Abstract— In this paper we apply an Adaptive Network-Based Fuzzy Inference System (ANFIS) with one input, the dependent variable with one lag, for the forecasting of four macroeconomic variables of US economy, the Gross Domestic Product, the inflation rate, six monthly treasury bills interest rates and unemployment rate. We compare the forecasting performance of ANFIS with those of the widely used linear autoregressive and nonlinear smoothing transition autoregressive (STAR) models. The results are greatly in favour of ANFIS indicating that is an effective tool for macroeconomic forecasting used in academic research and in research and application by the governmental and other institutions.

Keywords—Linear models, Macroeconomics, Neuro-Fuzzy, Non-Linear models

I. INTRODUCTION

FUZZY logic is an effective rule-based modelling in soft computing, that not only tolerates imprecise information, but also makes a framework of approximate reasoning. The disadvantage of fuzzy logic is the lack of self learning capability. The combination of fuzzy logic and neural network can overcome the disadvantages of the above approaches. In ANFIS, is combined both the learning capabilities of a neural network and reasoning capabilities of fuzzy logic in order to give enhanced prediction capabilities. ANFIS has been used by many researchers to forecast various time-series comparing with Autoregressive (AR) and Autoregressive Moving Average (ARMA) models finding superior results in favour of ANFIS [1]-[3]. On the other hand empirical analysis in macroeconomics as well as in financial economics is largely based on times series. This approach allows the model builder to use statistical inference in constructing and testing equations that characterize relationships between economic variables. There is a few number of researches made in the field of macroeconomics prediction with ANFIS, where artificial intelligence procedures are still not used by the national statistical services and central banks in many countries, while conventional econometric modelling is still in use, especially the Autoregressive models. Their argument is that economics and finance are described better by statistical modelling and properties. This is not absolute necessary, because macroeconomic and financial times-series can be characterized by non-linearities, which might be more efficient described by neural networks. Additionally, economics and finance are social sciences, so imprecision of the human behaviour is widely present, which can be described by fuzzy logic. For this reason we propose the neuro-fuzzy approach.

In section II we present the methodology of stationarity and unit root tests, as well as the estimating and forecasting procedure of the models examined. In section III the frequency and the type of data are described. In section IV the estimated and forecasting results are reported, while in the last section the concluding remarks of this study and some proposals are presented.

II. METHODOLOGY

A. Unit Root and Stationary Tests

It is possible that the variables are not stationary in the levels, but probably are in the first or second differences. To be specific we confirm this assumption by applying Augmented Dickey-Fuller-ADF [4] and KPSS stationary test [5]. The ADF test is defined from the following relation:

\[ \Delta y_t = \mu + \gamma y_{t-1} + \phi_1 \Delta y_{t-1} + \ldots + \phi_p \Delta y_{t-p} + \beta t + \varepsilon_t \]  (1)

where \( y_t \) is the variable we examine each time. In the right hand of (1) the lags of the dependent variable are added with order of lags equal with \( p \). Additionally, (1) includes the constant or drift \( \mu \) and trend parameter \( \beta \). The disturbance term is defined as \( \varepsilon_t \). In the next step we test the hypotheses:

\( H_0: \ |\phi|=1, \beta=0 \Rightarrow y_t \sim I(0) \) with drift

against the alternative

\( H_1: |\phi|<1 \Rightarrow y_t \sim I(1) \) with deterministic time trend

The KPSS statistic is then defined as:

\[ \text{KPSS} = T^{-1} \sum_{i=1}^{p} \delta_i^2 / \sigma^2 (p) \]  (2)

where \( T \) is the number of sample and \( \sigma^2 (p) \) is the long-run...
D. Autoregressive Moving Average (ARMA) Models

From the previous two sections we combine Autoregressive (AR) Moving Average (MA) Models and the Autoregressive Moving Average (ARMA) which encompasses (7) and (9) is defined as:

\[ y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \]

(11)

ARMA(p, q) process has p autoregressive, lagged dependent-variable, terms and q lagged moving-average terms. The series \( R_t \) is said to be integrated of order one, denoted I(1), because taking a first difference produces a stationary process. A nonstationary series is integrated of order d, denoted I(d), if it becomes stationary after being first differenced d times autoregressive integrated moving-average model, or ARIMA \((p, d, q)\) and will be [7]:

\[ \Delta^d y_t = \mu + \phi_1 \Delta y_{t-1} + \phi_2 \Delta^2 y_{t-2} + \ldots + \phi_p \Delta^p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \]

(12)

The forecasts for ARMA \((p, q)\) model is given by (13)

\[ \hat{y}_{t+j} = \mu + \phi_1 \hat{y}_t + \phi_2 \hat{y}_{t-1} + \ldots + \phi_p \hat{y}_{t-p+1} + \varepsilon_{t+1} \]

(13)

Various procedures have been suggested for determining the appropriate lag length in ARMA models. The first is the addition of Akaike Information Criterion (AIC) and Schwarz Criterion (SC), which are often used in ARIMA model selection and identification. That is, AIC and SC are used to determine if a particular model with a specific set of \( p, d, \) and \( q \) parameters is a good statistical fit. SC imposes a greater penalty for additional coefficients than the AIC but generally, the model with the lowest AIC and SC values should be chosen. Specifically for Autoregressive Moving Average, Autoregressive and Moving Average models we choose Akaike criterion which is defined as:

\[ AIC(p) = \ln \frac{e^T e}{T} + \frac{2p}{T} \]

(14)

where \( e \) denotes the residuals, \( T \) is the sample and \( p \) indicates the lag number. We use for the optimum setting up to 5 lags.

E. Smoothing Transition Autoregressive (STAR) Models

The smoothing transition auto-regressive (STAR) model was introduced and developed by Chan and Tong [8] and is defined as:
\begin{equation}
 y_t = \pi_{10} + \pi_{1} w_1 + (\pi_{20} + \pi_{2} w_2) F(y_{t-d}; \gamma, c) + u_t
 \end{equation}

(15)

where \( u_t \sim (0, \sigma^2) \), \( \pi_{10} \) and \( \pi_{20} \) are the intercepts in the middle (linear) and outer (nonlinear) regime respectively. \( w_i = (y_{t-1}, ..., y_{t-j}) \) is the vector of the explanatory variables consisting of the dependent variable with \( j = 1, ..., p \) lags, \( R_{2j} \) is the transition variable, parameter \( c \) is the threshold giving the location of the transition function and parameter \( \gamma \) is the slope of the transition function. We shall consider two transition functions, the logistic and the exponential \([9]\), which are defined by (16) and (17) respectively. Parameter \( d \) indicates the delay and we divide parameter \( \gamma \) with \( \sigma(r) \), which is the standard deviation of \( y_t \). We follow this procedure, recommended by Teräsvirta \([11]\) because the estimation of parameter \( \gamma \) may cause problems like overestimations. First we apply a test to examine if \( y_t \) is linear or not proposed by Teräsvirta et al. \([10]\). The STAR model estimation is consisted by three steps according to Teräsvirta \([11]\).

\begin{equation}
 F(y_{t-d}) = (1 + \exp[-\gamma(y_{t-d} - c)])^{1/\gamma}, \gamma > 0
 \end{equation}

(16)

\begin{equation}
 F(y_{t-d}) = 1 - \exp(-\gamma(y_{t-d} - c)^2), \gamma > 0
 \end{equation}

(17)

a) The specification of the autoregressive (AR) process of \( j = 1, ..., p \). One approach is to estimate AR models of different lag orders and the maximum value of \( j \) can be chosen based on the AIC information criterion. Besides this approach, \( j \) value can be selected by estimating the auxiliary regression (18) for various values of \( j = 1, ..., p \), and choose that value for which the \( P \)-value is the minimum, which is the process we follow.

b) The second step is testing linearity for different values of delay parameter \( d \). We estimate the following auxiliary regression:

\begin{equation}
 y_t = \beta_0 + \beta_1 y_{t-1} + \ldots + \sum_{j=1}^{p} \beta_{2j} y_{t-j} y_{t-d} + \sum_{j=1}^{p} \beta_{3j} y_{t-j} y_{t-d}^2 + \sum_{j=1}^{p} \beta_{4j} y_{t-j} y_{t-d}^3 + \epsilon_t
 \end{equation}

(18)

The null hypothesis of linearity is \( H_0: \beta_{2j} = \beta_{3j} = \beta_{4j} = 0 \). In order to specify the parameter \( d \) the estimation of (18) is carried out for a wide range of values \( 1 \leq d \leq D \) and we choose \( d = 1, ..., 6 \) in the cases where linearity is rejected for more than one value of \( d \), then \( d \) is chosen by the minimum value of \( p(d) \), where \( p(d) \) is the \( P \)-value of the linearity test. Additionally if there are more than one zero \( P \)-values we choose this one with the highest \( F \)-statistic. We examine for \( j = 1, 2, 3, 5 \) and we choose those values of \( j \) similarly with parameter \( d \). In order to compute parameters \( c \) and \( \gamma \), we apply a grid search procedure for equation (18) with non linear squares and Levenberg-Marquardt algorithm. The grid search for parameter \( c \) takes place in the interval between the minimum and maximum value of each variable with increment 0.01 and for parameter \( \gamma \) in the interval [1 10] with increment 0.05. The initial value for parameter \( c \) is the mean value of our data and for parameter \( \gamma \) is 1.

c) The third and last step is the specification of STAR model. We test the following hypotheses by \([9,11]\)

\begin{equation}
 H_{04}: \beta_{4j} = 0, j = 1, ..., p
 \end{equation}

(19)

\begin{equation}
 H_{05}: \beta_{5j} = 0 \mid \beta_{4j} = 0, j = 1, ..., p
 \end{equation}

(20)

\begin{equation}
 H_{02}: \beta_{2j} = 0 \mid \beta_{3j} = \beta_{4j} = 0, j = 1, ..., p
 \end{equation}

(21)

If we reject the (19) hypothesis then we choose LSTAR model. If (19) is accepted and (20) is rejected then ESTAR model is selected. Finally accepting (19) and (20) and rejecting (21) we choose LSTAR model. We examine both LSTAR and ESTAR models, to examine the forecasting performance and to show that the difference between their predicting performances might be almost zero.

F. Adaptive network-based fuzzy inference system (ANFIS)

Jang \([12]\) and Jang and Sun \([13]\) introduced the adaptive network-based fuzzy inference system (ANFIS). This system makes use of a hybrid learning rule to optimize the fuzzy system parameters of a first order Sugeno system. An example of a two input with two rules first order Sugeno system can be graphically represented by Fig. 1.

Fig. 1. Example of ANFIS architecture for a two-input, two-rule first-order Sugeno model

, where the consequence parameters \( p, q, r \) and \( r \) of the \( n \)'th rule contribute through a first order polynomial of the form:

\begin{equation}
 f_n = p_n x_1 + q_n x_2 + r_n
 \end{equation}

(22)

The ANFIS architecture is consisted of two trainable parameter sets, the antecedent membership function parameters and the polynomial parameters \( p, q, r \), also called the consequent parameters. The ANFIS training paradigm.
uses a gradient descent algorithm to optimize the antecedent parameters and a least squares algorithm to solve for the consequent parameters. Because it uses two very different algorithms to reduce the error, the training rule is called a hybrid. The consequent parameters are updated first using a least squares algorithm and the antecedent parameters are then updated by backpropagating the errors that still exist. We define five linguistic terms {very low, low, medium, high, very high}. Because we examine ANFIS with only one input, the dependent variable with one lag, and we have defined five linguistic terms the rules will be:

IF \( y_{t-1} \) is very low THEN \( f_1 = p_1 x + r_1 \)

IF \( y_{t-1} \) is low THEN \( f_2 = p_2 x + r_2 \)

IF \( y_{t-1} \) is medium THEN \( f_3 = p_3 x + r_3 \)

IF \( y_{t-1} \) is high THEN \( f_4 = p_4 x + r_4 \)

IF \( y_{t-1} \) is very high THEN \( f_5 = p_5 x + r_5 \)

, where \( r_1, r_2, \ldots, r_5 \) are constants for \( f_1, f_2, \ldots, f_5 \) in this layer.

The chain rule used in order to calculate the derivatives and update the membership function parameters are [17]-[19]:

\[
\frac{\partial E}{\partial a_{ij}} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_i} \cdot \frac{\partial y_i}{\partial w_i} \cdot \frac{\partial w_i}{\partial \mu_{ij}} \cdot \frac{\partial \mu_{ij}}{\partial a_{ij}}
\]

\[
\frac{\partial E}{\partial b_{ij}} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_i} \cdot \frac{\partial y_i}{\partial w_i} \cdot \frac{\partial w_i}{\partial \mu_{ij}} \cdot \frac{\partial \mu_{ij}}{\partial b_{ij}}
\]

\[
\frac{\partial E}{\partial a} = \frac{1}{2}(y - y')^2
\]

where \( y' \) is the target-actual and \( y \) is ANFIS output variable. The chain rule used in order to calculate the derivatives and update the membership function parameters are [17]-[19]:

\[
O^1_i = \mu_A(x)
\]

The adjustable parameters that determine the positions and shapes of these node functions are referred to as the premise parameters. In layer 2 we have:

\[
O^2_i = w_i = \prod_{j=1}^{m} \mu_{A_j}(x)
\]

Each node output represents the firing strength of the reasoning rule. In layer 3, each of these firing strengths of the rules is compared with the sum of all the firing strengths. Therefore, the normalized firing strengths are computed in this layer as:

\[
O^3_i = \frac{O^2_i}{\sum O^2_i}
\]

Layer 4 implements the Sugeno-type inference system, i.e., a linear combination of the input variables of ANFIS, \( x_1, x_2, \ldots, x_p \), plus a constant term, \( r_1, r_2, \ldots, r_p \), to form the output.

\[
O^4_i = y_i = \overline{w_i} f = \overline{w_i}(p_x + r_i)
\]

, where parameters \( p_1, p_2, \ldots, p_p \) and \( r_1, r_2, \ldots, r_p \) in this layer are referred to as the consequent parameters. In layer 5 we take:

\[
O^5_i = \sum_{i} w_i f_i = \frac{\sum_{i} w_i f_i}{\sum_{i} w_i}
\]

In the last layer the consequent parameters can be solved for using a least square algorithm as:

\[
Y = X \cdot \theta
\]

, where \( X \) is the matrix

\[
X = [w_1 x + w_2 x + w_3 + \ldots + w_5 x + w_6]
\]

, where \( x \) is the matrix of inputs and \( \theta \) is a vector of unknown parameters as:

\[
\theta = [p_1, q_1, r_1, p_2, q_2, r_2, \ldots, p_5, q_5, r_5]^T
\]

For the first layer and relation (23) we use the triangular membership function. The triangular function is defined as:

\[
\mu_y(x_j; a_y, b_y) = \begin{cases} 1 - \frac{|x_j - a_y|}{b_y/2}, & \text{if } |x_j - a_y| \leq b_y/2 \\ 0, & \text{otherwise} \end{cases}
\]

, where \( a_y \) is the peak or center parameter and \( b_y \) is the spread or support parameter. We use error back propagation algorithm with steepest descent method in order to find the optimum parameters \( a \) and \( b \). The peak parameter update for the triangle membership function is:

\[
a_{ij}(n + 1) = a_{ij}(n) - \eta_a \cdot \frac{\partial E}{\partial a_{ij}}
\]

\[
\frac{\partial E}{\partial a_{ij}} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_i} \cdot \frac{\partial y_i}{\partial w_i} \cdot \frac{\partial w_i}{\partial \mu_{ij}} \cdot \frac{\partial \mu_{ij}}{\partial a_{ij}}
\]

, where \( \eta_a \) is the learning rate for the parameter \( a_{ij} \) and \( E \) is the error functions which is:

\[
e = \frac{1}{2}(y - y')^2
\]

, where \( y' \) is the target-actual and \( y \) is ANFIS output variable.
After some partial derivatives computations, the update equations for $a_i$ are $b_j$ are respectively

$$
\alpha_i(n+1) = \alpha_i(n) - \eta_i \cdot e(p_i x + r_i) - y \cdot \sum_{j=1}^{w} \frac{w_j}{b_j} \left( \frac{2x_j - a_j}{b_j} \right)
$$

(35)

$$
\beta_j(n+1) = \beta_j(n) - \eta_j \cdot e(p_j x + r_j) - y \cdot \sum_{i=1}^{w} \frac{w_i}{b_i} \left( 1 - \frac{\mu_i(x_i)}{b_i} \right)
$$

(36)

In Table I we present the initial values for parameters $a_j$ and $b_j$ in each variable we examine. The learning rates for parameters $a$, $b$ and consequent parameters (RHS) are set up at 0.1, 0.5 and 0.5 respectively. The number of maximum epochs is 50.

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>VERY</th>
<th>LOW</th>
<th>MEDIUM</th>
<th>HIGH</th>
<th>VERY</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>a=-2.5</td>
<td>a=0.5</td>
<td>a=1.5</td>
<td>a=3.5</td>
<td>a=5.5</td>
</tr>
<tr>
<td>Inflation</td>
<td>a=0</td>
<td>a=0.2</td>
<td>a=0.4</td>
<td>a=0.6</td>
<td>a=0.8</td>
</tr>
<tr>
<td>TB rates</td>
<td>b=0.15</td>
<td>b=0.15</td>
<td>b=0.15</td>
<td>b=0.15</td>
<td>b=0.15</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>a=5</td>
<td>a=6</td>
<td>a=7</td>
<td>a=8</td>
<td>a=9</td>
</tr>
<tr>
<td></td>
<td>b=1.5</td>
<td>b=1.5</td>
<td>b=1.5</td>
<td>b=1.5</td>
<td>b=1.5</td>
</tr>
</tbody>
</table>

Table II: The initial values for parameters $a_j$ and $b_j$ in each variable.

IV. EMPIRICAL RESULTS

In Table II we present the results of ADF and KPSS tests, while in Table III their critical values are reported. The results are mixed. For gross domestic product we reject unit root in $a=0.05$ and 0.10 based on ADF test, while we accept stationarity only in $a=0.01$ based on KPSS test. For unemployment and inflation rates we reject unit root based on ADF statistic in all statistical significance levels, but we reject stationary hypothesis based on KPSS test. We accept that treasury bills interest rates are stationary in first differences, $I(1)$, based on both ADF and KPSS tests. Beside these results we examine both ARMA and ARIMA processes to compare the forecasts. In Table IV we present the Autoregressive and Moving average processes for the variables we examine.

<table>
<thead>
<tr>
<th>Indices</th>
<th>ADF-statistic</th>
<th>KPSS-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP Growth Rate</td>
<td>-4.008</td>
<td>0.1767</td>
</tr>
<tr>
<td>Levels</td>
<td>-5.040</td>
<td>0.3759</td>
</tr>
<tr>
<td>Inflation Rate</td>
<td>-5.040</td>
<td>0.3759</td>
</tr>
<tr>
<td>Levels</td>
<td>0.1102</td>
<td></td>
</tr>
<tr>
<td>First differences</td>
<td>-2.089</td>
<td>0.5547</td>
</tr>
<tr>
<td>Treasury Bills Levels</td>
<td>-8.461</td>
<td>0.0335</td>
</tr>
<tr>
<td>First differences</td>
<td>-4.245</td>
<td>0.2776</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>0.0531</td>
<td></td>
</tr>
<tr>
<td>First differences</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III: Critical values for ADF and KPSS statistics.

1 MacKinnon [20], 2 Kwiatkowski et al. [5]
TABLE IV
AR, MA AND ARMA PROCESSES FOR VARIABLES

<table>
<thead>
<tr>
<th>Indices</th>
<th>AR(p)</th>
<th>MA(q)</th>
<th>ARMA(p,q)</th>
<th>ARIMA(p,d,q)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>1</td>
<td>4</td>
<td>1,4</td>
<td>1,1,3</td>
</tr>
<tr>
<td>Growth Rate</td>
<td>5</td>
<td>4</td>
<td>5,4</td>
<td>3,1,3</td>
</tr>
<tr>
<td>Inflation Rate</td>
<td>5</td>
<td>5</td>
<td>4,5</td>
<td>5,1,5</td>
</tr>
<tr>
<td>Treasury Bills</td>
<td>5</td>
<td>5</td>
<td>5,5</td>
<td>2,1,0</td>
</tr>
</tbody>
</table>

In Tables V and VI we report the linearity tests for the four macroeconomic variables. We observe that we accept the linearity process for GDP and that it could not be found a non-linear process for the specific macroeconomic variable. Also the value of lag order \( p \) is chosen based on the minimum \( p \)-value and in the cases where there are more than one zero \( p \)-values lag order \( p \) is chosen based on the highest \( F \)-statistic.

In Table VII we present the results of STAR model specification tests and the hypotheses (19)-(21). In all cases we accept that LSTAR is the appropriate model. In the case of inflation rate we accept the LSTAR model in \( \alpha = 0.05 \) and \( \alpha = 0.10 \), while we could accept ESTAR marginally for \( \alpha = 0.01 \).

In Table VIII the estimated parameters \( \gamma \) and \( c \), with non-linear squares and Levenberg-Marquardt algorithm, are reported. Generally, from the value of parameter \( \gamma \) we observe that there is a smoothly transition between the regimes.

In Tables IX and X the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are reported. The results are mixed, but the general conclusion is that ANFIS outperforms significant all the models in out-of-sample periods. The exception is the case of unemployment rate where Logistic STAR model outperforms ANFIS only in the in-sample period. In the case of the gross domestic product (GDP), among the linear models, ARMA and ARIMA present the best forecasting performance in the in-sample period, but in the out-of-sample period Moving Average presents the best performance followed by Autoregressive (AR) model, indicating that the last models might be more appropriate for out of sample forecasts in GDP. In the case of inflation rate and the in-sample period, the linear models present a similar performance, as the worst performance is reported by Logistic STAR model. ANFIS present the best performance, as it was...
mentioned, followed by exponential STAR model. In the out-of-sample period and the inflation rate ANFIS presents the best forecasting performance followed by ARIMA and exponential STAR model.

For the 6-month treasury bills interest rates and linear models, ARMA and ARIMA present the best performance in both in-sample and out-of-sample periods, while the best performance among all models, presents ANFIS followed by Logistic STAR model.

Finally, in the last macroeconomic variable we examine, the unemployment rate, Exponential STAR model followed by Logistic STAR model and ANFIS, presents the best forecasting performance in the in-sample period. In the out-of-sample period ANFIS outperforms the other models, followed by Logistic STAR model. RMSE and MAE in the case of ANFIS and the out-of-sample period, are significantly lower relatively to the other models.

V. CONCLUSIONS

In this paper we proposed the Adaptive Network-Based Fuzzy Inference Autoregressive System (ANFIS) which is known also as Adaptive Neuro-Fuzzy Inference System. We examined the forecasting performance of linear and non-linear models, as the autoregressive, moving average, autoregressive moving average and smoothing transition autoregressive models and we have shown that ANFIS outperforms significant the other models in the most cases. Furthermore, we concluded that the non-linear smoothing transition autoregressive models present a superior forecasting performance to the linear models in some cases. We examined only one fuzzy membership function, the triangular, while other membership functions can be tested as well as the trapezoidal, the Gaussian or the Generalized Bell functions among others. Furthermore, genetic algorithms can be used instead to error backpropagation algorithm we used in the current study, for the training process. Generally, ANFIS outperforms significant the conventional linear and non-linear econometric modelling in all macroeconomic variables we examined in the out-of-sample period which is our main interest. For this reason ANFIS technology should be adopted and used from the national statistical services, the central banks and the financial industry in the countries, where the conventional econometric modelling is still mainly used. Further research applications should be made and the principles and techniques of artificial intelligence should be introduced in economic academic departments in the future.

REFERENCES


TABLE IX

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>MAE</th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>1.9216</td>
<td>1.4282</td>
<td>0.3506</td>
<td>0.2457</td>
</tr>
<tr>
<td>MA</td>
<td>1.9285</td>
<td>1.4291</td>
<td>0.3560</td>
<td>0.2498</td>
</tr>
<tr>
<td>ARMA</td>
<td>1.8344</td>
<td>1.3221</td>
<td>0.3491</td>
<td>0.2520</td>
</tr>
<tr>
<td>ARIMA</td>
<td>1.8528</td>
<td>1.3372</td>
<td>0.3681</td>
<td>0.2548</td>
</tr>
<tr>
<td>ESTAR</td>
<td>2.0907</td>
<td>0.2067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSTAR</td>
<td>0.3999</td>
<td>0.2658</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANFIS</td>
<td>0.8003</td>
<td>0.6403</td>
<td>0.2245</td>
<td>0.1478</td>
</tr>
</tbody>
</table>

TABLE X

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>MAE</th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GDP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>3.8269</td>
<td>2.1184</td>
<td>0.4999</td>
<td>0.2276</td>
</tr>
<tr>
<td>MA</td>
<td>2.5212</td>
<td>1.9123</td>
<td>0.2162</td>
<td>0.2010</td>
</tr>
<tr>
<td>ARMA</td>
<td>4.6297</td>
<td>2.7162</td>
<td>0.2251</td>
<td>0.2231</td>
</tr>
<tr>
<td>ARIMA</td>
<td>3.9451</td>
<td>1.7993</td>
<td>0.1922</td>
<td>0.1773</td>
</tr>
<tr>
<td>ESTAR</td>
<td>0.2130</td>
<td>0.1774</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSTAR</td>
<td>0.2781</td>
<td>0.2471</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANFIS</td>
<td>1.6351</td>
<td>1.0651</td>
<td>0.1759</td>
<td>0.1551</td>
</tr>
</tbody>
</table>

For the 6-month treasury bills interest rates and linear models, ARMA and ARIMA present the best performance in both in-sample and out-of-sample periods, while the best performance among all models, presents ANFIS followed by Logistic STAR model.