Jerzy Korczak
Department of Computer Science
University of Economics
Wrocław, Poland

Krzysztof Drelczuk
Department of Computer Science
University of Economics
Wrocław, Poland
kdrelczuk@hotmail.com

# EFFECT OF WAVELET COMPRESSION OF HIGH FREQUENCY TIME SERIES ON THE QUALITY OF INFORMATION AND PREDICTION

**ABSTRACT**

*In recent times the research works on the use of wavelet theory in data mining have increased significantly. In most cases, these works relate to specific applications. In this paper the general compression method of time series will be presented and adapted to financial time series analysis where dimensionality reduction is crucial. The hypothesis that a double compression using Daubechies 4 wavelet does not affect significantly the quality of information carried by a time series. The reduction of dimensionality significantly affects the algorithmic complexity and improve its quality of prediction. In order to verify the hypotheses the highly frequent time series will be evaluated in terms of forecasting quality where future value is predicted only on the basis of the past quotations. In the project as a predictive algorithm was used ARAR because of the good results in forecasting of the real financial time series.*

## 1. Introduction

Empirical evidence shows that the dimensionality reduction affects not only significantly the computing time of the classifiers, but also the quality of classification results. In Euclidean space while increasing number of dimensions diminishes the distance between vectors. This is particularly important impact on the process of clustering. Space of solutions that is divided into clusters has the same dimensionality as entering vectors. In the case of financial time series where data are taken from, for example, sixty recent observations, the space, that the algorithm will have to share, will be 60-dimensional. Clustering consists of assigning elements to the respective clusters by means of pre-defined metric (for example, the Euclidean norm). With a large number of dimensions of difference between the nearest and farthest neighbor is becoming less important and it is a serious obstacle to partition the space into the significant clusters [Beyene,1999]. Too much dimensions can also cause an overlapping multidimensional clusters impeding effective classification.

The hypothesis to prove is that double compression by Daubechies 4 wavelet does not affect significantly the quality of information carried by a time series in comparison to the original, raw time series. In other words, the double wavelet compression does not influence the deterioration of the time series, as the information source. In the case of clustering where the computational complexity is exponential, proving such an assumption has a significant impact on the usefulness of the clustering algorithms. Taking also into account the reduction of distance between clusters in the Euclidean space with increasing dimensionality, this demonstration would positively influence the quality of the prediction of time series.

In this work the authors made use of discrete wavelet transform for lossless compression of time series. Lossless in the sense of preservation of the same quality information as an untreated time series. Successful proving the hypothesis may be used in various systems for time series analysis, such as predictive systems, classifcation systems, archiving systems. In the next part of this paper the wavelet theory will be briefly presented and focused on Daubechies 4 wavelets. Then, the classification algorithm, ARAR, will be presented. In the third part, research methodology will be described and how the selection of time series. In the last part will be collected the results of empirical research on FOREX time series, which have been used to verify the hypothesis.

## 2. Discrete wavelet transform

Discrete wavelet transform, first described in [Mallat 1989], is very often applied in preliminary data analysis. With it one can either reduce the number of dimensions of input vector to the target system such as the classifier or predictive system, as well as remove some of the information considered as noise or data redundancies, in terms of Shanon's lossless data compression [Sharon 1948]. In most cases, the signal or function can be better explored, described and processed, when it can be defined as a linear combination of the form:

$$f(t) = \sum_{\ell} a_{\ell} \psi_{\ell}(t) \tag{1}$$

where $l \in Z$ denotes the index of a sum (finite or infinite), but $\psi_{\ell}$, $a_{\ell} \in \Re$ is a collection of real functions. If this distribution is well-defined then the base set of functions $\psi_t$ is called a functional space (instead of the vector space where the vectors have the features, while the scalars are real or complex numbers). If the scalar product of all the functions $\psi_{\ell}$ is equal to zero, then the base is called orthogonal, and can be written as:

$$\langle \psi_k(t), \psi_{\ell}(t) \rangle = \int \psi_k(t) \cdot \psi_{\ell}(t) dt = 0, \underset{k,\ell \in Z}{\forall} k \neq \ell \tag{2}$$

This allows us to determine the coefficients $a_{\ell}$ using scalar product as follows:

$$a_{\ell} = \langle f(t), \psi_{\ell}(t) \rangle = \int f(t) \cdot \psi_{\ell}(t) dt \tag{3}$$

Thus the defined space, spanned on the functions $\psi_k$, is called Hilbert space $L^2(\Re)$ or the space of squared integrable functions. To recall, the real variable function $f(t)$ belongs to the space $L^2(\Re)$ if and only if $|f(t)|^2$ it is integrable, that is, if:

$$\int_{t \in \Re} |f(t)|^2 dt < \infty \tag{4}$$

with the metric described as:

$$\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int_{t \in \Re} |f(t)|^2 dt} \tag{5}$$

A wavelet is defined as a real function $\psi(u) \in L^2(\Re)$ [Daubechies 1992] if the following constraints are satisfied:

$$\begin{aligned} &\int_{-\infty}^{\infty} \psi^2(u) du = 1 \\ &\int_{-\infty}^{\infty} \psi(u) du = 0 \end{aligned} \tag{6}$$

The wavelet transform will be called a set of functions $\psi_{a.b}$, such as:

$$\psi_{a,b}(t) = w(a) \cdot \psi\left(\frac{t-b}{a}\right)dt \tag{7}$$

where $a$ and $b$ are scaling parameters, *w(a)* is the weighting function to ensure that the wavelet energy does not change with the change of scale, ie. $\|\psi_{a,b}\| = \|\psi\| = a \in \Re^+$. In most cases, in formula (7) $a^{-1/2}$ replaces *w(a)*. Then the wavelet transform becomes:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)dt \tag{8}$$

However, the formula is suitable only for continuous signals. In this paper, the input signal is discrete. To apply this transform it is necessary to sample the input signal. This will be done using a logarithmic discretization of scale $a$ depending on the size of the step between the distances from the next parameter $b$. Such a discrete wavelet transform has the form:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} \cdot \left(\frac{t - nb_0 a_0^m}{a_0^m}\right) \tag{9}$$

where $a_0$ and $b_0$ are parameters defining the orthogonal base of function $\psi_{a,b}$ in space $L^2$. Most popular values of the parameters are 2 and 1 [Daubechies 1992]. We then say a dyadic sampling. The wavelet then takes the following form:

$$\psi_{m,n}(t) = 2^{-m/2} \cdot \psi\left(2^{-m}t - n\right) \tag{10}$$

Distribution of entry signal is named multi-resolution signal [Mallat 1989]. Signal $f(t) \in L^2$ is decomposed into the constituents localized in the sub-spaces spanned on the scaling functions. The scaling function has the same form as the wavelet:

$$\phi_{m,n}(t) = 2^{-m/2} \cdot \phi\left(2^{-m}t - n\right) \tag{11}$$

Signal belonging to the space $L^2$ can be defined by:

$$x_m(t) = \sum_{n=-\infty}^{\infty} S_{m,n} \cdot \phi_{m.n}(t) \tag{12}$$

where approximation coefficients $S_{m,n}$ are defined as follows:

$$S_{m,n} = \int_{-\infty}^{\infty} x(t) \cdot \phi_{m,n}(t)dt \tag{13}$$

The procedures of $\psi$ searching having $\phi$ are well defined in the literature ([Daubechies 1992] chap. 5.1). One possibility, described in ([Daubechies 1992] and applied in our approach, consists of definition of wavelet function in the following way:

$$\psi(t) = \sum_n (-1)^n h_{n-1}\sqrt{2} \cdot \phi(2t - n) \tag{14}$$

Signal included in the compliment of the sub-espace spanned on wavelet functions can be computed from:

$$d_m(t) = \sum_{n=-\infty}^{\infty} T_{m,n} \cdot \phi_{M.n}(t) \tag{15}$$

where the transform is defined as follows:

$$T_{m,n} = \int_{-\infty}^{\infty} x(t) \cdot \psi_{m,n}(t)dt \tag{16}$$

Complete signal in $L^2$ can be obtained from the following expression::

$$x(t)=\sum_{n=-\infty}^{\infty} S_{m,n}\cdot\phi_{m.n}(t)+\sum_{m=-\infty}^{m_0}\sum_{n=-\infty}^{\infty} T_{m,n}\cdot\phi_{m.n}(t) \quad (17)$$

Summing up, the signal is represented as the synthesis of approximation on a given level (12) and sum of details until a given level included (15). Approximations are a compressed time series that carry out the same information in terms of Shanon.

## 3. Algorithm ARAR

The chosen algorithm ARAR is a modification of previous algorithm ARARMA [Newton 1984]. Its characteristic is the application of *memory-shortening* transformation for each time series, and then fitting with a model ARMA [Brockwell 2002]. Suppose the we have to find w a given time series $\{Y_t,\ t=1,\ 2,\ \ldots,\ n\}$ its lengths of time dependency. It exists 3 possibilities:

1. D.Series $\{Y_t\}$ has a long time dependency.
2. W.Series $\{Y_t\}$ has a relatively long time dependency.
3. K.Series $\{Y_t\}$ is short memory.

If it is proven that the time series is a type D or W, then the transformation is executed until the time series becomes of type K.

The specification below describes the algorithm of classification.

**ARAR description**

1. For each $\tau=1,2,3,..,15$ find a value $\hat{\phi}(\tau)$ such that

$$E(\phi,\tau)=\frac{\sum_{t=\tau+1}^{n}(Y_t-\phi Y_{t-\tau})^2}{\sum_{t=\tau-1}^{n} Y_t^2} \quad (18)$$

is minimal

2. Let us define

$$E(\tau)=E(\varphi(\tau),\tau),$$

and delay $\hat{\tau}$ in a such way that value $\tau$, gives minimum of $E(\hat{\tau})$. Then, if:

   1. $E(\hat{\tau})\le 8/n$ time series is a type D.
   2. $\hat{\varphi}(\hat{\tau})\ge 0.93$ and $\hat{\tau}>2$ time series is a type *D*
   3. $\hat{\varphi}(\hat{\tau})\ge 0.93$ and $\hat{\tau}=1\vee 2$ describe the values $\hat{\varphi}_1,\hat{\varphi}_2$, where $\sum_{t=3}^{n}(Y_t-\phi_1 Y_{t-1}-\phi_2 Y_{t-2})^2$ is minimal; time series is a type W.
   4. $\hat{\varphi}(\hat{\tau})<0.93$ time series is a type K.

3. Define the time series after transformation $\{S_t, t=k+1,..,n\}$ and $\bar{S}$ as average of $S_{k+1},\ldots,S_n$. Autoregressive model is

$$X_t=\phi_1 X_{t-1}+\phi_{l_1}X_{t-l_1}+\phi_{l_2}X_{t-l_2}+\phi_{l_3}X_{t-l_3}+Z_t,$$

where $Z_t\sim N(0,\sigma^2)$, and for delays $l_1$, $l_2$ i $l_3$, coefficients $\phi_j$ and variance $\sigma^2$ compute from the Yule-Warker equasion:

$$\begin{bmatrix} 1 & \hat{\rho}(l_1-1) & \hat{\rho}(l_2-1) & \hat{\rho}(l_3-1) \\ \hat{\rho}(l_1-1) & 1 & \hat{\rho}(l_2-l_1) & \hat{\rho}(l_3-l_1) \\ \hat{\rho}(l_2-1) & \hat{\rho}(l_2-l_1) & 1 & \hat{\rho}(l_3-l_2) \\ \hat{\rho}(l_3-1) & \hat{\rho}(l_3-l_1) & \hat{\rho}(l_3-l_2) & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_{l_1} \\ \phi_{l_2} \\ \phi_{l_3} \end{bmatrix} = \begin{bmatrix} \hat{\rho}(1) \\ \hat{\rho}(l_1) \\ \hat{\rho}(l_2) \\ \hat{\rho}(l_3) \end{bmatrix}$$

and

$$\sigma^2 = \hat{\gamma}(0)\left[1-\phi_1\hat{\rho}(1)-\phi_{l_1}\hat{\rho}(l_1)-\phi_{l_2}\hat{\rho}(l_2)-\phi_{l_3}\hat{\rho}(l_3)\right]$$

where $\hat{\gamma}(0)$ and $\hat{\rho}(j)$, *j=0, 1, 2, …*,auto-covariance and auto-correlation of time series $\{Y_t\}$.

Described above algorithm will be served to test and as a reference for comparison of the amount of source information and compressed time-series.

## 5. Compression using Daubechies 4 wavelet

Daubechies wavelets are a family of orthogonal wavelets described in detail in [Daubechies 1992]. Wavelets, labeled D2-D20 (only the even index denotes nonzero coefficients of the scaling functions), are very often used, inter alia, because of very low computational complexity, that is $O(n)$. Comparatively, a widely used Fast Fourier Transformation (FFT) has the computational complexity $O(n \cdot \log n)$.

In our project, the applied wavelet is D4 wavelet with four coefficients of scaling function. The scaling function for $n$=4 can be defined as follows:

$$\phi(t) = \sqrt{2}\left(h_0\phi(2t) + h_1(2t-1) + h_2(2t-2) + h_3(2t-3)\right)$$

and its corresponding wavelet function:

$$\psi(t) = \sqrt{2}\left(h_3\phi(2t) - h_2(2t-1) + h_1(2t-2) - h_0(2t-3)\right)$$

Solving the above equation, assuming the orthogonality of wavelets, a normalizing parameter $\sqrt{2}^{-1}$, the following coefficients are obtained:

$$h_0 = \frac{1+\sqrt{3}}{4} \; h_1 = \frac{3+\sqrt{3}}{4} \; h_2 = \frac{3-\sqrt{3}}{4} \; h_3 = \frac{1-\sqrt{3}}{4}$$

and

$$h_0 = \frac{1-\sqrt{3}}{4} \; h_1 = \frac{3-\sqrt{3}}{4} \; h_2 = \frac{3+\sqrt{3}}{4} \; h_3 = \frac{1+\sqrt{3}}{4}$$

The first set of coefficients corresponds to the scaling function $\phi(t)$, the second one to $\phi(-t)$.

Decomposing the input signal into components, with each step one get rid of a certain amount of Signac interferences. Given the orthogonality space, we can be sure that there is no redundancy and every detail will be eliminated once. The input signal can be represented as the sum of the following details with the last approximation. Signal decomposition is performed recursively using the formula for synthesizing a given level of approximation and the amount of detail to the level of, inclusive. This is illustrated in fig.1.

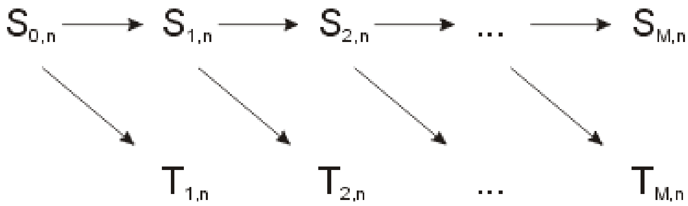


**Fig. 1. Time series decomposition**

Source: [Burrus 2001].

One may notice that $n$ is successively doubled at each iteration; for $m=1$ we have $2^M/2^1=N/2$ approximation coefficients, for $m=2$ there is $2^M/2^2=N/4$. Decomposing a vector of 8 elements ($S_{0,0},S_{0,1},S_{0,2},S_{0,3},S_{0,4},S_{0,5},S_{0,6},S_{0,7}$) after the first iteration a vector ($S_{1,0},S_{1,1},S_{1,2},S_{1,3},T_{1,0},T_{1,1},T_{1,2},T_{1,3}$) is obtained, after the next ($S_{2,0},S_{2,1},T_{2,0},T_{2,1},T_{1,0},T_{1,1},T_{1,2},T_{1,3}$), and so on. This imposes a constraint on the length of entry vector that must be a power of 2. Automatically, it results the length of the compressed vector (after removing details) that is also a power of 2 as illustrated on fig.2.

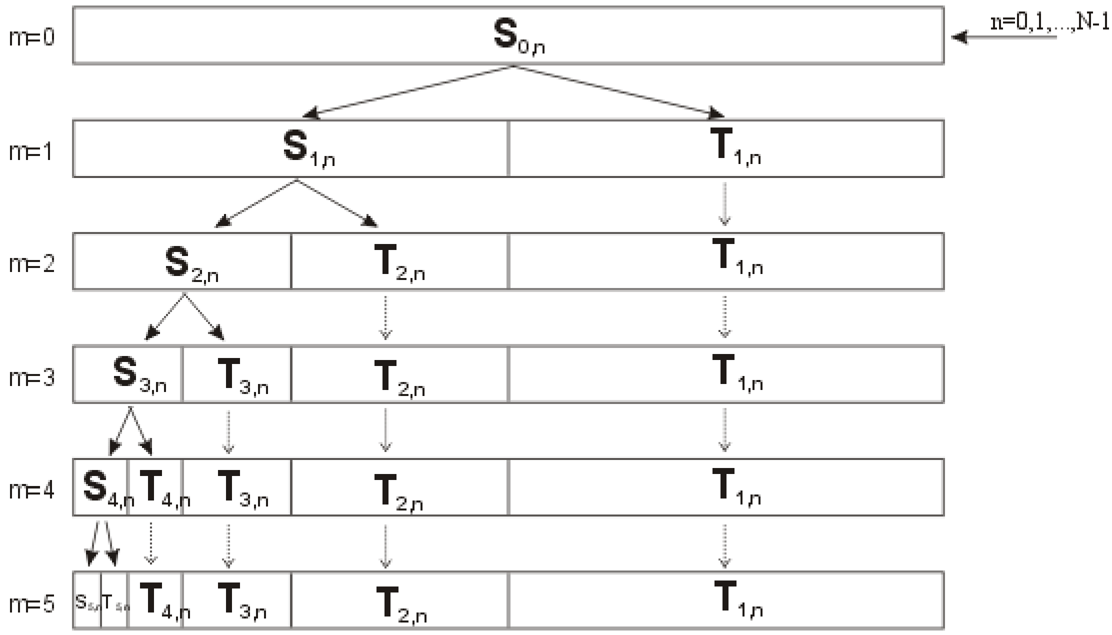


**Fig 2. Schema of vector decomposition**

Source: [Burrus, 2001].

Below the source code of the compression algorithm is presented to show its simplicity and low algorithmic complexity.

**Extract from the source code of discrete wavelet transformation**

```
public static double[] D4Transform(double[] input, int approximatonSize)
{
    double h0 = (1 + Math.Sqrt(3)) / 4 * Math.Sqrt(2);
    double h1 = (3 + Math.Sqrt(3)) / 4 * Math.Sqrt(2);
    double h2 = (3 – Math.Sqrt(3)) / 4 * Math.Sqrt(2);
    double h3 = (1 – Math.Sqrt(3)) / 4 * Math.Sqrt(2);
    double g0 = h3;      double g1 = -h2;
    double g2 = h1;      double g3 = -h0;
    int I = 0, j = 0;
    int half = approximatonSize >> 1;
    double[] tmp = new double[approximatonSize << 1];
```

```
        for (j = 0; j < approximatonSize – 3; j = j + 2)
        {
          tmp[i] = a[j] * h0 + a[j + 1] * h1 + a[j + 2] * h2 + a[j + 3] * h3;
          tmp[I + half] = a[j] * g0 + a[j + 1] * g1 + a[j + 2] * g2 + a[j + 3] * g3;
          i++;
        }
        tmp[i] = a[n – 2] * h0 + a[n – 1] * h1 + a[0] * h2 + a[1] * h3;
        tmp[I + half] = a[n – 2] * g0 + a[n – 1] * g1 + a[0] * g2 + a[1] * g3;
        return tmp;
    }
```

Input parameters are the input vector to be subjected to transformation (with a length which is the power of two) and the number of approximation coefficients, which are to be created from it (which is also the power of two). The result is an array of approximation and details. For example, if we introduce a time series $S_{0,0}, S_{0,1}, S_{0,2}, S_{0,3}, S_{0,4}, S_{0,5}, S_{0,6}, S_{0,7}$, and the second parametr 4, then the result will be a table of four approximations and four details: $S_{1,0}, S_{1,1}, S_{1,2}, S_{1,3}, T_{1,0}, T_{1,1}, T_{1,2}, T_{1,3}$.

The wavelet D4 has been chosen because of its high-speed (low computational complexity) and the simplicity of implementation. In addition, it has been well studied in the literature from the viewpoint of its usefulness in the preprocessing efficiency of highly frequent time series.

## 6. Experiments

The tests have been carried out in two stages. In the first stage was examined on a single financial time series randomly selected from a period of two months. The aim was to capture the different behavior of financial time series. In the second stage, it was selected one financial time series corresponding to 12 hours of quotations. The goal was to discover changes in time series that to be tested it was divided into equal size sliding window time series.

Studying the impact of wavelet compression on the time series using wavelets Daubechies 4 were carried out on the currency market FOREX. This market operates 24 hours a day, from Sunday 11.00 pm to 10.00 pm on Friday (according to Central European Time). According to statistics published in 2008, most transactions concerned the pairs [BIS 2008]:

- EUR/USD 27%
- USD/JPY 13%
- GBP/USD 12%

Given these observations the authors have chosen the time series describing historical transactions for these three pairs of currencies. Selection of samples for the testing was purely random. The intervals (the first day of the month and the last day of the shortest month) were generated randomly using the pseudo-generator platform. NET. The selected values indicate days of the time series to evaluate. To ensure the objectivity of research, we have drawn five time series of two consecutive months for the same days (when the number indicated on the day on which the FX is not a transaction was taken by the next closest date on which the transactions take place) and all pairs (EUR /USD, USD /JPY, GBP /USD) have been selected from those days. In this way, 30 time series has been chosen for the experiment. Time series of the length of 256 have been created from the aggregated data to one minute. The first value means the aggregate transactions within the first minute of the day (0:01) and the last transactions of 256 minutes later (at 4:17). The time series were grouped within the pairs of currencies. So the result was three sets of data, three sets of the average relative errors (computed for a single currency).

The test consisted of two phases. In the first we have examined the amount of information of time series, evaluating the effectiveness of the prediction algorithm ARAR. Each of the selected series was divided into two series of length of 128. The first 128 values served as a learning set, and 20 consecutive values (the first 20 values from the second series) were treated as a validation set. Then, the mean relative error (MRE) for each of the delays (from 1 to 20) was calculated according to the formula:

$$RE = \frac{\Delta x}{x} = \frac{x_0 - x}{x} = \frac{x_0}{x} - 1$$

where $x_0$ is an expected value, $x$ is a real value. MRE was obtained by dividing RE into a number of samples.

The second stage was to demonstrate if the quantity and quality of information afforded by the compressed time series significantly differs from the uncompressed one. To achieve this a series of length 256 was compressed by the discrete transform wavelet. The resulting series of length 128 (twice the compression) was divided into two series of length of 64. Values of the first series have been used as a learning set, and a set of next 20 values (the first 20 values from the second row), as a validation set. The last step was the calculation of the average relative errors of (MRE) for each of the delays using the same formula as for the series of the first stage.

As a result, three groups the test series were received, which included the values of the average relative errors ranked according to delay. Recall that the objective of this study was to demonstrate that the double wavelet compression does not significantly worsen the average prediction error. Such a result confirming the hypothesis put at the beginning of the work that the double compression using wavelet Daubechies 4 does not affect significantly the quality of information carried by a time series in comparison with to the original time series. To validate the hypothesis the compatibility the Kolmogorov-Smirnov test was applied. The test confirmed that the two populations have the same distribution, which is equivalent to saying that the two samples come from the same population.

In this project the algorithm ARAR for time series prediction and the authoring program to compress the time series were applied. In the experiments, the randomly selected quotes from January 8, 12, 16, 17, 19, 20; and February 9,12,14,17,19, 2009 were tested. For the second study, the quotes comes from February 25, 2009.

The average relative errors for the original and transformed time series are presented in fig. 3-5. The differences are practically negligible. Comparison of the cumulative average errors is shown in fig. 6-8. The solid lines represent the values for the transformed time series, while the dashed lines illustrate original ones. Values are almost similar; only in case of USD /GBP, the difference is greater. When comparing the maximum and minimum values we have noticed that in some cases better prediction results were achieved using the transformed time series, and in some, when predicting based on the original time series. Although in these cases the differences were small (except the USD/GBP where the wavelet compression significantly improved of the least effective prediction).

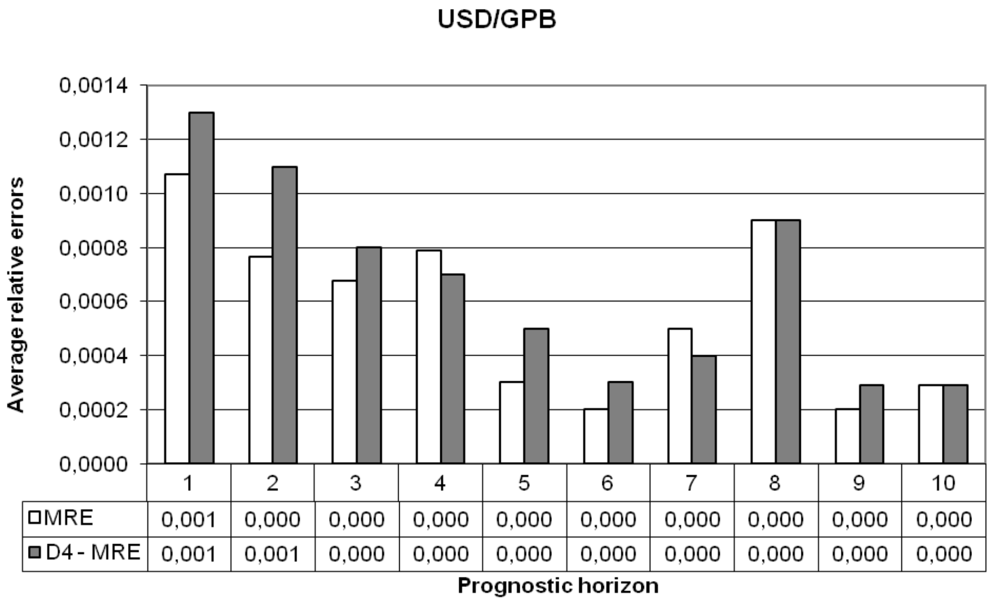


**Fig.3. Comparison of average relative errors for original and compressed time series. Pair USD/GBP**

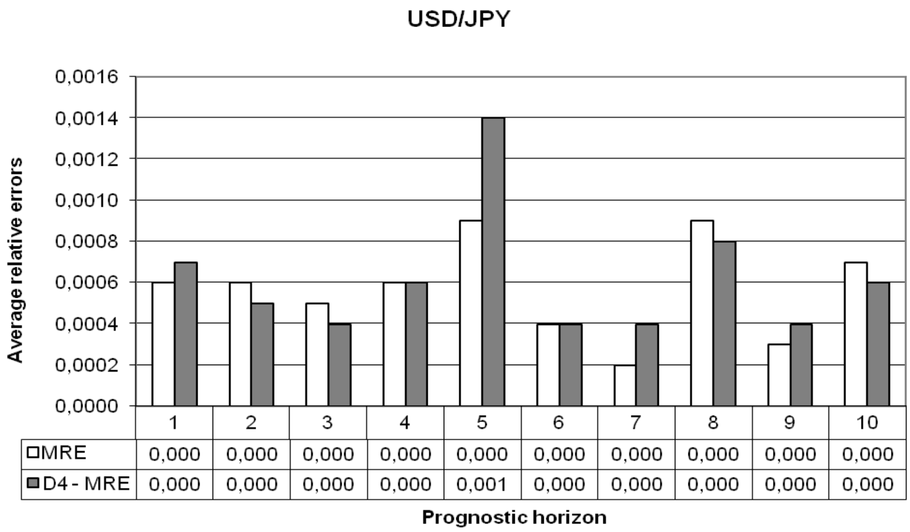


**Fig.4. Comparison of average relative errors for original and compressed time series. Pair USD/JPY**

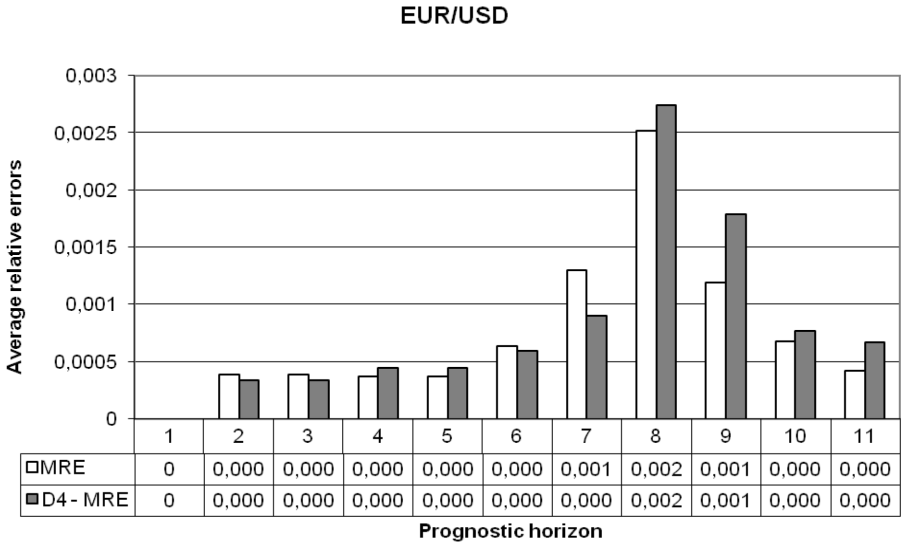


**Fig.5. Comparison of average relative errors for original and compressed time series. Pair EUR/USD**

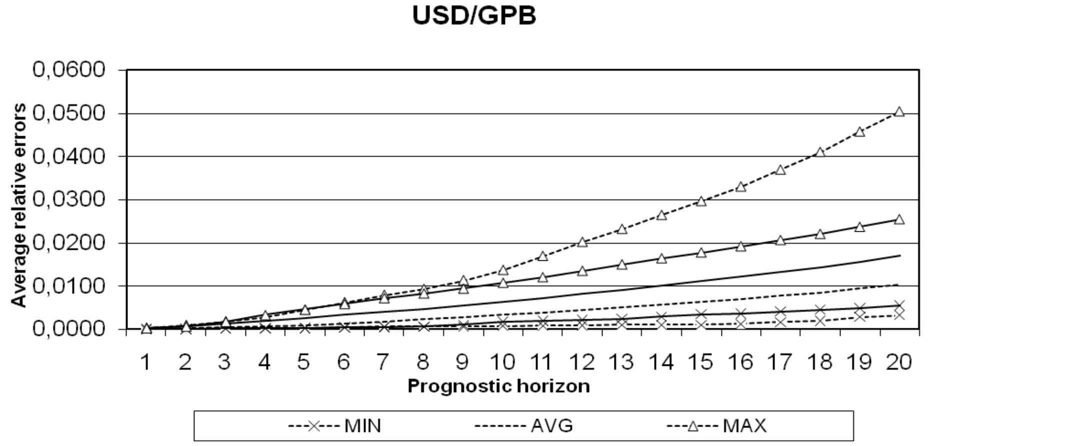


**Fig.6. Comparison of average cumulative errors (max, min and average )for original and compressed time series. Pair USD/GBP**

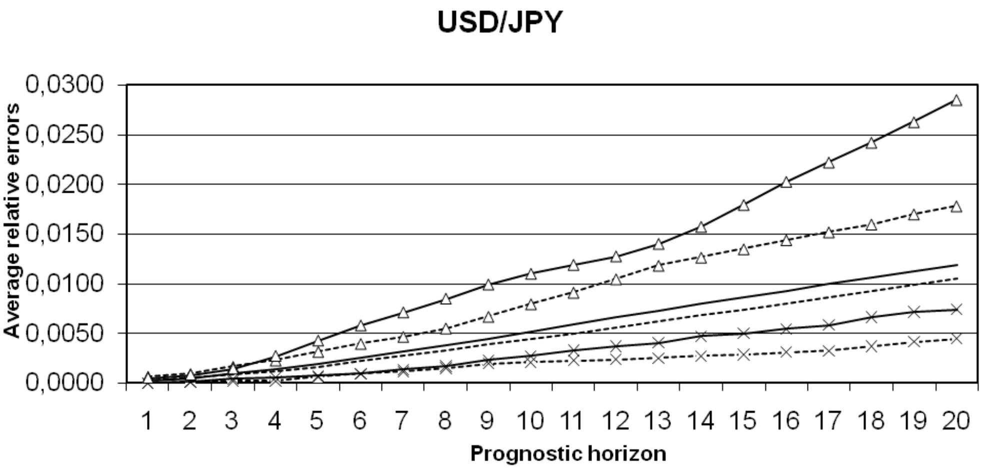

**Fig.7. Comparison of average cumulative errors (max, min and average )for original and compressed time series. Pair USD/JPY**

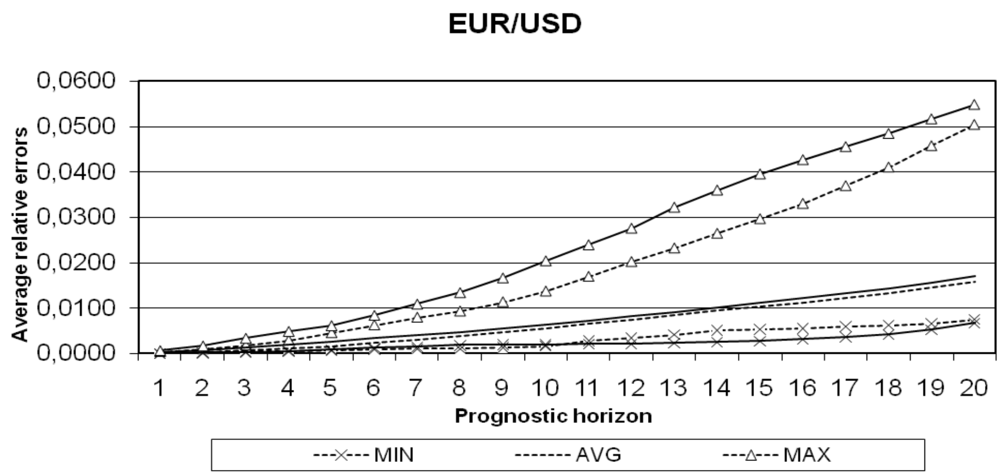


**Fig.8. Comparison of average cumulative errors (max, min and average ) for original and compressed time series. Pair EUR/USD**

The goodness-of-fit Kolmogorov-Smirnov test was also carried out. Null hypothesis was assumed that the distributions of the average relative prediction error in the original and compressed time series are the same. To validate the hypothesis the following statistic has been used:

$$\lambda_n = \sqrt{n} \cdot \sup\left|F_{n_1}(x) - F_{n_2}(x)\right|$$

where

$$n = \frac{n_1 \cdot n_2}{n_1 + n_2}$$

and $F_{n_1}(x), F_{n_2}(x)$ are the empirical distribution functions computed on the basis of samples. Graphical presentation of the distribuants is shown in fig.9.

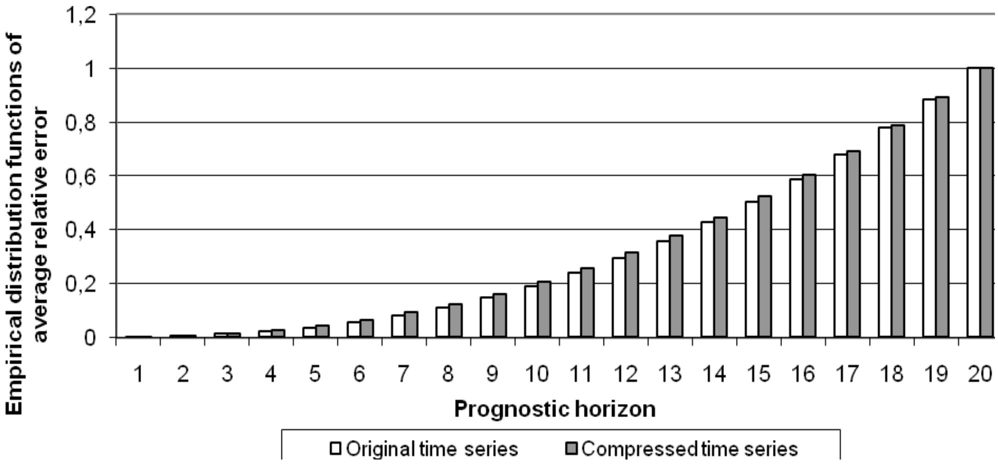


**Fig.9. Comparison of average error distributions for original and compressed time series.**

Values $n_1$, $n_2$ mean the sum of averages of relative errors of predictions. The value of empirical statistic was computed and it was equal to 0.004468. The limit λ-Kołmogorov distribution at the confidence level $\alpha = 0{,}01$ is equal to 1,63. So, based on the relation:

$$\lambda_e < \lambda_\alpha$$

there is no reason to reject the null hypothesis. So, we say with 99% probability that the average error distributions are the same in the case of time series transformed by Daubechies 4 wavelet and as the raw time series.

In the second experiment, the differences were less noticeable for the average and minimum values. However, in all three cases the results were better for compressed time series than the uncompressed one.

Taking into consideration that the predictions were made on double-compressed time series, it can be stated that result was very encouraging. Not only the computing time was reduced, but also the prediction accuracy was improved.

## 7. Conclusions and future works

The results of this research have confirmed the hypothesis established at the beginning of the work that the information carried by uncompressed time series is qualitatively identical to the information carried by double-compressed time series using Daubechies 4 wavelet. The consequences are importatnt. Colloquially speaking, it makes no sense to use the original time series since the use of a time series of two times shorter (after compressing by D4) assure the same results. Given the computing complexity of the classification algorithms, it is of utmost importance. It should be also noted that in the case of a long time series (covering 12-hours period) wavelet compression improved the quality of prediction. This would mean that, at least for these three examined time series, noise and redundant information have been eliminated by the compression process.

To determine the usefulness of wavelet compression in financial time series in general, it would be recommanded to test them on significantly greater empirical material coming fron various stock markets. One can also consider to try other wavelets of the Daubechies family. In the paper, we were focused on the computational complexity and its reduction in the context of prediction systems. It should be mentioned here that the wavelet D1 has lesser complexity than the D4, that could make it more useful for larger data sets or in real time systems. But one must have in mind that D1 is less sensitive to subtle, local changes in the original time series, and it is less efficient in analysis of highly frequent time series.

Suming up, although we have indicated that further research is required, the use of compression is fully justified if we are interested to reduce the multidimensional space and we do not want to lose any significant information contained in the original time series.

## References


[Bayen 1999] K. Beyen, J. Goldstein, R. Ramakrishnan, U. Shaft. When is nearest neighbor meaningful? In Proceedings of the 7th International Conference on Database Theory, pp.217–235, 1999.

[BIS 2007] Triennial Central Bank Survey (December 2007), Bank for International

Settlements. http://www.bis.org/publ/rpfxf07t.pdf

[Brockwell 2002] P. J. Brockwell, R. A. Davis. Introduction to Time Series and Forecasting, 2nd edition. Springer. 2002.

[Burrus 2001] C.S. Burrus. Introduction to Wavelets and Wavelet Transform. Pretntice Hall. 2001.

[Daubechies 1992] I.Daubechies. Ten Lectures on Wavelets, Society for Industrial and Apllied Mathematics. 1992.

[Mallat 1989] S.G. Mallat. A theory for multiresulution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), pp.647-693. 1989.

[Meyer 1995] Y. Meyer. Wavelets and Operators, Cambridge University Press. 1995.

[Newton 1984] Newton, H.J., Parzen, E. Forecasting and time series model types of 111 economic time series, The Forecasting Accuracy of Major Time Series Methods, S. Makridakis et al. (eds.), John Wiley and Sons, Chichester, 1984.

[Ostasiewicz 2006] S. Ostasiewicz, Z. Rusnak, U. Siedlecka. Statystyka, elementy teorii i zadania. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu. 2006.

[Shannon 1948] A.Shannon, A mathematical theory of communication, the Bell System Technical Journal, 27 (3), pp.379-423 & 623-656. 1948.