

3. Empirical analysis

The globalization theme was chosen for the application of the algorithm. Globalization is a very complex process, difficult to interpret unambiguously (see Hadaś-Dyduch, 2017). Numerous definitions indicate that it can be seen in several ways:

- as the most advanced form of international business,
- an increase in various types of connections between entities of international life,
- a form of mutual, mostly asymmetric influence in all spheres of social life, from international markets to culture,
- the state of the global economy, in which processes of internationalization and integration are intensifying,
- a certain system of management, penetration of structures, the growing role of international organizations and the growing importance of science, technology and information.

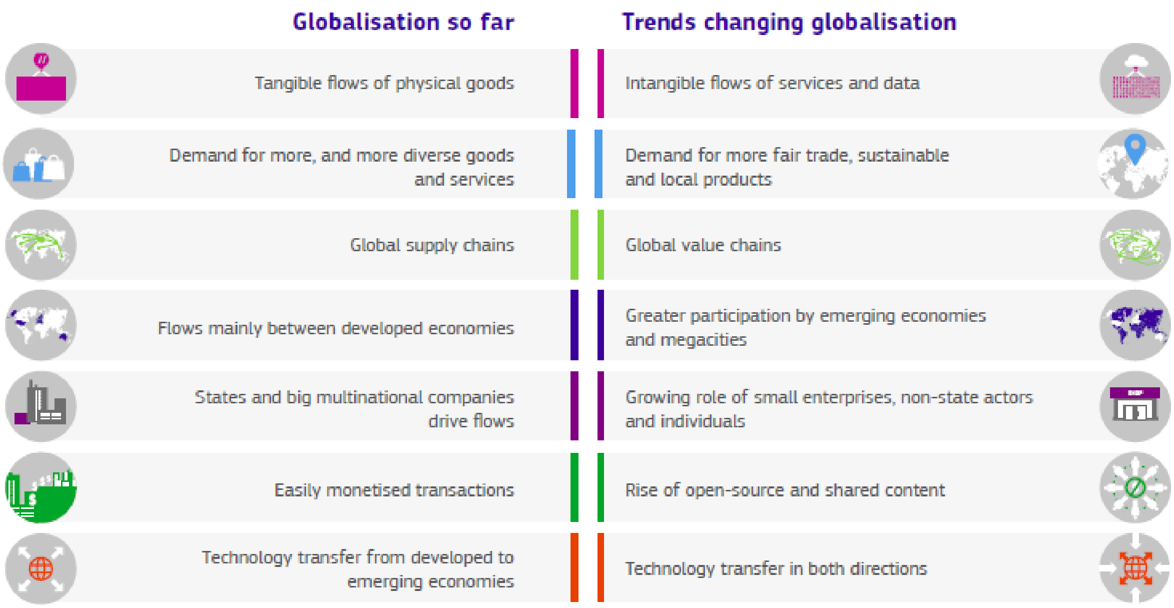
Globalization is a process that applies to everyone (see Fig. 5). There is a constantly growing process. According to European Comision virtually all sectors will be changed, such as (see Fig. 5, Fig. 6):

- transport with e.g. driverless and connected cars, drones and car-sharing;
- energy with e.g. smart grids, renewable energy and distributed generation;
- agri-food with e.g. climate-friendly farming and applications to reduce food waste;
- telecommunications with e.g. more powerful networks, virtual reality and virtual workspace;
- distribution with the growing importance of e-commerce;
- financial services with e.g. virtual banks and insurance and crowdfunding;
- factory production with automation;
- health care with e.g. online diagnosis and increased cross-border mobility of medical professionals.

Fig. 5 Global is today's reality

Source: McKinsey Global Institute, United Nations, World Tourism Organization, OECD, European Commission.

Fig. 6 Globalization then and now.



Source: McKinsey Global Institute, United Nations, World Tourism Organization, OECD, European Commission.

3.1. Materials

The study was based on the social globalization index. . Globalization in the socio-cultural dimension is the result of the development of mass tourism, increased migration, commercialization of cultural products and the spread of consumerism ideology. Social globalization consists of interpersonal, informational and cultural globalization, each of them contributing a third to the social globalization index. The indicator is a component with equivalent weights:

- interpersonal globalization,
- informational globalization,
- cultural globalization.

Interpersonal globalization includes:

- international voice traffic,
- transfers,
- international tourism,
- migration.

Informational globalization includes:

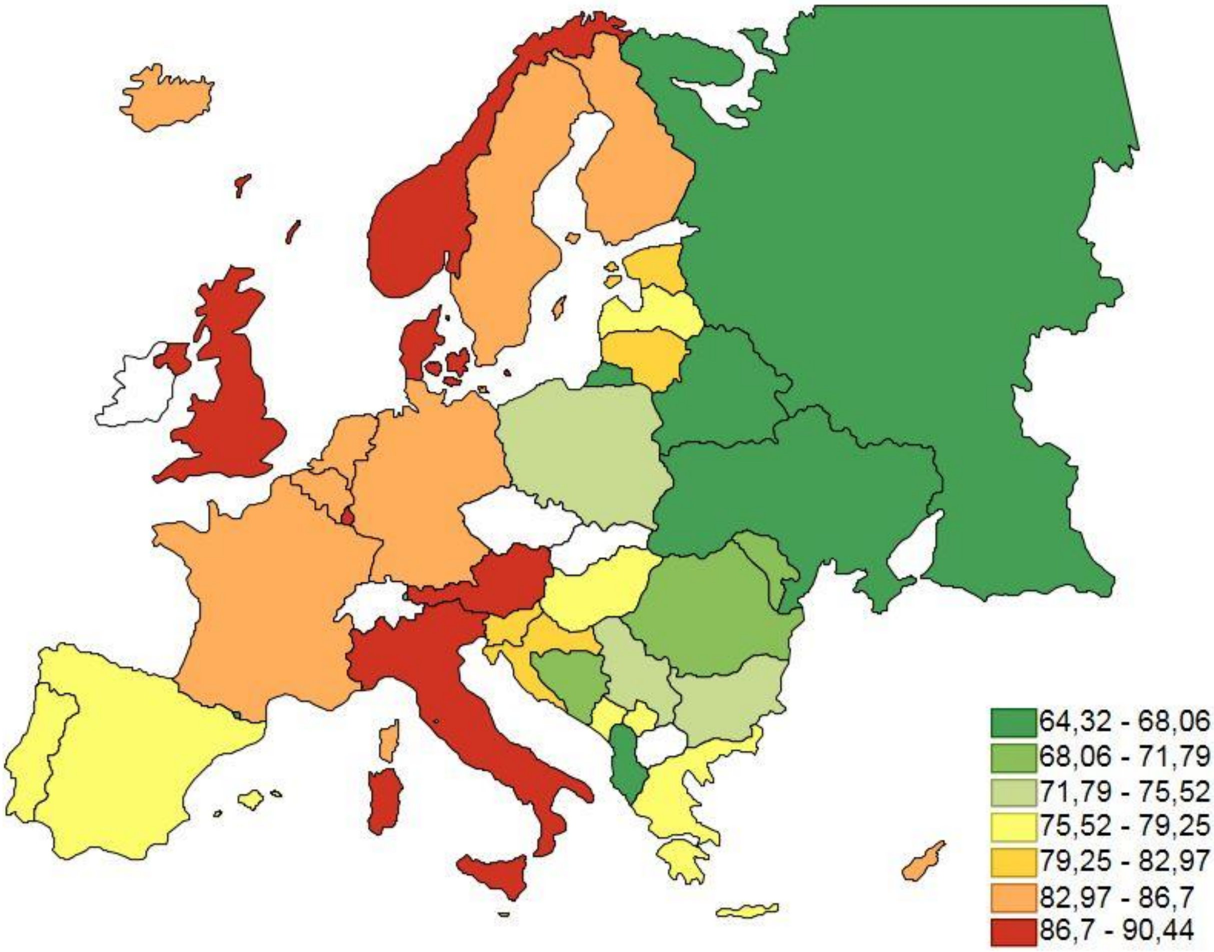
- patent applications,
- international students,
- high technology exports.

Cultural globalization includes:

- trade in cultural goods,
- trademark applications,
- trade in personal services,
- McDonald's restaurant,
- IKEA stores.

The index of social globalization in the world since 1970, increases on average by 0.97% annually. However, since 2008, it has increased on average by 0.61% annually. Currently, the highest value of the social globalization index is in the following countries: Albania 64.33; Ukraine 65.40; Belarus 65.70; Russia 65.81; Andorra 66.88; Moldavia 68.06; Bosnia and Hercegovina 68.25; Liechtenstein, 68.47; Romania 70.05; Serbia 72.08 (see Fig. 7). The lowest value of the social globalization index is in Sweden 85.68; Cyprus 85.81; Finland 85.85; Belgium 86.29; Austria 86.75; United Kingdom 88.05; Italy 88,12; Denmark 88.30; Luxembourg 89.89; Norway 90.43.

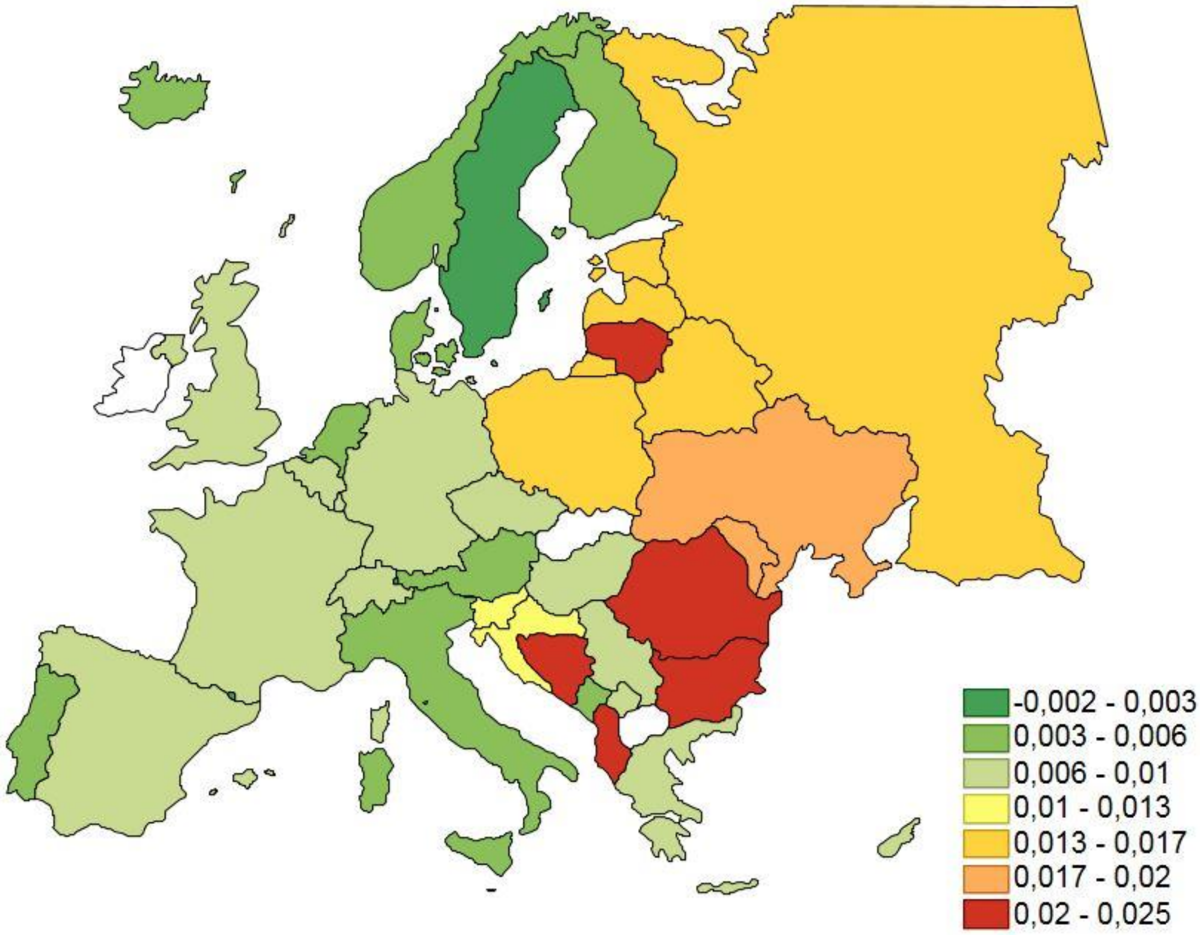
Fig. 7 Globalization index



Source: Own study based on data received from the Swiss Economic Institute.

The top ten countries with the highest growth rate of the social globalization index are: Estonia 1.43%, Russia 1.57%, Belarus 1.68%, Ukraine 1.77%, Moldavia 1.90%, Bulgaria 2.03%, Albania 2.09%, Lithuania 2.18%, Romania 2.20%, Bosnia and Hercegovina 2.41%. However, the countries with the lowest growth rate of the social globalization index are: Liechtenstein -0.10%, Andorra 0.04%, Sweden 0.29%, Denmark 0.41%, Ireland 0.44%, Austria 0.46%, Norway 0.49%, Iceland 0.49%, Portugal 0.55%, Finland 0.57% (see Fig. 8).

Fig. 8 The pace of change in social globalization since the 1970s.



Source: Own study based on data received from the Swiss Economic Institute.

4. Results

The proprietary model described above was used to predict the series presenting the social globalization index of countries belonging to the European continent. The prediction results can be made dependent on three factors: the method of mitigating edge effects, the alpha parameter and the length of the smoothing constant. Each of the above factors has a significant impact on the forecast, its effectiveness and error. The analysis includes the following methods of stretching the time series: Periodization, Zero Padding, Symmetrization (half-point), Symmetrization (whole-point), Antisymmetrization (half-point), Antisymmetrization (whole-point), Smooth Padding of order 1, Smooth Padding of order 0. Each method has its advantages and disadvantages influencing the value of the forecast. For example, the Periodic method is fast, and the Zeros method is very fast and accurate. In the presented proprietary prediction algorithm, we assume the value of the alpha parameter, which minimizes the error of expired forecasts. Table 1 shows the errors of one-period prediction. The errors obtained allow to conclude that the estimates estimated on the basis of the proprietary algorithm are permissible.

Tab. 1 Calculation results.

Extension	RMSPE
Periodization	1,1%
Zero Padding	1,7%
Symmetrization (half-point)	1.2%
Symmetrization (whole-point)	0,9%
Antisymmetrization (half-point)	2%
Antisymmetrization (whole-point)	0,8%
Smooth Padding of order 1	0,5%
Smooth Padding of order 0	0,7%

Source: Own elaboration.

In all the variants of forecast estimation presented, the smallest one-period prediction error, regardless of the length of the smoothing constant, was obtained for the extension of the series to determine the smooth transform coefficients of the "Smooth Padding of order 1" method. This method, ie. "Smooth Padding of order 1", assumes that signals or two can not be recycled to the simple first-order derivative extrapolation. The biggest prediction error for one period ahead. regardless of the length of the constant smoothing, we obtained using the method of stretching a series of methods: "Antisymmetrization (half-point)" and "Zero Padding".

4. Conclusions

The article presents a mixed model as a combination of wavelet anlise with econometric models. The purpose of the model is to predict social phenomena, for which there is no point in being predicted based on models based on archival data or it is impossible. The obtained results from the example application of the model for the prediction of social globalization are appropriate, with a small error.

References

Beylkin G., (1992), *On the Representation of Operators in Bases of Comapactly Supportef Wavelets*. SIAM J. Numer. Anal., tom 6, nr 6..

Bruzda, J. (2013). Prognozowanie metodą wyrównywania falkowego. Acta Universitatis Nicolai Copernici..

Freisleben B., Ripper K., (1997), Volatility Estimation with a Neural Network, IEEE Computional Inteligence for Financial Engineering.

Hadaś-Dyduch M., (2015a), Wavelets in prediction. Theory, Method, Simulation. Scholars' Press, ISBN 978-3-639-76844-2

Hadaś-Dyduch, M. (2015b). Prediction of wavelets analysis. In *Financial management of Firms and Financial Institutions, Proceedings (Part I.) 10th International Scientific Conference* (pp. 341-348).

Hadaś-Dyduch, M. (2016a). Econometric-wavelet prediction in spatial aspect. In *The 10th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings. Cracow: Foundation of the Cracow University of Economics* (pp. 45-52).

Hadaś-Dyduch M. (2016b). *Model - Spatial Aproach to Prediction of Minimum Wage*. [in] Proceedings of the International Scientific Conference Quantitative Methods in Economics Multiple Criteria Decision Making XVIII. Vratna: Letra Interactive, pp. 134-138.

Hadaś-Dyduch, M. (2016c). Predicting Changes in the Labor Market in Global Aspect. In *The 10th International Days of Statistics and Economics. Conference Proceedings*.

Hadaś-Dyduch M., (2017), *Analysis of globalization by wavelet through the prism of the selected euro area countries, 17th International Scientific Conference Globalization and Its Socio-Economic Consequences, University of Zilina, 4th–5th October 2017* Rajecke Teplice, Slovak Republic, *pp.605-612 (part II)*.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons.

Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J. M. (Eds.). (2013). *Wavelets and their Applications*. John Wiley & Sons.

Palit A.K., Propovic D., (2005), Computipnal Inteligence in Times Series Forcasting Theory and Engineering Applicatons. Springer.

Pełech-Pilichowski, T., & Duda, J. T. (2008). Adaptacyjne algorytmy detekcji zdarzeń w szeregach czasowych. *Rozprawa doktorska, WEAIiE AGH, Kraków.*

Ziółko M., (2000), Modelowanie zjawisk falkowych, Uczelniane Wydawnictwo Naukowo-Dydaktyczne, AGH, Kraków.