Neural Networks Optimization by Real Genetic Algorithms, Application to Exchange Rate Prediction

Raihane Mechgoug and Lakhmissi Cherroun

Abstract— **Neural networks have been successfully used for prediction and exchange rate forecasting. However, due to the large number of parameters to be estimated, the task of selecting an appropriate neural structure and its parameters is difficult. The performances of neural network are critically depending on learning algorithms for obtaining the best control parameters. Even when a suitable setting of parameters (weights) can be found, the ability of the resulting network to generalize the data during learning may be far from optimal. For these reasons it is suitable and attractive to apply an optimization method as genetic algorithms. In this paper, a Genetic algorithm is used for automating the design of neural network for exchange rates prediction. It provides a useful tool for neural parameters optimization. Simulation results show effectiveness of the proposed method for optimization of the predictor model using foreign exchange rate. This hybrid model exhibits effectively improved accuracy. Results are compared and discussed.**

Key words: **Prediction, Time Series, foreign exchange, Neural network, Genetic Algorithm.**

I. INTRODUCTION

Exchange rate forecasting is an important and challenging task for both academic researchers and business practitioners. Various theoretical models including both econometrics and time series approaches have been suggested to model and forecast exchange rates. During last twenty five years many different nonlinear models have been proposed in the literature to model and forecast exchange rates. The GARCH model has been used in the past for volatility estimation in U.S. dollar foreign exchange markets Baillie and Bollerslev [1], and in the European Monetary System (Neely, 1999). Initial studies into explanatory power of out-of-sample forecasts gave out disappointing results West and Cho [2]. Jorion [3] found that volatility forecasts for several major currencies from the GARCH model were outperformed by implied volatilities generated from the Black and Scholes option pricing model [4]. These studies used the squared daily return as the variable to be forecast. Since, the exchange rate may move around a lot during the day, it has been established that one can significantly improve the forecasting power of the GARCH model by using sum of intraday squared returns Andersen and Bollerslev [5]. This measure is referred to as integrated or realized volatility. Some authors claimed that exchange rates are rather unpredictable and so random walk model is better predictor Chang and Osler [6]; Meese and Rose [7] Gencay [8]. Kuan and Liu [9] estimate and select feedforward and recurrent networks to evaluate their forecasting performance of five exchange rates against USD. The networks performed differently for different exchange rate series. Yao and Tan [10] show that if technical indicators and time series data are fed to neural networks to capture the underlying process of the movement in currency exchange rates, then useful prediction can be made. Yang and Gradojevic [11] construct a neural network that never performs worse than a linear model but always performs better than the random walk model when predicting Canadian dollar/dollar exchange rate. Kiani and Kastens [12] have successfully employed neural networks to forecast the exchange rate.

The neural networks have particular's properties which make them suitable to solve a wide family of problems. However, they also have some limitations and disadvantages which do not allow their applications in some cases [13]. There are various forms for combining the neural network and genetic algorithms. The study of this combination can be summarizes in three general aspects: learn genetically the network neurons, genetic optimization of network topology and optimization of the control parameters [13]. The combination of neural network and genetic algorithms allows the merging all advantages of both individual approaches and having more the capacity to surmount difficulties and the limitations which characterize each approach [14].

The objective of this paper is to use a real genetic algorithm to automate de design of a neural network predictor; where a multi-layer feedforward neural network is used with back-propagation learning algorithm. Its topology and control parameters are selected by an optimization process. The designed predictor will be employed for predicting daily currency exchange rates of US Dollar (USD), against four other currencies such as Canadian Dollar(CAD), Great British Pound (GBP), Japanese Yen (JPY) and (EUR) using their historical exchange rates from 02/01/2003 to 30/06/2011 was found on the web site: http://pacific.commerce.ubc.ca/xr/. A total of 8 years historical exchange rates data, for each of four currency rates were collected and used as inputs to build the prediction model in our study, and then we add six month exchange rates data were used to evaluate the model. The prediction results of this method are compared with [15] based on two evaluation indicators such as Normalized Mean Square Error (NMSE) and Mean Absolute Error (MAE).

Raihane Mechgoug is with Electrical Engineering Department, University of Biskra, Algeria (e-mail: Raihane_2005@yahoo.fr).
Lakhmissi Cherroun is with Sciences and Technology Department, University of Djelfa, Algeria (e-mail: cherroun_lakh@yahoo.fr).

This paper is organized as follows: in section 2, we present briefly a review of the prediction and neural network theory. In section 3, the optimization of the neural predictor using the real genetic algorithms is explained and discussed. Section 4 gives the simulation results of application of this method for the exchange rates prediction. A comparison study with other model will be presented in section 5. Section 6 Conclude the paper.

II. PREDICTION AND NEURAL NETWORKS

Time series forecasting model assumes that there is an underlying process from which data are generated and the future value of a time series is solely determined by the past and the current observation. Neural network are able to capture the underlying pattern within a time series even when the underlying law governing the value of time series is unknown or too complex to describe. Therefore, for a time series forecasting problem the inputs of network are the past lagged observations of the time series and the output is the future value.

$$y_t = f\left(y_{t-d}, y_{t-2*d}, \ldots\ldots, y_{t-N*d}\right) \qquad (1)$$

Where N : the number of input and d: time lagged,

The NN architectures considered here is a feedforward network or multi-layer perceptron (*MLP*). In which the signal flow of input is from input layer towards output layer, and the training phase is obtained with Back propagation algorithm (BP).

The topology of the predictor requires the determination of:

- Number of hidden layers *(NHL)*.
- Number of nodes *(Nhl)* for each layer *l*
- The activation functions (f^1, f^{out}) for each layer.
- Presence or absence of the bias *(b)*.
- The learning rate (α).
- Momentum coefficient (η).
- The input number *(N)*.
- The times that separate two successive input *(d)*

The GA has been used to determine the appropriate set of the parameter listed above.

III. OPTIMIZING OF NEURAL NETWORK STRUCTURE

The neural network predictor (NP) is defined by genetic encoding in which the chromosome composed of the different characteristics of the NP. We consider a specific representation of the parameters related to the neural predictor structure; the chromosome is constituted of 17 genes. The number of hidden layers *(NHL)*, the numbers of neurons in each layer *l (Nhl),* and other genes representing the activation function type for the hidden layers and the output layer f^1, f^{out}, the number of neurons in the input layer *(N)*, the presence or absence of bias *(b)*, lag *(d)*, the learning rate (α), momentum coefficient (η), the variation interval of the synaptic weights for the 2 hidden layer and the output layer.

A. The Generation of the initial population

The chromosome in the initial population is produced by a stochastic generation (see Fig.1). Each allele is defined in a subset S_i where $i = 1\ldots7$ reported in the second row of the Table 1.

TABLE.I THE SUBSET OF EACH GENE

	NHL	η	α	b	N	Nh1	Nh2	δ	f^1	f^2	f^{out}
set	S_1	S_2	S_2	S_3	S_4	S_5	S_5	S_6	S_7	S_7	S_7

	$high_1$	low_1	$high_2$	low_2	$high_3$	low_3
set	S_2	S_2	S_2	S_2	S_2	S_2

Since it is sufficient in most applications that a network has few hidden layers, the value of *Nhl* has been allowed to vary from 1 to 2 while the maximum value for *Nhl* is calculate using the flowing equation:

$$\max\_\,Nhl = \frac{(N+1)}{2} + \sqrt{n\_data} \qquad (2)$$

Where n_data represents the number of the training data. All allele are defined in the following subsets:

$$S_1 = \{1,2\}, S_2 = [0,1], S_3 = \{0,1\}, S_4 = \{2,3,\ldots,30\},$$
$$S_5 = \{3,4,\ldots.,40\}, S_6 = \{2,3,\ldots.,10\}, S_7 = \{1,2,3\} \qquad (3)$$

S_7 reflects the activation function; we can assume that the value $1 \equiv f_1$ for the sigmoid, $2 \equiv f_2$ for the hyperbolic tangent, $3 \equiv f_3$ for the pure linear. S_3 means that $0 \equiv$ absence of the bias, $1 \equiv$ the presence of the bias. [low_i,$high_i$] presents the interval variation of the synaptic weights in the two hidden layers and output layer.

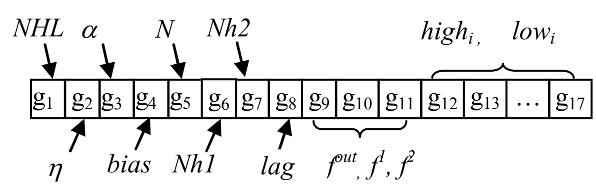


Fig.1. Chromosome representation.

B. Fitness Function:

To evaluate the goodness of an individual, the neural is trained with fixed number of data and then evaluated according to determinate their parameters. The parameters which seem to describe better the goodness of a neural configuration are the mean square error *MSE* at the end of training. This function is calculated by the following form:

$$f(x) = \frac{\sum_{i=1}^{N}(y_i - a_i)^2}{N} \quad (4)$$

Where a_i is the output of the neural predictor.

C. Genetics Operators:

1. Selection: Selection of a pair of individuals from the population. There are variety types of selection methods. In this work we use *remainder selection* that assigns parents are selected from the integer part of each individual's scaled value and then uses roulette selection on the remaining fractional part [13].

2. Crossover: is the reproduction process, in which two chromosomes exchanges some of their corresponding genes [13]. The crossover used is *heuristic crossover* that returns a child that lies on the line containing the two parents, a small distance away from the parent with the better fitness value in the direction away from the parent with the worse fitness value. We can specify how far the child is from the better parent by the parameter Ratio. If $parent_1$ and $parent_2$ are the parents, and $parent_1$ has the better fitness value, the function returns the child is :

$$child = parent_2 + R\,(parent_1 - parent_2) \quad (5)$$

Where R is the Ratio.

3. Mutation: Is an operation that applied for every gene of the new individuals with probability of P_m. The mutation used is *adapt feasible Mutation* witch randomly generates directions that are adaptive with respect to the last successful or unsuccessful generation.

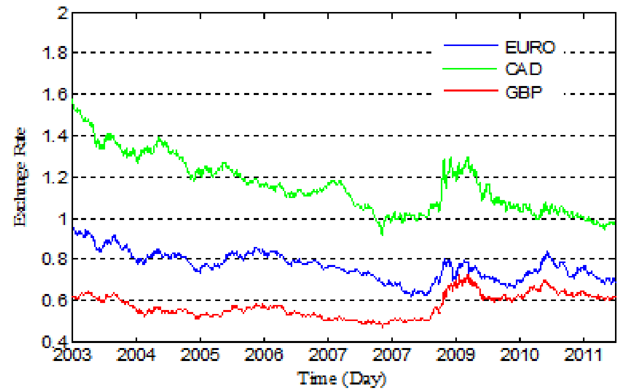
IV. SIMULATION RESULTS

In this section, we will describe the steps of optimizing the neural predictor and the simulation results for the exchange rate prediction.

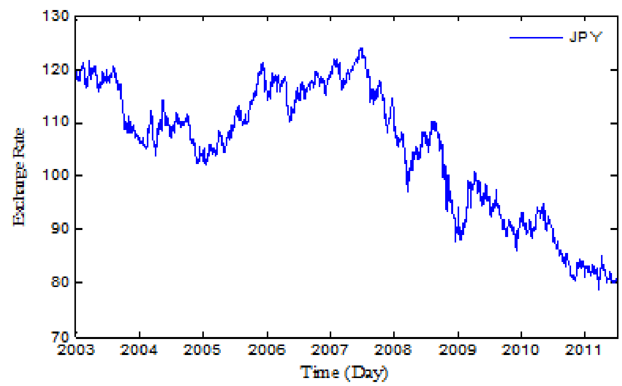
A. Data collection

The data used in this study is the foreign exchange rate of four different currencies against US Dollar (USD) from 02/01/ 2003 to 30/06/2011 found on the web site: http://pacific.commerce.ubc.ca/xr. We considered exchange rate European (EUR), British Pound (GBP), Japanese Yen (JPY), Canadian Dollar (CAD). The data used to train the four forecasters cover a period of 8 years from 02/01/2003 to 31/12/2010 whereas the test data covers seven months from 04/01/2011 to 30/06/2011.

Fig. 2 (a-b) shows the historical exchange rates for different currencies.



(a)



(b)

Fig 2: Historical exchange rates for (a) EUR, GBP, CAD and (b) JPY.

Evaluation indicators

The forecasting performance of the above model is evaluated against a two evaluation indicators namely; Normalized Mean Square Error (NMSE) and Mean Absolute Error (MAE). These criteria are defined in Equation 5 and 6. y_i and a_i are the actual and predicted values respectively. NMSE and MAE measure the deviation between actual and forecast value. Smaller values of these evaluation indicators signify higher accuracy in forecasting.

$$NMSE = \frac{\sum_i (y_i - a_i)^2}{\sum_i (y_i - \bar{a}_i)^2} \quad (6)$$

$$MAE = \frac{1}{N}\sum_i |y_i - a_i| \quad (7)$$

The main objective of this study is to select the best structure, control parameters and weights variation interval of neural network by using a real genetic algorithm. The neural network parameters used are: number of hidden layers, Number of nodes for each layer l, activation functions for each layer, presence or absence of the bias, learning and momentum coefficients, input number, time that separate two successive inputs, which are critical parameters in the design of neural network. These parameters are selected genetically by minimizing the mean square error. Table 2 shows the genetic algorithm parameters used in the optimization process.

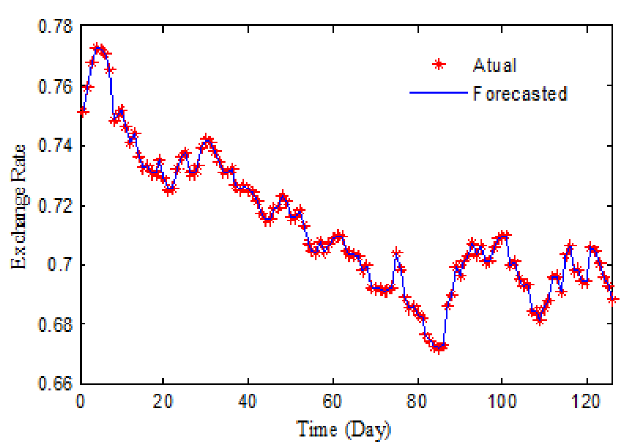
The neural model was trained with four inputs representing the four exchange rates, one or two hidden layers and an output unit to predict the exchange rate. The final set of weights to which a network settles down depends on the chosen initial weights. The training stage finished at max iteration number. The network that yielded the best results after the optimization process is presented in table 3. We measured the performance metrics on the test data to investigate how well the neural network forecasting model captured the underlying trend of the movement of each currency against USD.

TABLE II.

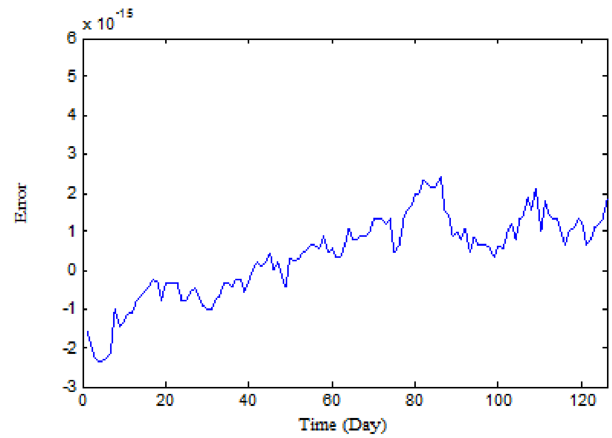
THE CONTROL PARAMETERS OF THE GA USED FOR OPTIMIZING THE NEURAL PREDICTOR.

No of generation	100
Size of population	50
Size of chromosome	17
Probability of crossover	1
Probability of Mutation	0.1

The comparative diagrams show the output forecasted by the neural network model and actual time series over 143 days for four currencies exchange rate are shown in figures (Fig. 3-a, Fig.4-a, Fig. 5-a and Fig.6-a). As depicted, a big similarity between actual and forecasted exchange rates indicating the effectiveness of the optimized model. Figures (Fig. 3-b, Fig.4-b, Fig. 5-b and Fig.6-b) show the forecasting error of EUR/USD, GBP/USD, CAD/USD and JPY/USD.

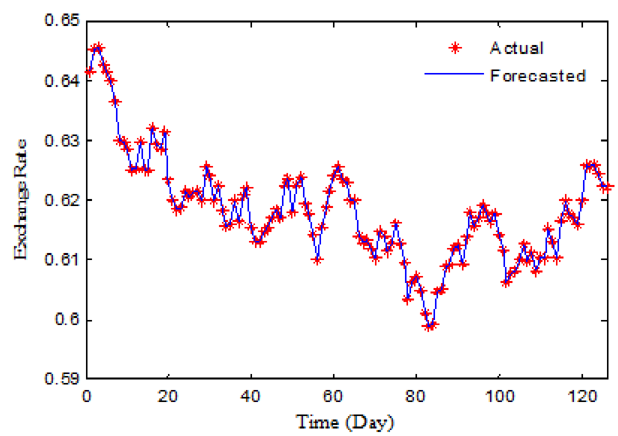


(a)

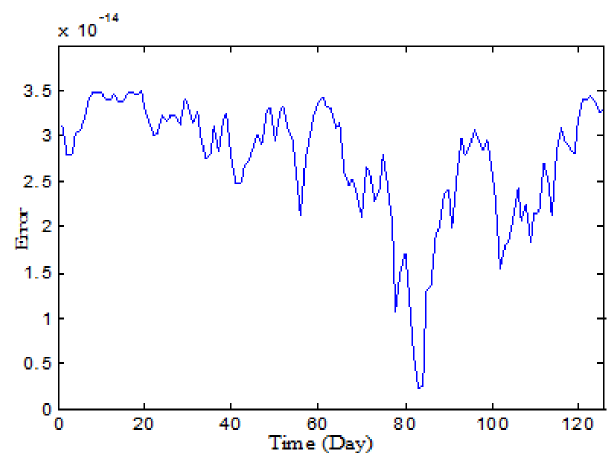


(b)

Fig. 3 (a) Actual and Forecasted exchange rate EUR/USD, (b) Forecasting Error EUR/USD



(a)

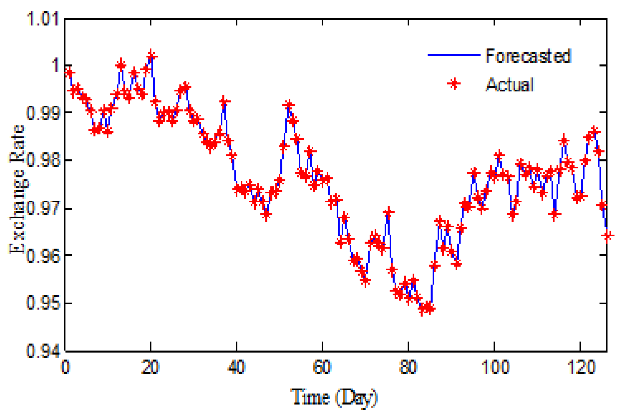


(b)

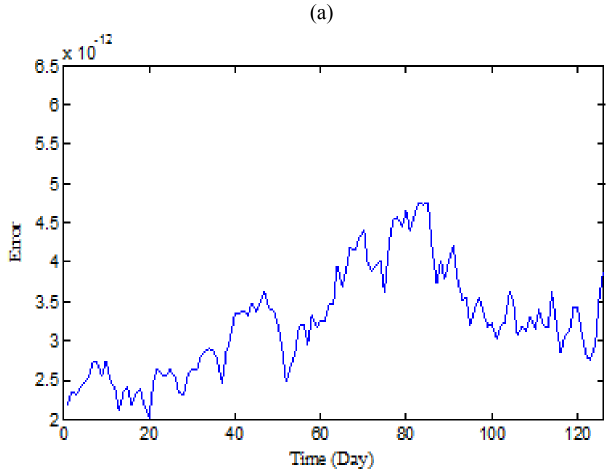
Fig. 4. (a) Actual and forecasted exchange rate GBP/USD, (b) Forecasting Error GBP/USD

TABLE III.
OPTIMUM STRUCTURE OF NEURAL PREDICTOR’S USING GAs

Neural network Parameters	EUR/USD	GBP/USD	CAD/USD	JPY/USD
N° of hidden layers	2	2	1	1
learning rate	0.7578	0.3524	0.9307	0.3475
Momentum coefficient	0.8421	0.8858	0.5457	0.4944
the presence/ absence of t he bias	Presence (1)	absence (0)	Presence (1)	Presence (1)
Number of the input	20	17	10	6
Lag time	1	1	4	3
Number of neurons in the hidden layer 1	16	29	12	14
Number of neurons in the hidden layer 2	16	8	7	9
the activation function in the hidden layer1	Pure linear 3	Pure linear 3	Pure linear 3	Pure linear 3
the activation function in the hidden layer2	Pure linear 3	sigmoid 1	sigmoid 1	Pure linear 3
the activation function in the output layer	hyperbolic 2	Pure linear 3	Pure linear 3	Pure linear 3
int_{out}	[0.0725, 0.2798]	[0.0901, 0.3240]	[0.1634, 0.4703]	[0.0240, 0.3264]
int1	[0.1035, 0.6864]	[0.0326, 0.5631]	[0.06360, 0.8137]	[0.0097, 0.3710]
int2	[0.1106, 0.3346]	[0.0773, 0.6963]	[0.2002,0.9350]	[0.04284, 04252]

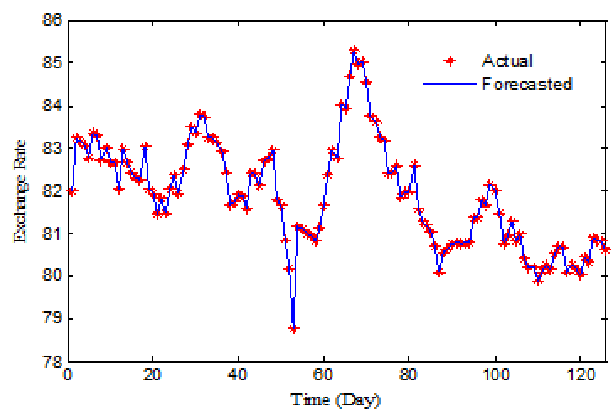


(a)

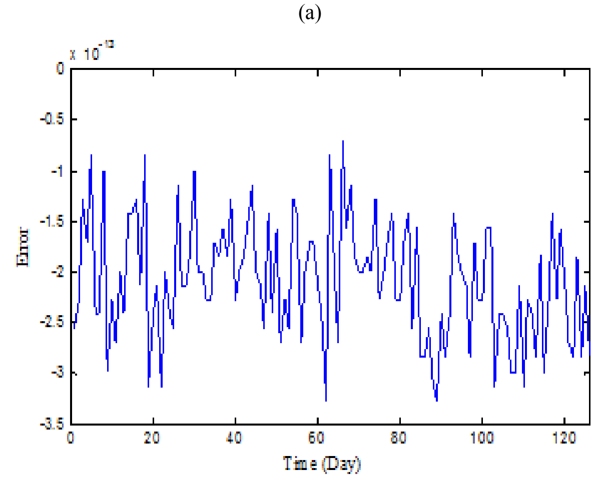


(b)

Fig. 5 (a) Actual and Forecasted exchange rate CAD/USD, (b) Forecasting Error CAD/USD



(a)



(b)

Fig. 6 (a) Actual and Forecasted exchange rate JPY/USD, (b) Forecasting Error JPY/USD

IV. COMPARISON WITH OTHER MODELS

Tables 4, 5, 6 and 7 show the Evaluation indicators achieved by each model over a forecasting period. The results show that the proposed Neuro-genetic model is consistently performs better than Network and Conditional heteroscedastic Models (GARCH, GARCH-M, EGARCH, TGARCH/GJR and IGARCH) in terms of the two Evaluation indicators in all the currency exchange rates. This means, the elaborated and optimized models are able of predicting exchange rates more closely than other models.

The main reason of better performances is due to the optimization phase; which allows them to search efficiently its structures, control parameters and weights variation interval space for the desired task. The results of the conditional heteroscedastic models and neural are taken from [15].

TABLE IV. COMPARATIVE ANALYSIS OF EUR/USD SERIES

Model	NMSE	MAE.10^2
GARCH(1,1)	0.99998	0.34930
GARCH(1,1)-M	0.93987	0.34918
EGARCH(1,1)	0.89833	0.30927
TGARCH/GJR(1,1)	0.89986	0.31414
IGARCH(1,1)	0.92064	0.34999
Neural network	0.88442	0.38644
Neuro-Genetic model	$7.1053\ 10^{-5}$	$5.5898.10^{-6}$

TABLE V. COMPARATIVE ANALYSIS OF GBP/USD SERIES

Model	NMSE	MAE.10^2
GARCH(1,1)	1.00011	0.27177
GARCH(1,1)-M	0.86999	0.26832
EGARCH(1,1)	0.89286	0.32988
TGARCH/GJR(1,1)	0.87995	0.26164
IGARCH(1,1)	0.80463	0.27175
Neural network	0.56853	0.2274
Neuro-Genetic model	$6.0868.10^{-4}$	$1.7504.10^{-5}$

TABLE VI. COMPARATIVE ANALYSIS OF USD/ JPY SERIES

Model	NMSE	MAE.10^2
GARCH(1,1)	0.89899	0.36216
GARCH(1,1)-M	0.84957	0.36010
EGARCH(1,1)	0.79654	0.32844
TGARCH/GJR(1,1)	0.80971	0.32207
IGARCH(1,1)	0.82374	0.36236
Neural network	0.74021	0.37706
Neuro-Genetic model	$9.8296.10^{-4}$	$1.3284.10^{-5}$

TABLE VII. COMPARATIVE ANALYSIS OF USD/ CAD SERIES

Model	NMSE	MAE.10^2
Neuro-Genetic model	$1.5483.10^{-3}$	$2.0568.10^{-4}$

V. CONCLUSION

In this paper, a genetic algorithm was used to optimize the parameters of a multilayer perceptron neural network. To evaluate the effectiveness of the proposed approach, we applied this optimized system for the prediction of exchange rates EUR/USD, GBP/USD, CAD/USD and JPY/USD. Simulation results show the effectiveness and ability of the intelligent hybrid method (evolutionary-connectionist hybrid model) to predict these exchange rates effectively and our method has better performances when it is compared with some other time series forecasting models. In future work we will improve the optimization method and apply it to other systems.

REFERENCES

[1] Baillie, R. and T. Bollerslev, "The message in daily exchange rates: A conditional variance tale". Journal of Business and Economic Statistics 7(3), 297-305, 1989.

[2] West, K. D. and D. Cho, "The predictive ability of several models of exchange rate volatility". Journal of Econometrics 69(2), 367-391, 1995.

[3] Jorion, P. "Predicting volatility in the foreign exchange market". Journal of Finance 50(2), 507-528, 1995.

[4] Black, F. and M. S. Scholes, "The pricing of options and corporate liabilities". Journal of Political Economy 81, 637-654, 1973.

[5] Andersen, T. and T. Bollerslev, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts". International Economic Review 39(4), 885-905, 1998.

[6] Chang, P. H. K. and C. L. Osler, "Methodical madness: Technical analysis and the irrationality of exchange-rate forecasts". Economic Journal 10, 636-661, 1999.

[7] Meese, R. A. and A. K. Rose, "Nonlinear, nonparametric, nonessential exchange rate estimation". American Economic Review Papers and Proceedings 80, 192-196, 1990.

[8] Gencay, R, "Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules". Journal of International Economics 47, 91-107, 1999.

[9] Kuan, C.-M. and T. Liu, "Forecasting exchange rates using feedforward and recurrent neural networks", Journal of Applied Econometrics 10(4), 347-364, 1995.

[10] Yao, J. T. and C. L. Tan, "A case study on using neural networks to perform technical forecasting of forex", Neurocomputing 34(1-4), 79-98, 2000.

[11] Yang, J. and N. Gradojevic, "Non-linear, non-parametric, non-fundamental exchange rate forecasting". Journal of Forecasting 25(4), 227-245, 2006.

[12] Kiani, K. M. and T. L. Kastens (2008). "Testing forecast accuracy of foreign exchange rates: Predictions from feed forward and various recurrent neural network architectures", Comput. Econ. 32(4), 383-06.

[13] P. Bonissone, T. Chen, K. Goebel, "Hybrid Soft Computing Systems: Industrial and Commercial Applications", IEEE Proceedings, Special Issue on Computational Intelligence, Vol. 87, No. 9, pp. 1641-1667, 1999.

[14] O. Castilo and P. Melin, "Hybrid Intelligent Systems for Time Series Prediction Using Neural Networks, Fuzzy Logic, and Fractal Theory", IEEE trans. Neural Networks. Vol .13, pp. 1395-1408, November 2002.

[15] AK. Dhamija and VK. Bhalla, "Financial Time Series Forecasting: Comparison of Neural Networks and ARCH Models", International Research Journal of Finance and Economics, Vol .49, pp.185-202, 2010.