A Study of Neuro-Fuzzy Inference System for Gross Domestic Product Growth Forecasting

E. Giovanis

Abstract—In this paper we present a Adaptive Neuro-Fuzzy System (ANFIS) with inputs the lagged dependent variable for the prediction of Gross domestic Product growth rate in six countries. We compare the results with those of Autoregressive (AR) model. We conclude that the forecasting performance of neuro-fuzzy-system in the out-of-sample period is much more superior and can be a very useful alternative tool used by the national statistical services and the banking and finance industry.

Keywords—Autoregressive model, Forecasting, Gross Domestic Product, Neuro-Fuzzy

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].

In this paper we examine the forecasting performance of Autoregressive (AR) models and Neuro-Fuzzy Systems in the growth rate of Gross Domestic Product in six countries. In section II we describe the methodology of Autoregressive and Neuro-Fuzzy system, while in section III we present the data frequency and the in-sample and out-of-sample period for the purpose of prediction examination. In section IV the empirical results are reported, while in the last section we discuss the main conclusion of the present study.

II. METHODOLOGY

A. Autoregressive (AR) Models

We consider a series $y_1, y_2, \ldots, y_n$. An autoregressive model of order p denoted AR(p), states that y_t is the linear function of the previous p values of the series plus an error term:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t \qquad (1)$$

, where $\varphi_1, \varphi_2, \ldots, \varphi_p$ are weights that we have to define or determine, and ε_t denotes the residuals which are normally distributed with zero mean and variance σ^2 [4]. Various procedures have been suggested for determining the appropriate lag length in a dynamic model such as based on information criteria Akaike, Schwartz and Hannan-Quinn or based on the t-student statistics indicating that the last added lagged dependent variable is significant. Specifically we choose Akaike criterion which is defined as:

$$AIC\,(p) = \ln \frac{e'e}{T} + \frac{2p}{T} \qquad (2)$$

, where e denotes the residuals, T is the sample and p indicates the lag number. We examine Akaike criterion up to 5 lags. Conditioned on the full set of information available up to time i and on forecasts of the exogenous variables, the one-period-ahead forecast of y_t would be

$$\hat{y}_{t+1|t} = \hat{\phi}_{0\,t+1|t} + \hat{\phi}_1\,y_t + \hat{\phi}_2\,y_{t-1} + \ldots + \hat{\phi}_p\,y_{t-p+1} + \hat{\varepsilon}_{t+1|t} \quad (3)$$

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

Jang [5] and Jang and Sun [6] 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.

Eleftherios Giovanis is with the Royal Holloway University of London, department of Economics, Egham, Surrey TW20 0EX, UK, e-mail: giovanis@freemail.gr, Eleftherios.Giovanis.2010@live.rhul.ac.uk
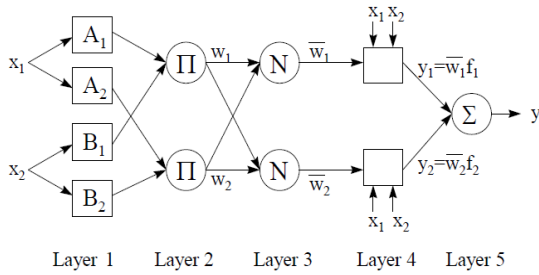.

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

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 on input, the dependent variable with one lag, we do on take the *AND-OR* operators. We could take more inputs, but one input is well enough to get very satisfying forecasts. The rules are:

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 y_{t-1} denotes the dependent or target variable with one lag, the gross domestic product growth rate.

The ANFIS architecture consists of five layers with the output of the nodes in each respective layer represented by O_i^l, where i is the ith node of layer l. Because we have five linguistic terms in the case we examine the steps for ANFIS system computation are:

$$O_i^1 = \mu_{A_i}(x) \tag{4}$$

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:

$$O_i^2 = w_i = \prod_{j=1}^{m} \mu_{A_i}(x) \tag{5}$$

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:

$$O_i^3 = \overline{w}_i = \frac{O_i^2}{\sum_i O_i^2} \tag{6}$$

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

$$O_i^4 = y_i = \overline{w}_i f = \overline{w}_i(p_i x + r_i) \tag{7}$$

, where parameters $p_1, p_2, ...,p_i$ and $r_1, r_2, ...,r_i$, in this layer are referred to as the consequent parameters. In layer 5 we take:

$$O_i^5 = \sum_i \overline{w}_i f_i = \frac{\sum_i \overline{w}_i f}{\sum_i \overline{w}_i} \tag{8}$$

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

$$Y = X \cdot \theta \tag{9}$$

, where X is the matrix

$$X = [w_1 x + w_1 + w_2 x + w_2 + ... + w_5 x + w_5] \tag{10}$$

, where x is the matrix of inputs and θ is a vector of unknown parameters as:

$$\theta = [p_1, q_1, r_1, p_2, q_2, r_2, ..., p_9, q_9, r_9]^T \tag{11}$$

, where T indicates the transpose. Because the normal least square method leads to singular inverted matrix we use the singular value decomposition (SVD) with Moore-Penrose pseudoinverse of matrix [7]-[9].

For the first layer and (4) we use the triangular and Gaussian membership functions. The triangular function is defined as:

$$\mu_{ij}(x_j; a_{ij}, b_{ij}) = \begin{cases} 1 - \frac{|x_j - a_{ij}|}{b_{ij}/2}, & if \ |x_j - a_{ij}| \le \frac{b_{ij}}{2} \\ 0, & otherwise \end{cases} \tag{12}$$

, where a_{ij} is the peak or center parameter and b_{ij} is the spread or support parameter. Gaussian function is:

$$\mu_{ij}(x_j; c_{ij}, \sigma_{ij}) = \exp\left(-\frac{(x_j - c_{ij})^2}{2\sigma_{ij}^2}\right) \tag{13}$$

, where c_{ij} is the center parameter and σ_{ij} is the spread parameter. In order to find the optimized antecedent

parameters we the backpropagation algorithm with the simple steepest descent method [10]-[12]

$$a_{ij}(n+1) = a_{ij}(n) - \frac{\eta_\alpha}{p} \cdot \frac{\partial e}{\partial a_{ij}} \qquad (14)$$

,where ηa is the learning rate for the parameter α_{ij} , p is the number of patterns and e is the error function which is:

$$e = \frac{1}{2}(y - y^t)^2 \qquad (15)$$

, where y^t 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 [10]-[12]:
:

$$\frac{\partial E}{\partial a_{ij}} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_i} \cdot \frac{\partial y_i}{\partial w_i} \cdot \frac{\partial w_i}{\partial \mu_{ij}} \frac{\partial \mu_{ij}}{\partial a_{ij}} \qquad (16)$$

After some partial derivatives computations, the update equations for a_{ij} are b_{ij} are respectively

$$\alpha_{ij}(n+1) = \alpha_{ij}(n) - \eta_\alpha \cdot e \frac{(p_i x + r_i) - y}{\sum_{i=1}^{n} w_i} \cdot \frac{w_i}{\mu_{ij}(x_j)} \cdot \frac{2|x_j - a_{ij}|}{b_{ij}} \qquad (17)$$

$$b_{ij}(n+1) = b_{ij}(n) - \eta_b \cdot e \frac{(p_i x + r_i) - y}{\sum_{i=1}^{n} w_i} \cdot \frac{w_i}{\mu_{ij}(x_j)} \cdot \frac{1 - \mu_{ij}(x_j)}{b_{ij}} \qquad (18)$$

Similarly the update equations for c_{ij} are σ_{ij} for Gaussian membership function are respectively:

$$c_{ij}(n+1) = c_{ij}(n) - \eta_c \cdot e \frac{(p_i x + r_i) - y}{\sum_{i=1}^{n} w_i} \cdot \left(\frac{(x_{ij} - c_{ij})}{\sigma^2_{ij}}\right) \mu_{ij}(x_j) \qquad (19)$$

$$\sigma_{ij}(n+1) = \sigma_{ij}(n) - \eta_\sigma \cdot e \frac{(p_i x + r_i) - y}{\sum_{i=1}^{n} w_i} \cdot \left(\frac{(x_{ij} - c_{ij})^2}{\sigma^3_{ij}}\right) \mu_i(x_j) \qquad (20)$$

Because each rule has one parameter and plus the constant there will be 2*5=10 parameters. The initial values for the triangular membership function and specifically for the spread parameters have been set up at 1.5. The center value parameters for the five rules have been set up respectively at -2.5, -0.5, 1.5, 3.5 and 5.5 On the other hand the parameter initialization for Gaussian membership function is quite different. First we took the values of GDP in specific intervals. In Table I we present the procedure.

TABLE I
SAMPLES OF GDP FOR GAUSSIAN FUNCTION

Very Low	If GDP>-2.5 and GDP<-0.5
Low	If GDP>-0.5 and GDP<1.5
Medium	If GDP>1.5 and GDP<3.5
High	If GDP>3.5 and GDP<5.5
Very High	If GDP>5.5

, where GDP denotes the gross domestic product growth rate. Based on the samples of Table I we obtain the respective average and standard deviations for each membership function in each rule the time. So the mean and the standard deviation correspond to parameters c and σ respectively. We follow this approach because generates very efficient forecasts. We use an ANFIS where the only input in the system is the dependent variable GDP with one lag. Furthermore, with less inputs we gain significant reduction in the computation time. The reason why we take one input is that in this case the forecasts are very satisfying in all the countries' GDP we examine. The forecasting performance of Autoregressive (AR) models and ANFIS in both in-sample and out-of- sample periods is measured based on the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) described respectively by (21) and (22).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \qquad (21)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (22)$$

The learning rate for parameters a and c is set up at 0.1, for spread parameters b and σ and for the consequent parameters is set up at 0.5. The number of maximum epochs is 50.

III. DATA

The data are in quarterly frequency and are referred in Gross Domestic Product (GDP) growth rates for quarter-by-quarter. The period examined is 1991-2009 for the countries of Canada, France, Italy, Japan, UK and USA. Moreover the period 1991-2006 is obtained as the in-sample for AR model or as the train period for the ANFIS model, while period 2007-2009 is taken as the out-of-sample period.

IV. EMPIRICAL RESULTS

In Table II the Autoregressive estimation results are reported. We conclude that the hypothesis of autocorrelation existence is rejected. In Table III we present the MAE and RMSE results for the forecasts generated by the two models we examine, for the in-sample and the out-of-sample period. We observe that ANFIS system slightly outperforms the

simple Autoregressive model in the in-sample period only in the cases of Canada, France and Japan. But the most significant fact is that the ANFIS system outperforms considerably very significant the AR model in the out-of-sample period, which is of greatest interest.

TABLE II
. AUTOREGRESSIVE (AR) MODEL ESTIMATIONS

	φ_1	φ_2	φ_3	φ_4	φ_5
	Estimated parameters				
Canada	0.9114	0.1542	-0.1429	-0.4483	0.5124
	(0.1072)	(0.1471)	(0.1470)	(0.1461)	(0.1064)
	[8.501]*	[1.047]	[-0.971]	[-3.068]*	[4.81]*
France	1.3021	-0.1914	-0.1955		
	(0.1231)	(0.2002)	(0.1243)		
	[10.577]*	[-0.956]	[-1.572]		
Italy	1.3204	-0.3020	-0.1743		
	(0.1231)	(0.2002)	(0.1243)		
	[11.197]*	[-1.550]	[-1.379]		
Japan	1.0447	-0.1479	-0.1001		
	(0.1220)	(0.1758)	(0.1219)		
	[8.565]*	[-0.841]	[-0.821]		
UK	1.631	-0.600	-0.1376	-0.2674	0.3503
	(0.1216)	(0.2293)	(0.2416)	(0.2514)	(0.1503)
	[13.423]*	[-2.620]**	[-0.571]	[-1.063]	[2.33]**
USA	1.316	-0.2154	-0.1892	-0.3396	0.4044
	(0.1186)	(0.1958)	(0.1980)	(0.1991)	(0.1233)
	[11.099]*	[-1.099]	[-0.955]	[-1.70]***	[3.28]*

	F-statistic	R^2_{adj}	Q-stat (2)	Standard Error of Estimate
	Diagnostic tests			
Australia	16.601	0.4713	4.719	0.8688
	{0.000}		{0.4510}	
France	202.969	0.8487	5.795	0.6506
	{0.000}		{0.3744}	
Italy	225.929	0.8620	5.083	0.7726
	{0.000}		{0.4230}	
Japan	88.907	0.7095	7.151	1.2168
	{0.000}		{0.1281}	
UK	234.785	0.9304	2.677	0.5540
	{0.000}		{0.6132}	
USA	113.928	0.8658	0.535	0.6852
	{0.000}		{0.9700}	

Standard errors in parentheses, t-statistics in brackets, p-values in {}, * denotes significance in α=0.01, ** denotes significance in α=0.05, *** denotes significance in α=0.10, Q-stat is the Ljung-Box test on squared standardized residuals with 2 lags

The forecasting superiority of ANFIS can be shown in Fig. 2-7, where the forecasts generated of the two models examined versus the actual values of GDP for the out-of-sample period are presented. This indicates that ANFIS is a good alterative choice for the economic policy makers and scientists working in great and central banks, as well as for those in financial and government institutions. In Fig. 2-7 ANFIS-AR denotes that we take as input the lagged dependent variable, so we have an autoregressive process.

TABLE III
FORECASTING PERFORMANCE OF AR AND ANFIS MODELS

|	In sample period 1991-2006					
	AR		ANFIS with triangular function		ANFIS with Gaussian function	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
Canada	0.4498	0.5689	0.3983	0.5167	0.6124	0.7071
France	0.5201	0.6371	0.5004	0.6233	0.5186	0.6577
Italy	0.5808	0.7566	0.6148	0.8552	0.6194	0.8243
Japan	0.9243	1.1916	0.7543	0.9535	0.8645	1.1513
UK	0.4196	0.5341	0.5955	0.6983	0.4586	0.6546
USA	0.5481	0.6606	0.5783	0.7310	0.5466	0.6998
	Out-of sample period 2007-2009					
	AR		ANFIS with triangular function		ANFIS with Gaussian function	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
Canada	1.8316	2.1814	0.9535	1.0982	0.9636	1.1568
France	2.1385	2.5840	1.8956	2.1290	1.9988	2.2126
Italy	2.3555	2.9334	2.1747	2.4904	2.3165	2.6035
Japan	2.2912	2.8301	1.4727	1.8296	1.9125	2.3580
UK	2.2230	2.8528	2.0451	2.6944	1.9257	2.3336
USA	2.0581	2.5179	1.9862	2.3702	1.9574	2.3558
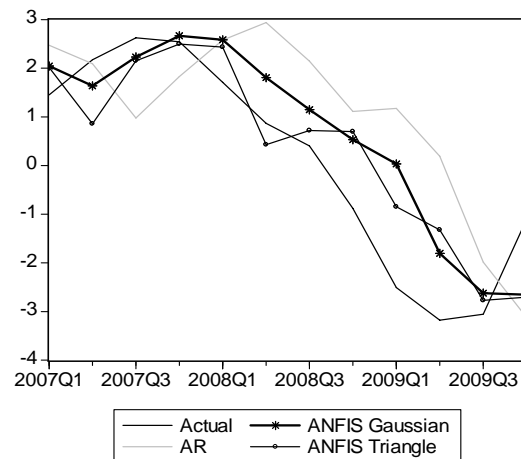
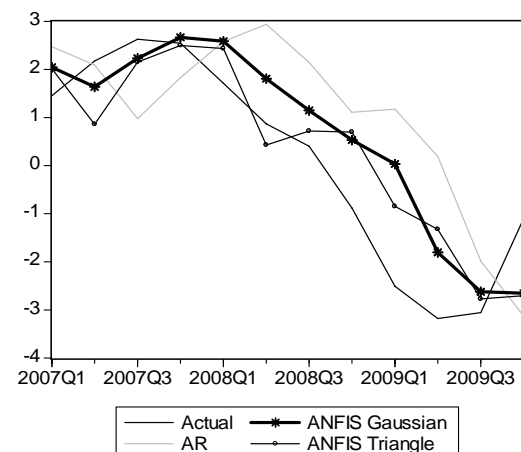Fig. 2. Out-of-sample period forecasts with AR and ANFIS for Canada

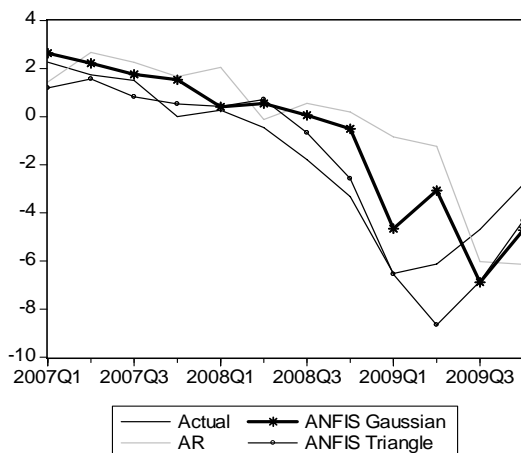Fig. 3. Out-of-sample period forecasts with AR and ANFIS for France

Fig. 4. Out-of-sample period forecasts with AR and ANFIS for Italy
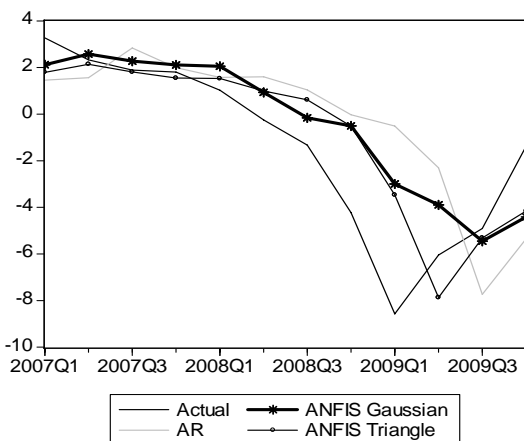

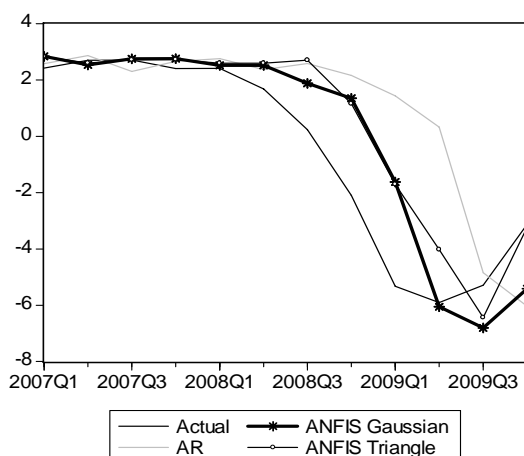Fig. 5. Out-of-sample period forecasts with AR and ANFIS for Japan

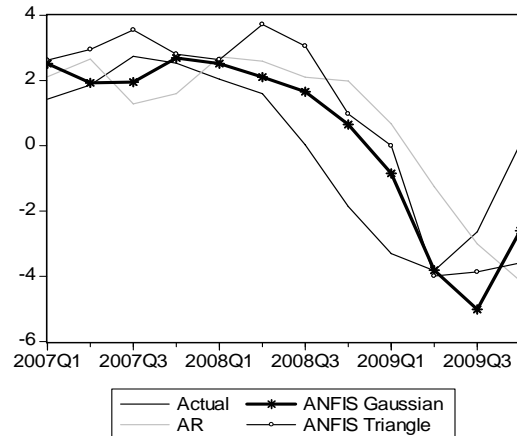Fig. 6. Out-of-sample period forecasts with AR and ANFIS for UK

Fig. 7. Out-of-sample period forecasts with AR and ANFIS for USA

V. CONCLUSIONS

In this paper we examined the forecasting performance of linear Autoregressive (AR) models and ANFIS. Our findings support ANFIS and this indicates the superiority of fuzzy logic and artificial intelligence models suggesting that is a powerful tool for the economic policy and decision makers. Furthermore, genetic algorithms can be applied instead to error backpropagation we used in this study and might have superior results. Additionally, we examined only two membership functions, while also other fuzzy membership functions can be applied, as the trapezoidal or the Generalized Bell function among others. Finally, more inputs can be obtained, but this is not absolutely necessary that it will improve the forecasts.

REFERENCES

[1] Atsalakis, S.G. and G.I. Atsalakis, "Fruit production forecasting by neuro-fuzzy techniques," Paper prepared for presentation at the 113th EAAE Seminar "*A resilient European food industry and food chain in a challenging world*", Chania, Crete, Greece, September, pp. 3-6, 2009

[2] Ballini, R., I. Luna, L.M. Lima and R.L.F. da Silveira, "A comparative analysis of neurofuzzy, ANN and ARIMA models for Brazilian stock index forecasting," *SCE - Computing in Economics and Finance,* 1995

[3] D. Marček, "Stock Price Forecasting: Autoregressive Modelling and Fuzzy Neural Network," *Proceedings of the 1999 EUSFLAT-ESTYLF Joint Conference*, Palma de Mallorca, Spain, September, pp. 22-25, 1995

[4] W.H Greene, *Econometric Analysis*, Sixth Edition, Prentice Hall, New Jersey, 2008, pp. 560-579

[5] J.-S.R. Jang, "ANFIS: Adaptive-Network-based Fuzzy Inference Systems," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 23, No. 3, pp. 665-685, 1993

[6] J.-S. R. Jang and C.-T. Sun, "Neuro-fuzzy Modeling and Control," *Proceedings of the IEEE*, vol. 83, no. 3, 378-406, March, 1995

[7] E.H. Moore, "On the reciprocal of the general algebraic matrix," *Bulletin of the American Mathematical Society,* vol. 26, pp. 394–395, 1920

[8] R. Penrose, "A generalized inverse for matrices," *Proceedings of the Cambridge Philosophical Society*, vol. 51, pp. 406–413, 1955

[9] M. Petrou, and P. Bosdogianni, *Image Processing: The Fundamentals*, John Wile & Sons, 2000

[10] L. Khan, S. Anjum and R. Bada, "Standard Fuzzy Model Identification using Gradient Methods," *World Applied Sciences Journal*, vol. 8, no. 1, pp. 01-09, 2010

[11] M.A. Denai, F. Palis and A. Zeghbib, ANFIS Based Modelling and Control of Non-Linear systems: A Tutorial, *IEEE Conf. on Systems, Man, and Cybernetics*, vol. 4, pp. 3433-3438, 2004

[12] L.H. Tsoukalas, and R.E. Uhrig, *Fuzzy and Neural Approaches in Engineering*, First Edition, John Wiley & Sons, 1997, pp. 445-470