Research Article

Deterministic Echo State Networks Based Stock Price Forecasting

Jingpei Dan,[1] Wenbo Guo,[1] Weiren Shi,[2] Bin Fang,[1] and Tingping Zhang[3]

[1] *College of Computer Science, Chongqing University, Chongqing 400044, China*
[2] *School of Automation, Chongqing University, Chongqing 400044, China*
[3] *College of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China*

Correspondence should be addressed to Jingpei Dan; danjingpei@cqu.edu.cn

Echo state networks (ESNs), as efficient and powerful computational models for approximating nonlinear dynamical systems, have been successfully applied in financial time series forecasting. Reservoir constructions in standard ESNs rely on trials and errors in real applications due to a series of randomized model building stages. A novel form of ESN with deterministically constructed reservoir is competitive with standard ESN by minimal complexity and possibility of optimizations for ESN specifications. In this paper, forecasting performances of deterministic ESNs are investigated in stock price prediction applications. The experiment results on two benchmark datasets (Shanghai Composite Index and S&P500) demonstrate that deterministic ESNs outperform standard ESN in both accuracy and efficiency, which indicate the prospect of deterministic ESNs for financial prediction.

1. Introduction

Prediction in financial or stock markets is a challenging research topic since the stock market is mostly complex (nonlinear) and volatile. Various technical, fundamental, and statistical indicators have been proposed; however, none of them or combination of techniques has been successful enough. The application of artificial neural network (ANN) in prediction problems is very promising for its general function approximation property and effective learning algorithms. Among them, recurrent neural networks (RNNs) offer arbitrary approximation for dynamical systems with arbitrary precision theoretically compared to other traditional ANNs. However, it is difficult for traditional RNNs based on gradient descent approach to develop in application because of slow convergence and high computational cost. A new form of RNN training methods, echo state network (ESN), has been proposed by Jaeger and Haas [1], which is simple and applicable for time series prediction with high accuracy and computational efficiency. During these ten years, ESNs have been applied in many areas, including time series forecasting [2–7], wireless communication [1], robot control [8], and

speech recognition [9]. Lin et al. investigated the effectiveness of ESN to predict the future stock prices in a short term. The experimental results on nearly all stocks in S&P 500 indicated that ESNs outperformed some conventional neural networks in most cases [10, 11]. Although it is verified that standard ESNs can predict future stock prices, the parameters have been determined by trials and errors due to their randomly constructed reservoir. Recently, a deterministically constructed form of ESN has been proposed, which has competitive performance with minimum complexity compared to standard model [12]. Since deterministic ESN model is relatively new, few application researches have been studied until now [13, 14]. This study investigates the performance of deterministic ESN for time series forecasting and its application of stock price prediction. The experiments are based on two benchmark datasets (Shanghai Composite Index and S&P 500) and three types of deterministic ESNs (DLR, DLRB, and SCR) are compared with standard ESNs in forecasting accuracy, stability, and efficiency. The experimental results demonstrate that deterministic ESNs outperform standard ESN in both accuracy and efficiency, which indicate the prospect of deterministic ESNs for financial prediction.
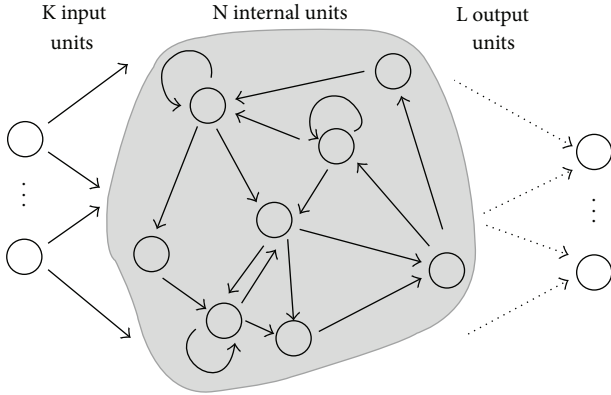
FIGURE 1: Architecture of standard ESN [1].

Section 2 briefly introduces standard ESNs and deterministic ESNs; Section 3 presents the experiments and the comparison results of applying deterministic ESNs for stock price forecasting on Shanghai Composite Index and S&P 500 stock datasets. Finally, it is concluded in Section 4.

2. Background

2.1. Echo State Networks. Echo state network is a recurrent discrete-time neural network with input, internal units (dynamic reservoir), and output. In the dynamic reservoir, there are large numbers of sparsely connected neurons where the weights are random and fixed. The states of the dynamic reservoir, called "echo states," reflect history information of inputs. The memoryless readout is the only part that needs to be trained linearly with echo states; thus the training becomes a simple linear regression, which makes the application of RNNs fast and easy.

2.1.1. Architecture. The basic architecture of ESN with K input units, N neurons in the dynamic reservoir, and L output units is shown in Figure 1. The solid arrows in Figure 1 represent the synaptic connections whose weights are randomly generated and fixed, while the dashed arrows are the trainable output weights.

2.1.2. Training Algorithm

Input. A training set of (input/output) $S = \{(u(n), d(n)), n = 1, 2, \ldots, T\}$, $u(n)$, and $d(n)$ are K-dimensional and L-dimensional vectors.

Output. A trained ESN $(W_{in}, W, W_{fb}, W_{out})$ is such that its output $y(n)$ of the input $u(n)$ is close to desired output $d(n)$, where W_{in}, W, W_{fb}, and W_{out} are input, internal, feedback, and output weight matrices, respectively, with suitable sizes.

(1) Initialize a random RNN (W_{in}, W, W_{fb}) which has the echo state property (ESP); spectral radius of W should be smaller than unity.

(2) Activate the reservoir to the dynamic N-dimensional states $x(n)$ of the RNN by

$$x(n+1) = f\left(Wx(n) + W_{in}u(n+1) + W_{fb}y(n)\right),$$
$$x(0) = 0, \tag{1}$$

where f is a state activation function which is usually a sigmoid function and W is the internal recurrent connection weight matrix, of which the spectral radius should be smaller than unity in order to ensure the ESN to work; in other words, W should satisfy the ESP.

(3) Compute the output weights as the linear regression weights of L-dimensional outputs $y(n)$ on the states $x(n)$ by

$$y(n) = g\left(W_{out}(x(n):u(n))\right), \tag{2}$$

where g is an output activation function.

2.2. Deterministic Echo State Networks. As explained previously, the reservoirs in standard ESNs are constructed randomly so that it is hard to construct ESN and unlikely to optimize the setup. To simplify the reservoir construction, deterministically constructed ESNs (Figure 2) have been proposed [12].

Three reservoir templates with fixed topologies concerned in this study are described as follows.

(1) Delay line reservoir (DLR): the units in reservoir are arranged in a line, as shown in Figure 2(a). Obviously, the reservoir matrix W has the property that nonzero values are located on the lower subdiagonal of W; that is, $W_{i+1,i} = r$ for $i = 1, \ldots, N-1$, where r is the weight of all the feedforward connections.

(2) DLR with backward connections (DLRB): each neuron in reservoir is connected to both the preceding neuron and the successive one, as shown in Figure 2(b). Thus, the lower subdiagonal and upper subdiagonal of W both have nonzero elements, $W_{i+1,i} = r$ and $W_{i,i+1} = b$, where b is the weight of the feedback connections.

(3) Simple cycle reservoir (SCR): neurons in reservoir like in Figure 2(c) are organized in a cycle, where $W_{i+1,i} = r$ and $W_{1,N} = r$.

Except the reservoir construction, the input layer in deterministic ESN is fully connected to the reservoir and all the input connections have the same absolute weight value $v > 0$ in contrast to randomly generated input weights in standard ESN. The sign of v is determined randomly by a random draw from Bernoulli distribution of mean 1/2.

Generally, the deterministic form of ESNs has the following properties, which simplifies the ESN construction by only setting two free parameters and enables a more thorough theoretical analysis of the ESN performance [12]:

(1) a simple fixed nonrandom reservoir topology with full connectivity from inputs to the reservoir,

Table 1: Comparative results of ESN and DLR by average forecasting accuracy (MSE).

| Stocks	ESN	DLR			Performance improvement (%)
		$r = 0.5$	$r = 0.7$	$r = 1$	
Minsheng Bank	**0.010935**	0.011121	0.011055	0.011175	−1.10
Sinopec	0.026413	0.011680	**0.011059**	0.037695	58.13
Baotou Steel Rare Earth	2.083440	2.307846*	2.238868*	**1.987902***	4.59
Yangquan Coal Industry	0.210954	**0.187878**	0.197090	0.225766	10.94
The Gemdale Group	0.025079	**0.023939**	0.024479	0.025767	4.55
China Shenhua	0.125612	**0.116748**	0.120193	0.128756	7.06
Western Mining	0.039255	**0.033440**	0.035893	0.042173	14.81
China Life	0.165341	**0.157840**	0.160240	0.169839	4.54
Luan Huanneng	0.326770	**0.320343**	0.322941	0.335372	1.97
PetroChina	0.014539	**0.007712**	0.009112	0.018609	46.96
China COSCO	0.046103	**0.015973**	0.020535	0.059996	65.35
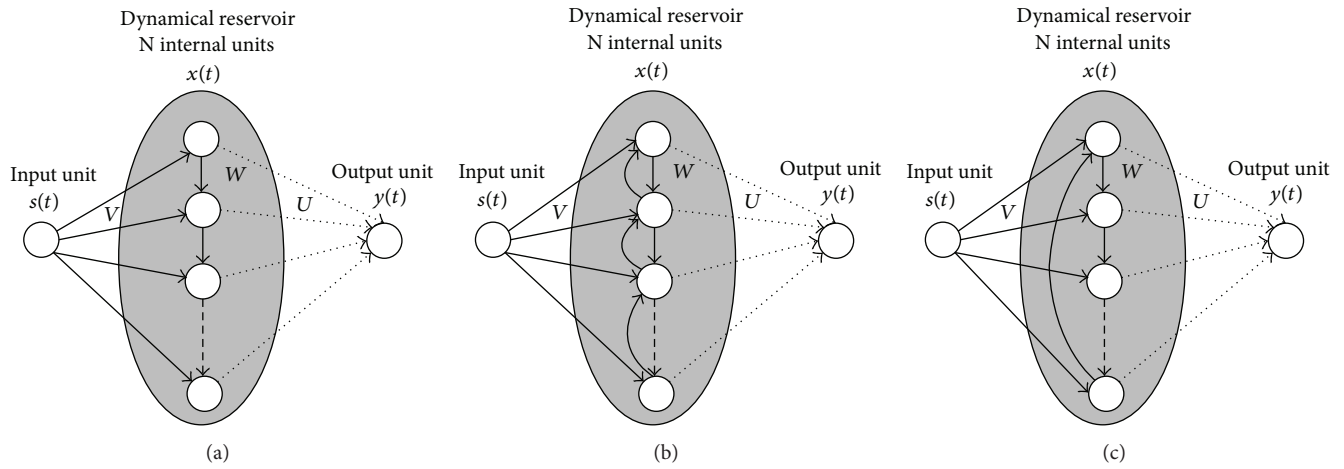
Figure 2: (a) DLR. (b) DLRB. (c) SCR. [12].

(2) a single fixed absolute weight value r for all reservoir connections, and

(3) a single weight value v for input connections, with aperiodic pattern of input signs.

3. Stock Price Forecasting Based on Deterministic Echo State Networks

In this study, to investigate performances of deterministic ESNs for stock price prediction, three types of deterministic ESNs described previously, DLR, DLRB, and SCR, are compared to standard ESN. Since stock price fluctuates violently, in order to make the experiments fair and sound, two benchmark datasets from Shanghai Composite Index and S&P 500 are adopted in this study.

3.1. Experimental Data

1. Shanghai Composite Index Stock Price Dataset. The trading price data from 14 December, 2007 to 31 December, 2012 (totally 1230 days) of fifty stocks in the benchmark Shanghai composite index is adopted.

2. S&P 500 Stock Price Dataset. S&P 500 is usually considered as the benchmark for the United States' equity performance; hence it is popular for stock price prediction research. The trading price data from 1 June, 2010 to 20 February, 2014 (totally 930 days) of 374 stocks in S&P 500 is also adopted in the experiments. All data come from Yahoo Finance.

In the experiments, 80% of data points in each stock price data are adopted as training set, and the remaining 20% are used for testing.

3.2. Data Preprocessing.
In most time series prediction researches, the original data are usually normalized in order to make the preprocessed data in a desired range. In this study, all the original data are normalized as follows:

$$X = k \times \frac{x - \min(x)}{\max(x) - \min(x)} \pm t, \tag{3}$$

where x is the original data, $\min(x)$ is the minimum data, and $\max(x)$ is the maximum value in the data, while k and t are two coefficients to make the preprocessed data X in a desired range.

TABLE 2: Comparative results of ESN and DLRB by average forecasting accuracy (MSE).

Stocks	ESN	DLRB			Performance improvement (%)
		$r = 0.5, b = 0.05$	$r = 0.7, b = 0.1$	$r = 1, b = 0.2$	
Minsheng Bank	0.010935	0.011139*	0.011047*	**0.010884***	0.46
Sinopec	0.026413	**0.012175**	0.017957	0.500804	53.91
Baotou Steel Rare Earth	2.083440	2.296547*	2.227776*	**1.889110***	9.33
Yangquan Coal Industry	0.210954	**0.188163**	0.202954	0.284849	10.80
The Gemdale Group	0.025079	**0.023920**	0.024640	0.030820	4.62
China Shenhua	0.125612	**0.117322**	0.123177	0.139961	6.60
Western Mining	0.039255	**0.033725**	0.037880	0.079748	14.09
China Life	0.165341	**0.158533**	0.161766	0.205709	4.12
Luan Huanneng	0.326770	**0.319993**	0.324581	0.373107	2.07
PetroChina	0.014539	**0.007990**	0.010373	0.317348	45.05
China COSCO	0.046103	**0.016929**	0.025623	2.403775	63.28

TABLE 3: Comparative results of ESN and SCR by average forecasting accuracy (MSE).

Stocks	ESN	SCR			Performance improvement (%)
		$r = 0.5$	$r = 0.7$	$r = 1$	
Minsheng Bank	0.010935	0.011165*	0.011061*	**0.010842***	0.85
Sinopec	0.026413	**0.011868**	0.014594	0.037969	55.07
Baotou Steel Rare Earth	2.083440	2.350138*	2.228629*	**1.975701***	5.17
Yangquan Coal Industry	0.210954	**0.187053**	0.196191	0.224106	11.33
The Gemdale Group	0.025079	**0.023933**	0.024412	0.025769	4.57
China Shenhua	0.125612	**0.116739**	0.120273	0.129783	7.06
Western Mining	0.039255	**0.033329**	0.035651	0.042433	15.10
China Life	0.165341	**0.157979**	0.160666	0.169405	4.45
Luan Huanneng	0.326770	**0.319483**	0.322704	0.334651	2.23
PetroChina	0.014539	**0.007842**	0.009102	0.018705	46.06
China COSCO	0.046103	**0.016163**	0.020780	0.063481	64.94

3.3. Prediction Accuracy Evaluation Measure.

In this study, a commonly used measure, mean squared error (MSE) shown as follows, is utilized to evaluate forecasting accuracy:

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^{n} (y_t - x_t)^2, \qquad (4)$$

where x_t denotes a given time series, $t = 1, \ldots, n, n \in N$, and y_t denotes the forecast of value at time t.

3.4. Experiments and Results

3.4.1. Experiments on Shanghai Composite Index Stock Price Dataset.

In this experiment, each ESN model (standard ESN, DLR, DLRB, and SCR) is built with 96 internal units and spectral radius is set to be 0.8. For each stock in this dataset, the next trading day closing price is predicted by its previous 30 closing prices. These forecasting models are investigated according to different values of input weight matrix W_{in}.

Random Input Weight Matrix W_{in}. The input weight matrix W_{in} is randomly set to be in the uniform distributed interval of $[-1, 1]$. The parameter r in internal weight matrix W of DLR, DLRB, and SCR models is set to be 0.5, 0.7, and 1.0, respectively. Accordingly, the parameter b in W of DLRB is set to be 0.05, 0.1, and 0.2.

Because of the randomness of W_{in}, each forecasting model is tested for 1000 times. Each deterministic ESN model is compared with standard ESN model by average forecasting accuracy. The experimental results are shown in Tables 1, 2, and 3. In these tables, the stocks of which the predicting performances are improved as the parameters increase are italic, while reduced ones are with asterisks. From the comparative results, the forecasting performances of DLR, DLRB, and SCR are increased by 19.90%, 19.48%, and 19.71%, respectively, which indicate that the deterministic ESNs outperform the standard ESN in forecasting accuracy. In addition, it is found that the forecasting performances of deterministic ESNs tend to change in one direction (increase or decrease) as the parameter values increase.

Deterministic Input Weight Matrix W_{in}. The forecasting performances of these ESN models are compared under the condition of the same W_{in} value. In this experiment, all the values in W_{in} are $v = 0.5$, and the sign of v is the probability

TABLE 4: Comparison of ESN and three deterministic ESN models by average forecasting accuracy (MSE) and stability improvement (%).

Stocks	Average forecasting accuracy				Stability improvement (%)		
	ESN	DLR	DLRB	SCR	DLR	DLRB	SCR
Minsheng Bank	**0.01097**	0.01118	0.01121	0.01118	−11.83	−30.99	−9.66
Sinopec	0.02155	0.01092	0.01125	**0.01089**	98.04	98.13	98.38
Baotou Steel Rare Earth	**2.18330**	2.36140	2.32180	2.35490	−42.67	−19.43	−42.68
Yangquan Coal Industry	0.20609	0.18508	0.18552	**0.18415**	69.47	66.43	67.29
The Gemdale Group	0.02499	0.02383	0.02385	**0.02382**	42.46	43.31	46.34
China Shenhua	0.12363	**0.11524**	0.11603	0.11537	67.26	60.85	65.96
Western Mining	0.03783	**0.03275**	0.03313	0.03278	81.06	78.36	79.57
China Life	0.16336	**0.15640**	0.15747	0.15717	39.28	42.17	36.40
Luan Huanneng	0.32304	0.31968	0.31931	**0.31875**	40.16	45.31	47.00
PetroChina	0.01200	**0.00745**	0.00756	0.00745	97.60	97.16	97.51
China COSCO	0.03419	**0.01448**	0.01492	0.01460	98.86	98.74	98.72

TABLE 5: Comparisons of average forecasting accuracy (MSE) and running time (s).

Stocks	Average forecasting accuracy ($r = 0.5$, $b = 0.05$, $v = \pm 0.5$)				Average running time (s)			
	ESN	DLR	DLRB	SCR	ESN	DLR	DLRB	SCR
"A"	3.814025	2.947236	**2.87385**	3.042726	0.1441	0.1103	0.1093	0.1134
"ALL"	0.729146	0.737923	**0.71832**	0.730556	0.1428	0.1092	0.1089	0.1087
"BAX"	**0.47736**	0.504663	0.504453	0.496313	0.1422	0.1097	0.1100	0.1094
"CAT"	1.227007	**1.21849**	1.221667	1.221545	0.1428	0.1091	0.1094	0.1092
"COH"	0.702417	0.670763	0.673214	**0.67075**	0.1424	0.1091	0.1099	0.1095
"DF"	**0.23828**	0.262445	0.262805	0.264985	0.1422	0.1095	0.1093	0.1089
"EIX"	**0.21518**	0.224842	0.225257	0.224794	0.1440	0.1094	0.1095	0.1098
"FMC"	**1.02408**	1.04685	1.034131	1.026145	0.1434	0.1092	0.1091	0.1091
"HD"	**4.05567**	4.809645	4.593392	4.835269	0.1396	0.1093	0.1092	0.1093
"IFF"	2.963587	2.798408	**2.70644**	2.977701	0.1420	0.1082	0.1082	0.1083
"KIM"	**0.04468**	0.045833	0.045762	0.045979	0.1440	0.1094	0.1094	0.1090
"LXK"	**0.51446**	0.522021	0.523765	0.521172	0.1446	0.1088	0.1087	0.1088
"MS"	**0.28087**	0.324216	0.323608	0.319095	0.1445	0.1094	0.1088	0.1082
"OI"	0.558538	0.476727	0.480012	**0.47534**	0.1450	0.1095	0.1100	0.1085
"PKI"	3.549767	2.748976	**2.64719**	2.657761	0.1446	0.1104	0.1094	0.1094
"RL"	**6.35157**	6.377153	6.40885	6.409925	0.1454	0.1101	0.1102	0.1102
"SRE"	7.084271	6.978886	**6.78874**	7.210348	0.1497	0.1095	0.1109	0.1099
"TER"	0.105475	0.104689	**0.10413**	0.106195	0.1421	0.1094	0.1099	0.1097
"VAR"	1.697554	**1.58872**	1.617942	1.613488	0.1444	0.1107	0.1103	0.1094
"WY"	**0.15677**	0.16673	0.165503	0.166299	0.1441	0.1094	0.1091	0.1093
"WYN"	1.538461	1.378661	**1.35106**	1.37561	0.1438	0.1095	0.1095	0.1092
"XEL"	**0.05319**	0.05395	0.053997	0.05375	0.1415	0.1084	0.1086	0.1091
"XOM"	1.870293	1.827589	1.824464	**1.80176**	0.1405	0.1089	0.1089	0.1086
"XRX"	**0.06333**	0.073928	0.073774	0.073895	0.1401	0.1086	0.1089	0.1086
"YUM"	2.818102	2.876777	2.806851	**2.79861**	0.1407	0.1089	0.1085	0.1086

of 50% randomly determined. In DLR, DLRB, and SCR, all the values in internal weight matrix W are $r = 0.5$ and the feedback connection values in W of DLRB are $b = 0.05$. The average forecasting accuracy in MSE for 1000 times tests for each ESN model is shown in Table 4, which indicates that SCR, DLR, and DLRB outperform standard ESN in average accuracy by 15.94%, 15.93%, and 15.51%, respectively.

During the experiment, for the first 500 tests, it is found that the values of forecasting accuracy (MSE) for the ESN models differ significantly, which indicate that the forecasting performances of the ESN models are unstable. To further investigate the stability, another 500 tests have been done, of which the results also show the instability. Therefore, the stability of each ESN model is evaluated by MSE for the 1000

tests. As shown in Table 4, compared to the standard ESN model, the stabilities of SCR, DLRB, and DLR models are improved by 53.17%, 52.73%, and 52.70%, respectively.

3.4.2. Experiments on S&P 500 Stock Price Dataset. The forecasting models in this experiment are constructed the same as the experiment described in Section 3.4.1. This experiment aims to further investigate the forecasting performance of each ESN model for various stocks since there are more representative stocks in S&P 500 dataset than in Shanghai Composite Index dataset. Meanwhile, time complexities are compared by computing the average running time. The experimental results are shown in Table 5 partly. Although the ESN models perform competitively in forecasting accuracy, the deterministic ESN models have obvious advantages in efficiency which is indicated by the average time saving of 23.46%, 23.45%, and 23.41% for SCR, DLRB, and DLR, respectively.

4. Conclusions

A new method for stock price forecasting by applying deterministic ESN models is presented in this paper. In standard ESNs, the reservoirs are randomly constructed, which makes researchers and practitioners rely on trials and errors. Differently, the deterministic ESNs have their reservoirs constructed deterministically so that the structures of ESN models are simple and easy to be applied. Therefore, deterministic ESNs (three typical types) have been applied to predict stock price. The forecasting performances have been investigated on two benchmark datasets (Shanghai Composite Index and S& P500). The experimental results on Shanghai Composite Index dataset show that deterministic ESNs outperform standard ESN in accuracy by about 20% and stability by 52% averagely. Further investigation on S&P 500 dataset shows that the deterministic ESNs have improvement in efficiency by about 23% while having insignificant improvement in forecasting accuracy. Since deterministic ESNs have advantages of higher forecasting accuracy and efficiency as indicated by the experimental results, they have great prospects in applications of stock price prediction and even other time series forecasting applications. In this work, the deterministic ESN models are constructed by hand. The problem of parameter settings will be considered in future work by applying optimization algorithms.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] H. Jaeger and H. Haas, "Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, 2004.

[2] Q. Song, X. Liu, and X. Zhao, "Short-term traffic flow and hourly electric load forecasting algorithm based on echo state neural networks," *International Journal of Digital Content Technology & Its Applications*, vol. 6, no. 4, pp. 166–172, 2012.

[3] Z. Shi and M. Han, "Support vector echo-state machine for chaotic time-series prediction," *IEEE Transactions on Neural Networks*, vol. 18, no. 2, pp. 359–372, 2007.

[4] C. Sheng, J. Zhao, Y. Liu, and W. Wang, "Prediction for noisy nonlinear time series by echo state network based on dual estimation," *Neurocomputing*, vol. 82, pp. 186–195, 2012.

[5] Y. Kara, M. Acar Boyacioglu, and Ö. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5311–5319, 2011.

[6] F. Zhai, X. Lin, Z. Yang, and Y. Song, "Financial time series prediction based on Echo State Network," in *Proceedings of the 6th International Conference on Natural Computation (ICNC '10)*, vol. 8, pp. 3983–3987, chn, August 2010.

[7] H. Zhang, J. Liang, and Z. Chai, "Stock prediction based on phase space reconstruction and echo state networks," *Journal of Algorithms & Computational Technology*, vol. 7, no. 1, pp. 87–100, 2013.

[8] M. Chang, A. Terzis, and P. Bonnet, "Mote-based online anomaly detection using echo state networks," *Distributed Computing in Sensor Systems*, vol. 5516, pp. 72–86, 2009.

[9] M. D. Skowronski and J. G. Harris, "Noise-robust automatic speech recognition using a predictive echo state network," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 5, pp. 1724–1730, 2007.

[10] X. Lin, Z. Yang, and Y. Song, "Short-term stock price prediction based on echo state networks," *Expert Systems with Applications*, vol. 36, no. 3, pp. 7313–7317, 2009.

[11] X. Lin, Z. Yang, and Y. Song, "Intelligent stock trading system based on improved technical analysis and Echo State Network," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11347–11354, 2011.

[12] A. Rodan and P. Tiňo, "Minimum complexity echo state network," *IEEE Transactions on Neural Networks*, vol. 22, no. 1, pp. 131–144, 2011.

[13] A. Rodan and P. Tiňo, "Simple deterministically constructed cycle reservoirs with regular jumps," *Neural Computation*, vol. 24, no. 7, pp. 1822–1852, 2012.

[14] P. Tiňo and A. Rodan, "Short term memory in input-driven linear dynamical systems," *Neurocomputing*, vol. 112, pp. 58–63, 2013.