A comparative analysis and forecasting of financial market in different domains using NARX neural networks

Stefan Nikolić, Petar Poznić
Enetel Solutions, Roaming Solutions Group
Belgrade, Serbia
stefan.nikolic@enetelsolutions.com, petar.poznic@enetelsolutions.com

Abstract - **Many successful companies are very interested in developing new methods for forecasting time series. It has been a challenging task for researchers over the years, given that investors want proven methods for multiplying money they invested. Financial market analysis using artificial neural network already has confirmed to be impressive in predicting financial market performance. In this paper, for analysis of time series, exploring their hidden insights and forecasting different values, Nonlinear Autoregressive Exogenous (NARX) neural network is used. Different time series were used both with an input in time domain, along with input in the frequency domain. Finally, results are compared to find potentially the best time series for predicting, also the suitability of the field in which better results are achieved.**

Keywords - financial market; time series; forecasting; NARX neural network; Fourier transform

I. INTRODUCTION

At this moment, modern technologies are the essential and reliable tools and that will be increased in the future. Clients will share with you their business needs and will expect from you to find the best solutions for them. Currently, organizations and companies are collecting big data, expecting to extract huge knowledge from them and to help decision makers to gain advantage over competition. Using conventional methods, accurate analysis of types of data is not easy task and sometimes is not even possible. Considering this, new horizons are opening in front of researchers to discover new methods to extract useful information from data they have [1].

One of the most important things in present-days is to discover user real needs and their behaviors and to analyze it in details. Using methods of machine learning to improve analysis gives a complete new dimension in combining of business and technology and allow to creating mechanism to meet and overcome increased needs and demands of users. This is the biggest competitive advantage of every organization in modern daily business. Neural networks are one of the essential parts of it in finding good solution for a traditional practical problem of different dynamic systems.

Neural networks are one of the models which learning is based on how neural networks works in the human brain. As changes on the market are very common, neural networks are perfect model for it because their learning model is adaptable to changes, and for that reason would have more success. This comes from the desire to build and artificial system intelligent enough to perform complex calculations and represents a valuable perspective in the future.

The goal of this paper is the comparative analysis of the financial time series and their forecasting using a neural network. Neural network has been trained and tested both in the foreign exchange market and the stock market. Historical data has been collected and analyzed to create a model that would establish a link between the corresponding variables.

II. RELATED METHODOLOGIES

There are two main directions of development of the neural networks. The first one is to develop more software tools that are user friendly and to increase accessibility of modern computers. This will help to accelerate development of neural networks by those who have only basic knowledge in this field. The second direction is evident success of neural networks in fields where traditional computer systems are struggling and have many disadvantages. There are many other methods and classic statistical methodologies that deal with similar problems and some of them will be listed below.

Support vector machines (SVM) is one of the methods that is progressively used for predicting financial time series. Many scientific papers are comparing accuracy, precision, advantages and disadvantages of this method over neural networks [2, 3].

Volatility has been the subject of many studies in financial markets, especially as an essential input to many financial decision making models. Investment decisions strongly depend on the forecast of expected returns and volatilities of the assets. The introduction of Autoregressive Conditional

Heteroskedastic (ARCH) model has created a new approach and has application for financial econometricians, becoming a popular tool for volatility modeling and forecasting [4].

Also known as econometric models for time series are Generalized Autoregressive Conditional Heteroskedastic (GARCH) and Exponential Generalized Autoregressive Conditional Heteroskedastic (EGARCH), but in other papers in comparative analysis they have proved less effective than NARX, so in this paper will not be considered or compared to the network [5].

Traditionally, Box Jenkins or Autoregressive Integrated Moving Average (ARIMA) model has been dominating over time series for forecasting the time series and includes the identification, evaluation and checking of the suitability of the selected time series model. Although it is rather flexible and can be used for a large number of time series, the main limitation is the assumption of the linearity of the model and it is used to model nonstationary time series. The model cannot explain nonlinear behavior, which is at the core of financial time series. The connection between conventional statistical approaches and neural networks for this use is complementary.

The neural network is not transparent. It is a black box model. It must be trained a certain number of times, and after that the average values is taken to see how stable solution is reached eventually. When it comes to nonlinearity of data, statistical predictive techniques have come to their limits, while neural networks are increasingly applied in the classification and pattern recognition [6, 7].

The main problem in finance is unstable nature of observed time series and its heteroscedasticity, and this makes it impossible to apply some time series models. Many studies have been empirically investigating the forecasting performance of GARCH model. One of them is example for NASDAQ-100 return during the period of six years, which prove to be a financial time series characterized by heteroscedasticity. Volatility performance is undoubtedly improved. ARCH and GARCH model along with their extensions provide a statistical stage on which many theories of asset pricing, portfolio analysis, value at risk or index volatility can be exhibited or tested [4].

One of the major issues threatening many companies and governments is bankruptcy, and its prediction is a complex process consisted of large number of factors. Financial distress begins when company is not able to meet its scheduled payments of when projected cash flow shows inability to meet payments in near future. This leads to business failure and following bankruptcy can be divided into economic, financial, fraud, disaster and others. Using more accurate bankruptcy detection techniques, organizations can minimize the risk of falling to bankruptcy taking some preventive measures [8].

III. EXPERIMENTAL DATA AND METHODS

A. Data

For the time series analysis, EUR/USD pair is used in this paper as one of the major currency pairs on Forex, because of its big share in the total trading volume (27%). Generally, cross currency pairs, which do not include US dollar, are less suitable for analysis due the smaller volume of trading and larger spreads than the major currency pairs. Based of Forex's characteristically large oscillations, it might be assumed that the S&P 500 index will show better features related to the prediction of the series.

Relevant historical currency pair data for more than ten years have been downloaded from the website of Fusion Media Limited [9]. In the analysis of time series from the stock exchange, a representative index S&P 500 was used with the historical data downloaded from the website of Yahoo! Finance [10]. The collected data are related to the prices (High, Low, Open, Close) in the period from 2003 until September 2018, for each day four prices, but the close price and then returns will be used in the analysis. S&P 500 is based on 3950 observations in the period 31/12/2002 - 07/09/2018 and EUR/USD currency pair is based on 4093 observations in the period 01/01/2003 - 07/09/2018. Stock exchange and Forex are closed during the weekend, so close price is not considered in that period of time.

B. Methods

For predicting time series analysis, previously were used models such as moving-average (MA), autoregressive (AR) or ARIMA. They were unable to solve problems related to nonstationary signals and signals whose mathematical model is not linear. Otherwise, when applied to problems where solutions require knowledge uneasy to specify, neural network is a very powerful tool.

NARX neural network is a dynamic neural architecture used to model nonlinear dynamic systems. The Nonlinear Autoregressive (NAR) network is distinct in that it has another additional time series with external data, which increases accuracy of the prediction. For applications related to the prediction of time series, it is designed as a feedforward neural network with time delay (TDNN).

Methods that are based Fourier transform have a huge application in all areas of science and engineering. Fourier transform is used in signal processing, for solving differential equations, or for analyzing the dynamics of the market. Nonetheless, Fourier transform is not often suitable for procession of nonstationary signals, or for signals whose frequency content changes over time, where the periodic signal should be centered around the integer multiplicity of selection frequencies. After that, signal is divided into smaller time segments and analyzes the frequency content of each individual part. Wavelet Transformation has the possibility of dilatation and translation of waves as the basic function of transformation [11].

In the given experiment, the prediction method applies to changes in the exchange rate or changes in the stock exchange index over a certain period of time. The idea is to find the specific pattern of observations along with the usual changes. These changes would mean that a certain inheritance or some kind of random variation occurred over period of time. At the end, based on the data, a series with the damped changes should be obtained, which indicates the long-term trend or trend present in the time series. After that, it is used to predict the future values of the time series.

A two-layered feedforward network is used, where the sigmoid function is in a hidden layer and that is the most common form of a transmission function, which is nonlinear. The linear transfer function is in the output layer.

Levenberg-Marquardt (LMA), a combination of gradient descent and Gauss-Newton algorithm, is used as an algorithm for learning, as opposed to Elman's recurrent networks, using gradient descent with a momentum. It is known as the advanced and fast algorithm for nonlinear optimization, whereby, unlike the Quasi-Newton algorithm, LMA does not need to compute Hesian matrix, so it has significantly better performance. The Jakobian matrix, which contains the first network error, is used, and it is expressed by a backpropagation algorithm, which is easier than calculation of the Hesian matrix. It is necessary to reach the proximity of the minimal error function and get closer as soon as possible [12].

IV. RESULTS AND DISCUSSION

The conclusion is that the time series of the prices are not stationary, while the returns are a stationary time series. Also, the conclusion is that prices don't have the normal distribution and deviate significantly from it, but returns have significantly better statistical characteristics. In this case, the time series of the returns are much closer to the normal distribution and the normal distribution with thick tails appears. Unexpected events appear more often than in the normal distribution, which is characteristic of the analysis of financial data and forecasts.

Figure 1 shows the deviation of the autocorrelation value beyond the confidence interval for the first 2 legs, and according to that, in the network architecture, the default value 2 should be used as a time delay. As a result of the lack of statistically significant autocorrelation in the data, the NARX neural network will be used for analyzing the time series. Observing variances of random errors and their differentiation by individual observations, there is the phenomenon of heteroskedasticity. The cause of this phenomenon may be specification errors, exclusion an important regressor whose influence will be covered by the error or the existence of extreme values in the sample. As a method of elimination, the method of the least squares is applied. The idea is that in the process of minimizing the sum of the quadrate of the residual, a smaller weight is given to those residues that are greater by absolute value, and vice versa.

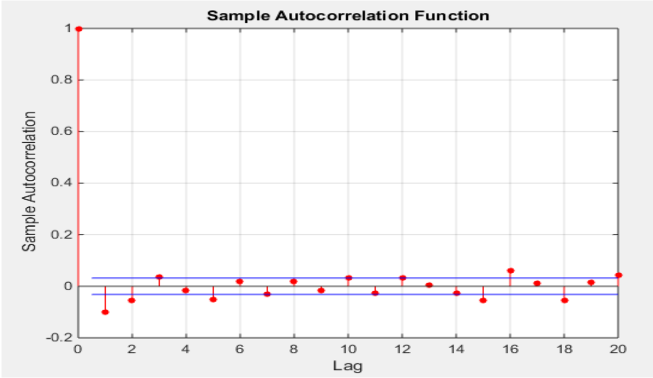

Figure 1. Autocorrelation function of returns

Engel's ARCH test allows seeing if there is heteroskedasticity or not. For the obtained value 1 as a result of the test, it was established for both time series, that the zero hypothesis is rejected (the residual series does not show heteroskedasticity), so it can be concluded that it exists in both time series.

The initial number of hidden neurons in the network architectures is set to 10 with 2 time delays. The network will be applied to returns instead of prices for both time series that are observed in the time and frequency domain. The smallest mean squared error occurred in the 3rd epoch and is $1.11455*10^{-4}$. It represents a deviation of the predicted value in relation to the actual value. If the number is closer to 0, the obtained result is more accurate.

After 10 consecutive training of the network, the smallest mean squared error after appeared in the 7th epoch and is $1.11092*10^{-4}$. As in the analysis of the previous time series, the same training algorithm was used and the subsets for training, validation and testing were obtained for the same percentile values. The network architecture is identical with sigmoid function in the hidden and linear function in the output layer.

The algorithm is also trained at 70% of the data, evaluated at 15% and tested at 15%. Each network consists of two hidden layers. The first hidden layer has ten neurons with a sigmoid transfer function and the other one is a neuron with a linear transfer function. In the second network, a smaller average mean squared error was detected than in the first one. Also, the standard deviation of the secondary squared error for the other network is lower than for the first one for all three stages of training, validation and testing. The results for each iteration and summary of mean squared error are presented in Tables I and II for S&P 500 and in Tables III and IV for EUR/USD currency pair respectively.

Table I. Mean Squared Error - S&P 500

Iterations	Mean Squared Error		
	Train	***Validation***	***Test***
1	$1.3568*10^{-4}$	$1.1455*10^{-4}$	$1.1280*10^{-4}$
2	$1.3680*10^{-4}$	$1.1922*10^{-4}$	$8.7396*10^{-4}$
3	$1.3512*10^{-4}$	$1.1848*10^{-4}$	$1.1948*10^{-4}$
4	$1.2437*10^{-4}$	$1.0698*10^{-4}$	$1.6513*10^{-4}$
5	$1.2820*10^{-4}$	$1.0336*10^{-4}$	$1.5894*10^{-4}$
6	$1.2941*10^{-4}$	$1.5599*10^{-4}$	$1.2687*10^{-4}$
7	$1.2601*10^{-4}$	$1.3396*10^{-4}$	$1.3046*10^{-4}$
8	$1.2619*10^{-4}$	$1.0994*10^{-4}$	$1.5612*10^{-4}$
9	$1.2308*10^{-4}$	$1.1070*10^{-4}$	$1.7836*10^{-4}$
10	$1.2748*10^{-4}$	$1.1092*10^{-4}$	$1.3480*10^{-4}$

Table II. Summary - S&P 500

Summary	Mean Squared Error		
	Train	***Validation***	***Test***
Min	$1.2308*10^{-4}$	$1.0336*10^{-4}$	$1.1380*10^{-4}$
Max	$1.3680*10^{-4}$	$1.5599*10^{-4}$	$8.7369*10^{-4}$
Average	$1.2923*10^{-4}$	$1.1841*10^{-4}$	$2.1569*10^{-4}$
Standard deviation	$4.9307*10^{-6}$	$1.5685*10^{-5}$	$2.3228*10^{-4}$

Table III. Mean Squared Error – EUR/USD

Iterations	Mean Squared Error		
	Train	***Validation***	***Test***
1	$3.6199*10^{-5}$	$3.7105*10^{-5}$	$4.1646*10^{-5}$
2	$3.7100*10^{-5}$	$3.7924*10^{-5}$	$3.8488*10^{-5}$
3	$3.8090*10^{-5}$	$3.6691*10^{-5}$	$3.7361*10^{-5}$
4	$3.7694*10^{-5}$	$3.4246*10^{-5}$	$3.8251*10^{-5}$
5	$3.6808*10^{-5}$	$3.7144*10^{-5}$	$3.8759*10^{-5}$
6	$3.8302*10^{-5}$	$3.5430*10^{-5}$	$3.4792*10^{-5}$
7	$3.7862*10^{-5}$	$3.4881*10^{-5}$	$3.7759*10^{-5}$
8	$3.6938*10^{-5}$	$3.7867*10^{-5}$	$3.7924*10^{-5}$
9	$3.8322*10^{-5}$	$3.7484*10^{-5}$	$3.6947*10^{-5}$
10	$3.8169*10^{-5}$	$3.5506*10^{-5}$	$3.5472*10^{-5}$

Table IV. Summary – EUR/USD

Summary	Mean Squared Error		
	Train	***Validation***	***Test***
Min	$3.6199*10^{-5}$	$3.4246*10^{-5}$	$3.4792*10^{-5}$
Max	$3.8302*10^{-5}$	$3.7924*10^{-5}$	$4.1646*10^{-5}$
Average	$3.7548*10^{-5}$	$3.6427*10^{-5}$	$3.7739*10^{-5}$
Standard deviation	$7.3840*10^{-7}$	$1.3108*10^{-6}$	$1.8784*10^{-6}$

In the analysis of this time series, the smallest mean squared error occurred in the 3rd epoch and is $1.11455*10^{-4}$ as shown in Figure 2. It represented the deviation of the predicted values in relation to the actual value.

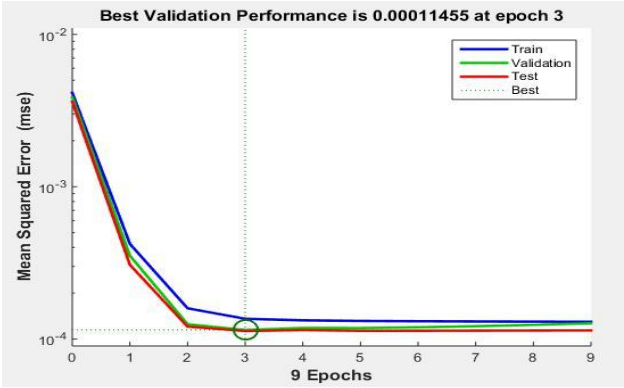

Figure 2. Mean squared error with best validation performance

The training error is significantly higher than the error during testing, which means that the model did not over fitting as shown in Figure 3.

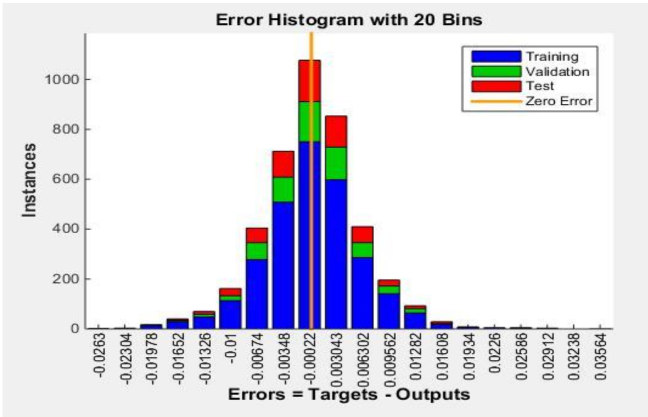

Figure 3. Histogram of time series errors

The first network for the stock exchange index S&P 500 was tested as a feed forward network. The smallest MSE for training was $1.23081*10^{-4}$, for validation $1.0336*10^{-4}$ and for testing $1.1380*10^{-4}$.

The network for the currency pair EUR/USD was tested also as a feed forward network. The smallest MSE was smaller than for the first network: $3.6199*10^{-5}$ for training, $3.4246*10^{-5}$ for validation and $3.4792*10^{-5}$ for testing.

Unlike the analysis of time series in the time domain, in the frequency domain, it is interesting to consider the spectrum of the amplitude (relative share of a certain frequency component relative to the other) of the historical price for the stock index S&P 500 and the currency pair EUR/USD in several different aspects. These analyses include the spectral analysis of time series, which are usually used for stationary time series. This is a good assumption for adjusted stock prices in the frequency domain statistics [13].

In order to better understand the shape of the spectrum, a log-log scale is used and logarithm of the amplitude values obtained after application of Fast Fourier Transform (FFT) is used. Observing the slope of such a curve, could be observed if the spectrum of the amplitude is close to the special power-law form 1/f. Using a logarithmic format is a good way to avoid overestimating high-frequency components.

After applying FFT on prices and returns, equivalent time series in the frequency domain are obtained. As in the above procedure, in order to better detect the spectrum, a modulus representing the amplitude was found, and then the result was logarithmic. The obtained values of the S&P 500 index and EUR/USD currency pair were used to train the NARX neural network. The average mean squared error obtained after 10 consecutive training is $1.5738*10^{-1}$ and $4.8713*10^{-1}$ respectively, which represents a significantly higher number than the one obtained in the time domain. Conclusion is that, regardless of the time series being analyzed, the results are significantly worse and the prediction is less reliable.

The simulation performed with the input that represents the logarithmic value of the amplitude and the frequency as an exogenous input did not show the possibility of good training and convergence even after the maximum possible 1000 iterations, nor the corresponding statistical characteristics, and hence its analysis would make no sense.

V. Conclusion

Analysis of time series and their statistical perspective is a very specific topic and it is necessary to use data science and statistical analysis for dealing with time series. By combining an analysis with a tool such as a neural network in a delicate area such as finance, we can say that in the future neural networks can have a global impact on financial market and trading. It is crucial to apply modern methods and continuously improve them to protect from losses. With assistance of existing platform with varied parameters and transactional data, this tool will be a necessity for forecasting fluctuation on financial market and to secure businesses.

The obtained results of the time series analysis confirmed the possibility of a good prediction. Therefore, this method of forecasting could be an essential advantage for early adopters, especially for those interested in researching and improving current methods. Better forecasting can be done for time series in Forex (EUR/USD), in the time domain without applying Fourier transform to input data. It is proved that NARX is a good method for solving the given type of problem in the time domain, but in the frequency domain it is recommended that the analysis be carried out by a classical feedforward neural network with the backpropagation algorithm. The results of the research indicated that NARX is capable of providing a certain amount of security to those entities that invest their funds, as well as to point out future expectations. Despite the results of the research indicate that NARX can provide valuable information for investors, this paper is not supposed to be a manual for investors, but only a proposal and advice on how to behave on the market during trading. Because of the market variability, it is always recommended to be cautious. News are also integral part of the price and can affect on the market fluctuations, because when particular news arrives on the market, then it reacts to certain changes. The news is then incorporated into the price and the market returns to the previous state where it was before the news arrived.

Improvements can be made with a new research making few changes. It will include new parameters for neural network and different data preparation for the training. Also, it is suggested that number of neurons in the hidden layer could be changed as well as time delay of activation function in the hidden and output layer.

Acknowledgement

This research was supported by Enetel Solutions and we want to show our gratitude for sharing their pearls of wisdom during of the course of this research. We are grateful to all colleagues from the Data Science team and General Manager Rade Vujović for support, inspiration and advice during writing this paper. We share the merits with professors from Faculty of Organizational Sciences - Laboratory of Systems with whom we have had the pleasure to work during this and other related projects.

References

[1] Z.J.M. Shen and X. Su, “Customer behavior modeling in revenue management and auctions: A review and new research opportunities,” *Production and operations management*, 16(6): 713-728, 2007.

[2] A. Kazem, E. Sharifi , F.K. Hussain, M. Saberi, and O.K.Hussain, “Support vector regression with chaos-based firefly algorithm for stock market price forecasting,” *Applied soft computing*, *13*(2), 947-958, 2013.

[3] Kim Kyoung-jae, “Financial time series forecasting using support vector machines,” *Neurocomputing* 55(1-2): 307-319, 2003.

[4] M. Dobrota, A. Poledica, M. Bulajić, B. Petrović, “Modelling volatility using GARCH model: NASDAQ-100 application,” XVIII International conferences on information and communication technologies, YU INFO, Kopaonik (Serbia), February 29 to March 03, 2012. Proceedings, pp. 18-23, ISBN 978-86-85525-09-4, 2012.

[5] T.D. Chaudhuri, I. Ghosh, “Artificial neural network and time series modeling based approach to forecasting the exchange rate in a multivariate framework,” *Journal of Insurance and Financial Management* 1(5): 92-123, 2016.

[6] B. Oancea, Ş.C. Ciucu, (2014). “Time series forecasting using neural networks,” *arXiv preprint arXiv:1401.1333*, 2014.

[7] T.J. Peter, K. Somasundaram, “An empirical study on prediction of heart disease using classification data mining techniques,” In *IEEE-International conference on advances in engineering, science and management (ICAESM-2012)* (pp. 514-518). IEEE.J, 2012.

[8] A. Poledica, D. Marković, S. Živančević, “Logical classification method for bankruptcy prediction,” Data science and business intelligence, XV International Symposium SymOrg 2016, Zlatibor (Serbia), June 10 to 13, 2016. Symposium proceedings, pp. 213-220, ISBN 8676803269, 2016.

[9] Fusion Media Limited. [Online]. Available: https://www.investing.com. [Accessed: 10.09.2018].

[10] Yahoo! Finance. [Online]. Available: https://finance.yahoo.com. [Accessed: 10.09.2018].

[11] Desanka Radunović, Talasići (Wavelets), Akademska misao, p. 159, ISBN: 86-7466-190-4, 2015.

[12] M. Ardalani-Farsa, S. Zolfaghari, “Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks,” *Neurocomputing* 73: 2540–2553, 2010.

[13] M.H. Izadi, Frequency-based analysis of financial time series, Chapter I-III, School of Computer and Communication Sciences, Lausanne, pp. 1-39, 2009.