IOSR Journal of Business and Management (IOSR-JBM)
e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 19, Issue 10. Ver. IX. (October. 2017), PP 59-66
www.iosrjournals.org

Wavelet Daubechies as a tool supporting stock index prediction in the author's multi-component and multi-stage model

Monika Hadaś-Dyduch, Adam Hadaś
University of Economics in Katowice, Department of Statistical and Mathematical Methods in Economics,
[1] Maja 50, 40-287 Katowice, Poland.

Corresponding Author: Monika Hadaś-Dyduch

Abstract: *This article presents the author's series model to predict the stock exchange. The algorithm is based on the artificial neural networks and multi-resolution analysis. However, the main feature of the algorithm, which gives a good quality of the forecasts, is all included in the division of the series analysis, a few partial under-series and prediction dependence of a number of other economic series. The algorithm used for prediction is an author's algorithm, labeled M.H-D_w1 in this article. When choosing the series for the model one should be led by the principle that the forecast index changes are dependent, to some extent, on the basic changes with some time delay. A feature of the algorithm, which allows to achieve good results, is the division of the series on sub-series and sliding time window. Application of the algorithm was performed on a series of presenting WIG. The WIG forecast was dependent on trading the Dow Jones, DAX, Nikkei, Hang Seng, taking into account the sliding time window, for example the application of the author's algorithm, there were appointed the WIG forecast for a period of two years, one year and six months.*

Keywords: *forecasting, neural network, stock index, wavelet, wavelet transform.*

Date of Submission: 16-10-2017 Date of acceptance: 02-11-2017

I. Introduction

The algorithm for the prediction of time series presenting the stock market indices, is based on neural networks and the wavelet analysis, Daubechies wavelets. However, the main feature of the algorithm is to divide the analyzed series into several partial under-series and prediction dependence of a number of other stock series with the appropriate sliding window of time.

Definition 1. *Wavelets we call function $\Psi(x) \in L^2(\mathbb{R})$, such that the system functions:*

$$B_\Psi = \left\{2^{\frac{j}{2}}\Psi\left(2^j x - k\right)\right\}; \quad j \in \mathbb{Z}, \quad k \in \mathbb{Z}$$

is an orthonormal basis in the space $L^2(\mathbb{R})$. Family B_Ψ will be called wavelet base.

The simplest wavelet is the Haar wavelet. "(…) In mathematics, the Haar wavelet is a sequence of rescaled 'square-shaped' functions which together form a wavelet family or basis. Wavelet analysis is similar to Fourier analysis in that it allows a target function over an interval to be represented in terms of an orthonormal function basis. The Haar sequence is now recognised as the first known wavelet basis and extensively used as a teaching example (…)" [1]. The Haar sequence was proposed in 1909 by Alfréd Haar. Haar used these functions to give an example of an orthonormal system for the space of squareintegrable functions on the unit interval [0, 1]. The study of wavelets, and even the term "wavelet", did not come until much later. As a special case of the Daubechies wavelet, the Haar wavelet is also known as D2 [2].

Definition 2. *Haar wavelets we call a function on the real line $\mathbb{R}$ defined by the formula:*

$$H(x) = \begin{cases} 1 & for \quad x \in \left[0, \frac{1}{2}\right) \\ -1 & for \quad \left[\frac{1}{2}, 1\right) \\ 0 & for \quad other\ x \end{cases}$$

Daubechies' wavelets first row (db1) is the Haar wavelets (Daubechies wavelets is wavelets created by Ingrid Daubechies6 in 1988 year).

II. Literature Review

Prediction of stock market indices can be determined on the basis of various models, for example, they may be determined on the basis of the forecast models based on autoregression VAR models, and Factor-Augmented Victor Autoregression (FAVAR). In addition, a tool providing synthetic information is a dynamic factor model Dynamic Factor Model (DFM) [3]. The technique of combining information from a large data set using factor models is used in macroeconomic analysis to solve many fundamental research issues. Examples include inference of a synthetic state of the market or the economy based on disaggregated data [4] and modeling monetary policy reaction to information from a large data set [5]. One of the methods of forecasting and series analysis is wavelet transform [6, 7]. The starting point for the analysis of wavelet analysis is multi-resolution analysis. Generally, the multi-resolution analysis is implemented based on Mallat's algorithm [8], which corresponds to the computation of the discrete wavelet transform. Several approaches have been proposed for time-series prediction by the wavelet transform, based on a neural network [10]. In views [9, 10] and [11], the undecimated Haar transform was used. This paper proposes a new combined prediction, using Daubechies wavelet with a sliding time window. Moreover, in contrast to previous work, a division of series into under-series is proposed. Interesting applications of wavelet analysis in economics are described in [12, 13, 14, 15, 16, 17, 18].

III. Method

Proposed a series prediction model M.H-D_w1 is composed of several essential steps:

- preparation of the series (Fig. 1),
- determine the length of the forecast (Fig. 1),
- application range of sliding time window (Fig. 1),
- division series at sub-series (Fig. 1),
- wavelet coefficients of sub-series (Fig. 2),
- wavelet coefficients of sub-series prediction (Fig. 3),
- prediction (Fig. 4),
- errors.

At the input into the model series $A_1, A_2, A_3, \cdots, A_n$ are introduced. The forecast is made for a specific time interval at the start of the series. The projected time period is determined by the length created at the start of the sub-series and the application of the sliding time window.

The first stage shown in Figure 1, is a step of preparing data for the model. During this stage all the operations that precede the process of building the model are proceeded including the standardization of the time series. This step involves examining the accuracy and nature of raw data and their operationalization. All the data must be transformed into a suitable form if necessary.

The series of $A_1, A_2, A_3, \cdots, A_n$ introduced to the first stage are not equipollent. Thus, time standardization of the series leading to comparable time series and equinumerosity should be performed. The result is equipotence and comparability of the time all included in the model of the series stock exchange.

Then, there is a shift time series data A_n with a t interval. This procedure makes the series: $A_1, A_2, A_3, \cdots, A_{n-1}$ introduced into the model are equinumerous. While the series of A_n is the number of observations reduced by the length of the forecast. The series $A_1, A_2, A_3, \cdots, A_{n-1}$ contains the $2N$ observation. The series A_n contains $2N - t$ observation, where: t - the length of the forecast, $2N = m$, m - number of observations in the series of the input. Thus, the series predicted, on initial observations $x_1, x_2, \ldots, x_m$, after moving has the form: $x_{t+1}, \quad x_{t+2}, \quad \ldots, \quad x_m$. You can write down for a series of A_n:

$$\underbrace{x_{t+1} \quad x_{t+2} \quad \cdots \quad x_m}_{data} \quad \underbrace{x_{m+1} \quad x_{m+2} \quad \cdots \quad x_{m+t}}_{\text{searched forecast}}$$

In the next step, the prepared series $A_1, A_2, A_3, \cdots, A_n$ are divided into sub-series of even numbers of observations. There are many possibilities, you can create several series of two-elements, four-elements, eight-elements, sixteen-elements and so on (Fig. 1).

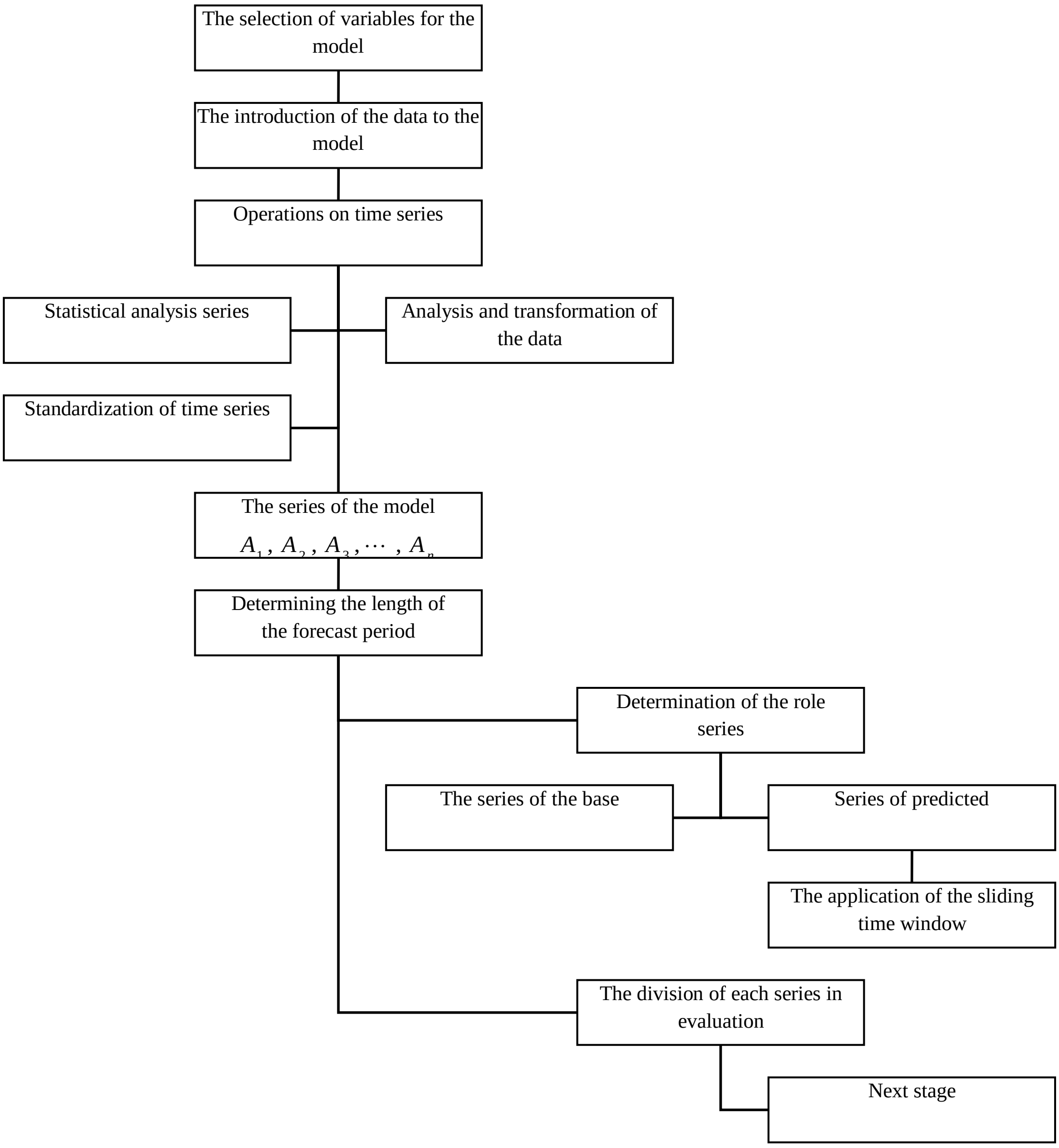

Figure 1 - Scheme of the first, second, third, fourth stage of the model

In stage two, properly prepared ranks, or more precisely under-series are subject to the operation of wavelet transform, after fixing the coefficients of wavelet transform. As a result of the wavelet transform for each under-series we obtain wavelet coefficients of selected under-series on different levels of resolution, which are necessary in learning process about artificial neural network.

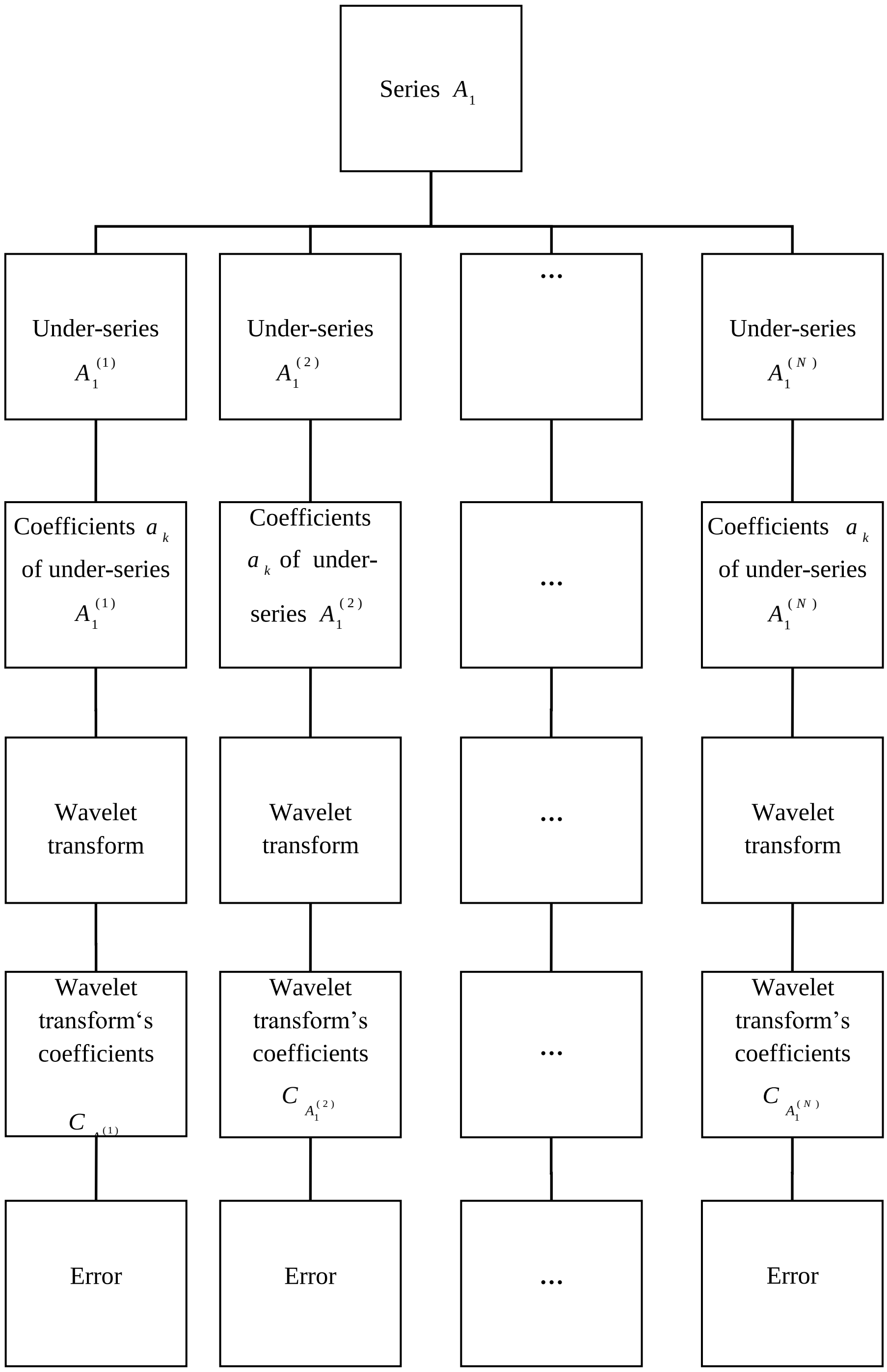

Figure 2 - Scheme of the fifth stage of algorithm for one series - A_1 .

In the next stage of the model we initialize the artificial neural network and act in accordance with the diagram shown in Figure 3. The starting point of the initialization of an artificial neural network is a division of the data into the training set and a test. The adopted breakdown of the data on these collections was arbitral, however, consistent with the rule that the training set is the most numerous and the manner of assigning the following items to the collections of the learner and the test is the same for each series of the data. At the output of an artificial neural network we get the coefficients of the wavelet transform for the future observations of the test series. The wavelet coefficients obtained via the inverse wavelet transform operation gain the values of the real numbers, i.e. the numbers of future values for a pointed time interval forecasts (Fig. 4).

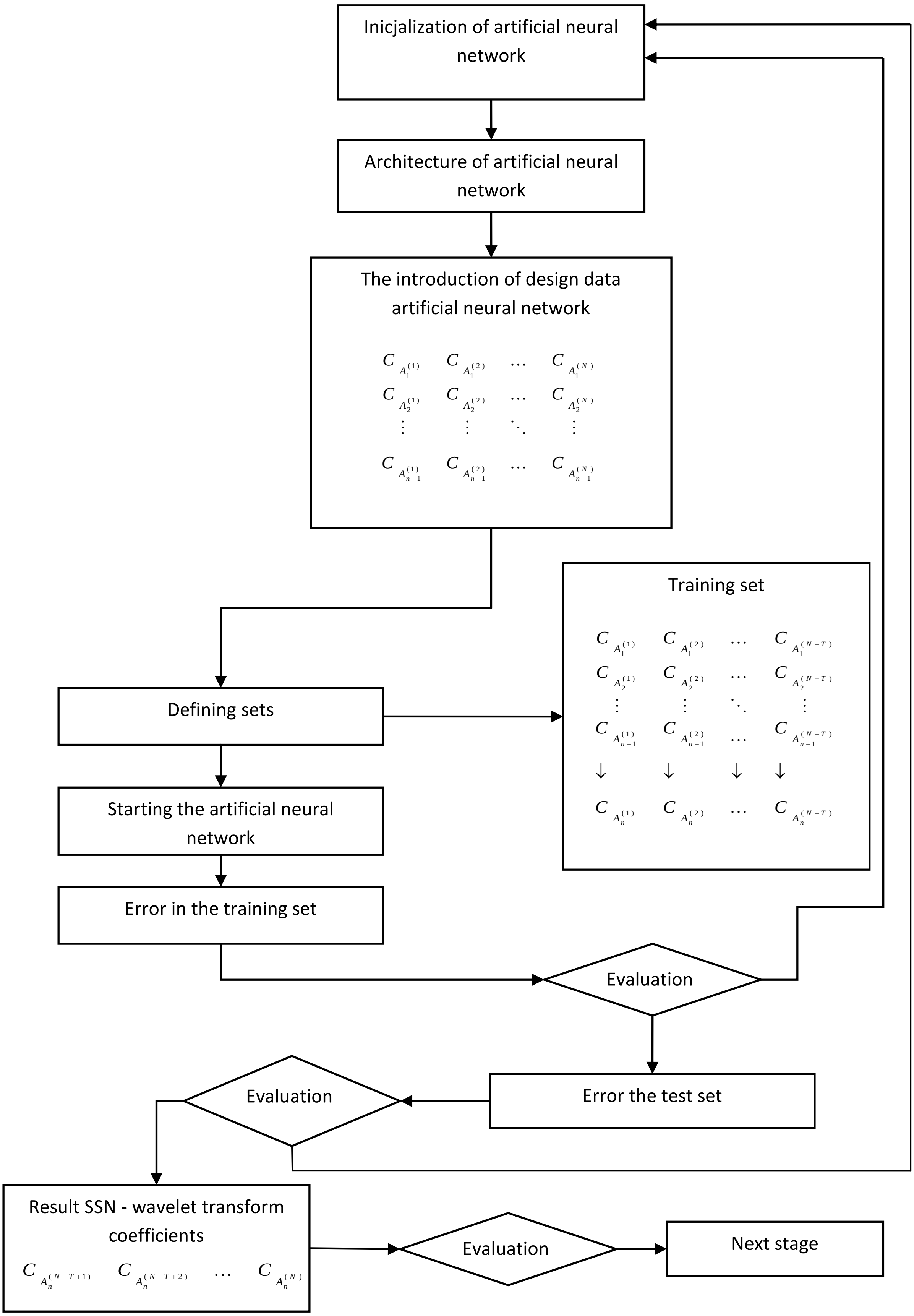

Figure 3 - Scheme of the sixth stage of the algorithm M.H-D_w1

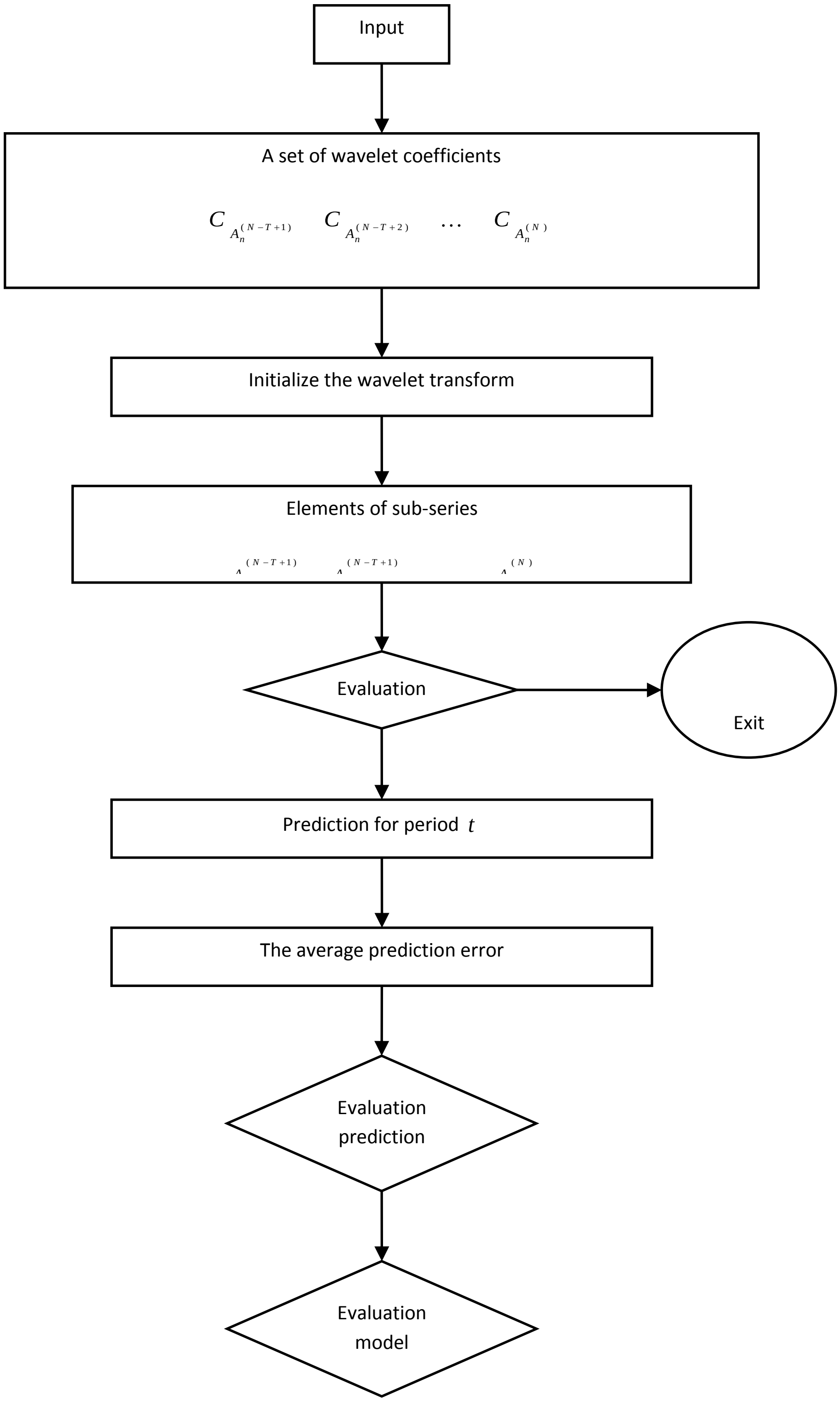

Figure 4 - Scheme of the last stage of the model

IV. Empirical Material

To the application of the proposed proprietary model such series stock exchanges as Dow Jones, DAX, Nikkei, Hang Seng and WIG were selected. The aim of the study is the WIG index prediction model. WIG index prediction is made based on the series Dow Jones, DAX, Nikkei, Hang Seng. The series was a random choice, but consistent with the principle that changes in the WIG index are dependent to some extent on changes in Dow Jones, DAX, Nikkei, Hang Seng, with a time delay. Ranks included in the study cover a period of more than 11 years. The ranks of Dow Jones, DAX, Nikkei, Hang Seng and WIG included in the study are daily quotations from period 23/04/1996 - 09/12/2016. The individual series have the following number of observations: Dow Jones – 5144 observations, DAX – 5166 observations, Nikkei – 5021 observations, Hang Seng – 5080 observations, WIG – 4624 observations.

Since the series selected for the study were equinumerous, the standardization of the time series has been listed. After the standardization the series are equipollent. Each series comprised 4116 observations. The aim of the study is to forecast WIG for a period of 6 months and 1 year. Thus, the number of WIG has two windows applications backward: the annual and semi-annual. In the first case, the first observation of a series of WIG is an observation about half a year later than the first observation of each series of Dow Jones, DAX, Nikkei, Hang Seng. In the second case, the first observation of a series of WIG is an observation about a year later than the first observation of each series of Dow Jones, DAX, Nikkei, Hang Seng.

V. Calculation – Application Of The Author's Model

According to this model the series the Dow Jones, DAX, Nikkei, Hang Seng, WIG should be divided into sub-series. Ranks can be divided into any number of series that are multiples of 2. The study adopted the division into 2-element series. In the division series: Dow Jones, DAX, Nikkei, Hang Seng down into 2-element sub-series receives the 2058 2-element sub-series of each series.

For each sub-series, formed from a series of original, wavelet coefficients were calculated, and then having values of wavelet coefficients for each sub-series artificial neural network was initialized. Wavelet transform coefficients were determined by applying Daubechies wavelet:

$$\psi(r) = -\frac{1+\sqrt{3}}{4}\varphi(2r-1) + \frac{3+\sqrt{3}}{4}\varphi(2r) - \frac{3-\sqrt{3}}{4}\varphi(2r+1) + \frac{1-\sqrt{3}}{4}\varphi(2r+2)$$

where:

$\psi(r) = 0$ for $r < -1$ or $r > 2$

$\varphi(r) = h_0 \cdot \varphi(2r) + h_1 \cdot \varphi(2r-1) + h_2 \cdot \varphi(2r-2) + h_3 \cdot \varphi(2r-3)$

$$h_0 = \frac{1+\sqrt{3}}{4},\ h_1 = \frac{3+\sqrt{3}}{4},\ h_2 = \frac{3-\sqrt{3}}{4},\ h_4 = \frac{1-\sqrt{3}}{4}.$$

$$\sum_{k \in Z} \varphi(k) = 1,\ \varphi : D \to R,\ \varphi(r) = 0 \text{ for } r \le 0 \vee r \ge 3,\ D_j = \{k2^j : k \in Z\},\ D = \bigcup_{j \in Z} D_j = \bigcup_{j=0}^{\infty} D_j.$$

As the input of the artificial neural network wavelet coefficients of the appropriate under-series of each of the five considered series stock exchange were taken That is, the network was taught on archived data Dow Jones, DAX, Nikkei and Hang Seng shifted in time of two years, a year, six months in relation to WIG. At the input of artificial neural network there are given semi-annual forecasts: wavelet coefficients under-series received from the ranks of the Dow Jones, DAX, Nikkei, Hang Seng, shifted with half a year, and wavelet coefficients under-series algorithm, therefore, M.H-D_w1 wavelet coefficients generated a series of WIG for a specified period of the forecast. The study was divided into sets of learners and the test was considered taking into account the percentage of the expected length of the forecast. Two strategies were adopted:

- for the annual WIG forecast :
  - a training set 94.19%,
  - a set of test 5.81%,
- for the six-month WIG forecast:
  - a training set 97.10%,
  - a set of test 2.90%,

placed into the input data to the neural network.

VI. Result

Average absolute percentage error of wavelet coefficients (the division series listed on two-factor under-series with one level of resolution, the network designed for 70 hidden layers):

- annual forecast for WIG:
  - for the test set: 0.00024%,
  - for the output file: 0.0027 %,
- for the six-month forecasts WIG:
  - for the test set: 0.368%,
  - for the output file: 0.005%.

Having generated the coefficients of the wavelet transform for the future value of the WIG index for the highlighted time periods (one year, half a year,) the algorithm was used for an inverse wavelet transform. The result of the inverse wavelet transform, Daubechies wavelets were future values, i.e. the value of the forecast range of WIG respectively for a period of one year and a half year.
An average absolute percentage error of each forecast was:

- 0.055% - for the annual WIG forecast ,
- 0.179% - for the six-month WIG forecasts.

VII. Conclusion

The paper presents an original method for time series forecasting based on artificial neural networks and wavelet transform - wavelet Daubechies, including a sliding time window. It also analyzes the distribution of ranks for under-series *n* - elements. The presented results show that the use of a model based on wavelet analysis and artificial neural networks is justified in the light of the analyzed data. The results show that the proposed M.H-D_w1 algorithm can be used for a long term prediction, as obtained forecast errors are relatively small. They are in the range from 0.055% to 0.179%. In comparison to other time series models, such as AR, MA, or ARMA, the precision of prediction is not a decline trend when the forecasting scale is extended. It can be concluded that presented model can be an effective tool for forecasting stock indices and macroeconomic indicators. The prediction is very difficult due to the complexity of the mechanism of these markets, especially the factors affecting the markets.

References

[1] M. Hadaś-Dyduch, Wavelets in prediction. Theory, Method, Simulation. Scholar's Press, Saarbrucken, Germany, 2015.

[2] A. Haar. Zur theorie der orthogonalen funktionensysteme. Mathematische Annalen, 1910, 69.3: 331-371.

[3] J. Stock, H. James, M. WATSON. Forecasting using principal components from a large number of predictors. Journal of the American statistical association, 2002, 97.460: 1167-1179.

[4] M. Forni, M. Lippi. Aggregation and the microfoundations of dynamic macroeconomics. Oxford University Press, 1997.

[5] R. Cristadoro, M. Forni, L. Reichlin & G. Veronese. A core inflation indicator for the euro area. Journal of Money, credit, and Banking, 2005, 37.3: 539-560.

[6] X. Wang, X. Shan. A wavelet-based method to predict Internet traffic. In: Communications, Circuits and Systems and West Sino Expositions, IEEE 2002 International Conference on. IEEE, 2002. 690-694.

[7] K. Papagiannaki, N. Taft, Z. Zhang, & C. Diot. Long-term forecasting of Internet backbone traffic. IEEE transactions on neural networks, 2005, 16.5: 1110-1124.

[8] S. Mallat. A wavelet tour of signal processing. Academic press, 1999.

[9] G. Zheng, J. Starck, J. Campbell, & F. Murtagh. Multiscale transforms for filtering financial data streams. Journal of Computational Intelligence in Finance, 1999, 7.18-35.

[10] M. Hadaś-Dyduch, Artificial neural networks as one of the methods to alleviate edge effects in wavelet analysis of macroeconomic indicators, 8th International Scientific Conference "Analysis of International Relations 2017. Methods and Models of Regional Development", 2017, 21-27.

[11] S. Soltani, D. Boichu, P. Simard, & S. Canu. The long-term memory prediction by multiscale decomposition. Signal Processing, 2000, 80.10: 2195-2205.

[12] T. Joo, S. Kim, Time series forecasting based on wavelet filtering. Expert Systems with Applications, 2015, 42.8: 3868-3874.

[13] M. Hadaś–Dyduch, Wavelets as Basis Functions in the Adaptation's Methods: Author's Model for Forecasting Short-Term, Chinese Business Review, 15(1), 2016. 8-18.

[14] M. Hadaś-Dyduch, A. Balcerzak, M. Pietrzak, Wavelet analysis of unemployment rate in Visegrad countries, Globalization and Its Socio-Economic Consequences, 16th International Scientific Conference, Conference Proceedings, University of Zilina, The Faculty of Operation and Economics of Transport and Communication, Department of Economics , 5th – 6th October 2016, 595-602.

[15] M. Hadaś-Dyduch, Econometric-wavelet prediction in spatial aspect. The 10th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings. 2016. 45-52.

[16] M. Hadaś-Dyduch, Alignment wavelets as main instruments in the short-time term prediction. Hradec Economic Days. Double-blind peer reviewed proceedings of the international scientific conference Hradec Economic Days. 2016. 62-68.

[17] M. Hadaś-Dyduch. Non-classical algorithm for time series prediction of the range of economic phenomena with regard to the interaction of financial market indicators. Chinese Business Review, 2014, 13.4.

[18] M. Hadaś-Dyduch.. Wavelets in the prediction of short-time series. Mathematical Economics, 2015, 11 (18): 43-54.

Monika Hadaś-Dyduch, Adam Hadaś. "Wavelet Daubechies as a tool supporting stock index prediction in the author's multi-Component and multi-Stage model." IOSR Journal of Business and Management (IOSR-JBM) , vol. 19, no. 10, 2017, pp. 59–66.