FINANCIAL TIME SERIES PREDICTION USING WAVELET AND ARTIFICIAL NEURAL NETWORK

GHASSANE BENRHMACH[1,*], KHALIL NAMIR[2], JAMAL BOUYAGHROUMNI[1], ABDELWAHED NAMIR[2]

[1]Laboratory of Analysis, Modelling and Simulation (LAMS), Faculty of Sciences Ben M'sik, Hassan II University, P.O. Box 7955, Sidi Othman, Casablanca, Morocco

[2]Laboratory of Information Technology and Modelling (LITM), Faculty of Sciences Ben M'sik, Hassan II University, P.O. Box 7955, Sidi Othman, Casablanca, Morocco

Abstract: This paper focuses on the modelling of financial time series using the coupling of the discrete wavelet transform and nonlinear autoregressive neural network. This hybrid modelling method is based on the use of decomposed time series using the discrete wavelet transform as inputs to the artificial neural networks. This method has been applied to three financial series (exchange rate EUR/USD, the Brent price and NASDAQ composite price). The simulation results using R software show the robustness of the proposed model compared to other modelling methods applied to the same financial time series.

Keywords: ARIMA-model; artificial neural network; prediction; time series forecasting; wavelet decomposition.

2010 AMS Subject Classification: 92B20.

---

*Corresponding author

E-mail address: ghassane.benrhmach-etu@etu.univh2c.ma

1. INTRODUCTION

Forecasting financial time series is an essential element of any investment activity in the entire investment industry. Traditional financial time series forecasting methods include sales percentage method and linear regression analysis method. One of the most popular method is the autoregressive integrated moving average (ARIMA). However, in order to improve the quality of financial forecasting, we need to rely on more accurate forecasting methods. Time series forecasting method is a method specially used to analyse a series of observations arranged in chronological order. It has extremely important significance in economics, especially in macroeconomics. Although there are already many machine learning and deep learning models applied to financial time series analysis, such as stacked auto-encoders (SAE) [1], artificial neural network (ANN) and long short-term memory (LSTM) [2]. In recent years, wavelet transformation has been proposed for time series analysis [3, 4].

In this paper, we propose a prediction method, based on the combination between the non-linear autoregressive neural network and the discrete wavelet transform. As with other prediction methods, the basic idea is that we use the data available, constituting the past evolution of the series, to determine, at least partially, its future development. To overcome the insufficient accuracy of traditional financial time series forecasting methods, the hybrid proposed model is established including three stages, which are data pre-processing stage, data prediction stage and prediction result analysis stage. In the first stage, we decompose the financial time series using the discrete wavelet transform into sub-series. In the second stage we use to pre-processed data as input for the ANN and in the final stage we analyse the predicted data using evaluation criteria. The rest of this paper is organized as follows. The section 2 briefly presents the artificial neural network. In section 3, the concept of wavelet analysis will be introduced, while in section 4 we describe the prediction method. In section 5, we apply the proposed model and ARIMA model to three financial time series (exchange rate EUR/USD, the Brent price and NASDAQ composite price). Finally, section 6 is dedicated to the conclusions.

2. NONLINEAR AUTOREGRESSIVE NEURAL NETWORK

Artificial neural network was first introduced to solve complex classification problems. But due to their universal approximation property [5, 6], they were quickly used as nonlinear regression models, and then for the modelling of time series and forecasting [7, 8]. In this paper we consider a family of multilayer perceptron (MLP see figure 1) models called nonlinear autoregressive neural network model (NAR), defined by:

$$Y_t = h(Y_{t-1}, Y_{t-2}, \dots, Y_{t-d}) + \varepsilon_t,$$

where $Y_t \in \mathbb{R}$, h represents a function implemented by a multi-layered perceptron with a single output unit, $Y_{t-j}, j = 1, \dots, d$ are the delays of the time series (Y_t) and ε_t is a white noise.

The NAR model can be defined exactly by the following equation:

$$Y_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i \psi \left(\sum_{j=1}^{d} \beta_{ji} Y_{t-j} + \beta_{0i} \right) + \varepsilon_t.$$

d	Number of entries
k	Number of hidden layer
ψ	Activation function
β_{ji}	The weight of the connection between the input unit i and the hidden unit j
α_i	The weight of the connection between the hidden unit j and the output unit
α_0	Constants that correspond to the output unit
β_{0i}	Constants that correspond to the hidden unit j

Table 1: Symbol definition

To achieve the approximation of a time series, from generally noisy samples, using a neural network, three successive steps are necessary. First we have to choose the network architecture (number of layers and neurons) in such a way that the network is able to reproduce what is deterministic in the data. the adjustment of the number of weights of the network to avoid the possession of bad interpolation properties (generalization of the network) [9]. Secondly, it is necessary to estimate the parameters of the nonlinear regression, by minimizing the approximation error on the points of the training set, this step constitutes the supervised learning for the neural

network. In the final stage, it is necessary to estimate the quality of the network obtained by testing it with examples which are not part of the training set.

There are many training functions used to train an ANN [10, 11]. Most of the learning algorithms of neural networks are optimization algorithms, they seek to minimize, by nonlinear optimization methods, a cost function which constitutes a measure of the difference between the desired responses and the real responses of the network. This optimization is done iteratively, by modifying the weights as a function of the gradient of the cost function: the gradient is estimated by a method specific to neural networks, called the backpropagation method. In our work, we chose the Levenberg-Marquardt method, characterized by its great computational and memory requirements. The Levenberg-Marquardt algorithm is given by:

$$W_{k+1} = W_k - (J^T J + \mu I)^{-1} J^t e,$$

where, 'W' is the weights of the neural network, 'J' is the Jacobian matrix of the performance criteria to be minimized, 'I' is the unit matrix, μ is a learning rate that controls the learning process and 'e' is the residual error vector. Hence, μ is decreased after each successful step and increased only when a step increases the error.

3. WAVELET ANALYSIS

The introduction to theory and practice of wavelet analysis are described by several authors in numerous books and articles [12, 13, 14]. Some other authors have described the application of wavelets in statistics and time series analysis [15, 16]. Unlike the Fourier transform, which does not work well when it has to locally describe a function that shows discontinuities, wavelet analysis offers a wide range of basic functions from which one we can choose the most appropriate for a given application. In his article, Strang make a more in-depth comparison between the Fourier transform and the wavelet transform [17]. The wavelet transform offers the possibility of analysing a signal simultaneously in the time domain and in the frequency domain (necessary for a non-stationary signal), it can be used to explore, denoise and smoothen any kind and size of time series [18]. A wavelet is an oscillating function φ of mean zero, possessing a certain degree of regularity and whose support is finite. The mother wavelet of a signal is defined by:

$$\psi(\tau, a) = a^{-\frac{1}{2}}\psi\left(\frac{t-\tau}{a}\right), \qquad (1)$$

where 't' stands for time, 'τ' is a translation parameter in time and 'a' is a dilation parameter, who referred to a scale parameter, with a value selected in the range of $0 < a < 1$. Regularity conditions also can be imposed such as multiple vanishing moments:

$$\int_{-\infty}^{+\infty}\psi(t)dt = 0,$$

second the integral of the square of the wavelet function is unity

$$\int_{-\infty}^{+\infty}\psi^2(t)dt = 1,$$

In addition, it is usually assumed that the following condition is true:

$$C_\psi = \int_{-\infty}^{+\infty}\frac{|\psi(s)|^2}{s}ds,$$

where C_ψ will be within the range of $0 < C_\psi < \infty$. We called the admissibility condition, which is necessary to reconstruct the signal. In this study, to analyse the financial time series, we choose the discrete wavelet transform (DWT), defined as follows:

$$W_\psi Y(\tau, a) = |a|^{-\frac{1}{2}}\int_{-\infty}^{+\infty}Y(t)\psi\left(\frac{t-\tau}{a}\right)dt,$$

where $\psi(t)$ is the basic wavelet with effective length (t) that is usually shorter than the target time series $Y(t)$. Using the DWT the original time series $Y(t)$ is decomposed to approximate series (A) and detailed (D) contains the rapid and sudden changes (see figure 2).

3. HYBRID SYSTEM DESIGN DWT-ANN

In this section, the prediction method will be described in detail. The main idea is to use as inputs for the MLP (NAR) network, the components of the subseries (details and approximations) which are derived from the use of the discrete wavelet transform on the original data of the time series according to the following steps:

- Step 1: Dividing the data into a training set and testing set (65%, 25%, 10%).

- Step 2: Data Pre-processing, using the discrete wavelet transform in order to obtain the approximation coefficients and detail coefficients of the original time series.

- Step 3: We use the approximation and detail coefficients as input to the nonlinear autoregressive neural network models. The network gives output as coefficients only.

- Step 4: We reconstructed the output coefficients from the ANN using inverse wavelet function.

4.1 Performance evaluation

This step consists in evaluating the models formed by comparing the difference between the estimated values and the actual values. The evaluation criteria are necessary to choose the best model. The indicators taken in this study are: the Mean Square Error (MSE) and Coefficient of determination (R^2). The definitions of these two indices are presented below:

- $MSE = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \bar{Y}_i)^2,$

- $R^2 = \frac{[\sum_{i=1}^{N}(Y_i - \bar{Y}_i)(\hat{Y}_i - \bar{\hat{Y}}_i)]^2}{\sum_{i=1}^{N}(Y_i - \bar{Y}_i)^2 \sum_{i=1}^{N}(\hat{Y}_i - \bar{\hat{Y}}_i)^2}$, with $\bar{Y}_i = \frac{1}{N} \sum_{i=1}^{N} Y_i$ and N is the number of sample.

5. RESULTS AND DISCUSSION

The data used in this study was collected from the "https://www.investing.com/" website. It represents the daily exchange rate EUR/USD, the Brent oil price and NASDAQ composite price between 02 January 2015 and 31 December 2020 (see figures 3, 4 and 5).

In a first part of the simulation, we apply the Box and Jenkins methodology [19] to the three financial time series, in order to determine the best ARIMA model, by following to three main step: Identification, Estimation, and Model checking according to the smaller AIC and BIC values and the higher likelihood value [20]. The most suitable model according to the smaller AIC and BIC values and the higher likelihood value for the daily exchange rate EUR/USD is ARIMA (1,1,1), for the Brent oil price is ARIMA (1,1,0) and for the NASDAQ composite price is ARIMA (0,1,2). In the second part of the simulation, we have chosen the Daubechies wavelet as the mother wavelet

and a decomposition number equal to 7 to represent our signals in the time scale domain. The level of decomposition depends on the sampling frequency of the signal to be analysed, it is determined by the application of to the following formula:

$$L = int(\ln(N)),$$

where 'L' is the level of decomposition and 'N' is the signal length. Note that the original time series Y_t can be synthesized by the approximation and detail sequences, i.e., $Y_t = A_7 + D_7 + D_6 + D_5 + D_4 + D_3 + D_2 + D_1$ (see figures 6, 7 and 8). After decomposing the time series, we use the sub series as inputs to our artificial neural network. In this study, the number of hidden nodes was determined by the learning algorithm [9]. The architecture and parameters of the NAR and DWT-NAR used for each data are described in the following table.

	Number of input neuron(s)		Number of output neuron(s)		Number of hidden layer(s)		Number of neurons in hidden layer(s)		Transfer function	
Time series	NAR	DWT-NAR	NAR	DWT-NAR	NAR	DWT-NAR	NAR	DWT-NAR	NAR	DWT-NAR
Brent oil	One	Seven	One		One		27		Sigmoid function	
EUR/USD	One	Seven	One		One		27			
NASDAQ	One	Seven	One		One		32			

Table 2: NAR and DWT-NAR parameter settings

In table 3, we find the MSE and Coefficient of determination (R^2) for the three models (ARIMA, NAR, DWT-NAR).

Time series	RMSE			R^2		
	ARIMA	ANN(NAR)	DWT-NAR	ARIMA	ANN(NAR)	DWT-NAR
Brent oil	1.267195	1.427403	0,116317	0.7531	0.7941	0.8865
EUR/USD	0.005742	0.005197	0.000556	0.8282	0.8852	0.9623
NASDAQ	97.68545	92.99349	44.94144	0.5427	0.7457	0.8429

Table 3: Models performance

The MSE and Coefficient of determination (R^2) for the three proposed methods are summarized in Table III. We find that the decomposition of the financial series in detail and approximation coefficients and use them as new entries of the studied system, optimizes the results obtained (see figures 9, 10 and 11).

6. CONCLUSION

In this article, a method for predicting financial time series was presented. This method was developed by the combination of the nonlinear autoregressive neural network and the wavelet transform. This combination was performed by taking as inputs to the neural network the approximations and details obtained by decomposing each input variable using the discrete wavelet transform. This method was applied to the series of exchange rate EUR/USD, the Brent price and NASDAQ composite price. The results clearly surpass those obtained by using the ARIMA model and nonlinear autoregressive neural networks (inputs undecomposed).

DATA AVAILABILITY

The data used to support the findings of this study are available from the corresponding author upon request.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

FIGURES
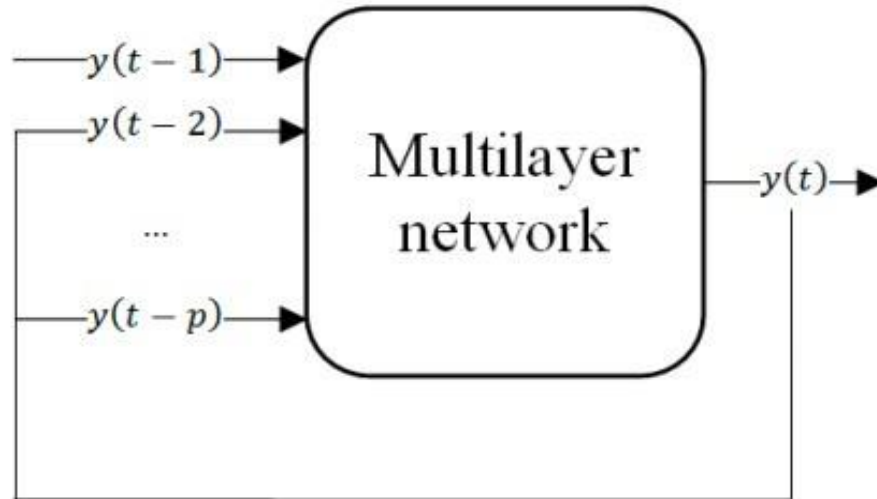
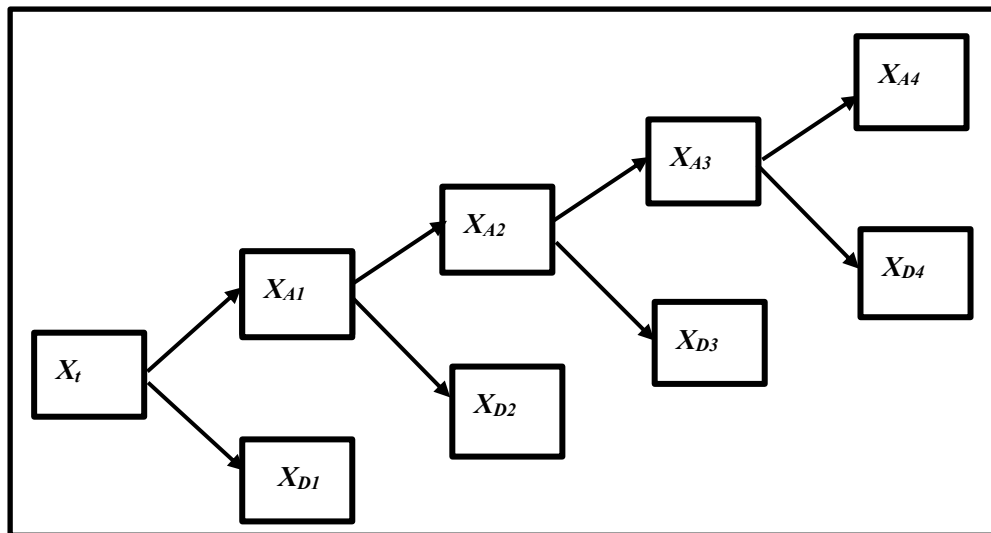*Figure 1: Architecture of the nonlinear autoregressive neural network*
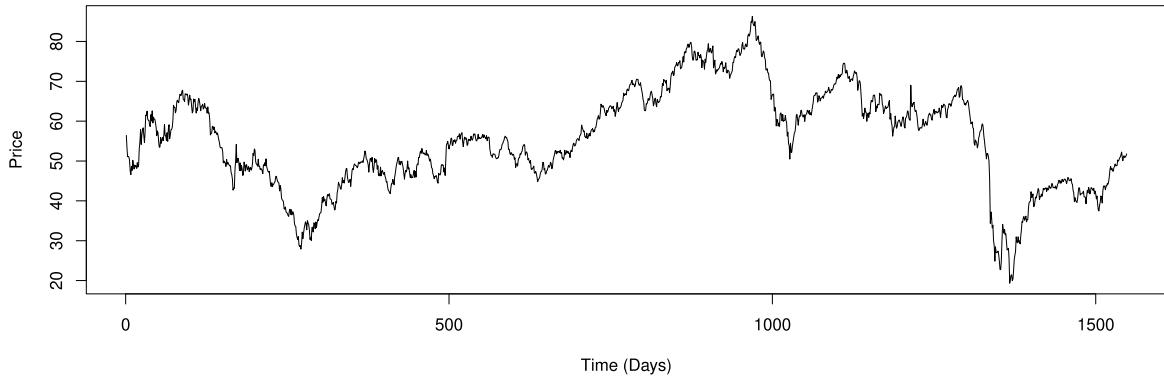

Figure 2: Wavelet decomposition tree

Figure 3: Daily price of Brent oil from 2015 to 2020
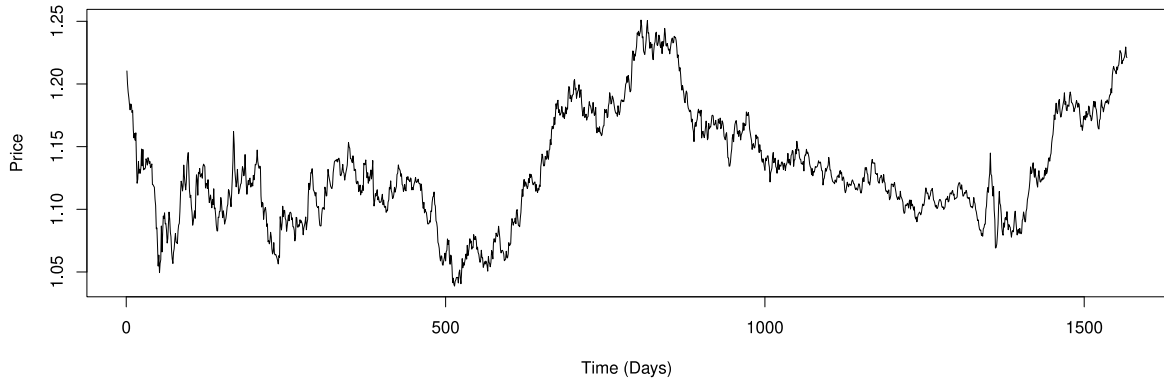

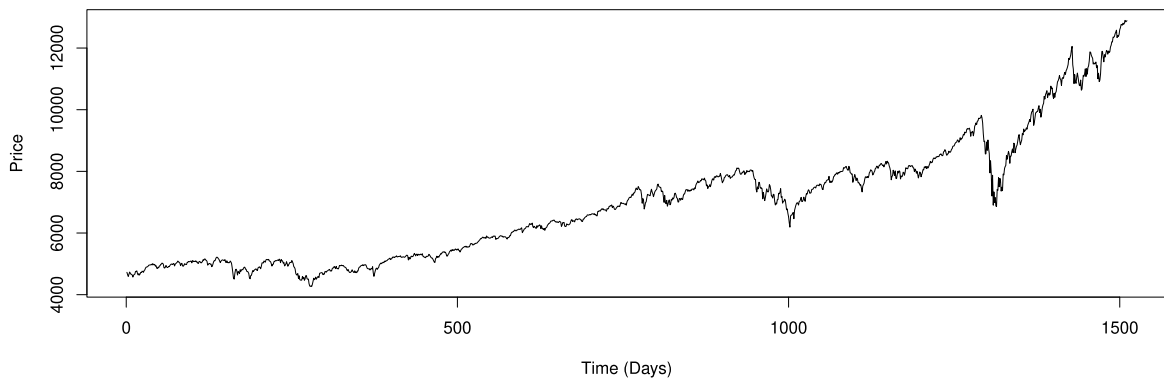*Figure 4: Daily exchange rate of EUR/USD from 2015 to 2020*

Figure 5: Daily price of NASDAQ composite from 2015 to 2020
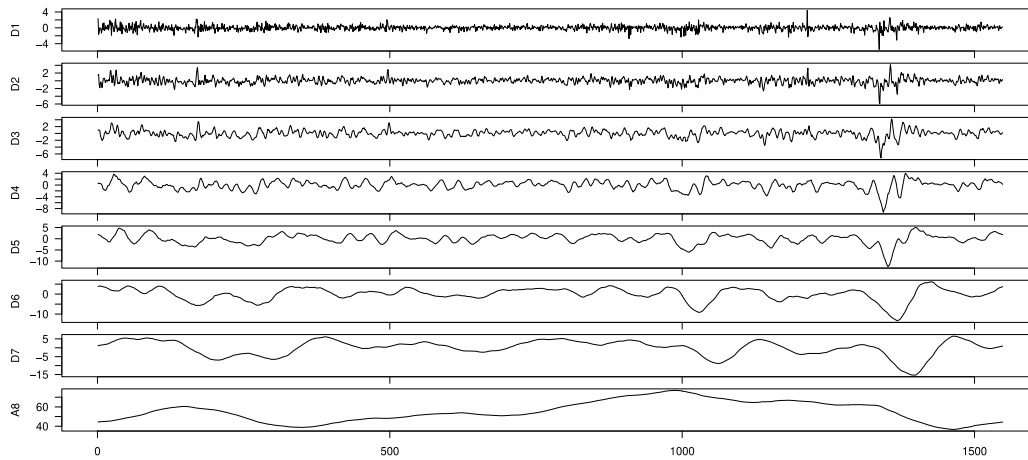
FINANCIAL TIME SERIES PREDICTION

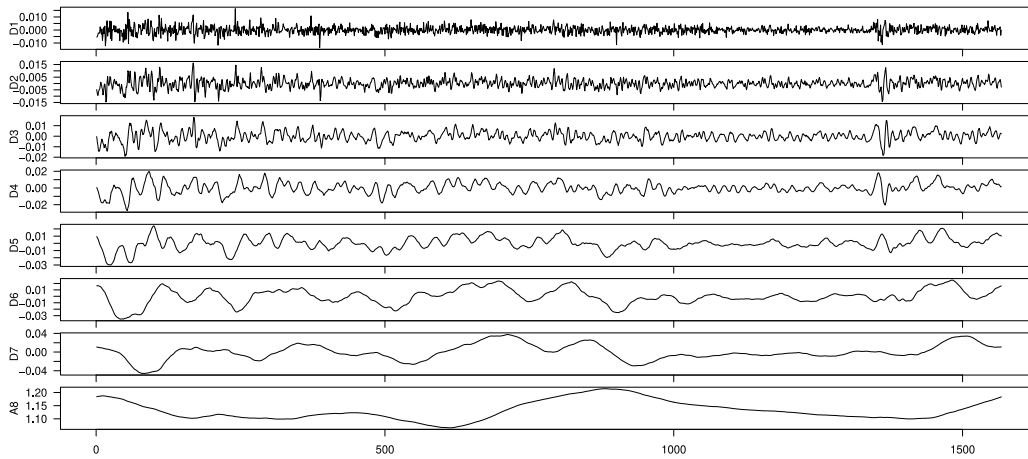*Figure 6: Decomposed wavelet sub-time series of daily Brent oil price from 2015 to 2020*

Figure 7: Decomposed wavelet sub-time series of daily exchange rate of EUR/USD from 2015 to 2020
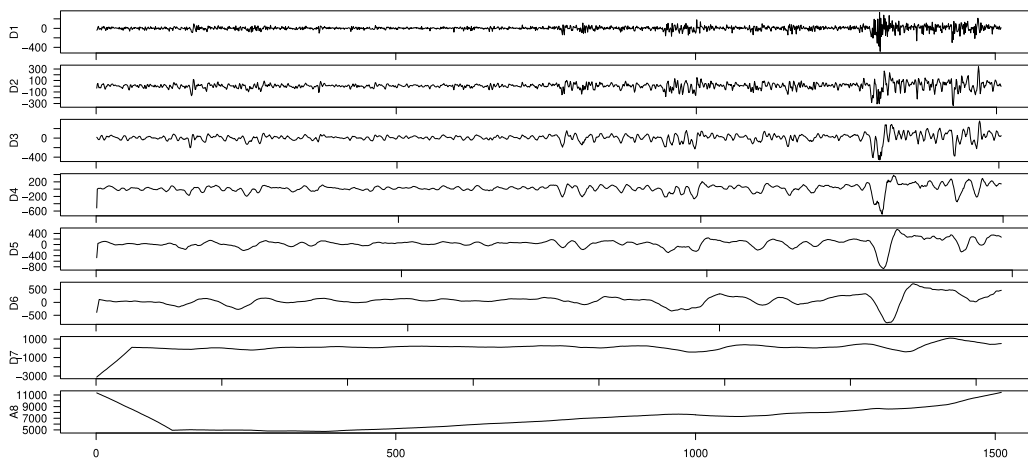

Figure 8: Decomposed wavelet sub-time series of daily NASDAQ composite price from 2015 to 2020
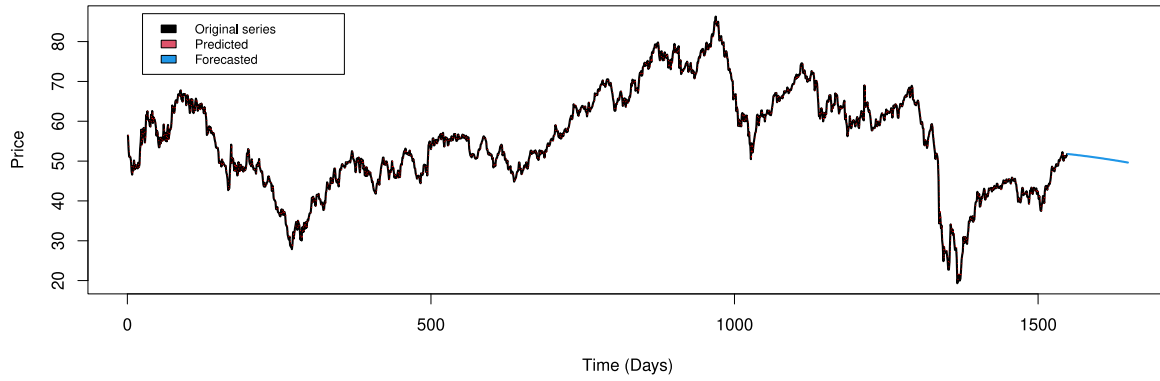
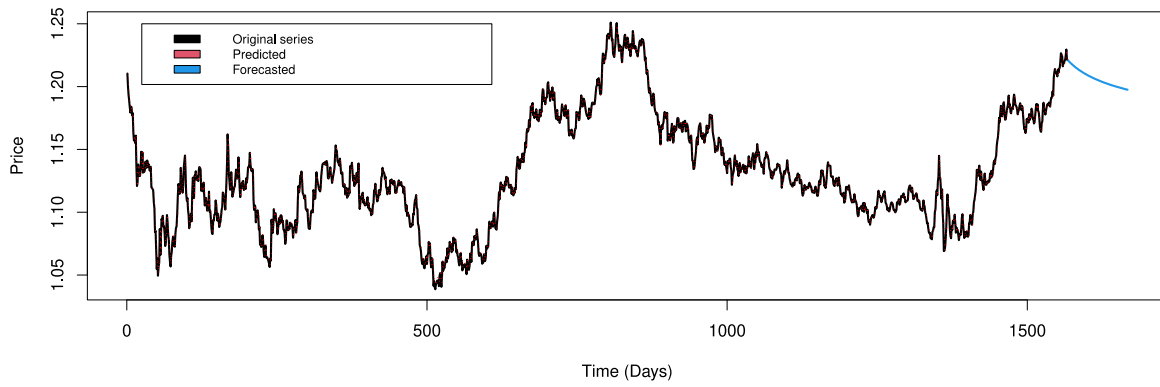Figure 9: Actual, predict and forecast values of the Brent oil price

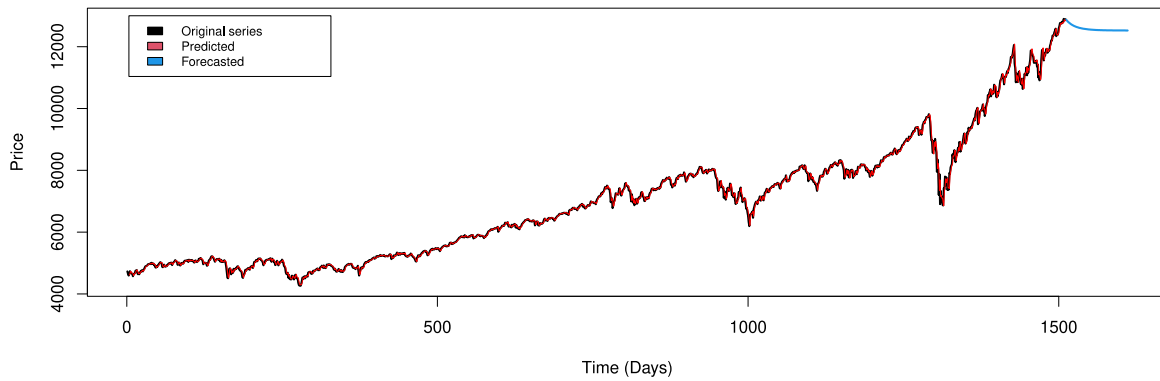Figure 10: Actual, predict and forecast values of the EUR/USD exchange rates

Figure 11: Actual, predict and forecast values of the NASDAQ composite price

REFERENCES

[1] W. Bao, J. Yue, and Y. Rao, A deep learning framework for financial time series using stacked autoencoders and long-short term memory, PLOS ONE, 12 (2017), e0180944.

[2] Y. Chen, Z. Lin, X. Zhao, G. Wang, Y. Gu, Deep Learning-Based Classification of Hyperspectral Data, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 7 (2014), 2094–2107.

[3] O. Renaud, J.-L. Starck, F. Murtagh, Prediction Based on a Multiscale Decomposition, Int. J. Wavelets Multiresolut. Inf. Process. 01 (2003), 217–232.

[4] S. Soltani, On the use of the wavelet decomposition for time series prediction, Neurocomputing. 48 (2002), 267–277.

[5] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Netw. 4 (1989), 359-366.

[6] K. Funahashi, On the approximate realization of continuous mappings by neural networks, Neural Netw. 2 (1989), 183-192.

[7] A. Weigend, N. A. Gershenfeld, Time series prediction, Addison-Wesley, 1993.

[8] M. Cottrell, B. Girard, Y. Girard, M. Mangeas, C. Muller, Neural modeling for time series: A statistical stepwise method for weight elimination, IEEE Trans. Neural Netw. 6 (1995), 1355–1364.

[9] Y. Bai, X. Jin, X. Wang, T. Su, J. Kong, Y. Lu, Compound Autoregressive Network for Prediction of Multivariate Time Series, Complexity. 2019 (2019), 9107167.

[10] Sheng-Tun Li, Shu-Ching Chen, Function approximation using robust wavelet neural networks, in: 14th IEEE International Conference on Tools with Artificial Intelligence, 2002. (ICTAI 2002). Proceedings., IEEE Comput. Soc, Washington, DC, USA, 2002: pp. 483–488.

[11] Y. He, F. Chu, B. Zhong, A Hierarchical Evolutionary Algorithm for Constructing and Training Wavelet Networks, Neural Comput. Appl. 10 (2002), 357–366.

[12] K. Blatter, Wavelet analysis. Fundamentals of the theory, M: Technosphere, (2006), 272.

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

[14] V.I. Vorobiev, V.G. Gribunin, Theory and Practice of Wavelet Transform. St. Petersburg. Publishing House of the Military Communication University, (1999).

[15] P.A. Morettin, Wavelets in Statistics. Resenhas do Inst. Matemática e Estatística da Univ. São Paulo, (1997).

[16] D.B. Percival, A.T. Walden, Wavelet Methods for Time Series Analysis, Cambridge University Press, Cambridge, 2000.

[17] G. Strang, Wavelet transforms versus Fourier transforms. Bull. Amer. Math. Soc. 28 (1993), 288-305.

[18] A. Antoniou, C.E. Vorlow, Recurrence quantification analysis of wavelet pre-filtered index returns, Physica A: Stat. Mech. Appl. 344 (2004), 257–262.

[19] D.J. Bartholomew, G.E.P. Box, G.M. Jenkins, Time Series Analysis Forecasting and Control., Operational Research Quarterly (1970-1977). 22 (1971), 199.

[20] P.J. Brockwell, R.A. Davis, Introduction to time series and forecasting. Springer Science & Business Media, 2006.