Do Dynamic Neural Networks Stand a Better Chance in Fractionally Integrated Process Forecasting?

Delavari, Majid and Gandali Alikhani, Nadiya and Naderi, Esmaeil

Department of Economics Science and Research Branch, Islamic Azad University, khouzestan-Iran., Department of Economics Science and Research Branch, Islamic Azad University, khouzestan-Iran., University of Tehran, Tehran, Iran.

11 September 2012

Do Dynamic Neural Networks Stand a Better Chance in Fractionally Integrated Process Forecasting?

Majid Delavari
Assistant Professor of Department of Economics Science and Research Branch,
Islamic Azad University, khouzestan-Iran. Email: Mjd_delavari@yhoo.com

Nadiya Gandali Alikhani
Corresponding Author**,** MA in Economics, Department of Economics, Islamic Azad University Khuzestan, Iran. Email: N.Alikhani@khouzestan.ac.ir

Esmaeil Naderi
MA in Economics, Faculty of Economic,
University of Tehran, Tehran, Iran. Email: Naderi.Ec@ut.ac.ir

ABSTRACT: The main purpose of the present study was to investigate the capabilities of two generations of models such as those based on dynamic neural network (e.g., Nonlinear Neural network Auto Regressive or NNAR model) and a regressive (Auto Regressive Fractionally Integrated Moving Average model which is based on Fractional Integration Approach) in forecasting daily data related to the return index of Tehran Stock Exchange (TSE). In order to compare these models under similar conditions, Mean Square Error (MSE) and also Root Mean Square Error (RMSE) were selected as criteria for the models' simulated out-of-sample forecasting performance. Besides, fractal markets hypothesis was examined and according to the findings, fractal structure was confirmed to exist in the time series under investigation. Another finding of the study was that dynamic artificial neural network model had the best performance in out-of-sample forecasting based on the criteria introduced for calculating forecasting error in comparison with the ARFIMA model.

Keywords: Stock Return, Forecasting, Long Memory, NNAR, ARFIMA.
JEL Classifications: C14, C22, C45, C53.

1. Introduction

Predicting macroeconomic variables has a high significance in scientific discussions related to economy. Different models have been developed for forecasting the future values in order to help economic policy-makers formulate and pursue appropriate monetary and financial policies. In addition, the issues related to Information Economics and the Asymmetric Information in financial markets and consequently, the complications involved in assessing the effects of different variables on the stock return under such conditions lend much further significance to the role of forecasting of variables (Sirucek, 2012; Asaolu and Ogunmuyiwa, 2011; Maysami and et al., 2004). On the other hand, one of the most important applications of different models can be forecasting future values of the variables; indeed, assessment of the accuracy of these forecasts is a means of comparing these models. Taking this fact into account, the very nature of forecasting and the tendency for making profits in financial markets has led to an increase in the number of studies on models and forecasting techniques during the recent decades in a way that today there are many models and methods in the literature on Econometrics and Applied Economics (Gooijer and Hyndman, 2006).

Although structural models have been rather successful in explaining the current situation, they have not been successful in forecasting; hence, during the recent years, economists have focused mainly on univariate time-series models rather than structural models for forecasting purposes (Haridy & Wu, 2009). In addition to univariate regression models, artificial intelligence network models have also become very popular because these models have yielded acceptable results in making accurate and reliable forecasts due to their identification and modeling of the complicated behavior of the financial markets .

Predictability of the stock return has a close relationship with the Efficient Markets Hypothesis (EMH). Efficiency is a fundamental concept in financial markets proposed in 1965 in the field of finance and many studies have been conducted in this regard According to this issue. After presenting these empirical evidences, EMH stated in its more full-fledged form that if return was predictable, many of the investors would gain huge benefits. Accordingly, we would be faced with a “money making machine” that could build unlimited wealth; this is, however, impossible in a stable economy (Granger & Timmermann, 2004).

EMH built based on Random Walk Models has been one of the basic challenges facing financial analysts since according to this hypothesis, the complex behavior of the financial markets cannot be modeled and predicted. By finding the roots of this issue, it can be found that the basic assumptions of the EMH cannot take into account all the elements involved in the financial markets. The most important assumption is that markets do not have memory in the sense that yesterday’s happenings will not influence today’s events and that investors are risk averse and always carefully consider all the information in the market (Burton, 1987). However, the results of many applied studies are indicative of the fact that majority of the investors are under the influence of the happenings in the market and form their expectations of the future stock prices in keeping with their experiences. This fact points to the conclusion that markets have memory (Granger and Joyeux, 1980). In addition, one cannot make a confident assertion that all the investors in the financial markets behave logically but that they may do trading and favor risking without paying attention to the market information because always some investors may make a profit and some may sustain losses. Therefore, although based on the assumptions of EMH, financial markets are apparently unpredictable, the fact is that this is not the case (Sowell, 1992). Thus, the assumptions of the EMH were faced with such criticisms and "Fractal Market Hypothesis" (FMH) was proposed which was able to provide a more comprehensive analysis of the markets. This hypothesis, in fact, implied the existence of a market composed of numerous investors pursuing their goals with different investment horizons. The types of information important to each one of these investors is different. On this basis, as long as the market sustains its fractal structure, it will stay stable without considering time scale of the investment horizons. On the other hand, when all the investors in the market have the same time horizon, the stability of the market will be undermined because people will do trading drawing on similar market information (Baillie, 1996).

Although rejecting the EMH implies non-randomness and, as a consequence, predictability of different series, this result is achieved because EMH has been formed based on the Random Walk Model and consequently the existence of a linear structure in the behavior of the market (Brock & et al., 1992). On the other hand, with regard to the financial markets which mainly have a complex and chaotic structure, FMH analyzes and assesses the issue of predictability from the perspective of nonlinear models (Vacha and Vosvrda, 2005). Although accepting the dependence of the behavior of a financial market on the FMH is a confirmation of the use of different non-linear models in consonance with the feature of long memory (e.g., Auto Regressive fractionally Integrated Moving Average or ARFIMA model) and also different types of neural network models (e.g., Nonlinear Neural network Auto Regressive or NNAR model as a dynamic model), it should also be noted that the fact that the inherent features of the mentioned markets (e.g., long memory) can improve the results of modeling should not be overlooked. Therefore, the present study attempts not only to consider fractal markets hypothesis in the return of TSE index but also compare different models based on the long memory (ARFIMA) and Dynamic Artificial Neural Network Model (NNAR) in terms of their out-of-sample forecasting performance using MSE and RMSE forecasting error measures. For this purpose, daily time series data were used from 25/3/2009 to 22/10/2011 (616 observations) out of which 555 observations (about 90% of the observations) were used for modeling and 60 observations for out-of-sample forecasting.

2. Methodology

2.1. Long Memory

After many important studies were conducted on the existence of Unite Root and Cointegration in time series starting in 1980, econometrics experts examined other types and subtypes of non-stationary and approximate persistence which explain the processes existing in many of the financial and economic time series. Today, different studies have been and are being conducted on these processes including "Fractional Brownian Motion" and "Fractional Integrated Process" and the "processes with long memory" (Lento, 2009). Hurst (1951) for the first time found out about the existence of processes with long memory in the field of hydrology. After that, in early 1980s econometricians such as Granger and Joyex (1980) and Hosking (1981) developed econometric models dealing with long memory and specified the statistical properties of these models. During the last three decades, numerous theoretical and empirical studies have been done in this area. For example, (Mandelbrot, 1999; Lee and et al. 2006; Onali and Goddard 2009)'s studies can be mentioned as among the most influential in this regard.

The concept of long memory includes a strong dependency between outlier observations in time series which, in fact, means that if a shock hits the market, the effect of this shock remains in the memory of the market and influences market activists' decisions; however, its effect will disappear after several periods of time (in the long term). Thus, considering the nature and the structure of financial markets such as the stock market, which are easily and quickly influenced by different shocks (economic, financial and political), it is possible to analyze the effects of these shocks and in a way determine the time of their disappearance by observing the behavior of these markets (Los and Yalamova, 2004). Meanwhile, the long memory will be used as a means of showing the memory of the market. By examining the long memory, the ground will also be prepared for improvement of financial data modeling.

2.2. ARFIMA Model

One of the most popular and most flexible models dealing with the long memory is the ARFIMA model in which fractional cointegration degree (*d*) is representative of the long memory parameter because it is indicative of the features of the long memory in the time series of the related variable. After making sure about the existence of this feature in a time series using ACF[1] tests, classic R/S[2] analysis and also semi-parametric methods such as GPH[3], MRS[4], etc. (Xiu and Jin, 2007), the most important stage in the process of estimation of these models is the "fractional differencing" stage; economists, however, used first-order differencing in their empirical analyses due to its ease of use (in order to avoid the problems of spurious regression in non-stationary data and the difficulty of fractional differencing). Undoubtedly, this replacement (of first-order differencing with fractional differencing) leads to over- or under-differencing and consequently loss of some of the information in the time series (Huang, 2010). On the other hand, considering the fact that majority of the financial and economic time series are non-stationary and of the Differencing Stationary Process (DSP[5]) kind, in order to eliminate the problems related to over differencing and to obtain stationary data and get rid of the problems related to spurious regression, we can use Fractional Integration. Another interesting point is that Fractional Integration can assume different values, but a specific value for this parameter (*d*) is indicative the long memory feature. Two conditions need to be met for assuming these values. Firstly, if $-0.5 < d < 0.5$, a series exhibits a stationary and invertible ARMA process with geometrically bounded autocorrelations. In other words, when $0 < d < 0.5$, the autocorrelation function decreases hyperbolically and the related process is a stationary long memory process meaning that the autocorrelations decay to zero and will not be summable. When $-0.5 < d < 0$, the long memory process will be invoked. The medium-term memory shows that the related variable has been over-

[1] Auto Correlation Function
[2] Rescaled Range Analysis
[3] Geweke and Porter-Hudak
[4] Modified Rescaled Range
[5] And some are also trend stationary processes

differenced and under such conditions, the reverse autocorrelation function decreases hyperbolically. The second condition is that a non-stationary is exhibited by the series if 0.5≤d<1 (Hosking, 1981).

Finally, it is worth mentioning that spurious long memory should not be overlooked; in fact, spurious long memory happens as the result of structural change and inattention to nonlinear transformations (Kuswanto and Sibbertsen, 2008). Therefore, based on the concepts introduced, we can correctly model the behavior of a variable using this model. The general form of the model *ARFIMA(p,d,q)* is as follows:

$$\phi(L)(1-L)^d(y_t-\mu_t)=\theta(L)\varepsilon_t \qquad t=1,2,3,...,T \tag{1}$$

In which $\phi(L)$ is polynomial autocorrelation, $\theta(L)$ represents moving average polynomial, L is Lag Operator, μ_t is the mean of y_t. Besides, in this equation, $Z_t = y_t - \mu_t$ and is cointegrated with rank *d*. Features of Z are dependent on the *d* value. If $d < 0.5$, covariance of the model will be fixed and if $d > 0$, it will have long memory feature (Hosking, 1981). *p* and *q* are integers and d is a long memory parameter. $(1-L)^d$ represents a fractional differencing operator which is calculated using the following formula:

$$(1-L)^d = \sum_{j=0}^{\infty}\delta_j L_j = \sum_{j=0}^{\infty}\binom{d}{j}(-L)^j \tag{2}$$

In the above equation, it has been hypothesized that $\varepsilon_t \sim N(0,\sigma_{\varepsilon_t}^2)$ and also ARMA section of the model are reversible (Aye and et al., 2012).

2.3. Nonlinear Neural Network Auto Regressive Model (NNAR)

Forecasting the behavior of a time series using econometric nonlinear models is constrained by many limitations. New models, however, enjoy more flexible structures and can get a better fitting of linear and non-linear econometric models. These models are a parallel distribution process with a natural structure and their most important feature is their ability to model nonlinear and complicated relationships without a need for prior hypothesis about the nature of relationships among the data. Generally, neural networks include two groups of dynamic and static networks (Dase and Pawar, 2010). Static networks such as the artificial neural network (ANN) do not have feedback factor; their output is calculated directly via inputs that have feedforward connections. But in dynamic neural networks (such as the Nonlinear Neural Network Auto Regressive (NNAR) and Nonlinear Neural Network Auto Regressive with exogenous variables (NNARX)), the value of the output is dependent on current and past input values, the outputs, and also the structure of the network (Georgescu and Dinucă, 2011; Khashei and Bijari, 2010).

These models have numerous applications in different areas such as prediction of financial markets, communication systems, power systems, classification, error detection, recognizing voices, and even in genetics. One of the most frequently used models among dynamic neural network models is the NNAR model. This model is developed by adding an AR process to a neural network model. Dynamic neural network (NNAR) has a linear and a nonlinear section; its nonlinear section is estimated by a Feed Forward artificial neural network with hidden layers and its linear section includes an autoregressive model (AR). The main advantage of using this model is that it is able to make more accurate long term predictions under similar conditions in comparison with the ANN model (Taskaya and Caseym 2005). The training approach in these models, which is consistent with Levenberg-Marquardt (LM) Training (Levenberg, 1944 and Marquardt, 1963) and the hyperbolic tangent activation function, is built on Error-Correction Learning Rule and starts the training process using random initial weights (Matkovskyy, 2012). After determining the output of the model for any of the models presented in the training set, the error resulting from the difference between the model output and the expected values is calculated and after moving back into the network in the reverse direction (from output to input), the error is corrected. The general form of the NNAR neural network models is:

$$\hat{Y}(t) = f[u(t), u(t-1), u(t-2), ..., u(t-n_u), y(t-1), y(t-2), ..., y(t-n_y)] \qquad (3)$$

In this formula, f represents a mapping performed by the neural network. The input for the network includes two u(t) exogenous variables (input signals) and target values (the lags of the output signals). The numbers for n_u and n_y include output signals and actual target values respectively which are determined by the neural network (Trapletti, and et al., 1998).

3. Empirical Results

In the present study, we are going to investigate the long memory feature in the returns of TSE index and to compare the performance of ARFIMA and NNAR models in forecasting this series. It should be mentioned that the abbreviation for the variables used in this study include *TEDPIX*, which is indicative of price index and dividend, *DLTED*, which shows *Logarithmic differential* of TEDPIX series.

3.1. Examining predictability of return of TSE

In this section, in order to explain the reasons for using non-linear models, two tests will be analyzed; first the non-randomness (and consequently predictability) of stock return series will be considered using the Variance Ratio Test and then its non-linearity will be examined using the BDS test.

3.1.1. Variance Ration test (VR Test)

This test (Lo and MacKinlay's, 1988) is used to examine whether the behavior of the components of stock return series is Martingale. In this test, when the null hypothesis is rejected, it can be concluded that the tested series will not be i.i.d. Overall, rejection of the null hypothesis in the VR test is indicative of the existence of linear or nonlinear effects among the residuals or the time series variable under investigation (Bley, 2011).

Table 1. The results of VR test in stock series

Test	Probability	Value
Variance ratio test	0.000	6.38

Source: findings of study

The results of the above test show that there is no evidence that the mentioned series (and the lag series) is of the Martingale; thus, the process of the data is not random. Accordingly, predictability of this series is implied in this way. The interesting point is that one cannot find out whether the data process in the stock return series is linear or non-linear as suggested by the results of this test, but can conclude that it is non-Martingale and predictable.

3.1.2. BDS Test

This test which was introduced in 1987 by Brock, Dechert and Scheinkman (BDS) acts based on the correlation integral which tests the randomness of the process of a time series against the existence of a general correlation in it. For this purpose, the BDS method first estimates the related time series using different methods. Then it uses correlation integral to test the null hypothesis on the existence of linear relationships between the series. Indeed, rejection of the null hypothesis indicates the existence of non-linear relationships between the related time series.

The statistics of this test (correlation integral) measures the probability that the distance between the two points from different directions in the fuzzy space is less than ε and like the fractal dimension in the fuzzy space when there is an increase in ε, this probability also changes in accordance with it (Olmedo, 2011). Accordingly, the general form of the test is $BDS_{m,T}(\varepsilon) = \dfrac{T^{\frac{1}{2}}[C_{m,T}(\varepsilon) - C_{1,T}(\varepsilon)^m]}{\sigma_{m,T}(\varepsilon)}$. In this equation, $\sigma_{m,T}(\varepsilon)$ is an estimation of the distribution of the asymptotic standard $C_{m,T}(\varepsilon) - C_{1,T}(\varepsilon)^m$. If a process is i.i.d, the BDS statistics will be normal distribution of the asymptotic standard. In this equation, if the BDS statistics is large enough, the null hypothesis will be rejected and the opposite hypothesis on the existence of a non-linear relationship in the process under investigation will be accepted (Moloney and Raghavendra, 2011). This test can be usefully applied for assessing the existence of a non-linear relationship in the observed time series. The results of this test have been provided in Table 2.

Table 2. The results of BDS test in the stock return series

Dimension	*BDS-Stat.*	*Standard division*	*Z-Stat.*	*Probability*
2	0.03678	0.003112	11.788	0.000
3	0.05957	0.004954	12.025	0.000
4	0.07071	0.005893	11.999	0.000
5	0.07201	0.006136	11.738	0.000

Source: findings of study

As it can be seen in Table 2, the null hypothesis, that means non-randomness of the stock return series, is rejected. So, this indicates the existence of a nonlinear process in the stock return series (there can also be a chaotic process as well). It is worth mentioning that whenever randomness of a series is rejected in more than two dimensions in the results of BDS test, the probability of the nonlinearity of this series will be high (because the opposite hypothesis is not clear in this test). So, this test can be a corroborative evidence of nonlinearity of the stock return series. Ergo, by confirming predictability and also nonlinearity of the related time series during the research, nonlinear models, i.e., ARFIMA and NNAR can be used for forecasting.

3.2. Stationary Test

As the next step, stationary of the dlted series (done to prevent creation of a spurious regression) will be assessed using different tests (see Table 3 for the results).

Table 3. The results related to stationary of the stock return series

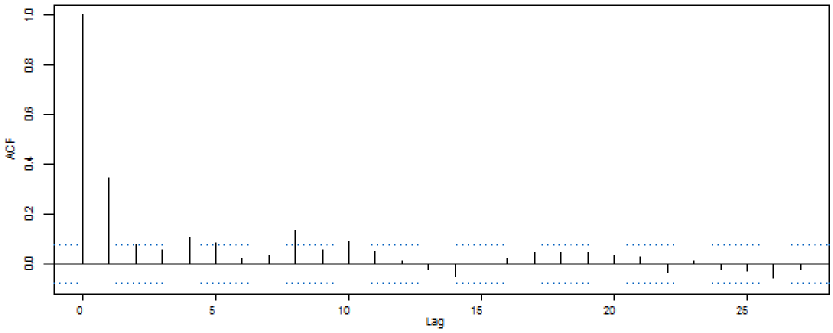
Test	Critical Stat.	Accounting Value	Result
ADF[1]	-1.9413	-16.586	Stationary
Ers[2]	3.2600	0.9403	Non-Stationary
Pp[3]	-1.9413	-17.543	Stationary
Kpss[4]	0.4630	0.590	Non-Stationary

Source: findings of study

If the long memory feature does not exist, it is expected that the series becomes stationary by first differencing, but the results of first differencing show that stock return series is stationary in ADF and PP tests while in the KPSS and also ERS test the results are indicative of non-stationary of the series (see Table 3 for the results). Such conditions might have been caused by the long memory feature in this series. For this reason, the long memory feature in the stock return series (by fractional differencing series) was further analyzed by the researchers. Besides, interpreting the Autocorrelation plot can also help to find if there is long memory in the stock return series; as shown in Fig. 1 below, the autocorrelation between different lags in the time series has not disappeared even after about 30 periods and, in fact, these autocorrelations in the series are decreasing at a very slow rate. This is anomalous to the behavior of autocorrelation of the stationary series in which the autocorrelations between different lags in the series decrease exponentially.

[1] Augmented Dickey–Fuller
[2] Elliott, Rothenberg and Stock
[3] Philips-Prone
[4] Kwiatkowski–Phillips–Schmidt–Shin

Fig. 1. ACF Graph for Stock Return Series



Source: findings of study

3.3. Examining the fractal market Hypothesis

Generally, dependence of the behavior of a market on the Efficient Markets Hypothesis depends on the significance of long memory parameter in its time series. In general, models that are based on long memory are highly dependent on the value of long memory parameter and also attenuation of the autocorrelation functions. On this basis, in the following subsections, the values of long memory parameter are estimated using the GPH. On the whole, this test is conforms to the frequency domain analysis and uses the Log-Period gram technique; this technique is a means for differentiating short and the long memory processes. It should also mentioned that slope of the regression line resulting from applying the Log-Period gram technique gives us the long memory parameter and if significant, the significance of the related feature in the stock return series can be inferred and the fractal markets hypothesis is confirmed. The results of this test have been provided in Table 4 below.

Table 4. Estimation of d parameter using GPH test based on the NLS method

Series	d-Parameter	t-stat.	Probability
Stock return series	1.04695	12.3	0.000
Stock series	0.14088	3.13	0.002

Source: findings of study

As shown in Table 4 above, the value for long memory parameter is non-zero (and also lower than 0.5) which is a confirmation of the existence of long memory in the stock return series. Therefore, two conclusions can be drawn from the above test: first, the fractal markets hypothesis is supported. The second conclusion is that this series should be fraction differenced once again so that modeling can be done in conformity with it. In the following sections, stock return series models will be focused upon using the models that are based upon long memory.

3.4. Estimation of the ARFIMA Model

There are different methods for estimation of the ARFIMA model and *d* parameter including Approximate Maximum Likelihood (AML), Exact Maximum Likelihood (EML), Modified Profile Likelihood (MPL), and Non Linear Least Square (NLS) (Ooms and Doornik, 1998). In the present study, EML, MPL, NLS methods have been selected for estimating these types of models using Ox-

Metrics software. Furthermore, based on the Akaike information criterion, a comparison was made between different models of ARFIMA and the model that is found to have the lowest score of the information criterion, will be the best model for explaining mean equation of the stock return series.

Table 5. The Results of Estimation of Different Models of ARFIMA

Models	Akaike Information Criterion		
	ML	NLS	EML
ARFIMA(1,0.14,1)	-7.2126	-7.3241	-7.3235
ARFIMA(1,0.14,2)	-7.2153	-7.3289	-7.3242
ARFIMA(2,0.14,1)	-7.2124	-7.3234	-7.3226
ARFIMA(2,0.14,2)	-7.2125	-7.3250	-7.3237

Source: findings of study

According to Table 5, it can be concluded that ARFIMA(1,0.14,2) has the lowest Akaike information criteria score and has the best performance (see Table 6 for specifications).

Table 6. The results of estimation for ARFIMA(1,0.14,2)

Variables	Coefficient	t-Stat.	Probability
Constant	0.0316	2.21	0.002
d-ARFIMA	0.1408	3. 13	0.002
AR(1)	0.8541	31.41	0.002
MA(1)	0.6163	18.67	0.002
MA(2)	0.2358	3.53	0.002
Dummy(1)	0.0796	7.28	0.002
Dummy(2)	0.0519	8.73	0.002

Source: findings of study

It is worth mentioning that, the dummy variables introduced in the above equation can be defined as the following: $Dum(1)$ are related to the financial crisis in 2007-2008 and $Dum(2)$ is related to transferring the shares of Telecommunication Company of Iran in the stock in line with the implementation of Article 44. Additionally, considering the fact that diagnostic tests conducted on residuals of the related model are indicative of the existence of conditional variance heteroscedasticity effects, Robust Regression was used for estimating this model.

3.5. Estimation of NNAR Model

Basically, the first step in modeling all non-linear models which is based on neural networks, determining the optimal combination of design elements of neural network with the same "Network Architecture". Hence, before comparing different models of dynamic neural network, some points related to the network architecture will be mentioned. First, for finding the number of optimal Neurons, an attempt was made to test and evaluate different networks using different neurons via encoding in the MATLAB software. Therefore, about 2 to 20 neurons were tested with two or three-layered networks; in this way, each one was trained 30 times. For comparing their performance, errors of the test data, which included 30% of the whole data, were randomly set as criterion in different models. Finally, the number of optimal neurons was found to be 10 and there were also two optimal hidden layers. The summary of information related to the network architecture has been provided in Table 7 below.

Table 7. Network Architecture

Design factor	Value
Network type	NNAR & NNARX
Number of neurons in the first hidden layer	10
Number of neurons in the second hidden layer	1
Preprocessing function	Feed Forward Network
Layer conversion function	Levenberg-Marquardt

According to the network architecture explained in Table 7 above, different models of NNAR will be estimated and compared:

Table 8. the results of estimation for different NNAR models

Models	MSE	RMSE
NNAR(1)	5.65*10^(-5)	7.52*10^(-3)
NNAR(2)	5.58*10^(-5)	7.47*10^(-3)
NNAR(3)	5.40*10^(-5)	7.35*10^(-3)
NNAR(4)	5.35*10^(-5)	7.31*10^(-3)
NNAR(5)	5.28*10^(-5)	7.26*10^(-3)
NNAR(6)	5.37*10^(-5)	7.33*10^(-3)
NNAR(7)	5.46*10^(-5)	7.39*10^(-3)
NNAR(8)	5.55*10^(-5)	7.45*10^(-3)
NNAR(9)	5.63*10^(-5)	7.50*10^(-3)
NNAR(10)	5.78*10^(-5)	7.60*10^(-3)
NNAR(11)	5.89*10^(-5)	7.67*10^(-3)
NNAR(15)	5.97*10^(-5)	7.73*10^(-3)
NNAR(20)	6.42*10^(-5)	8.01*10^(-3)
NNAR(30)	7.04*10^(-5)	8.39*10^(-3)

Source: findings of study

According to Table 8, NNAR(5) model (using 5 lags in stock series) had the best performance in comparison with other models based on the MSE and RMSE criteria.

3.6. Comparing the performance of models in accuracy of forecasts

MSE and RMSE are the most frequently used criteria for comparing models in accuracy of predictions among other criteria for assessing accuracy of prediction (Swanson & et al, 2011). Therefore, on the basis of the specified criteria, a comparison will be made between different models in their accuracy of out-of-sample forecasting (60 out-of-sample observations) (see Table 9 for the results of comparison).

Table 9. the results of comparison for ARFIMA and NNAR models

Models	MSE	RMSE
ARFIMA(1,0.14,2)	1.61*10^(-3)	4.01*10^(-2)
NNAR(5)	5.28*10^(-5)	7.26*10^(-3)

Source: findings of study

As shown in Table 8, the performance of NNAR(5) model is better than *ARFIMA(1,0.14,1)* in forecasting stock return series during the period under investigation.

4. Conclusions

In this study, Nonlinear Neural Network Auto Regressive Model (NNAR) and Autoregressive Fractionally Integrated Moving Average (ARFIMA) model were used to forecast TSE's Price and Dividend Index (TEDPIX). The results of this study showed that NNAR model yields more accurate forecasts about stock return index in the time series under investigation in comparison with the ARFIMA model. This result was not unexpected because considering the high flexibility of neural network models and especially dynamic neural network models in contrast the inflexible, imposed structure of regressive models such as ARFIMA model causes a change (adaptation) in their coefficients when there is a change in the time series under investigation. ARFIMA models might be inadequate for long memory time series as they might be both linear and nonlinear. This result has important implications for future studies in different fields. Therefore, future studies might be directed towards using Dynamic Neural Network models for forecasting purposes. Finally, these models can also be introduced to policy-makers and mass economy decision-makers and also investors as an appropriate method.

References

Asaolu, T.O., Ogunmuyiwa, M.S. (2011). An Econometric Analysis of the Impact of Macroecomomic Variables on Stock Market Movement in Nigeria. Asian Journal of Business Management, Vol. 3, No, 1, PP. 72-78.

Aye, G.C., Balcilar, M., Gupta, R., Kilimani, N., Nakumuryango, A., Redford, S. (2012). Predicting BRICS Stock Returns Using ARFIMA Models. University of Pretoria Working Paper, No. 2012-35, PP. 1-23.

Baillie, R.T. (1996). Long Memory Processes and Fractional Integration in Econometrics. Journal of Econometrics, Vol. 73, Issue. 1, PP. 5–59.

Bley, J. (2011). Are GCC Stock Markets Predictable?. Emerging Markets Review, Vol. 12, PP. 217–237.

Brock, W.A., Dechert, W.D., Sheinkman J.A. (1987). A Test of Independence Based on the Correlation Dimension. Working paper, University of Wisconsin at Madison, University of Houston, and University of Chicago, No. 8702, PP. 1-38.

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. Journal of Finance, Vol. 47, PP. 1731–1764.

Burton G.M. (1987). Efficient Market Hypothesis: The New Palgrave. A Dictionary of Economics, Vol. 2, PP. 120–23.

Los, C.A., Yalamova, R. (2004). Multifractal Spectral Analysis of the 1987 Stock Market Crash, Working Paper Kent State University, Department of Finance.

Dase R.K. and Pawar D.D. (2010). Application of Artificial Neural Network for Stock Market Predictions: A Review of Literature. International Journal of Machine Intelligence. Vol. 2, Issue. 2, PP. 14-17.

Geweke, J., Porter-Hudak, S. (1983). The Estimation and Application of Long-Memory Time Series Models. Journal of Time Series Analysis, Vol. 4, PP. 221–238.

Granger, C. W. J., Joyeux, R. (1980). An Introduction to Long Memory Time Series Models and Fractional Differencing. Journal of Time Series Analysis, No. 1, PP. 15-29.

Granger, C.W.J & Timmermann, A. (2004). Efficient Market Hypothesis and Forecasting. International Journal of Forecasting, PP. 15–27.

Gooijer, J. G., Hyndman, R. J. (2006). 25 Years of Time Series Forecasting. International Journal of Forecasting, Vol. 22, PP. 443– 473.

Georgescu, V., Dinucă, E. c. (2011). Evidence of Improvement in Neural-Network Based Predictability of Stock Market Indexes through Co-movement Entries. Recent Advances in Applied & Biomedical Informatics and Computational Engineering in Systems Applications, WSEAS Press, PP.412-417.

Haridy, S., Wu, Zh. (2009). Univariate and Multivariate Control Charts for Monitoring Dynamic Behavior Processes: A Case Study. Journal of Industrial Engineering and Management, Vol. 2, No. 3, PP.464-498.

Hosking, J. R. M. (1981). Fractional Differencing .Biometrika, No. 68, PP. 165-176.

Huang, S.C. (2010). Return and Volatility Contagions of Financial Markets over Difference Time Scales. International Research Journal of Finance and Economics, Vol. 42, PP. 140-148.

Hurst, H. (1951). Long-Term Storage of Reservoirs: An Experimental Study. Transactions of the American Society of Civil Engineers. Vol. 116, PP. 770 – 799.

Khashei, M., Bijari, M. (2010). An Artificial Neural Network (p,d,q) Model For Timeseries Forecasting. Expert Systems with Applications, Vol. 37, PP. 479–489.

Kuswanto, H., Sibbertsen. P. (2008). A Study on Spurious Long Memory in Nonlinear Time Series Models. Applied Mathematical Science, Ruse. 2, No. 53-56, PP. 2713-2734.

Lento, C. (2009). Long-term Dependencies and the Profitability of Technical Analysis. International Research Journal of Finance and Economics, Vol. 269, PP. 126-133.

Lee, J.W., Lee, K.E., Rikvold, P.A. (2006). Multifractal Behavior of the Korean Stock Market Index KOSPI. Physica A: Statistical Mechanics and its Applications, Vol. 364, PP. 355-361.

Lo, A.W., and MacKinlay, A.C. (1988). Stock Market Prices Do Not Follow Random Walks: Evidence From A Simple Specification Test. Review of Financial Studies, Vol. 1, No. 1, PP. 41-66.

Levenberg, K. (1944). A Method for the Solution of Certain Non-Linear Problems in Least Squares. Quarterly of Applied Mathematics, Vol. 2, No. 2, PP. 164-168.

Mandelbrot, B.B. (1999). A Multifractal Walk Down Wall Street. Scientific American, Vol. 280, No. 2, PP. 70-73.

Marquardt, D.W. (1963). An Algorithm for the Least-Squares Estimation of Nonlinear Parameters. Siam Journal of Applied Mathematics, Vol. 11, No. 2, PP. 431-441.

Matkovskyy, R. (2012). Forecasting the Index of Financial Safety (Ifs) of South Africa Using Neural Networks. Munich Personal Repec Archive (MPRA), No. 42153, PP. 1-19.

Maysami, R.C., Howe, L.Ch., Hamzah, M.A. (2004). Relationship between Macroeconomic Variables and Stock Market Indices: Cointegration Evidence from Stock Exchange of Singapore's all-S Sector Indices. Journal Pengurusan, Vol. 24, PP. 47-77.

Moloney, K., Raghavendra, S. (2011). Testing for Nonlinear Dependence in the Credit Default Swap Market. Emerging Economies, ISSN 2249-0949, Vol. 2, PP. 69-80.

Olmedo, E. (2011). Is There Chaos in the Spanish Labour Market?. Chaos, Solitons & Fractals, Vol. 44, PP. 1045-1053.

Onali, E., Goddard, J. (2009). Unifractality and Multifractality in the Italian Stock Market. International Review of Financial Analysis, Vol. 18, No. 4, PP. 154-163.

Ooms, M., Doornik, J.A. (1998). Estimation, Simulation and Forecasting for Fractional Autoregressive Integrated Moving Average Models. Discussion paper, Econometric Institute, Erasmus University Rotterdam, Presented at the Fourth Annual Meeting of the Society for Computational Economics, Cambridge, UK.

Sirucek, M. (2012). Macroeconomic Variables and Stock Market: US Review. Munich Personal RePEc Archive (MPRA), No. 39094, PP. 1-9.

Sowell, F. (1992). Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time-Series Models. Journal of Econometrics, Vol. 53, PP. 165-188.

Taskaya, T., Casey, M.C. (2005). A Comparative Study of Autoregressive Neural Network Hybrids. Neural Networks, Vol. 18, PP. 781–789.

Trapletti, A., Leisch, F., Hornik, K. (1998). Stationary and Integrated Auto Regressive Neural Network Processes. Working Paper, No. 24, PP. 1-15.

Vacha, L., Vosvrda, M.S. (2005). Dynamical Agents' Strategies and the Fractal Market Hypothesis. Prague Economic Papers, No. (2005-2), PP. 163-170.

Xiu, J., Jin, Y. (2007). Empirical Study of ARFIMA Model Based on Fractional Differencing. Physica:A, No. 377, PP. 137-184.