ESKİŞEHİR TECHNICAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
A- APPLIED SCIENCES AND ENGINEERING

2019, 20(2), pp. 153 - 160, DOI:10.18038/aubtda.443510

A NOVEL MULTIPLICATIVE NEURON MODEL BASED ON SINE COSINE ALGORITHM FOR TIME SERIES PREDICTION

Erdinc KOLAY[*]

Department of Statistics, Faculty of Art and Sciences, Sinop University, Sinop, Turkey

ABSTRACT

Time series prediction is a method to predict the system behavior in the future based on current given data. Neural Networks (NNs) approach is a well-known technique that is useful for time series prediction. In the literature many NN models such as Multilayer Perceptron (MLP), Pi-Sigma NN (PSNN), Recurrent NN etc. are proposed for solving time series prediction. In this paper, we use Multiplicative Neuron Model (MNM) to predict time series. For training this model, we propose to use newly developed evolutionary optimization algorithm called Sine Cosine algorithm (SCA), and this algorithm has not been used as far as we know in training the MNM. The proposed SCA-MNM model is employed for the well-known time series problems. In this paper, the application of the SCA-MNM on time prediction is illustrated using two mostly used datasets Mackey-Glass time series dataset, Box-Jenkins gas furnace dataset. To investigate the effect of the proposed SCA-MNM model, comparisons are made with some of the results given in the literature.

Keywords: Neural networks, Multiplicative neuron model, Sine Cosine algorithm, Time series prediction

1. INTRODUCTION

NNs are inspired by nervous system and mainly used for solving any scientific problem like pattern classification, time series prediction, etc. Scholars state that the nervous system can be modelled by means of a function although it cannot be fully explained. Therefore, NNs have been recognized as a simulation of biologic neuron cells. Transferring basic logic inputs into a certain function with specific mathematical operators is to lie behind data processing and machine learning. By this means, the machine learns from data and decides for new observations that it has never seen during training. The training of the NNs refers to the process of finding optimal network parameters via any optimization algorithm.

Optimization refers to finding optimal or near-optimal values for the parameters of a given system to find minimum or maximum point of its output. In the process of training NN, the weights of the network elements are adjusted by an optimization procedure. As is known, Back Propagation Algorithm (BPA) is based on the gradient descent method and depends on the initialized values and this method may easily fall into local optima. Recently, different meta-heuristic algorithms, including Particle Swarm Optimization (PSO) [1] or SCA [2], are playing a vital role in solving optimization problems and these algorithms can improve the training performance of the neural network.

Time series prediction is a method to predict the system behavior in the future based on current given data. Time series can contain linear and non-linear components frequently. Statistical models such as Autoregressive integrated moving average (ARIMA) [3] has been used for linear time series prediction. However, the ARIMA model include only linear component of time series and fails to including any nonlinearity. In this paper, we propose using MNM model because its multiplicative building block is promising non-linear time series prediction naturally.

In the literature, some papers propose several NN with various structure of NN like [4–9], various training algorithms [10–14] or both [15,16]. MNM is comprised of only single neuron and demands less

*Corresponding Author: ekolay@sinop.edu.tr
Recieving :13.07.2018 Accepted: 18.06.2019

training time. The MNM has been proposed and successfully applied for time series prediction in [7]. Its multiplicative building block is promising for non-linear time series prediction naturally. For training MNM, authors have used back-propagation algorithm (BPA) which based on gradient descent method to train the network. However, BPA can fall into the local optima easily. In the past few decades, a lot of optimization algorithms based on nature-inspired have been proposed in the literature, and to overcome above mentioned shortcomings of BPA, different evolutionary algorithms and its variants have been proposed in literature to train the MNM like PSO [14], glowworm swarm optimization [16], differential evolutionary algorithm [17] and harmony search algorithm [18]. Although a wide range of evolutionary algorithms are investigated in the literature for training NNs, the problem of local minima is still open. According to No-Free-Lunch (NFL) [19] theorem there is no optimization algorithm that can solve all optimization issues. Motivated by these reasons, in this work, a new MNM training method based on the SCA is proposed.

SCA is population based stochastic algorithm and firstly introduced in [2] to solve multi-object optimization problems. The SCA establishes multiple random candidate solutions and improves them to fluctuate towards or outwards the best solution by a mathematical model based on sine and cosine function. SCA and its variants was tested on unimodal, multimodal, composite benchmark functions, aircraft's wings and many other scientific problems [20]. Also, the SCA is used to solve feature selection problems [21], structural damage detection [22] and etc. [23,24]. Although the SCA has been used in many areas, it has not yet been used in MNM training as far as we know. In this paper, we use the SCA to train MNM model to obtain prediction of time series, and some comparisons of performance of different evolutionary algorithms are given.

The paper is organized as follows. In section 2 and section 3, we introduce the MNM model and SCA algorithm, respectively. In section 4, the proposed model is introduced and the algorithm based on SCA for training of this model is presented. Time Series Prediction problems are given in section 5, and conclusions are given in the last section.

2. MULTIPLICATIVE NEURON MODEL

MNM is based on the polynomial structure and firstly introduced in [7]. Contrary to MLP and other higher order NN like PSNN, MNM uses a simple aggregation function. The structure of MNM model which including n input is given in Figure 1. In this NN model, simple aggregation function is considered as product of linear functions like in (1). Also, MNM model reduces the computational time due to its simple architecture and involving fewer parameter contrary to MLP and PSNN. As shown in Figure 1, MNM model with n input includes $2\times n$ parameter where n parameter are weights and rest are the biases. Outputs of network are obtained by using (1) and (2),

$$\Omega(x, w, b) = \prod_{j=1}^{n} \left(x_j w_j + b_j\right), \tag{1}$$

where, x is the inputs, w and b are weights and biases, respectively. In this case, output of *i-th* learning sample with logistic activation is shown in (2).

$$y^{(i)} = \frac{1}{1+e^{-\Omega(x,w,b)}}. \tag{2}$$

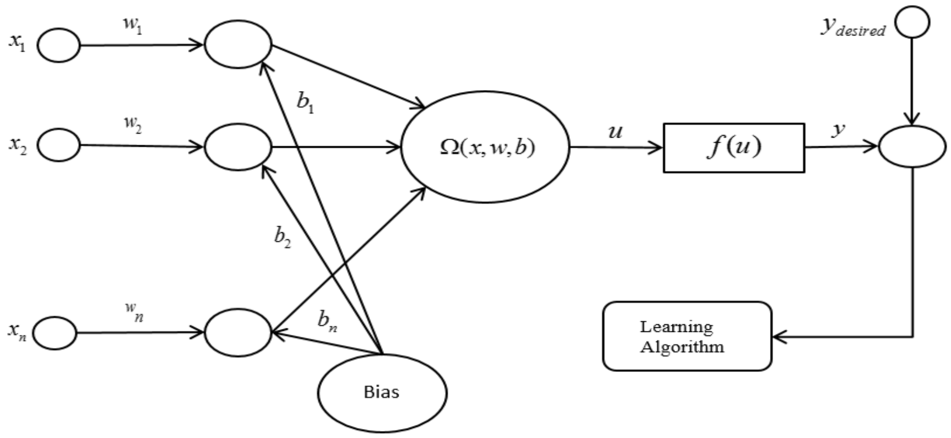

Figure 1. The structure of the MNM model

3. SINE COSINE ALGORITHM

SCA is a newly developed population-based algorithm that is used as global optimization approach to solve optimization problems. As all population-based algorithms, SCA creates multiple initial search agents randomly. Then, in the iterative process, the SCA searches for optimal solution with updating their positions outwards and towards the best solution according to sine and cosine functions. Some of the important features of SCA is that it searches for a wide range of search space, avoid local optima and quickly reaches an optimal solution. The search agents are guided toward an optimal solution in the search space via a fitness function that evaluate each search agent in each iteration of the algorithm [21]. In the SCA, each search agent is represented by *n* dimensional vector $X_i = (x_{i1}, x_{i2}, ..., x_{in})$ where X_i is the *i-th* search agent in the population. Also, the algorithm keeps the position of best individual solutions of each search agent in memory which is symbolized by *P* in the (3) during iterative process. In the SCA, each search agent X_i is updating by using (3-4).

$$\begin{aligned} X_i^{t+1} &= X_i^t + r_1 \times \sin(r_2) \times \left| r_3 P_i^t - X_i^t \right|, \qquad r_4 < 0.5 \\ X_i^{t+1} &= X_i^t + r_1 \times \cos(r_2) \times \left| r_3 P_i^t - X_i^t \right|, \qquad r_4 \geq 0.5 \end{aligned} \tag{3}$$

$$r_1 = a - t\frac{a}{Tmax}, \tag{4}$$

where *t* is the current iteration number, and r_1 is a control parameter that balances global search and local search of the algorithm which decreases linearly from constant value a to 0 by each iteration according to (4), and r2, r3 and r4 are the random numbers which ranged of (0,1) and T_{max} is the maximum iteration number.

4. THE PROPOSED SCA-MNM MODEL

The proposed method refers to using SCA for optimizing weights and biases of MNM which is showed in Figure 1. To be optimized vector of parameters can be shown in following vector called V_P.

$$V_P = [w_1 \quad w_2 \quad ... \quad w_n \quad b_1 \quad b_2 \quad ... \quad b_n], \tag{5}$$

where *w* and *b* are the network weights and biases, respectively. This vector can be considered as X_i (search agent) in the SCA. For beginning, random candidate weights and biases are generated *m* times where *m* is number of the generated population. Then according to optimal solution, the vector updates via (3) and (4). The algorithm of the proposed SCA-MNM model is given step by step in *"SCA-MNM Algorithm"*.

'SCA-MNM Algorithm'

**Normalization: First, time series data is divided into two group, train and test sets. Then, the data is converted from 0.9 to 0.1 with a formula similar to the following.*

$$x_{train} = 0.8\frac{x_{train}-\min(x_{train})}{\max(x_{train})-\min(x_{train})} + 0.1, x_{test} = 0.8\frac{x_{test}-\min(x_{train})}{\max(x_{train})-\min(x_{train})} + 0.1.$$

**Initialization: Random population matrix Par which rows created by m random V_p vectors is initialize. The size of this matrix is $m\times(2\times n)$, and m is the number of populations which is chosen arbitrarily. Each row of the matrix Par can be defined as search agent.*

$$Par = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} & b_{11} & b_{12} & \dots & b_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} & b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{m1} & w_{m2} & \dots & w_{mn} & b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}_{m\times(2\times n)}$$

The fitness function is determined as Minimum Square Error (MSE).

$MSE = \sum_{i=1}^{k} (y_{desired}^{i} - y^{i})^2$,*where k is the number of training sample.*

Using Eq. (1-2), MSEs are calculated for each row, and obtain m MSE values.

The matrix Par are updated by using (3) and (4). Then the new values of the P and MSEs are obtained.

This procedure is repeated until the maximum iteration number is achieved.

At the end, P is the best search agent corresponding to MNM's weights and biases based on minimum MSE. Now, predictions and MSE values of test sample can be obtained using P.

5. TIME SERIES PREDICTION PROBLEMS

In this paper, the application of the SCA-MNM on time prediction is illustrated using two mostly used datasets [7,14,16] Mackey-Glass (M-G) time series dataset, Box-Jenkins (B-J) gas furnace dataset. These data sets have been pre-processed by normalizing them between 0.1 and 0.9 like in '***SCA-MNM Algorithm'***. In all experiments, maximum number of iterations for SCA is set to be 1000, constant value *a* is set to 2, and the population number *m* is set to be 30. The proposed method is performed on MATLAB R2015a.

The results are summarized with the mean, the best and also the standard deviations of methods are in Table 1 and Table 2 for M-G time series dataset, B-J gas furnace time series dataset, respectively. Also, we compare the training and testing performance to some training algorithm in the literature like BPA [7], Genetic Algorithm (GA), PSO and Cooperative Random Learning Particle Swarm Optimization (CRPSO) [14] and hybrid of Glowworm Swarm Optimization and Differential Evolution Algorithm which called LWGSODE [16].

5.1. Mackey-Glass Time Series Prediction Problem

The M-G time series [18] is a chaotic time series, based on Mackey-Glass differential equation. This dataset is often regarded as a benchmark used for testing the performance of prediction methods. The M-G delay-difference equation is given in (6),

$$y(t+1) = a\frac{y(t-\tau)}{1+y^{10}y(t-\tau)} + y(t)(1-b) \tag{6}$$

where a=0.2, b=0.1 and $\tau = 17$. The objective is to obtain the prediction of $y(t+1)$ by using four measurements $y(t)$, $y(t-6)$, $y(t-12)$ and $y(t-18)$. The training is performed on 450 samples and the model is tested on 500 samples. The performances of the training and testing for different procedures have been given in Table. 1. From this table, using SCA for training MNM improves the performance of prediction results and SCA has better performance to BPA, PSO and LWGSODE.

Table 1. Performance comparison for prediction of MG time series over 50 independent runs

		BPA	PSO	CRPSO	GA	LWGSODE	SCA
Training	Mean	0.0038	0.0017	5.24e-04	5.95e-04	0.0022	8.13e-04
	Std	0.0037	0.0025	2.19e-06	7.17e-05	4.05e-04	5.71e-05
	Best	5.35e-04	5.25e-04	5.23e-04	5.38e-04	9.03e-04	7.08e-04
Testing	Mean	0.0046	0.0018	5.49e-04	6.22e-04	0.0019	7.10e-04
	Std	0.0040	0.0027	3.05e-06	7.25e-04	0.0012	4.14e-05
	Best	5.65e-04	5.31e-04	5.46e-04	5.18e-04	5.10e-04	6.22e-04

The best parameter vector V_p was found as in (7) using the SCA-MNM model.

$$V_p = [0.4490\ -0.0220\ 0.0967\ -2.2331\ -0.2224\ -1.4193\ -0.9442\ 9.0850] \tag{7}$$

Also, the results of training and testing performance and also errors are shown in Figure 2.

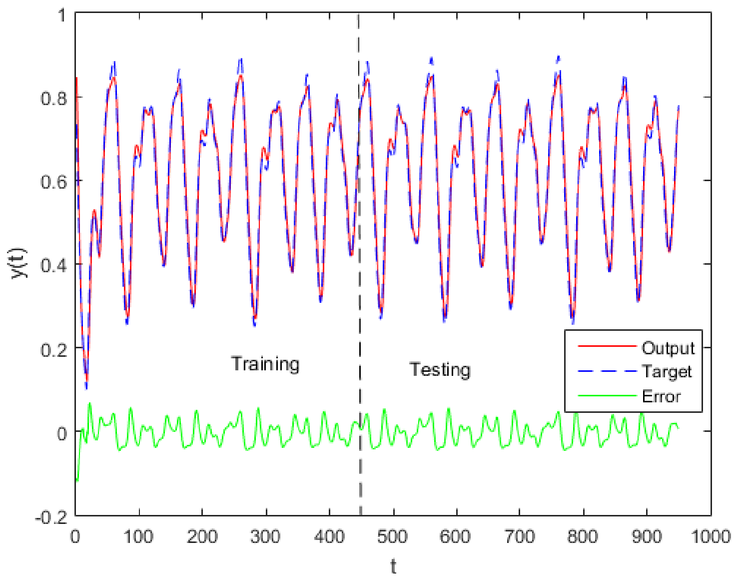

Figure 2. The prediction results of the M-G time series using the SCA-MNM model

5.2. Box- Jenkins Gas Furnace Time Series Prediction Problem

In the B-J gas furnace dataset [3] while the furnace input is assigned as the gas flow rate $u(t)$, the furnace output is assigned as CO_2 concentration $y(t)$. One combines air and methane in the gas furnace to get a mixture including CO_2 gas, and then we model the furnace output $y(t+1)$ in terms of the previous output functions $y(t)$ and $u(t-3)$. The data set is including 296 observations, and we use first 140 samples to train network and 150 samples for testing performance of the developed models. The training MSEs, testing MSEs for 50 independent runs are given in Table 2.

Table 2. Performance comparison for prediction of BJ time series over 50 independent runs

		BPA	PSO	CRPSO	GA	LWGSODE	SCA
Training	Mean	0.0030	0.0029	0.0017	0.0018	0.0018	0.0017
	Std	0.0019	0.0019	5.54e-04	7.72e-04	7.74-e-04	5.09-e05
	Best	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016
Testing	Mean	0.0056	0.0054	0.0021	0.0023	0.0026	0.0019
	Std	0.0049	0.0050	0.0015	0.0021	0.0014	7.39e-05
	Best	0.0019	0.0019	0.0018	0.0018	0.0018	0.0018

Additionally, The best parameter vector V_p was found as in (8) using the SCA-MNM model for B-J time series.

$$V_P= [-0.3787\ 0.9036\ 4.8492\ -0.4525] \tag{8}$$

Also, Figure 3 shows the performance of training and testing. As is seen in Table 2, SCA has the best performance for training MNM. The best MSE for training sample is 0.0016 for all methods. However, the results of standard deviation show that SCA has minimum deviation between all mentioned procedures, and is successful for obtaining the best solution.

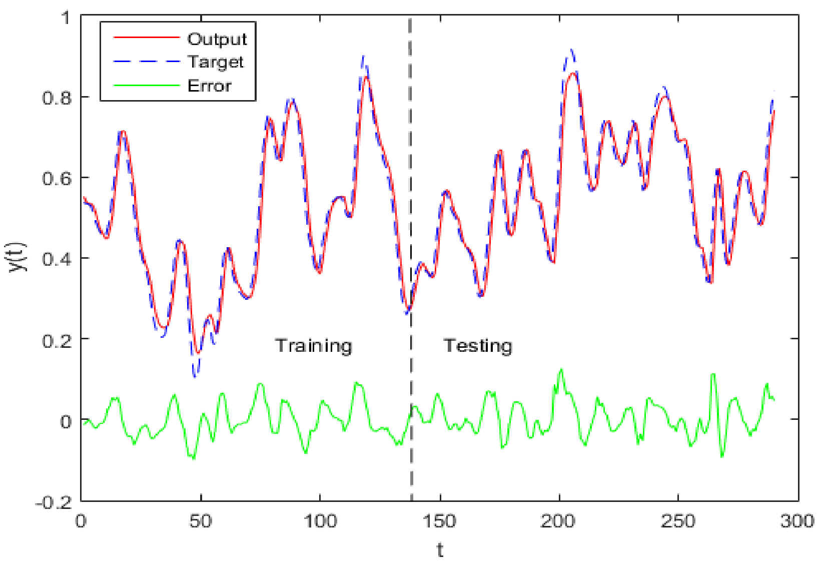

Figure 3. The prediction results of the B-J time series using the SCA-MNM model

6. CONCLUSIONS

For the last few decades, time series prediction has been an important research area. The validity of time series prediction is fundamental to making the right decision for any system. For this reason, studies are constantly continuing to make a better prediction. For this context, thanks to the nonlinear mapping capability, NNs are became mostly used tools for time series prediction. However, NN or MNM training is a hard optimization problem including many local optimum points.

Due to the its nature of the above-mentioned meta-heuristic algorithms, there is no guarantee to find best solution in the time series problems with one algorithm. Also, the No-Free-Lunch [19] theorem claim that there is no optimizer that is good enough to solve all optimization problems. Therefore, training algorithms will be upgraded consistently by researchers to reach the best algorithm for solving corresponding problems.

In this paper, we propose to consider using SCA algorithms to train the MNM network. Various training algorithms have been proposed in the literature. However, SCA or SCA-based algorithms are not used for training MNM up to now as far as we know. Experimental results have shown that using SCA for training MNM has satisfactory performance for given datasets.

REFERENCES

[1] Eberhart R., Kennedy J. A new optimizer using particle swarm theory, MHS'95. Proc. Sixth Int. Symp. Micro Mach. Hum. Sci. 1995; 39–43.

[2] Mirjalili S. SCA: A Sine Cosine Algorithm for solving optimization problems, Knowledge-Based Syst. 96 (2016) 120–133.

[3] Box GEP, Jenkins GM & Reinse GC. Time series analysis: Forecasting and control. Englewood Cliffs, New Jersey, USA: 1994; Prentice Hall.

[4] Karunasinghe DSK, Liong SY. Chaotic time series prediction with a global model: Artificial neural network, J. Hydrol., 2006; 323 92–105.

[5] Khashei M, Bijari M. An artificial neural network (p, d ,q) model for timeseries forecasting, Expert Syst. Appl., 2010; 37 479–489.

[6] Connor J R, Martin L. Atlas, Recurrent Neural Networks and Robust Time Series Prediction, Neural Networks, IEEE …. 5, 1994; 240–254.

[7] Yadav RN, Kalra PK, John J. Time series prediction with single multiplicative neuron model, Appl. Soft Comput., 2007; 7 1157–1163.

[8] Ghazali R, Hussain A, El-Deredy W. Application of Ridge Polynomial Neural Networks to Financial Time Series Prediction, IEEE Int. Jt. Conf. Neural Netw. Proc., 2006; 913–920.

[9] Gunn E, MacDonald C. Artificial Neural Networks in Finance and Manufacturing, 2006.

[10] Yu JJQ, Lam AYS, Li VOK. Evolutionary artificial neural network based on Chemical Reaction Optimization, in: 2011 IEEE Congr. Evol. Comput., IEEE, 2011: pp. 2083–2090.

[11] Hagan MT, Menhaj MB. Training Feedforward Networks with the Marquardt Algorithm, IEEE Trans. Neural Networks. 5, 1994; 989–993.

[12] Sheikhan M, Mohammadi N. Time series prediction using PSO-optimized neural network and hybrid feature selection algorithm for IEEE load data, Neural Comput. Appl., 2013; 23 1185–1194.

[13] Panigrahi S, Bhoi AK, Karali Y. A Modified Differential Evolution Algorithm trained Pi-Sigma Neural Network for Pattern Classification, 2013; 133–136.

[14] Zhao L, Yang Y. PSO-based single multiplicative neuron model for time series prediction, Expert Syst. Appl. 2009; 36 2805–2812.

[15] Egrioglu E, Yolcu U, Aladag CH, Bas E. Recurrent Multiplicative Neuron Model Artificial Neural Network for Non-linear Time Series Forecasting, Procedia - Soc. Behav. Sci. 2014; 109 1094–1100.

[16] Cui H, Feng J, Guo J, Wang T. A novel single multiplicative neuron model trained by an improved glowworm swarm optimization algorithm for time series prediction, Knowledge-Based Syst., 2015; 88 195–209.

[17] Bas E, The training of multiplicative neuron model based artificial neural networks with differential evolution algorithm for forecasting, J. Artif. Intell. Soft Comput. Res., 2016; 6 5–11.

[18] Worasucheep C. Training a single multiplicative neuron with a harmony search algorithm for prediction of S&P500 index - An extensive performance evaluation, Proc. 2012 4th Int. Conf. Knowl. Smart Technol. KST 2012. (2012); 1–5.

[19] Wolpert DH, Macready WG. No free lunch theorems for optimization. IEEE Trans Evol Comput, 1997; 1(1): 67-82.

[20] Singh N, Singh SB. A novel hybrid GWO-SCA approach for optimization problems, Eng. Sci. Technol. an Int. J. 2017; 20 1586–1601.

[21] Hafez AI, Zawbaa HM, Emary E, Hassanien AE. Sine cosine optimization algorithm for feature selection. In: International symposium on INnovations in intelligent systems and applications (INISTA), 2016; pp 1-5.

[22] Bureerat S, Pholdee N. Adaptive sine cosine algorithm integrated with differential evolution for structural damage detection" in proceeding international conference on computational science and its application, Lecture Notes Comput. Sci., 2017; (LNCS) 10404 71–86.

[23] Rizk-Allah RM. Hybridizing sine cosine algorithm with multi-orthogonal search strategy for engineering design problems, J. Comput. Des. Eng., 2017.

[24] Turgut OE. Thermal and economical optimization of a shell and tube evaporator using hybrid backtracking search-sine-cosine algorithm, Arab. J. Sci. Eng., 2017; 1–20.

[25] Mackey M, Glass L. Oscillation and chaos in physiological control systems, Science (80-.). 197 1977; 287–289.