Assessment and Comparison of linear and non-linear Methods for Forecasting Returns on Stock Market Index
Ali Raoofi1, Mansour ZarraNezhad2, Ozra Bayani3
1Phd. Candidate, faculty of Economic, Allameh Tabataba'i University, Tehran, Iran.
2 Prof., Faculty of Economic and Social Science, Shahid Chamran University, Ahvaz, Iran
3 Phd. Candidate, faculty of Economic, Allameh Tabataba'i University, Tehran, Iran

Available online at: www.IJSMD.Com
Received 16th July 2014, Revised 12th August 2014, Accepted 29th August 2014

Abstract
Nowadays, Investing in the stock exchange market is an important component of the countries economy. Being able to forecast the stock price variations is of great significance for the stockholders as it will enable them to obtain the highest return on their investment. The stock price index illustrates the general state of the stock market and could help stockholders make forecasts for their investments. This article assesses and forecasts returns based on Tehran Stock Exchange Index (TEPIX) using the daily data between 22/11/2010 and 18/03/2014 and different forecasting models such as ARIMA, GARCH, and Artificial Neural Networks (ANN) and Adaptive Network-based Fuzzy Inference System (ANFIS). Evaluating the accuracy of above mentioned models through their forecasting measures (e.g. RMSE, MAE, MAPE, TIC and CSP) reveals that ANFIS has a better performance of all models. The statistical comparison of the results through Diebold-Mariano Test also demonstrates a significant difference in models accuracy.

Keywords: ANFIS, Stock Price Return, Artificial Neural Network, GARCH, ARIMA
JEL: G10, C45, C22.

1. Introduction
Forecasting macroeconomic indices is of great and particular significance in all scientific discussions around economy and thus different models have been created to forecast the future value of these indices in order to help economic policy-makers to make proper monetary and financial policies. Therefore, the majority of governments and central banks today not only take into account the current conditions, but also the short-term and long-terms forecasts pertaining to the main economic indices when adopting and administrating policies. On the other hand, nowadays with the ever-increasing surge of the financial markets, the changes in these markets could have many implications for the global economy. Such events like September 11 and the financial scandals reported by the major American companies have been resulted in huge confusions for the investors and have caused uncertainties with regard to the performance of financial markets in different parts of the world. In addition, this has decreased investors’ confidence in these markets and has created a large number of negative implications for the global economy. This is evidence for the strong correlation between uncertainty of financial markets and public confidence of investors. As a result, financial policy-makers in different countries are usually in need of accurate forecasts and assessments of the changes in the prices of financial markets as a measure for adopting appropriate policies so to decrease the vulnerability of national and global economies. Hence, forecasting the changes in the price of financial assets is one of the most important responsibilities in the financial markets and has attracted a lot of attention by researchers and policy-makers in the past two decades so that they could apply such forecasts in assessment and pricing of assets, efficient allocation of financial resources, and assessment of the performance of risk management (Granger and Newbold, 1986).

Due to the significance of forecasting economic indices and particularly financial indices, this study forecasts the TEPIX using four different linear and non-linear time series models. These models include Autoregressive Integrated Moving Average (ARIMA), GARCH, and Artificial Neural Network (ANN) and Adaptive Network-based Fuzzy Inference System (ANFIS). Since stock prices and stock returns are highly complex, it seems that flexible non-linear models like ANN and ANFIS will have a better performance in terms of forecasts in comparison to standard linear models. This study aims to assess this hypothesis and answer if the soft computing model's has a better forecasting ability in comparison to ARIMA and GARCH with regard to making forecasts in the TEPIX.

This study first provides a review of the literature and then describes forecasting models and accuracy assessment methods of forecasting models in Part Three and Part Four respectively. Part Five introduces the data used in the study. Part Six provides the results of forecasts and compares the accuracy of models. The final part is the conclusion.

2. Review of Literature
Dunis and Jalilov (2001) used the ANN model to assess and forecast four main indices of the stock market, i.e. S&P 500, FTSE 100, EUROSTOXX 50, NIKKEI 225. In order to assess the model, they used the daily data for the period between 31rd of January 1994 and 4th of May 1994 and used the data between 5th of May 1999 and 6th of June 2000 to make forecasts outside the sample. They then compared the results obtained in the model with the forecasts of commercial models like purchase and maintenance, simple comparative expectations, moving average basket, and moving average basket using different measures like sharp ratio, t-statistic pertaining to stock revenue and demonstrated the superiority of the ANN model.

Yim (2002) used ANN model to forecast the return on the daily stock index of Brazil and then compared the results with the forecast results obtained from ARMA-GARCH and structural models using MAE and RMSE indices and Chong and Hendry Test, concluding that ANN was a stronger model. In order to assess the model, he used the daily data pertaining to
this index in the period between July 30, 1994 and June 30, 1998.

Grosan, Abraham, Ramos, and Han (2005) apply a genetic programming technique (called Multi-Expression programming-
MEP) for the prediction of Nasdaq-100 index of Nasdaq Stock MarketSM and the S&P CNX NIFTY stock index. The performance is compared with an artificial neural network trained using Levenberg-Marquardt algorithm, support vector machines, Takagi–Sugeno neuro–fuzzy inference system and a difference boosting neural network. The empirical results indicate that MEP is a novel promising technique for function approximation problems. MEP technique gives the lowest MAP values for both stock indices.

Abbasi and Abouec (2008) investigate the current trend of stock price of the Iran Khodro Corporation at Tehran Stock Exchange by utilizing an Adaptive Neuro–Fuzzy Inference System. The findings of the research demonstrate that the trend of stock price can be forecast with a low level of error.

Atsalakis and Valavanis (2009) develop a Neuro-fuzzy adaptive control system to forecast the next day’s stock price trends of the ASE and the NYSE index. The experimental results reveal that the proposed system performs very well in trading simulations, returning results superior to the buy and hold strategy. It also demonstrates solid and superior performance in terms of percentage of prediction accuracy of stock market trend.

The results of the study conducted by Wang (2009) with regard to the stock price index in Taiwan Stock Exchange indicate that the fluctuations of the data includes a trend correlating with ARCH family patterns and thus its combination with ANN could lead to an improvement in the forecasts.

Zarranzehad, Raoodf and kiyani (2012) conducted a study to compare the performance of ARIMA and ANFIS models for forecasting the price of gold. Their results showed that the ANFIS model was superior to the ARIMA.

In addition, in the review article by Atsalakis and Valavanis (2009), titled “Surveying Stock Market Forecasting Techniques – Part II: Soft Computing Methods”, a wide literature has been collected that could be beneficial for those interested. A number of studies in various financial markets conducted by the soft computing are seen in the following table:

**Table 1. Financial time series researches (Soft Computing Method)**

<table>
<thead>
<tr>
<th>Date</th>
<th>Researchers</th>
<th>Used method</th>
<th>Data years</th>
<th>Data type</th>
<th>Goal</th>
<th>Predicted period</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>Hazacebi &amp; Bayramoglu</td>
<td>ARIMA, ANN</td>
<td>2002-2006</td>
<td>ISE-XU100</td>
<td>To compare ARIMA &amp; ANN</td>
<td>Daily</td>
<td>ANN has better results</td>
</tr>
<tr>
<td>2007</td>
<td>Hassan, Nath &amp;</td>
<td>Hidden Markov Model(HMM), ANN, Genetic Algorithm(GA)</td>
<td>2003-2004</td>
<td>Stock; Apple, IBM, Dell</td>
<td>Exchange Prediction</td>
<td>5Weeks</td>
<td>Hybrid Model is better HMM &amp; ARIMA</td>
</tr>
<tr>
<td>2009</td>
<td>Lue, Lee &amp; Jou</td>
<td>Radial basis-function neural network(RBF NN)</td>
<td>2006-2007</td>
<td>TAIEX,NTD/USD,KRW/USD,CNY/USD,DJPY/USD</td>
<td>Index prediction &amp; To compare the methods</td>
<td>1.3,6&amp;1 2months</td>
<td>PNN, rough sets &amp; C4.5 classifiers to generate trading rule sets, which is helpful to construct a better predictive power trading system for stock market timing analysis</td>
</tr>
<tr>
<td>2009</td>
<td>Atsalakis &amp; Valavenis</td>
<td>ANFIS</td>
<td>1986-2005</td>
<td>Ten stock from Athens &amp; NYSE</td>
<td>TO determine the best stock trend prediction model</td>
<td>Daily</td>
<td>Proposed system clearly demonstrates the potential of neuro-fuzzy based modeling for financial market prediction</td>
</tr>
<tr>
<td>2010</td>
<td>Khashei &amp; Bijari</td>
<td>ARIMA, ANNs, Zhang's hybrid</td>
<td></td>
<td>Wolf's sunspot, Canadian lynx, GBP/USD</td>
<td>To Demonstrate the appropriateness &amp; effectiveness of the proposed models.</td>
<td>Daily</td>
<td>Zhang's hybrid model outperforms ARIMA &amp; ANNs</td>
</tr>
<tr>
<td>2010</td>
<td>Boyacioglu &amp; Avci</td>
<td>ANFIS,</td>
<td>1990-2008</td>
<td>the Istanbul Stock Exchange</td>
<td>Prediction of stock market return</td>
<td>Monthly</td>
<td>ANFIS can be a use-full tool for stock price prediction in emerging markets, like Turkey</td>
</tr>
<tr>
<td>2011</td>
<td>Guresen, Kayakutlu &amp; Daim</td>
<td>MLP, dynamic ANN(DNA2, GARCH-MLP)</td>
<td>2008-2009</td>
<td>stock exchange rates of NASDAQ</td>
<td>search for reducing the shortcomings of using ANN in predicting the market values With a new ANN model</td>
<td>Daily</td>
<td>classical ANN model MLP outperforms DAN2 and GARCH-MLP with a little difference.</td>
</tr>
</tbody>
</table>
3. Forecasting Models

In order to forecast the index of stock price return, the linear model of ARIMA and the non-linear model of GARCH, Artificial Neural Network (ANN) and Adaptive Network-based Fuzzy Inference System (ANFIS), have been used, which are to be reviewed in this part.

3.1 ARIMA Model

In general, a process is called ARMA (p, q) that includes p times auto-regression terms and q times of moving average

\[ r_t = \phi_1 r_{t-1} + \phi_2 r_{t-2} + \cdots + \phi_p r_{t-p} + \epsilon_t \sim i.i.d(0, \sigma^2) \]

where \( \epsilon_t \) white noise series, p and q are integers and \( \phi \) and \( \theta \) are the model’s coefficients.

3.2 Autoregressive conditional heteroskedasticity (ARCH)

In the traditional econometrics models such as ARMA, the fixed variance of the error terms is always considered one of the main hypotheses and classic assumptions. In order to get rid of this restrictive hypothesis, Robert Engle founded a new method called ARCH. One of the reasons behind the use of ARCH models is the existence of small and large forecast errors in economic clusters (such as value of foreign currency, inflation, stocks, etc.) in a way that the abovementioned series might have different behavior in different years. In fact, the advantage of ARCH model is in the fact that it could be used to describe the trend in conditional variance, taking into account previous information. A general model of ARCH (q) is as follows (Bollerslev, 1986):

\[ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + V_t \]

A problem that is practically faced when using these models is that when q is a large number, it usually violates the hypothesis that it is not negative for the variance equation to be valid. GARCH model was developed by Bollerslev (1986) and is considered a solution for this problem and is also a model that has a limited number of parameters (Johnston and Scott, 2000). Due to the large number of parameters that needed to be estimated in an ARCH model, Bollerslev (1986) presented a developed form of this model in which conditional variance equation includes p lags in conditional variances in addition to q past squares. In other words, conditional variance equation in a GARCH (p, q) is as follows (Bouchad, Matacz and Potters, 2001):

\[ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + V_t \]

3.3 Artificial Neural Network Model

The aim of Artificial neural networks is an attempt to build models that work like the human brain. The neural network’s job is to create an output pattern based on the input pattern provided to the network. Artificial Neural networks include a number of processing elements (artificial neurons), which receive and process the data and eventually provide an output for it. Input could consist of raw data or output of other processes elements. Output could be the final product or the input for another neuron. An artificial neural network consists of artificial neurons, which are the processing elements. The following diagram depicts a processing element.
In the above diagram, w’s are the weights allocated to each input, net is the sum function, f is the transformation function, x’s are the neuron inputs, and y’s are the neuron’s outputs.

Sum function calculates the weighted sum of the inputs and its formula is 
\[ \text{net} = \sum w_{ij}x_j \]

The Correlation between the surface of activity and output is described using the transformation function, which is of different types such as tangent, hyperbolic, Sigmoid, etc. (Wang, 2009).

3.4 Adaptive Network-based Fuzzy Inference System (ANFIS)

In recent years we have witnessed the widespread use of artificial neural networks. The idea of training to solve problems of identifying complex patterns using intelligent data agent perspective has been very challenging for academic researchers. Neural networks are simple computational tools for the data test and creation of a model from the data structures. The data used to generate models are known as the training data. Once neural networks use training data to learn the patterns in the data, they can use them to achieve different outputs and outcomes. Also, among the new modeling methods, fuzzy systems have gained a special place. This can be due to ability to implement human knowledge using the concept of language tags and fuzzy rules, nonlinearity and adaptation capabilities of these systems. In summary, a fuzzy system is a rule-based system based on if–then logic.

The starting point for building a fuzzy system is to obtain a set of fuzzy if-then rules from the knowledge of an expert or knowledge of the relevant area. The most important and most difficult step is to obtain these rules, because it requires knowledge of an expert and its implementation in a correct manner. In general, artificial neural networks do not have great ability to develop a model during a logical time for our purposes to use them. On the other hand, fuzzy modeling for application of decision integration from different variables requires an approach to learn from experiences (collected data). Artificial neural networks and fuzzy models have been used in many application areas and there are advantages and disadvantages in each of them. So combining these two approaches successfully and also artificial and fuzzy neural networks modeling have been the subject of future studies. One of the most common combined methods is ANFIS system which was introduced by Zhang in 1996.

This model implements a Sugeno fuzzy system in a neural structure and uses the error propagation method or a combination of error propagation and the least squares error methods for the training process. Rules are fixed in ANFIS and what are optimized are the membership function agents. To determine agents of Membership functions (or the shape of membership functions) the training features of neural network is used. Types of membership functions (such as triangular, Gaussian functions, etc.) and the number of membership functions for inputs and outputs are determined by trial and error method. In the first layer, ANFIS is necessary to determine the number and the type of membership function. The common definition of this is to identify a \( f \) function (so that it can almost be used instead of the \( f \) function). Fuzzy system containing \( N \) fuzzy rules to predict the price of gold is expressed as follows.

\[
 f(x) = \frac{\sum_{i=1}^{N} E_t \left( \prod_{j=1}^{n} \mu_{A_i(j)} (E_t - p_j) \right)}{\sum_{i=1}^{N} \left( \prod_{j=1}^{n} \mu_{A_i(j)} (E_t - p_j) \right)}
\]

In the above equation \( E_t, p, \mu, A \) and \( i \) are the gold price breaks up to the rank \( p \), membership level, fuzzy sets and fuzzy rules, respectively. Also \( \mu_{A_i(j)}(x_i) \) represents the membership degree from the input \( X \), corresponding to the \( i \)th fuzzy rule \( A_i(j) \) (Figure 2).

![Fig. 2. Structure of ANFIS with two inputs](image-url)
In this layer, x and y are non-fuzzy inputs to the node i and A_i and B_i are the names of the linguistic variables corresponding to this node.

\[ O_{2,i} = W_i = \mu A_i(x_i) \ast \mu B_i(y_i) i = 1,2. \]

Each output node represents the firing strength of a rule.

\[ O_{3,i} = \overline{W}_i = \frac{W_i}{W_1 + W_2} i = 1,2. \]

Fourth layer: node function calculates the fourth distribution layer of the rule to the total output which is defined as follows:

\[ O_{4,i} = \overline{W}_i f_i = \overline{W}_i (p_i x + q_i y + r_i) i = 1,2. \]

Where \( \overline{W}_i \) is a firing strength normalized from the third layer and \( p_i, q_i, \) and \( r_i \) are a set of parameters of this node. Also, the parameters of this layer are called inductive parameters.

\[ O_{5,i} = \sum \overline{W}_i f_i = \frac{\sum w_i f_i}{\sum w_i} i = 1,2. \]

### 4. Precision Assessment Methods in Forecasting Models

In order to investigate the efficiency of forecasting models of time series, Haykin (1994) introduced a number of measures, of which some of the most important are to be presented here:

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{n}}
\]

Other common measures of assessment in forecasts are Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

The advantage of using MAPE measure is that it could be used to compare the forecasts of the series that have different scales, as this measure is not dependent on the scale. This measure is defined as follows (Marcellinio, Stock, & Watson, 2006):

\[
\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \cdot 100
\]

If the time series has very small values, it is recommended not to use this index as dividing the error by small values will inflate the index. Mean Absolute Error Index is as follows:

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \hat{Y}_i|
\]

Theil inequality coefficient (TIC) is another measure that is used to compare the performance of forecasting models. This measure balances RMSE in a way that it is always between 0 and 1. The lower the value of these indices, the better would be the forecast (Leung, Daouk and Chen, 2000).
International Journal of Scientific Management and Development

Vol. 3 (10), 595-603 October (2015)

In which n is the data used, y is the actual value, and Ŷ is the predicted value.

Most econometrics studies with regard to forecasts have investigated the accuracy of forecasts by the models using the abovementioned statistic. Meanwhile, even when these measures have confirmed the accuracy of the forecasts, the models have not necessarily provided an accurate forecast. Gerlow, Irowin and Liu (1993) have demonstrated that the accuracy of the forecasts based on traditional statistical measures (like the abovementioned measures) provide little guidance on the benefits of using those forecasts in the exchanges in the capital market. Therefore, if the models that have statistically poor performance are used in the stock market, they could be either beneficial or harmful. In the capital market, the models that could accurately forecast the signs of future returns or could predict the turning points (change of direction of the prices) of a series will be more beneficial (Leitch and Tanner, 1991). Two possible indicators of the ability of a model to predict direction changes irrespective of their magnitude are those suggested by Pesaran and Timmerman (1992) and Refenes (1995). The relevant formulae to compute these measures are, respectively:

The first index is the Correct Sign Predictions, which is calculated in the following formula:

\[
CSP = \frac{1}{T - (T_1 - 1)} \sum_{t=T_1}^{T} Z_{t+s}
\]

\[Z_{t+s} = \begin{cases} 
1 & \text{if } (y_{t+s} \cdot f_{t,s}) > 0 \\
0 & \text{otherwise}
\end{cases}
\]

That:

T: the total size of the sample (sample data and out of sample data)
T_1: the first observation out of the sample

f_{t,s}: forecasting s steps ahead of the variance in time t
Y_t: the actual value of variance in time t

The second index is the Correct Direction Change Prediction index, which is calculated using the following formula:

\[
CDCP = \frac{1}{T - (T_1 - 1)} \sum_{t=T_1}^{T} Z_{t+s}
\]

\[Z_{t+s} = \begin{cases} 
1 & \text{if } (y_{t+s} - y_t)(f_{t,s} - y_t) > 0 \\
0 & \text{otherwise}
\end{cases}
\]

This article uses CSP to assess the models

5. Data and Descriptive Statistics

This study uses daily stock return index of Tehran Stock Exchange in the period between 22/11/2010 and 18/03/2014. Therefore, the size of the samples used in this study is 804 observations for the daily index. Of this number, 644 were used for teaching and the rest of the observations (160 observations) were used for forecasts.

Fig. 3. daily stock return index of Tehran Stock Exchange
Table (2) provides the descriptive statistics such as mean, variance, and standard deviation of stock price and other statistical features of the data related to indices of return in Tehran Stock Exchange. Kurtosis of these indices is greater than the kurtosis of the normal density function. Therefore, their curves have a narrow and wide continuance and high peak. In addition, the evidence is indicative of negative skewness in distribution of daily return of gold price, which is most probably for large decreases in comparison to increases. Jarque-Bera statistic, which is used for normality test, confirms the same. On the other hand, Dickey-Fuller (ADF) statistic is greater than the critical values related to the level of significance. As a result, the null hypothesis pertaining to the existence of a Unit root is rejected.

Table 2. Descriptive Statistics

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Observations</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Jarque-Bera</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td>stock return index</td>
<td>804</td>
<td>100.122</td>
<td>97.476</td>
<td>105.401</td>
<td>0.661</td>
<td>0.568</td>
<td>9.169</td>
<td>1316.66</td>
<td>-17.281</td>
</tr>
</tbody>
</table>

6. Precision Assessment of Forecasting Models

Assessment measures of forecast for different models have been presented in Table 3.

Table 3. Assessment measures of forecast for different models

<table>
<thead>
<tr>
<th>Forecasting Method</th>
<th>CSP</th>
<th>TIC</th>
<th>MAPE</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(1, 0, 1)</td>
<td>0.484</td>
<td>0.8980</td>
<td>291.7724</td>
<td>0.4640</td>
<td>0.6101</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.490</td>
<td>0.8084</td>
<td>260.2073</td>
<td>0.4599</td>
<td>0.6036</td>
</tr>
<tr>
<td>ANN</td>
<td>0.620</td>
<td>0.6070</td>
<td>192.0615</td>
<td>0.3944</td>
<td>0.5300</td>
</tr>
<tr>
<td>ANFIS</td>
<td>0.671</td>
<td>0.6342</td>
<td>181.362</td>
<td>0.3163</td>
<td>0.4491</td>
</tr>
</tbody>
</table>

As it could be seen in the table, ANFIS method is superior to all other methods in terms of all performance measures. For instance, ANFIS has been correct in 67 percent of the forecasts of the signs. This is while ARIMA, GARCH and ANN respectively show 48, 49 and 62 percent of success in accurate forecast of the sign of the variance. In addition, MAPE index for the ANFIS is significantly lower than the other models under study. TIC index for all four models is indicative of the superiority of ANFIS. The actual values and forecast based on the all four models of ANFIS, ANN, ARIMA, and GARCH could be seen in the following figure. The forecast values based on ANFIS are more or less similar to the actual values.

Fig. 4. The comparasion of actual and ARIMA predicted Values

Fig. 5. The comparasion of actual and GARCH predicted Values
6.1 Investigation of the meaningfulness of the difference in the Forecasts

Although the assessment measures above are among the useful and practical measures with regard to studying the forecast strength of different methods, none of them is able to investigate the statistical superiority of one single method. Therefore, in order to statistically test the hypothesis pertaining to the similarity of forecasts in these models, Morgan-Granger-Newbold (MGN) test was used (Diebold and Mariano, 2002).

Based on this formula, first the forecasting error of two different methods that are shown by $e_{1,t}$ and $e_{2,t}$ is calculated. Then the sum of $S_t$ and difference of $D_t$ of forecasting error is calculated based on the following correlations.

$$S_t = e_{1,t} + e_{2,t}$$
$$D_t = e_{1,t} - e_{2,t}$$
$$e_{1,t} = x_m - x_{1p}$$
$$e_{2,t} = x_m - x_{2p}$$

$X_{1p}$ and $X_{2p}$ are respectively the output values forecasted by the first and second models (rival models) and $X_m$ is the output values measured. The test for equality of mean of error squares of two different methods showed that the low error in the first model could be investigated using the MGN statistic.

$$MGN = \frac{\hat{\rho}_{sd}}{\sqrt{1 - \hat{\rho}_{sd}^2}}$$

$\hat{\rho}_{sd}$ is the correlation coefficient between $S_t$ and $D_t$ and $n$ is the number of observations. The value of MGN calculated with $t$ is compared with the degree of freedom of N-1 (Diebold and Mariano, 2002). The results of this test are provided in Table 4.

<table>
<thead>
<tr>
<th>Forecasting Method</th>
<th>ARIMA</th>
<th>GARCH</th>
<th>ANN</th>
<th>ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>------</td>
<td>------</td>
<td></td>
<td>------</td>
</tr>
<tr>
<td>GARCH</td>
<td>-2.35463819</td>
<td>------</td>
<td></td>
<td>------</td>
</tr>
<tr>
<td>ANN</td>
<td>-2.83040827</td>
<td>-2.723986661</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>ANFIS</td>
<td>-3.21087242</td>
<td>2.87543178</td>
<td>2.21435641</td>
<td>------</td>
</tr>
</tbody>
</table>

As it could be seen in the results reported in Table 4, all values of Diebold-Mariano statistic for different states is at a meaningful level. As a result, the precision of forecasting models is statistically and significantly different Hence, it could
be said that forecasts in linear and non-linear methods are significantly different and it is better to use non-linear methods for making forecasts on stock exchange returns specially ANFIS model.

7. Conclusion

Due to the significance of making forecasts on macro variables of economy, the majority of governments and central banks today not only take into account the current conditions, but also the short-term and long-terms forecasts pertaining to the main economic indices when adopting and administrating policies. Among economic variables, stock exchange prices, due to their wide, sudden and drastic changes and due to being affected by different economic, social, political, and even natural factors, are among the most complex and difficult variables for making forecasts. In fact, one of the famous economic theories in financial markets is the theory of unpredictability of the changes in stock prices, which is known in statistics as the Random Walk Hypothesis. The forecast models that have been developed to forecast stock prices are in fact a challenge to the above mentioned hypothesis as they attempt to show that despite the high complexity of price trends, their future trend could be predicted with an acceptable level of error. Among such models, one could mention the non-linear GARCH model, flexible non-linear models such as ANN and ANFIS. These models have demonstrated that they could be successful with regard to predicting the variables that have high complexity.

This study used linear and non-linear models, which are more or less new, for predicting TEPIX. The results of the dynamic forecasts done by ANFIS, ANN, ARIMA, and GARCH indicate that ANFIS model has a higher degree of precision and accuracy. Since the results of the forecasts were close in many of the cases, the statistical test of difference in prediction error was used, which indicated that prediction errors were significantly different. Based on the results, it could be recommended that non-linear models such as ANFIS and ANN are more suitable for making forecasts for Tehran stock prices.

References: