

# Forecasting Financial Time Series Using Multiple Regression, Multi Layer Perception, Radial Basis Function and Adaptive Neuro Fuzzy Inference System Models: A Comparative Analysis

Arindam Chaudhuri<sup>1</sup>

<sup>1</sup> Faculty of Post Graduation Studies & Research in Engineering & Technology, Marwadi Education Foundation's Group of Institutions, Rajkot, State Gujarat, India

Correspondence: Arindam Chaudhuri, Faculty of Post Graduation Studies & Research in Engineering & Technology, Marwadi Education Foundation's Group of Institutions, Rajkot, State Gujarat, India. E-mail: arindam\_chau@yahoo.co.in

Received: December 11, 2011    Accepted: January 10, 2012    Online Published: September 27, 2012

doi:10.5539/cis.v5n6p13

URL: <http://dx.doi.org/10.5539/cis.v5n6p13>

## Abstract

In the last few decades, techniques such as Artificial Neural Networks and Fuzzy Inference Systems were used for developing predictive models to estimate the required parameters. Since the recent past Soft Computing techniques are being used as alternate statistical tool. Determination of nature of financial time series data is difficult, expensive, time consuming and involves complex tests. In this paper, we use Multi Layer Perception and Radial Basis Functions of Artificial Neural Networks, Adaptive Neuro Fuzzy Inference System for prediction of S% (Financial Stress percent) of financial time series data and compare it with traditional statistical tool of Multiple Regression. The accuracies of Artificial Neural Network and Adaptive Neuro Fuzzy Inference System techniques are evaluated as relatively similar. It is found that Radial Basis Functions constructed exhibit high performance than Multi Layer Perception, Adaptive Neuro Fuzzy Inference System and Multiple Regression for predicting S%. The performance comparison shows that Soft Computing paradigm is a promising tool for minimizing uncertainties in financial time series data. Further Soft Computing also minimizes the potential inconsistency of correlations.

**Keywords:** artificial neural network, adaptive neuro fuzzy inference system, multiple regression, soft computing, financial stress

## 1. Introduction

Forecasting (Armstrong, 2001; Chen, 2002) and predicting financial time series (Box, 2008; Brockwell et al., 2009; Chatfield, 2000; Fuller et al., 1999) has been a topic of active research since past few decades. It is a key element of financial and managerial decision making. It is highly utilized in predicting economic and business trends for improved decisions and investments. The inherent challenge involves in accuracy of predicting the financial market data (Altman, 1993). The central aspect of improving prediction accuracy is to have good and efficient forecasting techniques. The problem has initially been handled using different Statistical techniques. However, after the emergence of Computational Intelligence techniques such as Artificial Neural Network (ANN), Fuzzy Sets, Evolutionary Algorithms, Rough Sets etc. (Altun et al., 2004; Benardos et al., 2007; Chaudhuri et al., 2009; Dash et al., 2008; Jang et al., 1995; Kosko, 2008; Simpson, 1990; Zadeh, 1994; Zhang et al., 2005; Zimmermann, 2001) as alternative techniques to conventional Statistical techniques with better performance have paved the road for increased usage of these techniques in areas of financial time series forecasting. The stock traders have come to rely upon various types of Intelligent Systems to make trading decisions. Several Information Systems have been developed in recent years for modeling expertise, decision support and complicated automation tasks.

Every financial time series data is characterized by a unique financial cycle. Despite its apparent uniqueness from conditions that lead to boom times to triggers that result in reversals, historical narratives (Kindleberger, 2005) suggest that most cycles display common features. Boom times are associated with periods of credit expansion and persistent increases in asset prices often followed by rapid reversals. These commonalities confirmed by different empirical work (Bordo et al., 2001) suggest that developments in credit and asset markets

of individual countries provide an early warning indicator of vulnerability in financial system that would be useful in assessing current situation and in discussions of possible policy actions. In light of this it is somewhat surprising that the empirical work in this area is scarce. Whatever reasons there may be at general level, the problem in doing this type of analysis for developed countries is compounded by the scarcity of events that would qualify as a situation of financial crises resulting in financial stress. Absence of financial crises does not however mean that financial systems of developed countries have not or cannot come under stress but it does raise the issue of the best way to proceed. Financial stress is characterized as a situation in which large parts of financial sector face prospects of large financial losses. These situations are usually accompanied by an increased degree of perceived risk i.e., widening of distribution of probable losses and uncertainty i.e., decreased confidence in shape of that distribution.

In this work, we use a set of methodologies that can be used to assess the role of credit and asset prices as early warning indicators of vulnerability in the financial system of countries that have experienced very few or no financial crises over the sample period of interest. A typical example is Canada which is the basis of empirical work in this research. Bordo et al. (2001) illustrates that Canada has not experienced twin crises viz. banking and currency crises since the beginning of 1883 and has experienced only four currency crises since 1945. These features of sample preclude a meaningful country level analysis based on binary indicators of crises. Instead we suggest that in such circumstances one focuses on incidences of financial stress. Here we use the Financial Stress Index (SI), a continuous measure of financial stress developed by Illing and Liu (2006). The measure was originally developed for Canada, but the underlying approach can be applied to any country. In our examination of the role of credit and asset price in episodes of financial stress we consider both linear and nonlinear models, since the latter may be more suitable in capturing any behavioral asymmetries of financial market participants. The working hypothesis is that movements in credit and asset prices are indicators of the health of the system and its ability to withstand various types of shocks. Since the impact of a shock depends not only on the state of the system but also on the magnitude of the shock one would expect that everything else being the same, excessive growth of credit and persistent increases in asset prices reduce the ability of system to withstand shocks. The study aims to determine comparative empirical relationships for estimation of Financial Stress percent of time series data by using Multiple Regression (Fuller et al., 1999; Lund et al., 2002) Artificial Neural Network (ANN) models such as Multi Layer Perception (MLP) and Radial Basis Functions (RBF) (Kenneth et al., 2001) and Adaptive Neuro Fuzzy Inference System (ANFIS) (Jang, 1993). 239 time series data samples are tested for determination of Financial Stress percent (S%) (Bordo et al., 2001; Kindleberger, 2005) in terms of four major explanatory variables viz. Credit Measures (CM), Asset Prices (AP), Macroeconomic Variables (MV) and Foreign Variables (FV) to establish predictive models using Statistical and Machine Learning and Soft Computing techniques (Kosko, 2008; Zadeh, 1994). It is found that the relationships developed here allows CM, AP, MV and FV to be used as rapid, easy to determine, low cost means to estimate the stress potential with sufficient accuracy to allow for adequate design in situations where financial crisis situations can be prevented. Moreover the comparison of performance indices and coefficient of correlations for predicting Financial Stress percent revealed that prediction performance of RBF is higher than that of Multiple Regression, MLP and ANFIS. This paper is presented as follows. In the next section, experimental framework is highlighted. In section 3, data analysis is presented using Multiple Regression, Artificial Neural Network (ANN) models such as Multi Layer Perception (MLP) and Radial Basis Functions (RBF) and Adaptive Neuro Fuzzy Inference System (ANFIS). This is followed by illustration of experimental results in section 4. Finally, conclusions are given in section 5.

## 2. Experimental Framework

In this study, the data was captured from extensive studies of financial stress developed by Illing and Liu (2006). They constructed a weighted average of various indicators of expected loss, risk and uncertainty in financial sector. The resulting SI is a continuous, broad based measure that includes indicators from equity, bond, and foreign exchange markets such as: (a) the spread between yields on bonds issued by Canadian financial institutions and yields on government bonds of comparable duration (b) spread between yields on Canadian nonfinancial corporate bonds and government bonds (c) inverted term spread (d) beta derived from total return index for Canadian financial institutions (e) Canadian trade weighted dollar GARCH volatility (f) Canadian stock market GARCH volatility (g) the difference between Canadian and U.S. government short term borrowing rates (h) average bid-ask spread on Canadian Treasury bills (i) spread between Canadian commercial paper rates and Treasury bill rates of comparable duration. In constructing SI, Illing and Liu (2006) considered several weighting options that reflect relative shares of credit for particular sectors in economy. The resulting index is shown in Figure 1 is most effective in correctly signaling events that are widely associated with high financial

stress, for example, stock market crash in October 1987, peso crisis in 1994, long term capital management crisis in 1998 etc. This is not surprising given that Canada is a small open economy whose markets are well integrated internationally. As such it is not insulated from international financial developments. Turmoil in international financial markets will be reflected in increased stress in Canadian markets. This does not mean that financial stress is not or cannot be domestically generated, but it may indicate that the level of internal stress is secondary to level of external stress that spills over into Canadian financial markets. To assess importance of external factors in predicting financial stress in Canada, the sets of international explanatory variables are: (a) *Credit Measures*: the growth rate of total household credit (HouseCR), total business credit (BusCR) and total credit/GDP (CR/Y); (b) *Asset Prices*: stock prices (TSX), commercial real estate indices (real (ComREI) and nominal (real ComREI)), residential real estate indices, average price to personal disposable income ratio (AvgP/PDI) and Canadian dollar price of gold; (c) *Macroeconomic Variables*: Investment/GDP (I/Y), GDP growth rate, money (M1++ and M2++) and inflation (Total CPI and Core CPI); (d) *Foreign Variables*: crude oil, asset price indices (United States, Australia, Japan), world gold price, U.S. bank credit, US federal funds rate and world GDP.

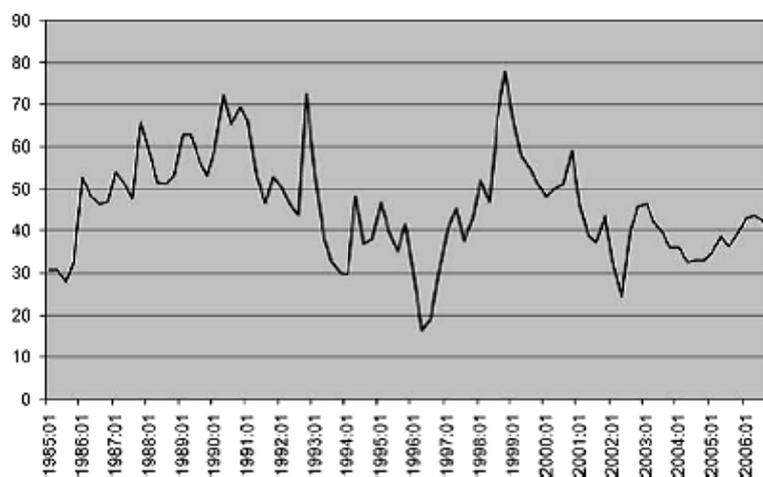


Figure 1. Financial stress index

Sufficient quantity of data having high quality is required in order to construct reliable predictive model. To achieve this quarterly data containing 239 samples spanning over the period 1984 – 2006 is considered. The forecasting experiment is performed over the period 1996 – 2006. The last observation is 2006 quarter 4. The explanatory variables are converted into growth rates such that all variables are stationary. We consider both quarterly and annual growth rates since it is possible that longer run cumulative growth rates in explanatory variables may contain more information about financial stress than quarterly growth rates. The results obtained by considering explanatory variables as basic test statistics are tabulated in Table 1.

Table 1. Explanatory variables as basic test statistics

	S %	CM	AP	MV	FV
Maximum	1.1	5.5	0.11	5.1	3.9
Minimum	15.1	112	1.86	96	69
Average	6.76	56.6	0.86	46.6	36.66
Standard Deviation	3.439	26.566	0.369	24.32	26.99

### 3. Data Analysis

In order to establish the predictive models among the parameters obtained in this work, simple regression analysis was performed in initial stage of the analysis. The relations between S with other parameters were analyzed employing linear, power, logarithmic and exponential functions. Statistically significant and strong correlations were found to be linear and regression equations were established among index parameters with S as given in Table 2. All obtained relationships were found to be statistically significant according to *Student's t test* at 95 % level of confidence (Brockwell et al., 2009).

Table 2. Predictive models for assessing S%

	Predictive Model	R <sup>2</sup>
S – CM	S = 0.1239 CM + 0.2686	0.90
S – AP	S = 6.6766 AP + 0.2276	0.69
S – MV	S = 0.1436 MV + 0.456	0.86
S – FV	S = 0.8686 FV + 0.5666	0.55

In following subsections the analysis of the financial time series data analysis is performed using Multiple Regression, MLP and RBF ANN models and ANFIS.

### 3.1 Multiple Regression

Multiple Regression is a commonly used statistical technique dating back to 1908 (Fuller et al., 1999) due to Pearson. It is employed to predict variance in an interval dependent, based on linear combinations of interval, dichotomous or dummy independent variables. However, the real life data is subjected to different distortions and therefore it is not possible to produce totally accurate predictions. Keeping this in view multiple regression allows the identification of a set of predictor variables which together provide a useful estimate of likely value on a criterion variable. The multiple regression equation takes the form  $y = a_1x_1 + a_2x_2 + \dots + a_nx_n + b$  where  $a_1, a_2, \dots, a_n$  are regression coefficients representing the amount the dependent variable  $y$  changes when corresponding independent variable changes one unit;  $c$  is a constant where the regression line intercepts  $y$  axis representing the amount the dependent  $y$  will be when all independent variables are zero. The standardized versions of  $a$  coefficients are beta weights and ratio of beta coefficients is the ratio of relative predictive power of independent variables. The major drawback of all regression techniques is that only relationships are ascertained but it is difficult to validate the underlying mechanism. Multiple regression analysis is conducted to correlate the measured Financial Stress percent to four financial indices, viz. Credit Measures (CM), Asset Prices (AP), Macroeconomic Variables (MV) and Foreign Variables (FV) as given in Table 3. The multiple regression equation to predict S percent is given as follows:

$$FS = (9.226 \times 10^{-2})CM + (2.500 \times 10^{-2})AP + (5.339 \times 10^{-2})MV + (6.600 \times 10^{-2})FV - 0.155 \quad (1)$$

Table 3. Summary of Multiple Regression for predicting S %

Independent Variables	Coefficient	Standard Error	t – Value	Sig. Level
Constant	-0.155	0.180	0.860	0.396
CM	$9.226 \times 10^{-2}$	0.011	8.626	0.000
AP	$2.500 \times 10^{-2}$	0.332	0.072	0.946
MV	$5.339 \times 10^{-2}$	0.012	3.066	0.002
FV	$6.600 \times 10^{-2}$	0.016	4.001	0.001

The correlation coefficient between measured and predicted values is good measure to verify prediction performance of model. Figure 2 shows the relationships between measured and predicted values obtained from multiple regression equation for S % with good correlation coefficient. In this work, indices Variance Factor (VAF) and Root Mean Square Error (RMSE) given by Equations 2 and 3 respectively account for measured and predicted values. These indices were calculated to control performance of prediction capacity of the predictive model developed (Fuller et al., 1999; Lund et al., 2002).

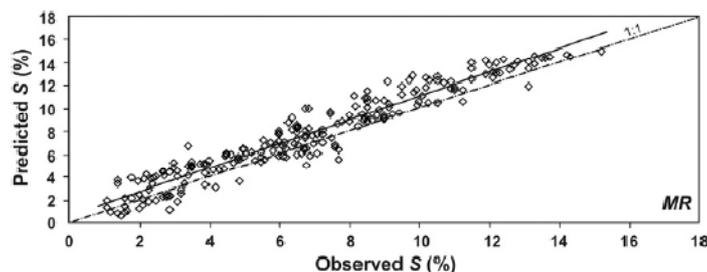


Figure 2. Cross correlation of predicted and observed values of S % for Multiple Regression

$$VAF = \left[ 1 - \frac{\text{var}(y - y')}{\text{var}(y)} \right] \times 100 \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y - y')^2} \quad (3)$$

Here  $y$  and  $y'$  are measured and predicted values respectively. The calculated indices are given in Table 4. The prediction model is excellent if RMSE is zero and VAF is 100.

Table 4. Performance indices such as RMSE, VAF, MAPE and  $R^2$  for different models

Model	RMSE	VAF %	MAPE	$R^2$
MR	1.36	90.55	22.6	0.921
MLP	0.81	93.11	12.6	0.939
RBF	0.75	93.55	12.4	0.946
ANFIS	0.82	93.86	14.6	0.936

Another measure of accuracy in a fitted series value in statistics viz. Mean Absolute Percentage Error (MAPE) is also used for comparison of prediction performances of the models. MAPE which usually expresses accuracy as a percentage is given by the following Equation:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{A_i - P_i}{A_i} \right| \times 100 \quad (4)$$

Here  $A_i$  and  $P_i$  denote the actual and predicted values respectively. High prediction performances are indicated from the calculated values of RMSE, VAF and MAPE as illustrated in Table 4.

### 3.2 Artificial Neural Networks

ANN is a mathematical model (Kosko, 2008; Zadeh, 1994) that is inspired by structure and functional aspects of biological neural systems. ANN consists of an interconnected group of artificial neurons which process information using connectionist approach. ANN is infact an adaptive system that changes its structure based on external or internal information that flows through the network during learning phase. Modern neural networks are non linear Statistical data modeling tools. They usually model complicated relationships between inputs and outputs in data to extract patterns or trends that are too complex to be observed by humans or other computational techniques. A trained ANN can be considered as an *expert* in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and solve them. The particular network can be defined by three fundamental components, viz. transfer function, network architecture and learning law (Simpson, 1990). It is essential to define these components, to solve the problem satisfactorily. Neural networks consist of a large class of different architectures. Keeping this point in view ANN is often used as a direct substitute for auto correlation, multivariable regression, linear regression, trigonometric and other statistical analysis and techniques.

In this work, MLP and RBF (Kenneth et al., 2001; Kosko, 2008) are used as two ANN models to estimate the Financial Stress percent of the Financial Time Series data. Both MLP and RBF are the two most widely used ANNs for classification and regression problems as they produce appreciable results in pattern classification (Loh et al., 2000). They are robust classifiers with generalization ability for imprecise data. The main difference between MLP and RBF is that later is based on localist type of learning which is responsive to a limited section of input space. On the other hand, MLP relies on more distributed approach. The output of MLP is produced by linear combinations of outputs of hidden layer nodes in which every neuron maps a weighted average of inputs through a sigmoid function. In one hidden layer RBF network hidden nodes map distances between input vectors and center vectors to outputs through a nonlinear kernel or radial function. The entire data is first normalized and divided into three sets in the proportions of: (a) 60% for training; (b) 20% for testing and (c) 20% for verification. The implementation is performed using MATLAB. The ANNs used are decomposed into three layer feed-forward network that consists of: (a) one input layer having 3 neurons; (b) one hidden layer consisting of 2 neurons for MLP and 16 neurons for RBF and (c) one output layer with one neuron represented schematically in Figure 3. The optimum number of neurons in the hidden layer is decided after a series of trial runs in networks

having minimum error. In this process network parameters of learning rate and momentum were set at 0.01 and 0.10 respectively. The variable learning rate with momentum *trainlm* as networks training function and *tansig* as an activation or transfer function are used for all layers here.

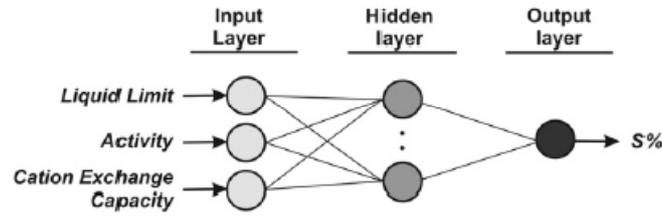


Figure 3. MLP and RBF ANN used in this work

### 3.2.1 Multi Layer Perception

MLP is an ANN with feed forward topology (Kenneth et al., 2001; Kosko, 2008) that maps set of weighted sum of input data and bias term into a set of desired outputs. MLP consists of multiple layers of nodes in a directed graph with each layer fully connected to the next one. Except for input nodes each node is a neuron with a nonlinear activation function. MLP utilizes a supervised learning technique called back propagation for training the network. MLP is a modification of standard linear perception which distinguishes data that is not linearly separable. MLP networks consist of an input layer, one or more hidden layers and an output layer. Each layer has a number of processing units and each unit is fully interconnected with weighted connections to units in subsequent layer. MLP transforms  $n$  inputs to  $l$  outputs through some nonlinear functions. The output of the network is determined by activation of units in output layer as follows:

$$x_n = f\left(\sum_p x_p \tilde{w}_{pn}\right) \quad (5)$$

Here  $f(\bullet)$  is activation function;  $x_p$  is activation of  $p^{th}$  hidden layer node and  $\tilde{w}_{pn}$  is interconnection weight between  $p^{th}$  hidden layer node and  $n^{th}$  output layer node which is modeled using the following  $\pi_2$  fuzzy membership function (Zadeh, 1994; Zimmermann, 2001):

$$\mu_{\tilde{w}_{i,j}}(w_{i,j}) = \begin{cases} \frac{a_{i,j}}{b_{i,j} + a_{i,j} - w_{i,j}}, & w_{i,j} < b_{i,j} \\ 1, & b_{i,j} \leq w_{i,j} \leq c_{i,j} \\ \frac{d_{i,j}}{w_{i,j} - c_{i,j} + d_{i,j}}, & w_{i,j} > c_{i,j} \end{cases} \quad (6)$$

The activation level of nodes in hidden layer is determined in a similar fashion. Based on differences between calculated output and target value an error function is defined as follows:

$$E = \frac{1}{2} \sum_m^M \sum_n^N (v_n^{(m)} - x_n^{(m)})^2 \quad (7)$$

Here  $M$  is the number of pattern in data set and  $N$  is the number of output nodes. The objective is to reduce error by adjusting interconnections between layers. The weights are adjusted using gradient descent Back Propagation (BP) algorithm. A training data consisting of a set of corresponding input and target pattern values  $v_n$  is required by the algorithm. During training MLP starts with random set of initial weights distributed through fuzzy membership function given Equation (6). The process is continued until set of values  $\tilde{w}_{ip}$  and  $\tilde{w}_{pn}$  is optimized so that predefined error threshold is met between  $x_n$  and  $v_n$  (Chaudhuri et al, 2009; Cohen et al., 2003; Lee, 1990). Each interconnection between the nodes is adjusted by the amount of weight update value according to BP algorithm as follows:

$$\Delta \tilde{w}_{pn} = -\eta \frac{\delta E}{\delta \tilde{w}_{pn}} = -\eta \delta_n x_p \quad (8)$$

$$\Delta \tilde{w}_{ip} = -\eta \frac{\delta E}{\delta \tilde{w}_{ip}} = -\eta \tilde{\delta}_p x_i \quad (9)$$

Here,  $\delta_n = x_n(v_n - x_n)$ ;  $\tilde{\delta}_p = x_p = \sum \delta_n \tilde{w}_{pn}$ ;  $x_n = x_n(1 - x_n)$  and  $x_p = x_p(1 - x_p)$  when bipolar sigmoid activation function is used. The cross correlation between predicted and observed values indicate that MLP is highly favorable for prediction of S % as indicated in Figure 4. The RMSE, VAF, MAPE and R<sup>2</sup> values are presented in Table 4.

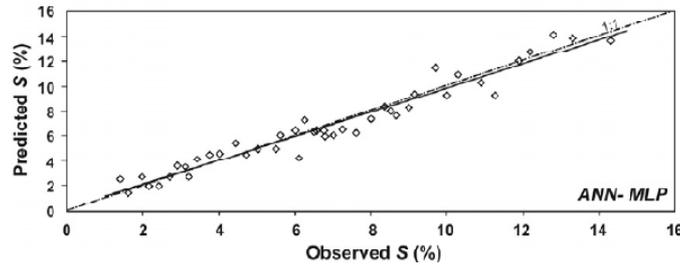


Figure 4. Cross correlation of predicted and observed values of S% for MLP

### 3.2.2 Radial Basis Function

RBF based on supervised learning (Cohen et al., 2002; Looney, 2002) emerged as a variant ANN in late 1980s. Their roots are entrenched in much older Pattern Recognition techniques such as potential functions, clustering, functional approximation, spline interpolation etc. They act as good alternative to MLP. RBF networks model nonlinear data effectively and can be trained in one stage rather than using an iterative process (Pao, 1994). RBF is similar in structure to MLP. It has a hidden layer which contains nodes called RBF units. Each RBF has two key parameters that describe location of function's center and its deviation or width. The hidden unit measures the distance between an input data vector and center of its RBF. RBF has its peak when distance between its center and input data vector is zero and declines gradually as this distance increases. There is a single hidden layer in RBF network and there are only two sets of weights, one connecting hidden layer to input layer and other connecting hidden layer to output layer. Those weights connecting to input layer contain parameters of basis functions. The weights connecting hidden layer to output layer are used to form linear combinations of activations of basis functions or hidden units to generate network outputs. Since hidden units are nonlinear, the outputs of hidden layer may be combined linearly and so processing increases.

RBF used here is a variant of radial basis functional link net, functional link net or radial basis function ANN (Cohen et al., 2002; Jang, 1995; Kenneth, 2001; Looney, 2002). It is more general than RBF because it consists of both non-linear and linear links. RBF shown in Figure 3 has input layer of  $N$  nodes, hidden layer of  $M$  units and output layer of  $P$  units. The  $M$  units represent RBF network architecture and are connected to output units by weight vector  $\tilde{w}_{mp}$  which is modeled using the following gaussian fuzzy membership function (Zadeh, 1994; Zimmermann, 2001):

$$\mu_{\sigma,c}(x) = \exp \frac{-(x-c)^2}{2\sigma^2} \quad (10)$$

Here  $x$  and  $\mu$  are the input and the center of RBF unit respectively;  $\sigma$  is the spread of gaussian basis function. Total output of one of the output units of RBF network is given by the expression:

$$\tilde{z}_p^{(q)} = \frac{1}{M+N} \left[ \sum_{m=1}^M \tilde{w}_{mp} \exp \left\{ \frac{\|x^{(q)} - v^{(m)}\|^2}{2\sigma_m^2} \right\} + \sum_{n=1}^N \tilde{w}_{np} x^{(q)} \right] \quad (11)$$

Here,  $\{x^{(q)} : q=1, \dots, Q\}$  are training vector;  $\{v^{(m)} : m=1, \dots, M\}$  is centre of a specific RBF;  $\sigma_m$  is spread parameter and  $\tilde{w}_{np}$  is weight vector from input to output layer. The gaussian function centers,  $v^{(m)}$  are initialized using fuzzy  $c$ -means clustering algorithm (Zadeh, 1994) and standard deviation widths  $\sigma_m$  are initialized using following equations:

$$\sigma_m = \sigma = \frac{D_{\max}}{\sqrt{2M}}, m=1, \dots, M; \quad (12)$$

$$D_{\max} = \max_{i,k=1, \dots, M} \|v^{(i)} - v^{(k)}\|$$

The weights are updated by steepest descent through fuzzy membership functions as given by following

equations (Zadeh, 1994; Zimmermann, 2001):

$$\tilde{w}_{mp} = \tilde{w}_{mp} + \frac{\eta_1}{M + N} \left[ \sum_{q=1}^Q (\tilde{t}_j^{(q)} - \tilde{z}_t^{(q)}) \exp \left\{ \frac{\|x^{(q)} - v^{(m)}\|^2}{2\sigma_m^2} \right\} \right] \tag{13}$$

$$\tilde{w}_{mp} = \tilde{w}_{mp} + \frac{\eta_2}{M + N} \left[ \sum_{q=1}^Q (\tilde{t}_j^{(q)} - \tilde{z}_t^{(q)}) x_n^{(q)} \right] \tag{14}$$

In the above expressions  $\eta_1$  and  $\eta_2$  are the learning rates. As evident from Table 4 and Figure 5 of cross correlation between predicted and observed values RBF model is highly acceptable for prediction of S %.

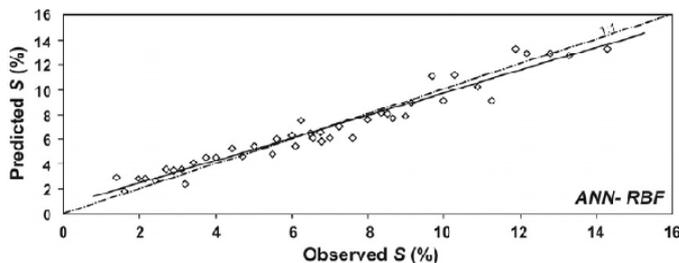


Figure 5. Cross correlation of predicted and observed values of S% for RBF

### 3.3 Adaptive Neuro Fuzzy Inference System

Roger Jang (1993) suggested Adaptive Neuro Fuzzy Inference System (ANFIS) in which both the learning capabilities of an ANN and reasoning capabilities of Fuzzy Logic were combined in order to give enhanced prediction capabilities. The objective of ANFIS is to find an exemplar that will correctly map input values with target values. ANFIS can serve as a basis for constructing a set of fuzzy *if-then* rules with appropriate membership functions to generate stipulated input output pairs. Here, the membership functions are so tuned to input output data such that excellent results are obtained. ANFIS thus takes an initial Fuzzy Inference System (FIS) and tunes it with Back Propagation algorithm based on collection of input output data. FIS thus acts as a knowledge representation system where each fuzzy rule describes local behavior of the system. The basic structure of FIS consists of three conceptual components, viz. (i) rule base containing a selection of fuzzy rules; (ii) database defining membership functions used in fuzzy rules and (iii) reasoning mechanism which performs inference procedure upon the rules and given facts to derive a reasonable output. The fuzzy component of ANFIS takes care of inherent vagueness and impreciseness present in real life data. The bell membership function is used in ANFIS (Zadeh, 1994; Zimmermann, 2001):

$$\mu_{a,b,c}(x) = \frac{1}{1 + \left| \frac{x-a}{c} \right|^{2b}} \tag{15}$$

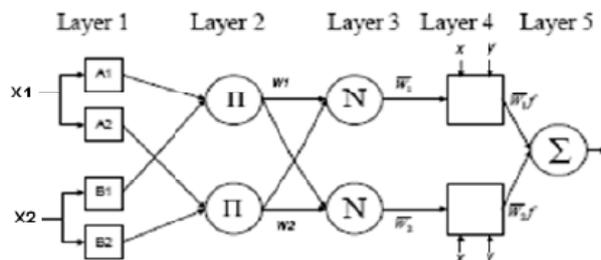


Figure 6. Five layered ANFIS architecture

Parameter  $b$  is usually positive. The desired bell membership function is obtained by a proper selection of parameter set  $\{a, b, c\}$ . During the learning phase these parameters are changing continuously in order to minimize error function between target output values and the calculated ones (Lee, 1990). ANFIS model used here is a multilayer ANN based Fuzzy system. Its topology is represented in Figure 6 and the system has a total of five layers, viz. fuzzification layer, product layer, normalized layer, defuzzification layer and total output layer.

In the directed graph structure, input and output nodes represent training values and predicted values respectively. In hidden layers, there are nodes which take care of membership functions and corresponding rules. This architecture eliminates the disadvantage involved in normal feed forward multilayer network where it is difficult for an observer to understand or modify the network. We assume that FIS has two inputs  $x$  and  $y$  and one output. For a first order Sugeno Fuzzy model, a common rule set with two fuzzy if-then rules is defined as follows:

$$\text{Rule 1: If } (x \text{ is } A_1) \text{ and } (y \text{ is } B_1) \text{ then } t_1 = p_1x + q_1y + r_1$$

$$\text{Rule 2: If } (x \text{ is } A_2) \text{ and } (y \text{ is } B_2) \text{ then } t_2 = p_2x + q_2y + r_2$$

Here,  $p_1, p_2, q_1, q_2, r_1, r_2$  are linear parameters and  $A_1, A_2, B_1, B_2$  are non-linear parameters. In ANFIS, layer 1 is the fuzzification layer in which  $x$  and  $y$  are input of nodes  $A_1, B_1$  and  $A_2, B_2$  respectively.  $A_1, A_2, B_1, B_2$  are linguistic labels used in Fuzzy Set Theory for dividing the membership functions. The membership relationship between the output and input functions of this layer is expressed as follows:

$$\text{Output}_{1,i} = \mu_{A_i}(x); i = 1,2 \quad \text{Output}_{1,j} = \mu_{B_j}(y); j = 1,2$$

Here  $\text{Output}_{1,i}$  and  $\text{Output}_{1,j}$  denote the output functions;  $\mu_{A_i}$  and  $\mu_{B_j}$  denote the membership functions. Layer 2 is the product layer consisting of two nodes labeled as  $\Pi$ . The output  $W_1$  and  $W_2$  are weight functions of next layer. The output of this layer is product of input signal, which is defined as follows:

$$\text{Output}_{2,i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y); i = 1,2$$

Here  $\text{Output}_{2,i}$  denotes the output of layer 2. The 3<sup>rd</sup> layer is normalized layer whose nodes are labeled  $N$ . Its function is to normalize the weight function in the process of following conditions, where  $\text{Output}_{3,i}$  denotes the output of layer 3.

$$\text{Output}_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}; i = 1,2$$

Layer 4 is the defuzzification layer whose nodes are adaptive. The output is  $\bar{w}_i(p_i x + q_i y + r_i)$  where linear parameters  $p_i, q_i, r_i$  are also called consequent parameters of the node. The defuzzification relationship between input and output of this layer is defined as follows:

$$\text{Output}_{4,i} = \bar{w}_i t_i = \bar{w}_i(p_i x + q_i y + r_i); i = 1,2$$

Here  $\text{Output}_{4,i}$  denote the output of layer 4. The 5<sup>th</sup> layer is total output layer whose node is labeled as  $\Sigma$ . The output of this layer is total of input signals which represent the results of Financial Stress percent and is represented as follows:

$$\text{Output}_{5,i} = \sum \bar{w}_i t_i = \frac{\sum_i w_i t_i}{\sum_i w_i}; i = 1,2$$

Here  $\text{Output}_{5,i}$  denote the output of layer 5.

Here ANFIS uses a hybrid learning algorithm which is a combination of gradient descent and least squares method. In forward pass of hybrid learning algorithm node outputs go forward until layer 4 and consequent parameters are identified by least squares method (Jang, 1993). In backward pass error signals propagate backwards and premise parameters are updated by gradient descent. The consequent parameters are optimized under the condition that the premise parameters are fixed. The major advantage of hybrid approach is that it converges much faster since it reduces search space dimensions of original back propagation methods. The overall output is expressed as linear combination of consequent parameters. The error measure to train the ANFIS is defined as follows (Jang, 1993):

$$E = \sum_{i=1}^n (t_i - t'_i)^2 \quad (16)$$

Here  $t_i$  and  $t'_i$  are the  $i^{\text{th}}$  desired and estimated output respectively and  $n$  is total number of input output pairs of data in the training set. In this work ANFIS is trained with the help of MATLAB. The RMSE and Statistical calculations were performed using Excel. The different parameter and their values used in ANFIS are presented in Table 5. According to RMSE, VAF, MAPE and  $R^2$  values from Table 4 and cross correlation between predicted and observed values from Figure 7, ANFIS constructed to predict S % has a high performance of prediction.

Table 5. Different parameter values for training ANFIS

ANFIS Parameters	Values
Membership Function	Gaussian Function
Number of Membership Functions	9
Output Function	Linear
Number of Linear parameters	14
Number of Non-linear parameters	18
Total Number of parameters	32
Number of Training data pairs	132
Number of Verifying data pairs	46
Number of Validating data pairs	46

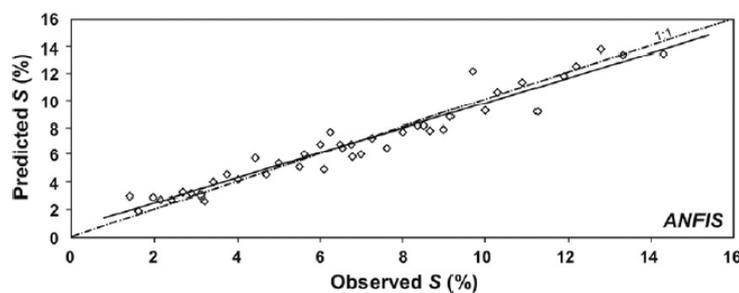


Figure 7. Cross correlation of predicted and observed values of S% for ANFIS

#### 4. Experimental Results

This section illustrates the results obtained towards prediction of financial stress percent of financial time series data using Multiple Regression, MLP, RBF and ANFIS models. A comparative analysis among different stated methods is also highlighted. According to the results of simple regression analyses, there are statistically meaningful relationships between Financial Stress percent with Credit Measures, Asset Prices, Macroeconomic Variables and Foreign Variables. The models of Multiple Regression, MLP, RBF and ANFIS for prediction of Financial Stress percent were then constructed using three inputs and one output. Based on experiments performed we have the following results:

- The result obtained from prediction of Financial Stress percent showed that Multiple Regression has high prediction performance.
- ANFIS for prediction of Financial Stress percent revealed a more reliable prediction when compared with Multiple Regression.
- In order to predict Financial Stress percent RBF having three inputs and one output was applied successfully and exhibited more reliable predictions than Multiple Regression and ANFIS.

As a result of the comparison of VAF, RMSE and MAPE indices and  $R^2$  for predicting Financial Stress percent, it was obtained that prediction performance of RBF is higher than those of MLP, ANFIS and Multiple Regression. In order to show deviations from observed values of Financial Stress percent, the distances of predicted values using Multiple Regression, MLP, RBF and ANFIS models constructed from observed values were also calculated and represented schematically in Figure 8. The Figure 8 indicates that deviation interval (-1.239 to +1.302) of predicted values from RBF is smaller than deviation interval of MLP (-1.536 to +2.116), ANFIS (-2.136 to +2.106) and Multiple Regression (-2.721 to +1.752).

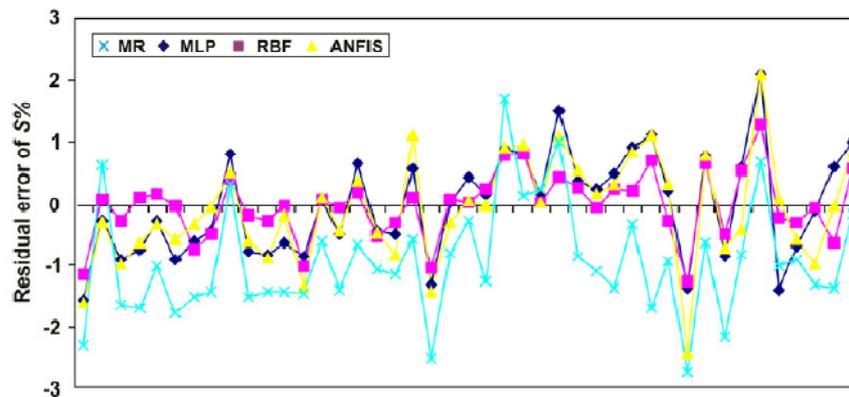


Figure 8. Variation of values predicted by Multiple Regression, MLP, RBF and ANFIS from observed values

## 5. Conclusions

In this work, we have used multiple regression, MLP, RBF and ANFIS models to predict Financial Stress percent of financial time series data. The accuracies of ANN and ANFIS are evaluated to be relatively similar. It is evident from above discussion that RBF and MLP exhibit a higher performance than ANFIS and Multiple Regression for predicting Financial Stress percent. The comparison performance shows that the Machine Learning and Soft Computing paradigms are good tools for minimizing uncertainties in financial time series data. The usage of Machine Learning and Soft Computing will also provide new approaches and methodologies to minimize the potential inconsistency of correlations. The potential benefits of Machine Learning and Soft Computing extend beyond high computation rates. Higher performance of Machine Learning and Soft Computing were sourced from greater degree of robustness and fault tolerance than traditional Statistical models because there are many more processing neurons each with primarily local connections. However, comparison of RBF and MLP indicates good predictive capabilities of RBF model. Their accuracies are almost the same. It was found that the time taken by RBF is less than that of MLP in this research. But limitation of RBF is that it is more sensitive to dimensionality and has greater difficulty if number of units is large. The population of analyzed data is relatively limited in this research. Therefore, the practical outcome of models could be used with acceptable accuracy at the preliminary stage of design.

## References

- Altman, E. L. (1993). *Corporate Financial Distress and Bankruptcy*. John Wiley and Sons.
- Altun, H., & Gelen, G. (2004). *Enhancing performance of MLP/RBF Neural Classifiers via Multivariate data distribution scheme* (pp. 1-6). Proceedings of International Conference on Computational Intelligence, Nicosia, North Cyprus.
- Armstrong, J. S. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Boston: Kluwer Academic.
- Benardos, P. G., & Vosniakos, G. C. (2007). Optimizing Feed-forward Artificial Neural Network Architecture. *Engineering Applications of Artificial Intelligence*, 20(3), 365-382. <http://dx.doi.org/10.1016/j.engappai.2006.06.005>
- Box, G. (2008). *Time Series Analysis: Forecasting and Control* (3rd ed.). Dorling Kindersley India.
- Bordo, M., Eichengreen, B., Klingebiel, D., & Martinez-Peria, M. S. (2001). Is the Crisis Problem Growing More Severe? *Economic Policy*, 16(32), 51-82. <http://dx.doi.org/10.1111/1468-0327.00070>
- Brockwell, P. J., & Davis, R. A. (2009). *Time Series: Theory and Methods* (2nd ed.). Springer Series in Statistics, Springer Verlag.
- Chatfield, C. (2000). *Time Series Forecasting*. Chapman and Hall, Boca Raton, Florida. <http://dx.doi.org/10.1201/9781420036206>
- Chaudhuri, A., & De, K. (2009). Time Series Forecasting using Hybrid Neuro-Fuzzy Regression Model. In H. Sakai et al. (Editors), *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*, Lecture Notes in Artificial Intelligence, Springer Verlag, Berlin, LNAI 5908, pp. 369-381.
- Chen, S. M. (2002). Forecasting Enrollments based on Higher Order Fuzzy Time Series. *Cybernetics and*

- Systems*, 33(1), 1-16. <http://dx.doi.org/10.1080/019697202753306479>
- Cohen, S., & Intrator, N. (2003). A study of ensemble of Hybrid Networks with strong Regularization (pp. 227-235). Proceedings of Multiple Classifier Systems.
- Cohen, S., & Intrator, N. (2002). Automatic model selection in a Hybrid Perceptron/Radial Network. *Information Fusion: Special Issue on Multiple Experts*, 3(4), 259-266.
- Dash, P. K., Behera, H. S., & Lee, I. W. C. (2008). Time Sequence Data Mining using Time – Frequency Analysis and Soft Computing techniques. *Applied Soft Computing*, 8(1), 202-215. <http://dx.doi.org/10.1016/j.asoc.2007.01.001>
- Fuller, W. A., & Stephens, V. W. (1999). Introduction to Statistical Time Series (3rd ed.). New York: John Wiley and Sons.
- Illing, M., & Liu, Y. (2006). Measuring Financial Stress in a Developed Country: An Application to Canada. *Journal of Financial Stability*, 2(3), 243-265. <http://dx.doi.org/10.1016/j.jfs.2006.06.002>
- Jang, J. R. (1993). ANFIS: Adaptive Network Based Fuzzy Inference System. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), 665-685. <http://dx.doi.org/10.1109/21.256541>
- Jang, J. S. R., & Chuen, T. S. (1995). Neuro Fuzzy Modeling and Control. *Proceedings of IEEE*, 83(3), 378-406. <http://dx.doi.org/10.1109/5.364486>
- Kenneth, J., Wernter, S., & MacInyre, J. (2001). Knowledge extraction from Radial Basis Function Networks and Multi Layer Perceptions. *International Journal of Computational Intelligence and Applications*, 1(3), 369-382.
- Kindleberger, C. P., & Aliber, R. (2005). Manias, Panics and Crashes: A History of Financial Crises (5th ed.). John Wiley and Sons. <http://dx.doi.org/10.1057/9780230628045>
- Kosko, B. (2008). *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall of India.
- Lee, C. C. (1990a). Fuzzy Logic in Control Systems: Fuzzy Logic Controller – Part I. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2), 404-418. <http://dx.doi.org/10.1109/21.52551>
- Lee, C. C. (1990b). Fuzzy Logic in Control Systems: Fuzzy Logic Controller – Part II. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2), 419-435. <http://dx.doi.org/10.1109/21.52552>
- Loh, W., & Tim, L. (2000). A Comparison of Prediction Accuracy, Complexity and Training Time of Thirty Three old and new Classification Algorithm. *Machine Learning*, 40(3), 203-238. <http://dx.doi.org/10.1023/A:1007608224229>
- Looney, C. G. (2002). Radial Basis Functional Link Nets and Fuzzy Reasoning. *Neurocomputing*, 48(1-4), 489-509. [http://dx.doi.org/10.1016/S0925-2312\(01\)00613-0](http://dx.doi.org/10.1016/S0925-2312(01)00613-0)
- Lund, R., Brockwell, P. J., & Davis, R. A. (2002). Introduction to Time Series and Forecasting (2nd ed.). New York: Springer Verlag.
- Pao, Y. H., Park, G. H., & Sobajic, D. J. (1994). Learning and generalization characteristics of the random vector Functional Link Net. *Neurocomputing*, 6, 163-180. [http://dx.doi.org/10.1016/0925-2312\(94\)90053-1](http://dx.doi.org/10.1016/0925-2312(94)90053-1)
- Simpson, P. K. (1990). *Artificial Neural System – Foundation, Paradigm, Application and Implementation*. New York: Pergamon Press.
- Zadeh, L. A. (1994). Fuzzy Logic, Neural Networks and Soft Computing. *Communications of the ACM*, 37(3), 77-84. <http://dx.doi.org/10.1145/175247.175255>
- Zhang, P., & Min, Q. G. (2005). Neural Network Forecasting for Seasonal and trend Time Series. *European Journal of Operations Research*, 160(2), 501-514. <http://dx.doi.org/10.1016/j.ejor.2003.08.037>
- Zimmermann, H. J. (2001). Fuzzy Set Theory and its Applications (4th ed.). Boston: Kluwer Academic Publishers. <http://dx.doi.org/10.1007/978-94-010-0646-0>