



# ECHO STATE NETWORKS USAGE FOR STOCK PRICE PREDICTIONS

Lidija Barjaktarović<sup>1\*</sup>,  
Marko Barjaktarović<sup>2</sup>,  
Snežana Konjikušić<sup>3</sup>

<sup>1</sup>Singidunum University,  
Belgrade, Serbia

<sup>2</sup>School of Electrical Engineering,  
University of Belgrade,  
Belgrade, Serbia

<sup>3</sup>FEFA Faculty,  
University Metropolitan,  
Belgrade, Serbia,

## Abstract:

The subject of research is the consideration of potential use of Echo State Networks (ESN) for prediction of stock prices. The reasons are: 1) stock prices show non stationary behaviour and ESN is well suited to time series prediction of chaotic systems; 2) implementation in optical domain can bring very fast inference and financial market transactions requires prompt brought decisions; 3) training time of ESN is short and doesn't require special hardware like training deep Feedforward Neural Network (NNs). Although there is only a few studies about using ESN for prediction financial data (pricing of securities subject of trade, and their volumes), it can be concluded than ESNs have great potential to be used in prediction of the stock prices. Having in mind the intention of the Serbian government to stimulate issuing of bonds, ESN can be applied on Belgrade Stock Exchange.

## Keywords:

Echo State Networks, Reservoir Computing, predictions, stock prices.

## 1. INTRODUCTION

Time series analysis is widely used in many domains, such as speech recognition, cyber-attack detection, structural health monitoring, medical diagnostics, economics, risk management, etc.

Outcome prediction has been very popular recently, even more so with aid of deep learning algorithms. However, traditional algorithms based on Feedforward Neural Network (NN) are not suited for dynamic systems, where historical data must be incorporated in predictions (Gao et.al, 2020). This is a special problem with nonlinear and nonstationary behaviour, which can be hidden in the signals, and feed forward NNs can fit very well to historical data, but poorly performing with contemporary data when there is inherent uncertainty in the signal.

There are a lot of algorithms for time series forecasting, with recurrent neural networks (RNN) being among the most commonly used. In RNNs, signals passing through recurrent connections can be viewed as a memory element in the network while memory contents offer additional gain for predicting future values. However, RNN is difficult to train (Schaetti et.al, 2016) due to the vanishing gradient problem. This is the result of optimizing procedure used to train neural network called backpropagation. The backpropagation uses derivatives of current prediction error between each layer to "learn" from data. In spite of that, with too many layers (deep NNs can have more than 100) and with orbiting in returning branches, error derivatives become very small and eventually updating factor disappears, reducing in poor network performance. In order to cope with this problems, several method exists like LSTM (Long-short term memory), residual deep NNs with skip connections, Echo State Networks (ESN) and others.

Correspondence:  
Lidija Barjaktarović

e-mail:  
lbarjaktarovic@singidunum.ac.rs



Among them ESNs gain lot of attention, especially recently when researcher implement ESN architecture in photonic domain (Photonic Computing) (Tan et.al, 2019). Traditionally, after training NN, which can be time consuming for deep architectures, trained NN is used for prediction (or inference in Machine Learning vocabulary). Although inference is much faster than training, it requires computation power for real time application. Specialized hardware for edge computing like TPU (Tensor Processing Unit, Google Coral /Coral, 2020/), NPU (Neural Processing Unit, Intel Movidius /Intel, 2020/), and GPU (Graphics Processing Unit, like Nvidia Jaton Nano /Nvidia Developer, 2020/) are optimized for computation and more energy efficient than standard CPU, but they still have delay and use some energy, both being proportional to the number of nodes in NN. Using photonic computing, inference time is proportional to propagation of optical signal which is at order of ps ( $10^{-12}$  second) and power dissipation is negligible to standard edge devices (Brunner et.al, 2013). Additionally ESN can be implemented in photonics using only single dynamics node (like nonlinear optical amplifier) and one feedback line (optical fiber) and there is lot of research attention to optimized optical ESN (Tanaka et.al, 2019), as a potential candidate for all optical NN, but that is out of scope of this paper. The subject of research is consideration of potential use of ESN in stock price prediction. There are different reasons for it. Firstly, stock prices show non stationary behaviour and ESN is well suited to time series prediction of chaotic systems. Secondly, implementation in optical domain can bring very fast inference and financial market transactions requires prompt brought decisions. Finally, training time of ESN is short and doesn't require special hardware like training deep NNs.

The article consists of 4 chapters. Introduction is the first chapter. Basic theory of ESN is the second chapter. The application of ESNs for prediction of stock prices is the third chapter. The conclusion is the last chapter.

## 2. BASIC THEORY OF ESN

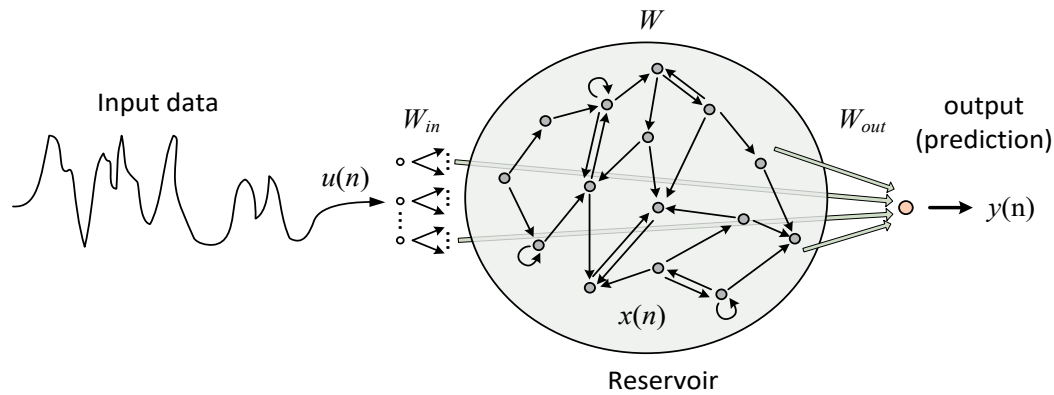
ESN is well suited for non-stationary signals. For a signal it is said that it is stationary if its statistical properties like mean and standard deviation do not change over time (remain constant). Additionally, standard methods for analysing stationary signals require that signal autocorrelation is not varying over time. Stationarity is required for building time series model and prediction of future values. Some earlier ways to deal with non-stationary signals it to "stationarize". It means to decouple non stationarity into several different signal which are stationary, make prediction for each decoupled signal and join all predictions in final outcome. However, sometimes mixing of individual predictions is not straightforward, due to unknown underlying chaotic properties of signal, like stock prices.

In order to overcome the non-stationarity problems, recently ESN has been developed. Essentially ESN is a RNN with a loosely connected hidden layer. This hidden layer has a few hundred nodes, standard values are between 400 and 800, but that number depends on prediction problem and optimal value is found during optimization process. Each node contain one value, which is constant between two time samples. Current node's value depends nonlinearly on contemporary input of network and values from previous moment of that and neighbouring nodes. Inputs and nodes are connected by weights which are randomly initialized. Those weights are not changed during training process and this is one of advantages of ESN. The hidden layer is called a "reservoir" and for that name, ESN is also stated as reservoir computing. The role of reservoir can be observed as nonlinear function which expand dimensionality of input vector, making predictions more accurate.

The prediction (classification or regression, depends on task) is achieved using linear combination of all nodes and current inputs. In figure 1, which shows basic architecture of ESN, the prediction is done by output layer. The weights in output layer are also randomly initialized, but they change during training. The training performs quite fast, due to small numbers of trainable parameters compared to deep NNs (sometimes more than hundred million). If the task is to predict one value in future, number of parameters is equal to sum of number of nodes in reservoir and number of inputs. If classification should be accomplished or prediction of several values in future is a required number of parameters, it is sum of number of nodes and inputs multiplied by number of classes or future values, which is not more than few thousands.



Figure 1. Basic architecture of ESN



Source: Authors' prepared

Due to the fact that only one layer needs to be trained, there is no problem with vanishing gradient. Also, ESN is very computationally efficient and easy to train and above all, several researches shows that ESN works surprisingly well when input is chaotic time series, outperforming other methods for modelling dynamical nonlinear systems.

Mathematically, the prediction  $y$  at current time sample  $n$  can be written as (Lukoševičius, 2012):

$$y(n) = W_{out}[u(n); x(n)] \quad (1)$$

where  $u(n)$  is vector of current inputs,  $x(n)$  represents current state of reservoir (value at each of  $N$  nodes, where  $N$  is number of nodes) and  $W_{out}$  is output weight matrix whose coefficients are trainable. If dimension of output vector  $y$  (column vector) is  $s$  and dimension of input vector  $u$  is  $p$  then dimension of weight matrix is  $s \cdot (p+N)$ . For example, as input stock price from  $p = 100$  previous days are used, reservoir has  $N = 400$  nodes and ESN predicts stock prices for next  $s = 5$  days. In equation 1  $[u(n); x(n)]$  represents column vector obtained by concatenation of vector  $u(n)$  and  $x(n)$ .

Updating reservoir  $x(n)$  is achieved using following equation:

$$x'(n) = \tanh[W_{in} \cdot u(n) + W \cdot x(n-1)] \quad (2)$$

$$x(n) = (1 - \alpha) \cdot x(n-1) + \alpha \cdot x'(n) \quad (3)$$

where  $x(n)$  is vector with state (numerical value) of all nodes in reservoir,  $\tanh$  is hyperbolic tangent function, but other function can be used as nonlinear activation,  $W_{in}$  is input weight matrix with dimension  $N \cdot p$ ,  $W$  is reservoir state  $N \cdot N$  matrix and  $\alpha$  is leaking parameter which defines how fast memory is fading out.

If true value for output is  $y_{target}$  than Root Mean Square Error (RMSE) is commonly used as an estimation of ESN accuracy:

$$E(y, y_{target}) = \frac{1}{s} \sum_{i=1}^s \sqrt{\frac{1}{k} \sum_{n=1}^k (y_i(n) - y_{target,i}(n))^2} \quad (4)$$

where  $k$  is dimension of training set (number of pairs with input  $u(n)$  and corresponding output  $y_{target}(n)$ ). If RMSE is divided by variance of  $y_{target}$  Normalized RMSE (NRMSE) is obtained.

Trainable weight  $W_{out}$  if determined by minimizing RMSE.

### 3. THE APPLICATION OF ESNS FOR THE PREDICTION OF STOCK PRICES

As stated, ESN is quite novel algorithm for time series prediction and only few researches have been published for estimation of financial data set. Table 1 represents review of the related researches.



**Table 1.** Review of researches related to the application of ESNs for the prediction of financial data state

Research	Method	Period of analysis	Financial data set	Results
Ilies et.al, 2007	ESN	18 months	Unknown source	ESN must be optimized for each application;
Fangwen et.al, 2010	ESN	05/16/03-10/01/07;	the daily prices of securities (opening, the highest, the lowest, closing) incorporated in six major stock indexes in the world DJIA, S&P 500, NASDAQ, HSI, FTSE 100 and NIKKEI 225;	ESN achieves smallest prediction error
Bernal et.al, 2014	ESN	Late 2004-early 2009;	the prediction of S&P 500 index;	the proposed architecture showed good accuracy
Liu et.al, 2017	ESN	04/01/00-01/21/16;	Daily prices of securities of Shanghai Composite index (opening, the highest, the lowest, closing) and especially created the ratio of the average closure in the previous 5 or 10 days;	Provides more reliable prediction of securities' pricing;
Liu et.al, 2018	ESN (first optimized on task-independent test)	Period is not stated	the dataset from Shanghai stock market (with 4,579 sets);	the proposed model achieves a better prediction accuracy;
Taehwan and King, 2020	Deep ESN - instead of one, 6 reservoir are used in serial configuration	14/08/1999 - 01/08/2019;	S&P 500 index	Slightly better than standard ESN, but much more complicated and slower to train.

*Source:* Authors' created on the basis of available researches

Ilies et.al (2007) investigated the possibility of ESN usage for predicting of financial data. As dataset of 118 time series from 18 months period is used from unknown source (it was part of competition). They showed that depending on data, for the best results the size of reservoir varies between 45 and 110 nodes, indicating than ESN must be optimized for each application.

Fangwen et.al (2010) presented a study of using an ESN to estimate the next value in financial time series. Authors chosen to use daily prices of securities incorporated in six major stock indexes in the world DJIA, S&P 500, NASDAQ, HSI, FTSE 100 and NIKKEI 225. Firstly, they proved that selected time series manifest predictable behaviour using Hurst exponent. They extracted several parameters like Row Stochastic Value (reflects price changes in short-term), K line (indicates strength or weakness on the market in the future period of time), MTM parameters (shows how quickly is respond to change) and Volume's Rate of Change (measure of investors' emotional state in short-term period). Secondly, they investigated whether it was possible to reduce number of features using PCA (Principle Component Analysis). Moreover, they trained ESN using financial data of 1100 days long stock indices

set and then compared ESN with two other NN algorithms (Elman network and backpropagation network). Finally, they found that ESN achieves the smallest prediction error.

Bernar et.al (2012) used ESN for prediction of S&P 500 index (based on 500 biggest companies in the United States of America). ESN had  $N = 100$  nodes with  $\alpha = 0.8$  and used 5 day history to predict next day value. The proposed architecture showed good accuracy, better than moving average (next value is predicted as average from resent ones) and Kalman filter (standard algorithm in control theory commonly used for prediction true value from sensors).

Lui et al (2017) proposed a way how to improve the performance of ESN averaging prediction results from 18 ESNs using voting (weighing each result by estimated accuracy of each ESN). They got 4% increase in accuracy using this method over standard ESN. As dataset they used values of Shanghai Composite Index for 16 years.

Liu et.al (2018) suggested optimisation of using of improved fruit fly optimization algorithm (IFOA). IFOA is basically an iterative search through the parameters with the goal to find optimal ones for specific tast.





Basically, parameters space of ESN was searched to find the optimal ones. That optimized ESN is then used to predict financial time series. The data set comes from Shanghai stock market and it has 4,579 sets where 80% is used for training and 20% is used for testing. The experiment is a one-step prediction, where six financial indicators (i.e. opening, the highest, the lowest and closing price of securities subject of trade, daily trading volume, and daily turnover of securities subject of trade) are used as input features, and the opening price (of securities subject of trade) on the next day is the output. Compared to the backpropagation (BP), Elman neural networks and standard ESN, the proposed model achieves a better prediction accuracy.

Taehwan and King (2020) research is the most recent research about using ESN in predictions of financial data set i.e. prices of securities and volumes of trade. S&P 500 index was used as dataset. Using paradigm from deep neural network, authors suggested applying 6 reservoirs in a row. Output from the current reservoir is input for the next one, except for the first reservoir, where the input is the input of network. The final output of the network is linear combination of outputs from all six reservoirs and only those weights are trainable. This deep ESN manifested smaller error than gated recurrent unit (GRU) neural network, LSTM (Long short-term memory) and standard (one reservoir) ESN, but with longer training time.

#### 4. CONCLUSION

Although there is only a few studies about using ESN for predicting financial data (pricing of securities subject of trade, and their volumes), it can be concluded that ESNs have great potential to be used in prediction of the stock prices, due to the fact that they can foresight future value of time series well, when the underlying mechanism demonstrates chaotic behaviour. Additional ESN is easy and fast to train, so it can be implemented in photonic circuit which guarantees instant inference with negligible power dissipation and this is crucial for analysing big amount of data like in economics, where not only trading with securities, but also in enterprise risk management (especially in the fraud risk management, credit risk management – early warning system, operational risk management – business continuity and disaster recovery plan, market risk management, liquidity risk management, etc.).

For future work, optimisation of ESN parameters could be considered, such as finding number of  $N$  nodes, leaking rate  $\alpha$  and length  $p$  of input sequence. This parameters should be adjusted for each task, like to predict prices of securities and volume of trading on Belgrade Stock Exchange (BELEX). Additionally, the usage of other activation functions like sinus function might be researched due to

fact that it recently showed drastically faster training of deep NNs (Stizmann et.al, 2020). Finally, recent efforts of using deep ESNs should be also investigated in more details.

#### 5. ACKNOWLEDGMENT

This paper is: a) part of the project COST CA19130 - Fintech and Artificial Intelligence in Finance - Towards a Transparent Financial Industry; b) supported by the Ministry of Science and Technological Development of the Republic of Serbia.

#### 6. LITERATURE

- Bernal, A., Fok, S., Pidaparathi, R. (2012). Financial Market Time Series Prediction with Recurrent Neural Networks, *Stanford University*, December 2012.
- Brunner, D., Soriano, M. C., Mirasso, C. R., Fischer, I. (2013). Parallel photonic information processing at gigabyte per second data rates using transient states, *Nature Communications*, Vol. 4, article number: 1364.
- Coral (2020) Products, available on the site: <https://coral.ai/>, date of the access 09/25/20
- Gao, P. Zhang, R., Yang, X. (2020). The Application of Stock Index Price Prediction with Neural Network. *Mathematical and Computational Applications MDPI*, Vol. 25 (3), 53, doi:10.3390/mca25030053.
- Ilies, I., Jaeger, H., Kosuchinas, O., Rincon, M., Sakenas, V., Vaskevicius, N. (2007). Stepping forward through echoes of the past: forecasting with Echo State Networks, *University Bremen*.
- Intel (2020) Content, available on the site: <https://www.intel.com/content/www/us/en/products/processors/movidius-vpu/movidius-myriad-x.html>, date of the access 09/25/20.
- Kim, T., King, B.R. (2020). Time series prediction using deep echo state networks, *Neural Computing and Applications*, April 2020, <https://doi.org/10.1007/s00521-020-04948-x>.
- Liu, Z., Liu, Z., Song, Y., Gong, Z., Chen, H. (2017). Predicting stock trend using multi-objective diversified Echo State Network, *Seventh International Conference on Information Science and Technology (ICIST)*, April 2017, 10.1109/ICIST.2017.7926754.
- Liu, J., Sun, T., Luo, Y., Fu, Q., Cao, Y., Zhai, J., Ding, X. (2018). Financial Data Forecasting Using Optimized Echo State Network, *International Conference on Neural Information Processing ICONIP 2018*. Published in Springer's *Lecture Notes in Computer Science*, Vol 11305, 138-149.



- Lukoševičius, M. (2012). A Practical Guide to Applying Echo State Networks, chapter in: Montavon , G., Orr, G.B., Müller K.R. (eds), “Neural Networks: Tricks of the Trade”. Lecture Notes in Computer Science, Vol 7700. Springer, 2012, [https://doi.org/10.1007/978-3-642-35289-8\\_36](https://doi.org/10.1007/978-3-642-35289-8_36).
- Nvidia Developer (2020) Autonomous Machines, available on the site: <https://developer.nvidia.com/embedded/jetson-nano-developer-kit>, date of the access 09/25/20.
- Schaetti, N., Salomon, M., Couturier, R. (2016). Echo State Networks-Based Reservoir Computing for MNIST Handwritten Digits Recognition, IEEE Intl Conference on Computational Science and Engineering, 10.1109/CSE-EUC-DCABES.2016.229.
- Sitzmann, V., Martel, J. N. P. , Bergman, A. W. , Lindell, D. B., Wetzstein, G. (2020). *Implicit Neural Representations with Periodic Activation Functions*, Preprint - under review, Jun 2020, arXiv:2006.09661.
- Tan, X., Hou, Y., Wu, Z., Xia, G. (2019). Parallel information processing by a reservoir computing system based on a VCSEL subject to double optical feedback and optical injection, Optics Express, Vol. 27, Issue 18, 26070-26079.
- Tanaka, G., Yamane, T., Héroux, J.B., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D., Hirose, A. (2019). Recent advances in physical reservoir computing: A review, Neural Networks, Vol. 115, 100-123.
- Zhai, F., Lin, X., Yang, Z., Song, Y. (2010). Financial time series prediction based on Echo State Network, Sixth International Conference on Natural Computation, Aug. 2010, 10.1109/ICNC.2010.5584802.