Exchange rate prediction: a wavelet-neural approach

Irena Šindelářová[1]

Abstract. The paper deals with the use of artificial neural networks (ANN) for predictions of economic time series. First, we revise the basic existing ANN architectures for time series forecasting and describe their application on CZK/EUR exchange rate prediction. Next, we introduce a hybrid version of the ANN that builds upon the same network strategy but tries to enhance the prediction accuracy by first decomposing the signal into individual frequency bands and then training the perceptron parameters on each frequency band separately. The time-frequency signal analysis is obtained using discrete wavelet transform. Different signal segments in various levels of decomposition are reconstructed separately. The resulting signals (in time domain again) are samples for ANNs as above. Finally, we compare the accuracy of the hybrid approach to that of the traditional ANN setting for the case of CZK/EUR exchange rate data.

Keywords: artificial neural network, wavelet transform, digital signal processing, CZK/EUR exchange rate prediction

JEL Classification: C45
AMS Classification: 65T60, 82C32

1 Introduction

In the paper the time series of CZK/EUR exchange rate is predicted. The daily data from the year 2001 to 2012 is considered. Time series forecasting is based on ANN with one hidden layer, discussed in detail in Section 2. The discrete wavelet transform (DWT) is used as a filter bank for various frequency bands obtaining. DWT is presented in Section 3. The essential idea of the hybrid wavelet-neural system presented is to forecast higher signal frequencies only. It is proposed that slow components corresponding to the low frequencies do not influence the predicted value. Only chosen components of the original signal go to the ANN of the same architecture. All computations were performed in the Matlab software.

2 Neural networks

The neural network is one of the computational models used in artificial intelligence. The pattern is the behavior of the relevant biological structures. ANN consists of formal neurons, which are interconnected and transform signals using certain transmission functions. Neuron has any number of inputs, but only one output. One of the most widely used models is the following [3]: $a = f(\mathbf{wx} + b)$. First, the vector input $\mathbf{x}$ is multiplied by the vector of synaptic weights $\mathbf{w}$. Second the product $\mathbf{wx}$ is added to the scalar bias b to form the net input. Finally, the net input is an argument for the transfer function f, which produces the scalar output a. Weights express experience store to the neuron. The higher the value, the more important given input is. Bias denotes the neuron activation threshold. Depending on the type of neuron an appropriate transfer function is used (linear or sigmoid).

Multilayer perceptron

The data prediction is based on *multilayer perceptrons* (MLP) with three layers: six neurons are in the input layer, one to six neurons in the hidden layer with a log-sigmoid transfer function and one neuron in the output layer with a linear transfer function. The suggested perceptron is illustrated in Figure 1. The MLPs realize the function $x_m \approx f(x_{m-6}, \ldots, x_{m-1})$, where $x_1,\ldots,x_N$ denotes an original signal. The networks were trained with the momentum back-propagation algorithm, for more details see [3].

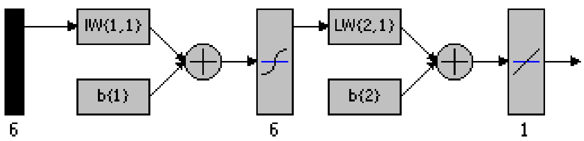

Figure 1 Multilayer feed-forward network

[1] University of Economics in Prague, Department of Econometrics, W. Churchill sq. 4, Prague, isin@vse.cz.

3 Wavelet transform

Mathematical Background

A *wavelet* is an oscillating function which can be used as a sine wave in Fourier analysis. It expands signals in terms of sinusoids having infinite energy. On the other hand a wavelet has finite energy concentrated in time around a point (Figure 2). While Fourier transform provides a given signal frequency analysis, a wavelet transform gives both time and frequency simultaneously [2].

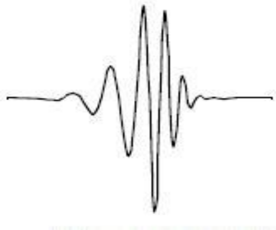

Figure 2 Example of wavelet

Wavelets are functions $w_{j,k}(t)$ with parameters of translation k a and scale j, where j, $k = 1, 2, \ldots$ The family of functions derived from the *initial wavelet* creates an orthonormal system:

$$w_{j,k}(t) = \frac{1}{\sqrt{j}} w\left(\frac{t-k}{j}\right)$$

The scale decreasing is referred to the dilation of a wavelet and translation is a shifting along the time axis. The *discrete wavelet transform* of a given signal f is the comparison of the signal and wavelets with varying k and j. The correlation rate of this comparison is expressed by *wavelet coefficients* ($a_{j,k}$). Therefore the wavelet expansion represents signal components with a more accurate local description than Fourier coefficients.

$$f(\mathrm{t}) = \sum \mathrm{a}_{j,k}\, 2^{j/2}\, w(2^j t - k)$$

Signal processing

Wavelet coefficients $a_{j,k}$ as the given signal representation in the system of wavelets can be organized into levels (various signal scales). This *multiresolution* representation [4] separates signal frequency information into selected number of levels of decomposition. The input signal length (N) has to be of power of 2. The number of decomposition levels can be equal to maximal n where $N = 2^n$. The described process is a signal filtering in fact. It is possible to develop most of the wavelet theory results using *filter banks* [2].

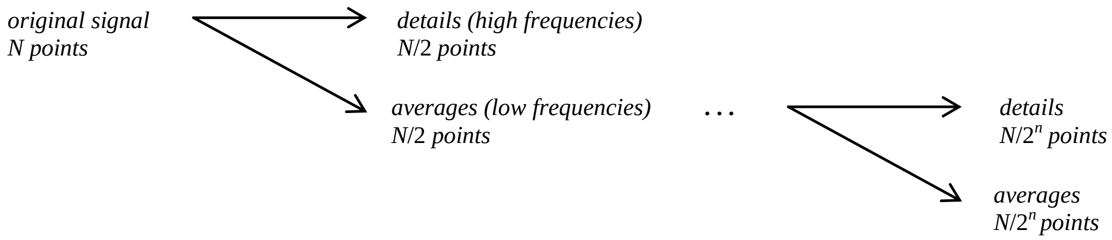

Figure 3 Principle of multiresolution

Real data processing

The time series of CZK/EUR exchange rate from October 2000 to April 2012 is considered. The number of observations is 2048, i.g. $n = 11$. We denote the vector of details in level k as $\mathbf{d}^{k}$, and averages as $\mathbf{a}^{k}$. Lengths of $\mathbf{d}^{k}$ and $\mathbf{a}^{k}$ are 2^{n-k}. The given signal discrete wavelet transform for n as a number of levels results in the vector:

$$\mathbf{r}^* = [\mathbf{a}^{n},\mathbf{d}^{n},\mathbf{d}^{n-1},\ldots,\mathbf{d}^{1}].$$

The Daubechies $D4$ wavelet in the decomposition is proposed for our purposes. $D4$ wavelet function is defined by the dilation equation in the form

$$w(x) = -\,c_3 w(2x) + c_2 w(2x-1) - c_1 w(2x-2) - c_0 w(2x-3),$$

where $c_0 = (1+\sqrt{3})/4$, $c_1 = (3+\sqrt{3})/4$, $c_2 = (3-\sqrt{3})/4$, $c_3 = (1-\sqrt{3})/4$.

4 Hybrid wavelet-neural system

This approach is based on the ANN prediction of the signal frequency bands separately. The input data of the proposed MLP defined in Section 1 are preprocessed data by means of DWT and iDWT. The number of decomposition level is 11 and $r^* = [a^{11}, d^{11}, d_1^{10}, d_2^{10}, d_1^9, d_2^9, d_3^9, d_4^9, \dots, d_1^1, \dots, d_{1024}^1]$. Then the signal transformation back to the time domain is following but it will not be perfect reconstruction. The modified wavelet coefficients enter to the inverse DWT. Signal details contained in vectors $d^1, \dots, d^{11}$ are zero-padded so the resulting length of signal in each frequency band is 2^{11}. Vectors with undesirable frequencies set to zeros are following:

$r^{(11)*} = [0, \dots, 0, d_1^1, \dots, d_{1024}^1]$,

$r^{(10)*} = [0, \dots, 0, d_1^2, \dots, d_{512}^2, 0, \dots 0]$,

…

$r^{(1)*} = [0, d^{11}, 0, \dots, 0]$,

$r^{(0)*} = [a^{11}, 0, \dots, 0]$.

Inverse discrete wavelet transform is applied to modified vectors:

$$s^{(i)} = \mathrm{iDWT}(r^{(i)*}).$$

MLP with three layers designed above is used for $s^{(i)}$ prediction. The patterns are given by the following way:

$([s_1^{(i)}, s_2^{(i)}, \dots, s_6^{(i)}], s_7^{(i)})$,

$([s_2^{(i)}, s_3^{(i)}, \dots, s_7^{(i)}], s_8^{(i)})$,

…,

$([s_{p-6}^{(i)}, s_{p-5}^{(i)}, \dots, s_{p-1}^{(i)}], s_p^{(i)})$,

where $i = 1, \dots, n$ and p is the number of learning set patterns. Predicted values in given frequency bands are $\hat{s}_{p+1}^{(i)}, \hat{s}_{p+2}^{(i)}, \hat{s}_{p+3}^{(i)}$, where $i = 1, \dots, n$. Then the m-steps forecast of original value x is given by the evaluation of

$$\sum_{i=1}^{n} \hat{s}_{p+m}^{(i)} + a^n$$

where a^n ($n = 11$ in our case) is the average in the last level of the wavelet decomposition.

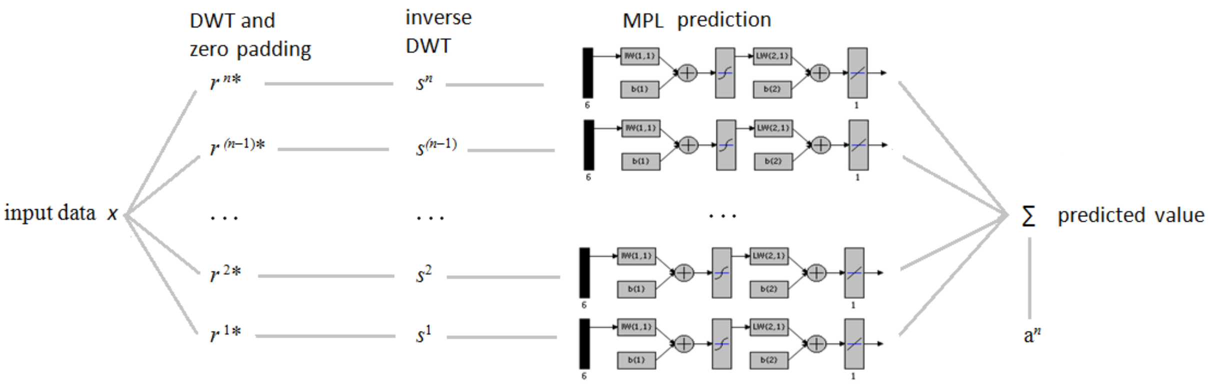

Figure 4 Structure of wavelet-neural system

As [1] shows the influence of low signal frequencies is minimal. The suggested algorithm contains the prediction of the four fastest frequency bands and estimation of additional values can be substituted by the previous values. The evaluation of vectors $r^{(8)*}, \dots, r^{(11)*}$ is sufficient. Figure 5 illustrates this situation.

Figure 4 contains the time series processing scheme: iDWT of preprocessed data enters to the MLP designed according to Figure 1. Number of perceptrons in hidden layers for chosen decomposition levels (8, 9, 10, 11) are 6, 2, 1, and 1 respectively. Number of perceptrons in the first ANN layer is 6 for each level as Figure 1 illustrates. The last layer is single element usually.

The learning set for back-propagation algorithm contains the first 1800 patterns for each frequency range, and the testing set contains 100 patterns for each range from 8 to 11.

5 Conclusion

The Table 1 provides summary of exchange rate prediction errors (*m* = 1, 2, 3, 4). The predicted exchange rates were divided into four groups of the 1st, 2nd, 3rd and 4th forecasted rate. All these values were compared to the real exchange rates with the following error measures:

- RASE – root average square error,
- MAPE – mean absolute percentage error,
- DIR – the percentage the correct identification of the next value direction.

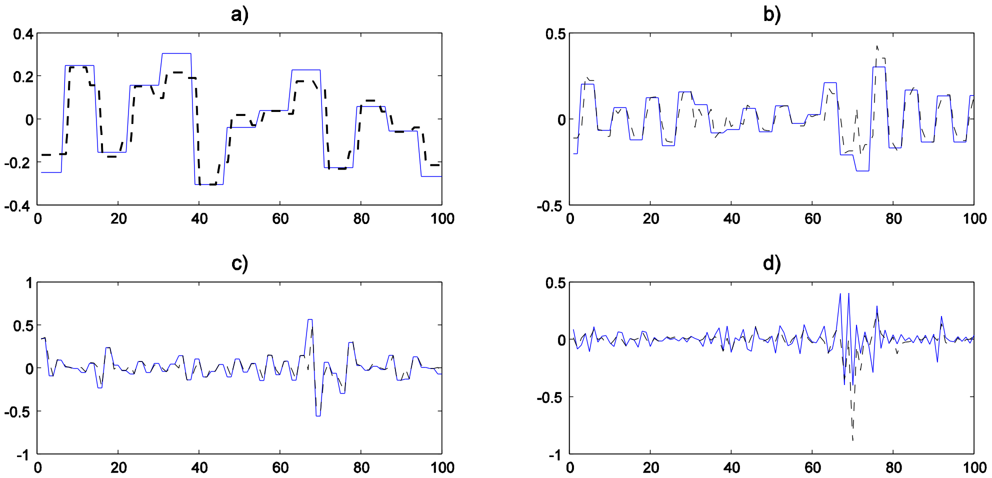

Figure 5 Signal prediction (Day 1) for the four highest frequency bands: a) *k* = 8, b) *k* = 9, c) *k* = 10, d) *k* = 11, (solid line – iDWT of zero padded wavelet coefficients, dashed line – its prediction by the MLP)

The graphs presented in Figure 5 indicate that the ANNs predict frequency bands pretty well. Surprisingly the RASE error for the whole signal decreases as the forecast period is extended to 3rd day. The MAPE, RASE errors are relatively high as well as direction coefficients from the 1st and 2nd forecasted exchange rate. DIR achieves the maximum value at the 3rd rate similar to [1].

Measure	Day 1	Day 2	Day 3	Day 4
RASE	0.438	0.433	0.424	0.587
MAPE	1.32 %	1.32 %	1.34 %	1.70 %
DIR	45 %	43 %	55 %	45 %

Table 1 Results of forecasting

The applied methods provide the results presented. The wavelet-neural system is the efficient tool for time series forecasting. Further optimization of ANNs training could improve the final prediction results.

Acknowledgements

The paper was supported by the grant F4/14/2011 from University of Economics in Prague.

References

[1] Bielecki, A., Hajto P., and Schaefer R.: Hybrid neural systems in exchange rate prediction. *Studies in Computational Intelligence* (*SCI*) **100** (2008), 211–230. Springer-Verlag Berlin Heidelberg, 2008.
[2] Burrus, C. A., Gopinath, R. A., and Guo, H.: *Wavelets and wavelet transform*. Prentice Hall, New Jersey, 1998.
[3] Haykin, S.: *Neural networks. A comprehensive foundation*. Pearson Education, Singapore, 2005.
[4] Mallat, S. G.: A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence* **11** (1998), 674–693.