Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 9 (20): 3641-3651, 2009
ISSN 1812-5654

Comparison of ANFIS, ANN, GARCH and ARIMA Techniques to Exchange Rate Forecasting

[1]S.M. Fahimifard, [1]M. Homayounifar, [1]M. Sabouhi and [2]A.R. Moghaddamnia
[1]Department of Agricultural Economics,
[2]Department of Natural Sciences, University of Zabol, Zabol, Iran

Abstract: The need of exchange rate forecasting in order to preventing its disruptive movements has engrossed many policy makers and economists for many years. The determinants of exchange rate have grown manifold making its behavior complex, nonlinear and volatile so that nonlinear models have better performance for its forecasting. Nonlinear models estimated by various methods can fit a data base much better than linear models. Beside they can learn from examples, are fault tolerant in the sense that they are able to handle noisy and incomplete data, are able to deal with non-linear problems and once trained can perform prediction and generalization at high speed. In this study, the accuracy of ANFIS and ANN as the nonlinear models and GARCH and ARIMA as the linear models for forecasting 2, 4 and 8 days ahead of daily Iran Rial/ and Rial/US$ was compared. Using three forecast evaluation criteria (R^2, MAD and RMSE) we found that nonlinear models outperform linear models, GARCH outperforms ARIMA model and ANFIS outperforms ANN model. And consequently the effective role of ANFIS model to improve the Iran's exchange rate forecasting accuracy can't be denied.

Key words: ANFIS, ANN, GARCH, ARIMA, exchange rate, forecasting

INTRODUCTION

Assessing future changes in exchange rates has been of long interest to economists as well as policy makers (Groen, 2005a). Exchange rate plays a principal conduit through which monetary policy affects real activity and inflation. In order to keep inflation stable at appropriate level and economic activity at higher level, the monetary authority must have confidence, which will come through the better understanding of the movements of exchange rate, in conducting the monetary policy (Pandaa and Narasimhanb, 2007). Also, in order to intervene efficiently in the foreign exchange market, the policy makers in the central bank must be very much aware of the movement of exchange rate and its consequences (Pandaa and Narasimhanb, 2007). Multinational corporations in order to gain a competitive advantage over their rivals are extending in the fast growing emerging markets. Although, these corporations have enjoyed many benefits from economic growth in these regions, business operations in the developing economics, the recent financial turmoil in the developing economics highlights the instability of these growing economies and stresses the firms' need to closer scrutinize the foreign exchange rates. This notion has been echoed by many industrial leaders to call for greater transparency of the foreign exchange markets and enhancing the predictability of the currency exchange movements (Chen and Leung, 2003). Perhaps, these are few reasons why monetary authority, policy makers and corporations might wish to forecast exchange rates (De Grauwe *et al*., 1993).

Recently, it is well documented that many economic time series observations are non-linear while, a linear correlation structure is assumed among the time series values therefore, the ARIMA (Auto-Regressive Integrated Mmoving Average) model can not capture nonlinear patterns and approximation of linear models to complex real-world problem is not always satisfactory. While nonparametric nonlinear models estimated by various methods such as Artificial Intelligence (AI), can fit a data base much better than linear models and it has been observed that linear models, often forecast poorly which limits their appeal in applied setting (Racine, 2001). Artificial Intelligence (AI) systems are widely accepted as a technology offering an alternative way to tackle complex and ill-defined problems (Kalogirou, 2003). They can learn from examples, are fault tolerant in the sense that they are able to handle noisy and incomplete data, are able to deal with non-linear problems and once trained can perform prediction and generalization at high speed

Corresponding Author: Seyed Mohammad Fahimifard, Department of Agricultural Economics, University of Zabol, Postal Code 532-98615, Zabol, Iran Fax: 0098-0542-2220633

(Kalogirou, 2003). They have been used in diverse applications in control, robotics, pattern recognition, forecasting, medicine, power systems, manufacturing, optimization, signal processing and social/psychological sciences. AI systems comprise areas like expert systems, ANNs, genetic algorithms, fuzzy logic and various hybrid systems, which combine two or more techniques (Kalogirou, 2003). Among the mentioned AI systems, according to Hykin (1994), a neural network is a massively parallel-distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. Also, the greatest advantage of a neural network is its ability to model complex nonlinear relationship without a priori assumptions of the nature of the relationship like a black box (Karayiannis and Venetsanopoulos, 1993).

Concerning the application of neural nets to economic time series forecasting, there have been mixed reviews. For instance, Mark and Sul (2001) and Groen (2005b) use panels of between 3 to 17 OECD countries to first test for cointegration between the exchange rate and monetary fundamentals and secondly use this cointegrating relationship to successfully predict exchanges rates at horizons of three to four years. Cifter (2008) proposed wavelet network to investigate the effects of international F/X markets on emerging markets currencies. They used EUR/USD parity as input indicator and three emerging markets currencies as output indicator. Using wavelet networks, they found that the effects of international F/X markets increase with higher timescale. They also find that the effects of EUR/USD parity on Turkish lira are higher on 9-16 days and 33-64 days scales. Fahimifard *et al*. (2009) compared the ANFIS and ARIMA models for forecasting 1, 2 and 4 week(s) ahead of Iran's poultry retail price. Their study stated that ANFIS outperforms ARIMA in all three horizons and the effective role of ANFIS model to improve the Iran's poultry retail price forecasting accuracy can't be denied.

In this study, we compare the accuracy of Adaptive Neuro-Fuzzy Interface System (ANFIS) and Artificial Neuro Network (ANN) as the nonlinear models and GRCH and ARIMA as the linear models to forecasting 2, 4 and 8 days ahead of daily Iran Rial/US$ and Rial/using data collected from the Central Bank of Iran (CBI) website and forecast evaluation criteria include; R^2, MAD and RMSE.

MATERIALS AND METHODS

Auto-Regressive Integrated Moving Average (ARIMA) Model: Introduced by Box and Jenkins (1978), in the last few decades the ARIMA model has been one of the most popular approaches of linear time series forecasting methods. An ARIMA process is a mathematical model used for forecasting. One of the attractive features of the Box-Jenkins approach to forecasting is that ARIMA processes are a very rich class of possible models and it is usually possible to find a process which provides an adequate description to the data. The original Box-Jenkins modeling procedure involved an iterative three-stage process of model selection, parameter estimation and model checking. Recent explanations of the process (Makridakis *et al*., 1998) often add a preliminary stage of data preparation and a final stage of model application (or forecasting).

The ARIMA (p, d, q) model is as follow:

$$y_t = f(t) + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \ldots + \theta e_{t-q} \quad (1)$$

where, y_t and e_t are the target value and random error at time t, respectively, $\phi_i(i = 1, 2,...,p)$ and $\theta_j(1, 2,...,q)$ are model parameters, p and q are integers and often referred to as orders of autoregressive and moving average polynomials.

Generalized Auto-Regressive Conditional Hetroskedasticity (GARCH) model: Heteroscedasticity is often associated with cross-sectional data, whereas time series are usually studied in the context of homoscedastic processes. In analysis of macroeconomic data, Engle (1982, 1983) and Cragg (1982) found evidence that for some kinds of data, the disturbance variances in time-series models were less stable than usually assumed. Engle's results suggested that in models of inflation, large and small forecast errors appeared to occur in clusters, suggesting a form of heteroscedasticity in which the variance of the forecast error depends on the size of the previous disturbance. He suggested the autoregressive, conditionally heteroscedastic, or ARCH, model as an alternative to the usual time-series process. More recent studies of financial markets suggest that the phenomenon is quite common. The ARCH model has proven to be useful in studying the volatility of inflation (Coulson and Robins, 1985), the term structure of interest rates (Engle *et al*., 1985, 1987), the volatility of stock market returns and the behavior of foreign exchange markets (Domowitz and Hakkio, 1985; Bollerslev and Ghysels, 1996), to name but a few.

The simplest form of this model is the ARCH (1) model,

$$y_t = \beta' x_t + \varepsilon_t, \qquad \varepsilon_t = u_t \sqrt{\alpha_0 +_{\alpha_1} \varepsilon_{t-1}^2} \quad (2)$$

where, u_t is distributed as standard normal. It follows that $E[\varepsilon_t|x_t, \varepsilon_t-1] = 0$, so that $E[\varepsilon_t|x_t]$ and $E[y_t|x_t] = \beta' x_t$. Therefore, this model is a classical regression model.

The model of Generalized Auto-Regressive Conditional Heteroscedasticity (GARCH) is defined as follows. The underlying regression is the usual one in (Eq. 2). Conditioned on an information set at time t, denoted Ψ_t, the distribution of the disturbance is assumed to be:

$$\varepsilon_t = |\Psi_t \approx N[0, \sigma_t^2] \quad (3)$$

Where, the conditional variance is:

$$\sigma_t^2 = \alpha_0 + \delta_1\sigma_{t-1}^2 + \delta_2\sigma_{t-2}^2 + ... + \delta_p\sigma_{t-p}^2 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + ... + \alpha_q\varepsilon_{t-q}^2 \quad (4)$$

Define

$$z_t = [1, \sigma_{t-1}^2, \sigma_{t-2}^2, ..., \sigma_{t-p}^2, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, ..., \varepsilon_{t-q}^2]' \quad (5)$$

and

$$\gamma = [\alpha_0, \delta_1, \delta_2, ..., \delta_p, \alpha_1, \alpha_2, ... + \alpha_q]' \quad (6)$$

Then

$$\sigma_t^2 = \gamma' z_t \quad (7)$$

The model in Eq. 4 is a GARCH (p,q) model, where p refers to the order of the autoregressive part and q refers to the order of the moving average part.

Artificial Neural Network (ANN) model: Many economic time series observations are non-linear while, a linear correlation structure is assumed among the time series values therefore, ARIMA and GARCH models can not capture nonlinear patterns and, approximation of linear models to complex real-world problem is not always satisfactory. Therefore, the ANN and ANFIS nonlinear models will be introduced follows.

The major advantage of neural networks is their flexible capability of nonlinear modeling. With ANN, there is no need to specify a particular model. Rather, the model is adaptively based on the features presented from the data (Haoffi *et al.*, 2007). This data-driven approach is suitable for many empirical researches where no theoretical guidance is available to suggest an appropriate data generating process. The most common types of ANN models have been shown in Fig. 1.

For the purposes of this study, the feed-forward backpropagation neural network (also known as a MLP (Multilayer Perceptron) network) is the neural network model most widely used in time series forecasting, because it is capable of resolving a wide variety of problems (Sarle, 2002). MLP network is made up of an input layer, an output layer and one or more hidden layers of neurons. As the Fig. 2 shows, each input is weighted with an appropriate w. The sum of the weighted inputs and the bias forms the input to the transfer function f.

Neurons can use any differentiable transfer function f to generate their output. In general, transfer function introduces a degree of nonlinearity that is valuable for most ANN applications and ideally, it should be continuous, differentiable and monotonic. Feed-forward networks often have hidden layer(s) of sigmoid neurons followed by an output layer of linear neurons. Two stages may be considered in the MLP network: the running stage, in which an input pattern is presented to the trained network and transmitted through successive layers of neurons until reaching an output and the training or learning stage in which the weights or parameters of the network are iteratively modified on the basis of a set of input–output patterns known as a training set, in order to minimize the deviance or error between the output obtained by the network and the user's desired output.

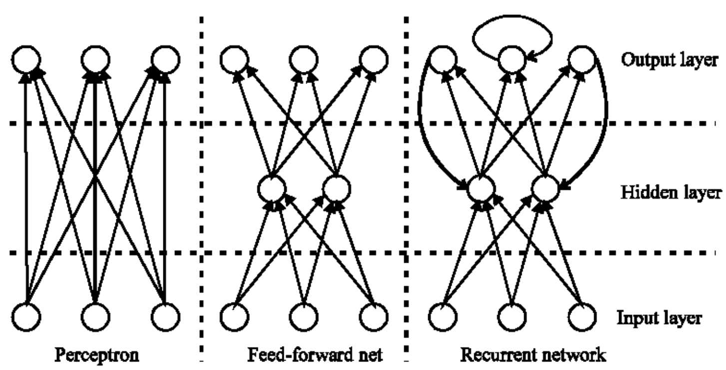


Fig. 1: Most common types of ANN models

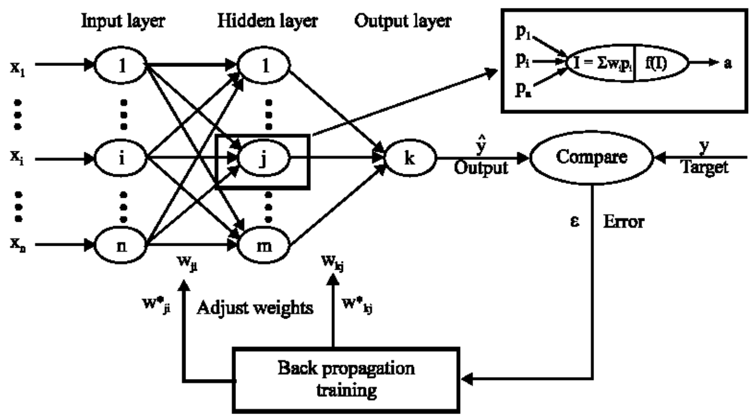


Fig. 2: A typical back-propagation neural network

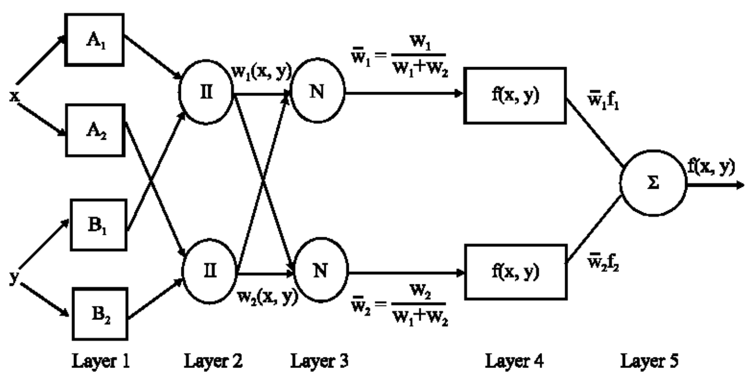


Fig. 3: The scheme of adaptive neural fuzzy inference system

This is why MLP network learning is said to be supervised. The learning rule commonly used in this type of network is the back propagation algorithm or gradient descent method, developed and disseminated by Rumelhart *et al.* (1986). In this study, we use the following three-layer feed-back networks:

$$F = F\left[\beta_0 + \sum_{j=1}^{J} \beta_j G\left[\sum_{k=1}^{K} \gamma_{kj} x_J\right]\right] \quad (8)$$

where, F is the output function of the output layer unit, β_0 is the bias unit (equal to 1), G is the output function of the hidden layer units j, γ_{kj} denotes the weight for the connection linking input k to the hidden unit j, β_j is the weight of outputs from the hidden layers in the output layer unit and x is the input vector.

Adaptive Neuro-Fuzzy Inference System (ANFIS) model: Fuzzy logic is a form of multi-valued logic derived from fuzzy set theory to deal with reasoning that is approximate rather than precise. In contrast with binary sets having binary logic, also known as crisp logic, the fuzzy logic variables may have a membership value of not only 0 or 1. Just as in fuzzy set theory with fuzzy logic the set membership values can range (inclusively) between 0 and 1, in fuzzy logic the degree of truth of a statement can range between 0 and 1 and is not constrained to the two truth values {true (1), false (0)} as in classic propositional logic (Von Altrock, 1995). Considering these advantages in contrast with pre-described models the ANFIS model has been carried out in this study.

The ANFIS is a neuro-fuzzy system developed by (Jang and Sun, 1995). It has a feed-forward neural network structure where each layer is a neuro-fuzzy system component (Fig. 3).

It simulates TSK (Takagi-Sugeno-Kang) fuzzy rule of type 3 where the consequent part of the rule is a linear combination of input variables and a constant. The final output of the system is the weighted average of each rule's output (Sugeno and Kang, 1988). The form of the type 3 rule simulated in the system is as follows:

IF x_1 is A_1 AND x_2 is A_2 AND... AND x_p is A_p
THEN $y=c_0+ c_1 x_1+ c_2 x_{2+...+} c_p x_p$

where, x_1 and x_2 are the input variables, A_1 and A_2 are the membership functions, y is the output variable and c_0, c_1 and c_2 are the consequent parameters. The neural network structure contains 6 layers.

- Layer 0 is the input layer. It has n nodes where n is the number of inputs to the system
- The fuzzy part of ANFIS is mathematically incorporated in the form of Membership Functions (Mfs)

A membership function can be any continuous and piecewise differentiable function that transforms the input value x into a membership degree, that is to say a value between 0 and 1. The most widely applied membership function is the generalized bell (gbell MF), which is described by the three parameters, a, b and c (Eq. 9). Therefore, layer 1 is the fuzzification layer in which each node represents a membership value to a linguistic term as a Gaussian function with the mean;

$$\mu_{A^i}(x) = \frac{1}{1+[(\frac{x-c_i}{a_i})^2]^{b_i}} \tag{9}$$

where a_i, b_i and c_i are parameters of the function. These are adaptive parameters. Their values are adapted by means of the back-propagation algorithm during the learning stage.

As the values of the parameters change, the membership function of the linguistic term, A_i changes. These parameters are called premise parameters. In that layer there exist n×p nodes where n is the number of input variables and p is the number of membership functions. For example, if size is an input variable and there exist two linguistic values for size which are SMALL and LARGE then two nodes are kept in the first layer and they denote the membership values of input variable size to the linguistic values SMALL and LARGE.

- Each node in layer 2 provides the strength of the rule by means of multiplication operator. It performs AND operation

$$w_i = \mu_{A^i}(x_0)\times\mu_{B^i}(x_1) \tag{10}$$

Every node in this layer computes the multiplication of the input values and gives the product as the output as in the above equation. The membership values represented by $\mu_{A^i}(x_0)$ and $\mu_{A^i}(x_0)$ are multiplied in order to find the firing strength of a rule where the variable x_0 has linguistic value A_i and x_1 has linguistic value B_i in the antecedent part of Rule l.

There are p^n nodes denoting the number of rules in layer 2. Each node represents the antecedent part of the rule. If there are two variables in the system namely x_1 and x_2 that can take two fuzzy linguistic values, SMALL and LARGE, there exist four rules in the system whose antecedent parts are as follows:

IF x_1 is SMALL AND x_2 is SMALL
IF x_1 is SMALL AND x_2 is LARGE
IF x_1 is LARGE AND x_2 is SMALL
IF x_1 is LARGE AND x_2 is LARGE

- Layer 3 is the normalization layer which normalizes the strength of all rules according to the equation:

$$\bar{w}_i = \frac{w_i}{\sum_{j=1}^{R} w_j} \tag{11}$$

where, w_i is the firing strength of the ith rule which is computed in layer 2. Node i computes the ratio of the *i*th rule's firing strength to the sum of all rules' firing strengths. There are p^n nodes in this layer.

- Layer 4 is a layer of adaptive nodes. Every node in this layer computes a linear function where the function coefficients are adapted by using the error function of the multi-layer feed-forward neural network.

$$\bar{w}_i f_i = \bar{w}_i(p_0x_0 + p_1x_1 + p_2) \tag{12}$$

where, p^i's are the parameters where i = n + 1 and n is the number of inputs to the system (i.e., number of nodes in layer 0). In this example, since there exist two variables (x_1 and x_2), there are three parameters p_0, p_1 and p_2) in layer 4 and $\bar{w}_i$ is the output of layer 3. The parameters are updated by a learning step. Kalman filtering based on least-squares approximation and back-propagation algorithm is used as the learning step.

- Layer 5 is the output layer whose function is the summation of the net outputs of the nodes in layer 4. The output is computed as:

$$\sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{13}$$

where, $\bar{w}_i f_i$ is the output of node i in layer 4. It denotes the consequent part of rule i. The overall output of the neuro-fuzzy system is the summation of the rule consequences.

The ANFIS uses a hybrid learning algorithm in order to train the network. For the parameters in the layer 1, back-propagation algorithm is used. For training the parameters in the layer 4, a variation of least-squares approximation or back-propagation algorithm is used.

DATA DESCRIPTION AND FORECAST EVALUATION CRITERIA

The exchange rate data used in this study are daily Iran Rial/U$ and Rial/€, covering the period from 20 Mar. 2002 to 21 Nov. 2008 with a total of 2436 observations, as shown in Fig. 4. Although, there is no consensus on how to split the data for neural network applications, the general practice is to allocate more data for model building and selection. Most studies in research use convenient ratio of splitting for in-and out-samples such as 70:30, 80:20, or 90:10%. This investigation selects the 70:30% one. We take the daily data from 20 Mar. 2002 to 20 Oct. 2006 as in-sample data set with 1706 observations for training and the remainder as out-sample data set with 730 observations for testing purposes. The study was carried out in Iran through the 2009:1-2009:4. We obtained the daily Iran's exchange rate time series from the website of Central Bank of Iran (www.CBI.ir). For space reasons, the original data are not listed hear and detailed data can be obtained from the website www.CBI.ir.

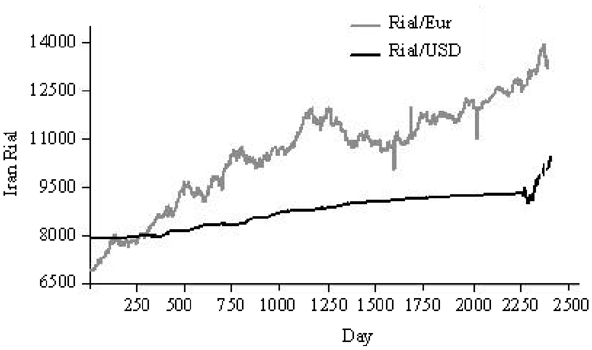


Fig. 4: Daily Iran Rial/US$ and Rial/Changes from 20 Mar. 2002 to 21 Nov. 2008

Table 1: Forecasting evaluation criteria

Criteria	Formulation
R-squared (R^2)	$R^2 = 1 - \frac{\sum(\hat{y}_t - y_t)^2}{\sum \hat{y}_t^2}$
Mean Absolute Deviation (MAD)	$MAD = \frac{\sum \lvert \hat{y}_t - y_t \rvert}{n}$
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{\sum(\hat{y}_t - y_t)^2}{n}}$

Where y_t, $\hat{y}_t$ and n are the actual value, output value and the number of observations, respectively

In order to evaluate and compare the forecasting performance, it is necessary to introduce forecasting evaluation criteria. In this research, three criteria include; R-squared, Mean Absolute Deviations (MAD) and Root Mean Square Error (RMSE) are used. Table 1 shows the R^2, MAD and RMSE formulation:

RESULTS AND DISCUSSION

Linear models performance to exchange rate forecasting: In ARIMA model we identified the degree of integration (d) by augmented Dickey-Fuller and Schwarz criteria and the degree of autoregessive (p) and moving average (q) by Log-likelihood function and Akaike Information Criterion. In GARCH model, the Lagrange Multiplier (LM) test was used to identifying the ARCH effects. In order to model specification the degrees of autoregressive (P) and moving average (q) were identified using normal distribution for conditional error terms.

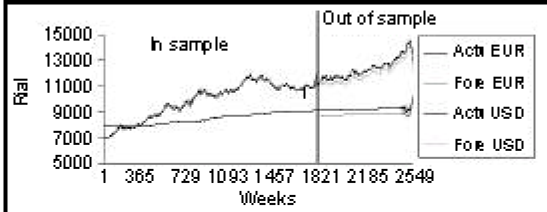
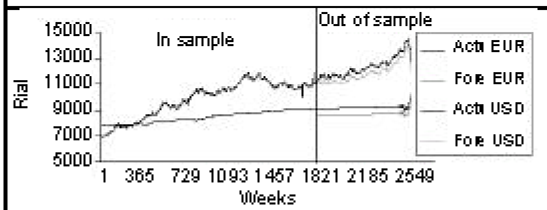
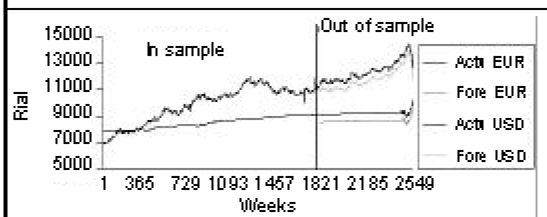
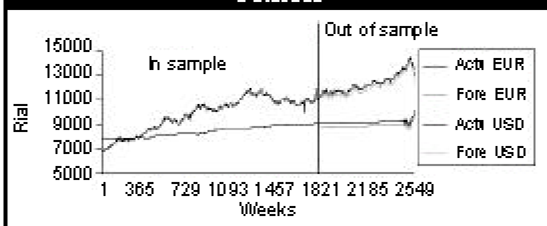
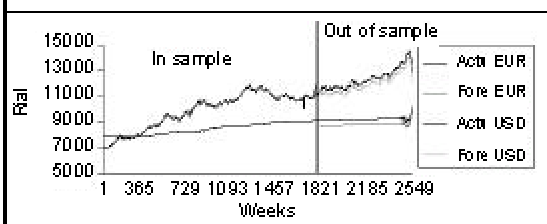
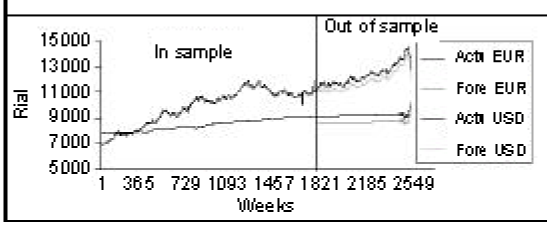
The forecasting performance of Rial/USD and Rial/EUR exchange rates obtained by the ARIMA and GARCH models is shown in Table 2.

The left side of Table 2 shows the out-sample fitness of ARIMA and GARCH models for forecasting 2, 4 and 8 days ahead of Rial/USD and Rial/EUR exchange rates in comparison with the actual observations. And its right side presents the quantity of evaluation criterions to forecast the considered horizons of Rial/USD and Rial/EUR exchange rates.

The results of Table 2 shows that ARIMA and GARCH models provide the better forecasting results for Rial/USD in contrast with Rial/EUR by all three performance measures. Also, Table 2 shows that the forecasting accuracy of ARIMA and GARCH models will reduced through the horizon increscent.

Nonlinear models performance to exchange rate forecasting: In ANN a single-hidden-layer feedforward network is used for the training in which sigmoid transfer function is used in the hidden layer and linear transfer function is used in the output layer. The weights are initialized to small values based on the technique of Nguyen and Widrow (1990) and mean square error is the taken as the cost function in our study. We train the network by using Levenberg-Marquardt backpropagation. The number of input nodes, in this work, corresponds to the number of lagged past observations. However, we only experiment with four levels of hidden nodes 2, 4, 6 and 8 across each level of input node by following previous findings (Hu *et al*., 1999) that the forecasting performance of neural networks is not as sensitive to the number of hidden nodes as to the number of input nodes.

Table 2: ARIMA and GARCH performance to exchange rate forecasting

Fitness	Model	Horizon	Structure	RMSE $	RMSE €	MAD $	MAD €	R^2 $	R^2 €
	ARIMA	2 days ahead	Structure (2,1,2)	0.0101	0.0105	0.0117	0.0128	0.9228	0.9135
	ARIMA	4 days ahead	Structure (4,1,1)	0.0101	0.0112	0.0117	0.0130	0.9219	0.9126
	ARIMA	8 days ahead	Structure (8,1,1)	0.0101	0.0114	0.0118	0.0131	0.9210	0.9120
	GARCH	2 days ahead	Structure (1,1)	0.0097	0.0098	0.0115	0.0123	0.9413	0.9319
	GARCH	4 days ahead	Structure (4,1)	0.0098	0.0106	0.0116	0.0125	0.9402	0.9310
	GARCH	8 days ahead	Structure (8,1)	0.0099	0.0109	0.0117	0.0125	0.9395	0.9304

In ANFIS the hybrid learning algorithm is used to identify the membership function parameters of single-output, Sugeno type Fuzzy Inference Systems (FIS). A combination of least-squares and backpropagation gradient descent methods are used for training FIS membership function parameters to model a given set of input/output data. In Genfis1 which, Generates an initial Sugeno-type FIS for ANFIS training using a grid partition the gauss and gauss 2 types of membership function are used for each input and linear membership function is used for output. Also, 3 and 4 numbers of membership functions are used for each input.

Table 3: ANN performance to exchange rate forecasting

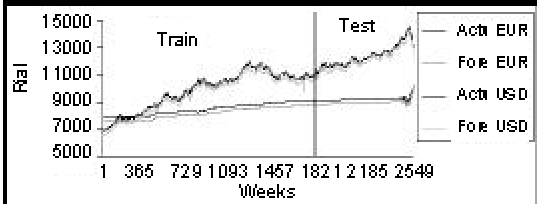
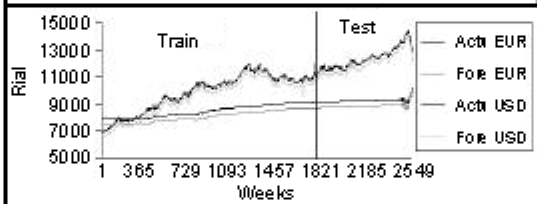
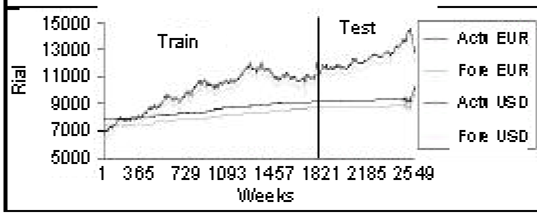
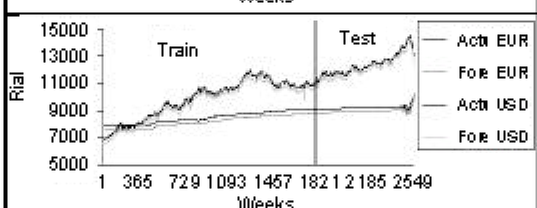
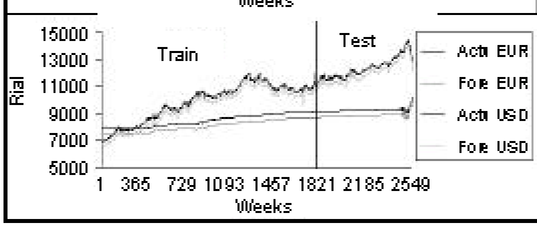
Fitness	ANN							
	Horizon	Structure	RMSE $	RMSE €	MAD $	MAD €	R^2 $	R^2 €
	2 days ahead	Structure (5-6-5-4-3-2-1-1)	0.0094	0.0095	0.0109	0.0115	0.9607	0.9506
	4 days ahead	Structure (5-2-1-1)	0.0094	0.0102	0.0111	0.0120	0.9589	0.9499
	8 days ahead	Structure (5-8-7-6-5-4-3-2-1-1)	0.0096	0.0105	0.0112	0.0121	0.9583	0.9490

Table 4: ANFIS performance to exchange rate forecasting

Fitness	ANFIS							
	Horizon	Structure	RMSE $	RMSE €	MAD $	MAD €	R^2 $	R^2 €
	2 days ahead	Structure (gauss-4-100)	0.0090	0.0090	0.0082	0.0110	0.9840	0.9698
	4 days ahead	Structure (gauss-3-100)	0.0092	0.0097	0.0084	0.0115	0.9791	0.9690
	8 days ahead	Structure (gauss2-4-100)	0.0093	0.0100	0.0085	0.0116	0.9662	0.9658

The forecasting performances of Rial/USD and Rial/EUR exchange rates obtained by the ANN and ANFIS models are shown in Table 3 and 4.

Similarly, the left side of Table 3 and 4 demonstrates the fitness of the best structures of ANN and ANFIS models for forecasting 2, 4 and 8 days ahead of Rial/USD and Rial/EUR exchange rates in comparison with the actual observations. And its right side represents the quantity of evaluation criterions to forecast the considered horizons of Rial/USD and Rial/EUR exchange rates.

In ANN structure e.g. structure (5-6-5-4-3-2-1-1) for forecasting 2 days ahead of Rial/USD or Rial/EUR the 101 terms 5, 6-5-4-3-2-1 and 1 represent the number of

Table 5: Comparison of ARIMA, GARCH, ANN and ANFIS Performance

R² €	R² $	MAD €	MAD $	MSE €	MSE $	RMSE €	RMSE $	Structure		Week(s) ahead
GARCH criterions to ARIMA criterions								ARIMA	GARCH	
1.020	1.020	0.961	0.980	0.968	0.980	0.937	0.960	(2,1,2)	(1,1)	2
1.020	1.020	0.962	0.991	0.975	0.985	0.950	0.970	(4,1,1)	(4,1)	4
1.020	1.020	0.954	0.992	0.978	0.990	0.956	0.980	(8,1,1)	(8,1)	8
ANN criterions to GARCH criterions								GARCH	ANN	
1.019	1.020	0.943	0.959	0.960	0.959	0.980	0.979	(1,1)	5-2-1-1	2
1.020	1.020	0.951	0.968	0.980	0.979	0.990	0.990	(1,1)	5-4-3-2-1-1	
1.020	1.021	0.935	0.951	0.940	0.939	0.969	0.969	(1,1)	5-6-5-4-3-2-1-1	
1.020	1.020	0.943	0.959	0.980	0.979	0.990	0.990	(1,1)	5-8-7-6-5-4-3-2-1-1	
1.020	1.020	0.960	0.960	0.919	0.922	0.959	0.960	(4,1)	5-2-1-1	4
1.018	1.017	0.976	0.974	0.955	0.960	0.977	0.980	(4,1)	5-4-3-2-1-1	
1.016	1.016	0.968	0.961	0.974	0.980	0.987	0.990	(4,1)	5-6-5-4-3-2-1-1	
1.018	1.018	0.968	0.966	0.937	0.940	0.968	0.969	(4,1)	5-8-7-6-5-4-3-2-1-1	
1.012	1.012	0.992	0.991	0.964	0.980	0.982	0.990	(8,1)	5-2-1-1	8
1.018	1.018	0.976	0.966	0.946	0.960	0.972	0.980	(8,1)	5-4-3-2-1-1	
1.016	1.016	0.992	0.983	0.964	0.980	0.982	0.990	(8,1)	5-6-5-4-3-2-1-1	
1.020	1.020	0.968	0.957	0.928	0.941	0.963	0.970	(8,1)	5-8-7-6-5-4-3-2-1-1	
ANFIS criterions to ANN criterions								ANN	ANFIS	
1.014	1.022	0.965	0.826	0.938	0.937	0.968	0.968	5-6-5-4-3-2-1-1	Gauss-3-100	2
1.020	1.024	0.957	0.752	0.903	0.917	0.950	0.957	5-6-5-4-3-2-1-1	Gauss-4-100	
1.013	1.021	0.974	0.844	0.979	0.979	0.989	0.989	5-6-5-4-3-2-1-1	Gauss2-3-100	
1.012	1.021	0.991	0.853	0.979	1.000	0.989	1.000	5-6-5-4-3-2-1-1	Gauss2-4-100	
1.020	1.021	0.958	0.754	0.903	0.956	0.950	0.978	5-2-1-1	Gauss-3-100	4
1.018	1.018	0.975	0.790	0.923	0.977	0.961	0.989	5-2-1-1	Gauss-4-100	
1.016	1.017	0.983	0.808	0.961	0.977	0.980	0.989	5-2-1-1	Gauss2-3-100	
1.014	1.016	0.992	0.817	0.980	0.998	0.990	0.999	5-2-1-1	Gauss2-4-100	
1.013	1.007	0.992	0.777	0.981	1.000	0.990	1.000	5-8-7-6-5-4-3-2-1-1	Gauss-3-100	8
1.015	1.006	0.967	0.804	0.944	1.000	0.971	1.000	5-8-7-6-5-4-3-2-1-1	Gauss-4-100	
1.014	1.008	0.975	0.786	0.962	0.979	0.981	0.990	5-8-7-6-5-4-3-2-1-1	Gauss2-3-100	
1.018	1.008	0.959	0.759	0.903	0.938	0.950	0.969	5-8-7-6-5-4-3-2-1-1	Gauss2-4-100	

input nodes, the number of neuron(s) in each hidden node and the number of output node, respectively. In ANFIS structure e.g., structure (gauss-4-100) for forecasting 2 days ahead of Rial/USD or Rial/EUR the terms gauss, 4 and 100 represent the type of membership function, the number of membership function and the number of training epochs, respectively.

As can bee seen in Table 3 and 4, the ANN and ANFIS nonlinear models perform very well in Rial/USD and Rial/EUR forecasting. Also, similarly the results of Table 3 and 4 state that ANN and ANFIS models provide the better forecasting results for Rial/USD in contrast with Rial/EUR by all three performance measures. Also, Table 3 and 4 shows that the forecasting accuracy of ANN and ANFIS models will reduced through the horizon increscent.

Comparison of linear and nonlinear models performance to exchange rate forecasting: In order to comparison the performance of considered linear and nonlinear models to Rial/USD and Rial/EUR exchange rates forecasting, we divided the quantity of forecast evaluation criterions of GARCH to ARIMA model, various structures of ANN to GARCH model and various structure of ANFIS to the best structure of ANN model per each horizon. Table 5 shows the results of these comparisons.

As can bee shown in Table 5, the ANN and ANFIS nonlinear models forecasting performance is better in contrast with the ARIMA and GARCH linear models because (1) the RMSE, MSE and MAD divided are less than 1 and (2) the R^2 divided is more than 1.

Also, as to quantity of divided forecast evaluation criterions, Table 5 indicates that ANFIS model provides the best forecasting results for forecasting judging by all four performance measures.

CONCLUSION

The researches previously done in the same field are usually based on traditional econometrics or static neural network models which their outputs are negligent in contrast with the assimilated ANFIS model. In this study, the accuracy of ANFIS and ANN as the nonlinear models and GARCH and ARIMA as the linear models has been compare for forecasting 2, 4 and 8 days ahead of daily Iran Rial/_ and Rial/US$ exchange rates. Results indicated that nonlinear models especially ANFIS model forecasts are considerably more accurate than either the linear traditional GARCH or specially ARIMA models which used as benchmarks in terms of error measures, such as RMSE, MSE and MAD. On the other hand, as the R^2 criterion is concerned, nonlinear models especially ANFIS model are absolutely better than linear models especially ARIMA model.

Briefly using forecast evaluation criteria we found that nonlinear models outperform linear models, GARCH outperforms ARIMA model and ANFIS outperforms ANN model. And we cannot deny that the ANFIS model is an effective way to improve the Iran Rial/ and Rial/US$ exchange rates forecasting accuracy.

REFERENCES

Bollerslev, T. and E. Ghysels, 1996. Periodic autoregressive conditional heteroscedasticity. J. Bus. Eco. Stat., 14: 139-151.

Box, G.P.F. and G.M. Jenkins, 1978. Time Series Analysis: Forecasting and Control. 3rd Edn., Holden Day, San Francisco, ISBN-10: 0130607746.

Chen, A.S. and M.T. Leung, 2003. A bayesian vector error correction model for forecasting exchange rates. Comput. Operat. Res., 30: 887-900.

Cifter, A., 2008. A signal processing model for time series analysis: The effects of International F/X markets on domestic currencies using wavelet networks. Int. Rev. Elect. Eng., 3: 580-591.

Coulson, N. and R. Robins, 1985. Aggregate economic activity and the variance of inflation: Another look. Econ. Lett., 17: 71-75.

Cragg, J., 1982. Estimation and testing in testing in time series regression models with heteroscedastic disturbances. J. Econ., 20: 135-157.

De Grauwe, P., H. Dewachter and M. Embrechts, 1993. Exchange Rate Theory: Chaotic Models of Foreign Exchange Markets. Blackwell, Oxford, ISBN: 0-631-18016-8.

Domowitz, I. and C. Hakkio, 1985. Conditional variance and the risk premium in the foreign exchange market. J. Int. Econ., 19: 47-66.

Engle, R., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. Econometrica, 50: 987-1007.

Engle, R., 1983. Estimates of the variance of U.S. Inflation based on the ARCH model. J. Money, Credit, Bank., 15: 286-301.

Engle, R., D. Hendry and D. Trumble, 1985. Small sample properties of ARCH estimators and tests. Can. J. Econ., 18: 66-93.

Engle, R., D. Lilen and R. Robins, 1987. Estimating time varying risk premia in the term structure: The ARCH-M model. Econometrica, 55: 391-407.

Fahimifard, S.M. M. Salarpour, M. Sabouhi and S. Shirzady, 2009. Application of ANFIS to agricultural economic variables forecasting (Case study: Poultry retail price). J. Arti. Intel., 2: 65-72.

Groen, J.J., 2005a. Fundamentals based exchange rate prediction revisited. Working Paper of Bank of England, London. http://www.eea-esem.com/files/papers/EEA-ESEM/2006/129/erpwp.pdf.

Groen, J., 2005b. Exchange rate predictability and monetary fundamentals in a small multi-country panel. J. Money Credit Bank., 37: 495-516.

Haoffi, Z., X. Guoping, Y. Fagting and Y. Han, 2007. A neural network model based on the multi-stage optimization approach for short-term food price forecasting in China. Expert Syst. Appl., 33: 347-356.

Hu, M.Y., G. Zhang, C.Y. Jiang and B.E. Patuwo, 1999. A cross validation analysis of neural networks out-of-sample performance in exchange rate forecasting. Decision Sci., 30: 197-216.

Hykin, S., 1994. Neural Networks. Macmillan, UK., ISBN: 0-02-352761-7.

Jang, J.S.R. and C.T. Sun, 1995. Neuron-fuzzy modeling and control. Proc IEEE., 83: 378-406.

Kalogirou, S.A., 2003. Artificial intelligence for the modeling and control of combustion processes: A review. Prog. Energy Combustion Sci., 29: 515-566.

Karayiannis, N.B. and A.N. Venetsanopoulos, 1993. Artificial Neural Networks: Learning Algorithms. Performance Evaluation and Applications. Kluwer Academic Publishers, Boston, USA., ISBN-13: 978-0792392972.

Makridakis, S., S.C. Wheelwright and R.J. Hyndman, 1998. Forecasting: Methods and Applications. 3rd Edn., John Wiley and Sons, Inc., New York, ISBN: 0471532339.

Mark, N.C. and D. Sul, 2001. Nominal exchange rates and monetary fundamentals; Evidence from a small post-bretton woods panel. J. Int. Econ., 53: 29-52.

Nguyen, D. and B. Widrow, 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. Proc.Int. Joint Con. Neural Networks, 3: 21-26.

Pandaa, C. and V. Narasimhanb, 2007. Forecasting exchange rate better with artificial neural network. J. Policy Modeling, 29: 227-236.

Racine, J.S., 2001. On the nonlinear predictability of stock returns using financial and economic variables, forthcoming. J. Business Econ. Stat., 19: 380-382.

Rumelhart, D.E., G.E. Hinton and R.J. Williams, 1986. Learning Internal Representations by Error Propagation. In: Parallel Distributed Processing, Rumelhart, D.E. and J.L. McClelland (Eds.). MIT Press, Cambridge, ISBN: 0-262-68053-X, pp: 318-362.

Sarle, W.S., 2002. Neural network FAQ. ftp://ftp.sas.com/pub/neural/FAQ.html.

Sugeno, M. and G.T. Kang, 1988. Structure identification of fuzzy model. Fuzzy Sets Syst., 28: 15-33.

Von Altrock, C., 1995. Fuzzy Logic and NeuroFuzzy Applications Explained. Prentice Hall, Upper Saddle River, New Jersey, ISBN: 0-13-368465-2.