

Design and analysis of experiments in ANFIS modeling for stock price predictionMeysam Alizadeh^{a*}, Mohsen Gharakhani^b, Elnaz Fotoohi^c and Roy Rada^d^a*School of Information Sciences, University of Pittsburgh, PA, USA*^b*Department of Industrial Engineering, Iran University of Science & Technology, Tehran, Iran*^c*Department of Industrial Engineering, Amirkabir University of Technology, Tehran, Iran*^d*Department of Information Systems, University of Maryland Baltimore County, MD, USA***ARTICLE INFO***Article history:*

Received 10 September 2010

Received in revised form

4 January 2011

Accepted 6 January 2011

Available online

6 January 2011

Keywords:

ANFIS

Neuro-fuzzy systems

Design of experiment

Stock price prediction

ABSTRACT

At the computational point of view, a fuzzy system has a layered structure, similar to an artificial neural network (ANN) of the radial basis function type. ANN learning algorithms can be employed for optimization of parameters in a fuzzy system. This neuro-fuzzy modeling approach has preference to explain solutions over completely black-box models, such as ANN. In this paper, we implement the design of experiment (DOE) technique to identify the significant parameters in the design of adaptive neuro-fuzzy inference systems (ANFIS) for stock price prediction.

© 2011 Growing Science Ltd. All rights reserved

1. Introduction

Fuzzy systems and neural networks (NN) are considered as two most widely used techniques in intelligent systems. Automatic control, pattern recognition, human-machine interaction, expert systems, modeling, medical diagnosis, economics, etc. are some of these systems' application areas (Echanobe et al., 2008). Obviously, each technique has its own advantages and drawbacks. Fuzzy systems have the ability to represent comprehensive linguistic knowledge and perform reasoning through fuzzy rules. However, fuzzy systems do not provide a mechanism to tune those rules. On the other hand, NN are adaptive systems that can be trained and tuned from a set of input-output data set. Nevertheless, it is very difficult to understand and represent the obtained knowledge.

In general, hybrid systems focus on combination of the advantages of different paradigms in order to overcome their shortcomings. Among these systems, neuro-fuzzy systems are characterized by the combination of neural networks with techniques from fuzzy sets and systems. Neuro-fuzzy systems exhibit the noise robustness and learning capabilities of neural networks together with the ability of fuzzy systems to explicitly model uncertainty, linguistic concepts, and the knowledge of human experts. These systems can combine both fuzzy and neuro paradigms in two different ways (Tsoukalas & Uhrig, 1997): 1) by introducing the fuzzification into the NN structure (i.e., fuzzy NN)

* Corresponding author. Tel: +1 443 255 8736
E-mail: mea62@pitt.edu (M. Alizadeh)

and 2) by providing the fuzzy systems with learning ability by means of neural-network algorithms (i.e., NN-driven fuzzy reasoning techniques). In the first case, fuzzification can be introduced in any of the network aspects: neuron inputs, output, weights, aggregation operations, transfer functions, etc. In the second case, NN methods are used both with the aim of identifying rules and membership functions and for tuning the system (Echanobe et al., 2008).

By combining information from different sources, such as empirical models, heuristics, and data, neuro-fuzzy modeling has been recognized as a powerful tool which can facilitate the effective development of models (Babuška and Verbruggen, 2003). Neuro-fuzzy models can describe systems by means of fuzzy *if-then* rules represented in a network structure, in which learning algorithms known from the area of artificial neural networks (i.e. gradient descent, Levenberg-Marquardt, etc.) can be applied. Interestingly, neuro-fuzzy models tend to gravitate toward meeting a high accuracy requirement that happens at a substantial expense of lowering their transparency. This is somewhat inevitable considering the underlying black-box processing paradigm and various topologies existing in neuro-computing (Pedrycz & Reformat, 2003). Considering that many of the parameters in fact are fuzzy variables and that most often these systems operate in real time, the problem of neuro-fuzzy systems topology and parametric configuration becomes worse (Zanchettin et al., 2005). Determining the ANFIS parameters such as the number and the shape of input membership functions (MFs), the initial rule based construction approach, the number of data points; the clustering algorithms become a difficult designing task. Even in automatically rule based constructive models, the performance of the system still depends on the careful selection of sensitivity threshold, error threshold, and learning rates (Zanchettin et al., 2005). Mostly, the tuning and configuration of ANFIS are performed experimentally. These decisions are usually made in terms of the most popular and prevalent parameters, operators, and algorithms accomplished.

In this way, it is quite worthwhile to determine which variables have the most relevance and the greatest influence on the behavior and the performance of the ANFIS. As a result, the system designer could pay more attention in selection of the parameters which are statistically significant. Zanchettin et al. (2005) performed design of experiment (DOE) technique to verify the interactions and interrelations among parameters in the design of ANFIS and evolving fuzzy neural networks (EFuNN). They considered six factors for each of ANFIS and EFuNN and accomplished their tests to predict points of the time series which are the result of the Mackey-Glass equation integration. The aim of this paper is to extend their work by considering more factors in experimental design and to test their influence on the behavior ANFIS. In this regard, while they considered six factors each in two levels, we make use of nine factors mostly in three levels in experiments. There are three common factors with Zanchettin et al. (2005): input MF number, output MF shape, and training epochs. Besides, we consider six additional important factors which have been tuned for ANFIS modeling. Since stock price prediction is a popular problem and soft computing methods are widely used in order to solve this problem, we performed our experiments in input-output data of an automotive part-making manufacturer company for an Asian stock market.

The rest of the paper is organized as follows: Section 2 presents an overview of the ANFIS and DOE. Section 3 systematically discusses the problem description and experimental design of our problem. Section 4 presents the results of the statistical experiment. Finally, some conclusions and future directions are appeared in section 5.

1. Background

1.1. Adaptive Neuro-Fuzzy Inference System

Adaptive neuro-fuzzy inference systems (ANFIS) represent a neural network approach to the design of fuzzy inference systems (Jang, 1993). Since its introduction, ANFIS networks have been widely considered in the technical literature and successfully applied to classification tasks, rule-based expert systems, prediction of time series, and so on. There are also some revisions and different versions of

ANFIS (Panella & Gallo, 2005; Buragohain & Mahanta, 2008; Echanobe et al., 2008; Riverol & Di Sanctis, 2009). ANFIS is a fuzzy inference system that can be trained to model the collection of input-output data. This network makes use of a supervised learning algorithm to determine a nonlinear relationship among inputs and output. According to Kosko (1994), an ANFIS network is particularly suited to solve function approximation problems in several engineering fields.

There are two approaches used by ANFIS: Artificial neural network and fuzzy modeling. Suitable reasoning in quality and quantity might be achieved through composing these two approaches (Teshnehlab et al., 2008). In ANFIS, fuzzy logic is used to determine decision surfaces rather than the uncertainty associated with particular linguistic terms. Also the rule-based representation of neuro-fuzzy systems offers transparency. For pedagogical purposes, one can imagine a fuzzy inference system with two inputs x and y and one output z . Fig. 1 shows the equivalent ANFIS architecture (Type-3 ANFIS). The node functions in the same layer are from the same function family. The first layer implements a fuzzification, the second layer executes the T-norm of the antecedent part of the fuzzy rules, the third layer normalizes the membership functions, the fourth layer calculates the consequent parameters, and finally the last layer computes the overall output as the summation of all incoming signals (Jang, 1993). The feed forward equations of this ANFIS are as follows:

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), \quad (1)$$

$$\bar{w}_i = \frac{w_i}{w_1 + w_2}, i = 1, 2. \quad (2)$$

$$\begin{cases} f_1 = p_1x + q_1y + r_1z \\ f_2 = p_2x + q_2y + r_2z \end{cases} \quad (3)$$

$$f = \frac{w_1f_1 + w_2f_2}{w_1 + w_2} = \bar{w}_1f_1 + \bar{w}_2f_2, \quad (4)$$

where,

- x is the input to node i ,
- $\mu_{A_i}(x)$ is the membership function of A_i ,
- A_i is the linguistic label associated with node functions,
- \bar{w}_i is the firing strength of the i th rule,
- \bar{w}_i is the ratio of the i th rule's firing strength to the sum of all rules' firing strength,
- $\{p_i, q_i, r_i\}$ is the parameter set, and
- f_i is the consequent value.

Note that the network's output y is nonlinear in terms of w and the training of this ANN is thus a *nonlinear optimization problem* (Babuška & Verbruggen, 2003).

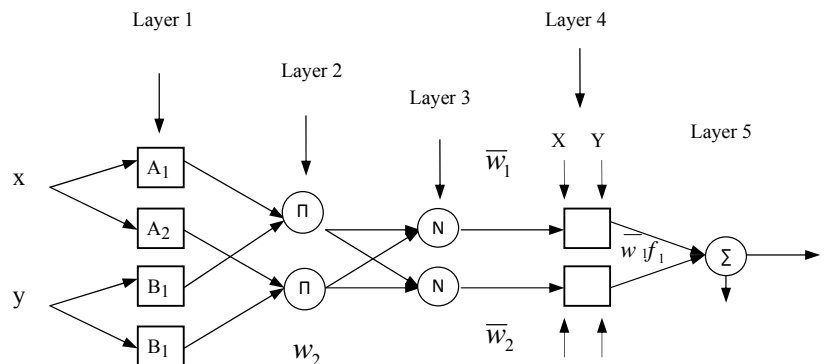


Fig. 1. The equivalent ANFIS structure

The neuro-fuzzy inference system is optimized by adapting the antecedent parameters and consequent parameters so that a specified objective function (usually a difference between the model output and the actual output) is minimized. A number of methods have been proposed for learning rules. Jang (1993) proposed different methods to update the ANFIS parameters involving gradient descent and least square error (LSE) and high complexity is one of these methods' features. Mascioli et al. (1997) proposed merging of min-max and ANFIS models to determine the optimal set of fuzzy rules. Jang and Mizutani (1996) presented an application of the Lavenberg-Marquardt method, which is essentially a nonlinear least-squares technique, for learning in an ANFIS network. Chen (1999) compared several popular training algorithms for tuning parameters of ANFIS membership functions. Tang et al. (2005) proposed a hybrid system combining a fuzzy inference system and genetic algorithms to tune the parameters in the TSK fuzzy ANN. Shoorehdeli et al. (2009) proposed a novel hybrid learning algorithm with stable learning laws for ANFIS as a system identifier and studied the stability of this algorithm. Their hybrid learning algorithm is based on particle swarm optimization for training the antecedent part and forgetting factor recursive least square for training the conclusion part.

1.2. Design of Experiments

Design of experiments (DOE) or experimental design is a method for tuning the input parameters. DOE has very broad applications across all the natural, social, and engineering sciences. Assuming that true experiments are our basic concern, there are three important phases which could be used to make a meaningful study (Hicks, 1999): 1) the experimental or planning phase, 2) the design phase, and 3) the analysis phase. The main steps of each of these three phases are presented in Table 1. The analysis of experiments usually consists of the well-known analysis of variance (ANOVA) test. It is used to perform a systematic decomposition of the variability in the observed response values and to assign portions of the variability either to the effect of an independent variable or to experimental error. The analysis provides information regarding how much each factor and factors interaction contribute to the total variance of the data (Zanchettin et al., 2005).

Table 1
Phases in design of experiments

| Phase | Steps |
|------------|---|
| Experiment | <ol style="list-style-type: none"> 1. Statement of problem 2. Choice of response or dependent variable 3. Selection of factors to be varied 4. Choice of levels of these factors <ol style="list-style-type: none"> a) Quantitative or qualitative b) Fixed or random 5. How factor levels are to be combined |
| Design | <ol style="list-style-type: none"> 1. Number of observations to be taken 2. Order of experimentation 3. Method of randomization to be used 4. Mathematical model to describe the experiment 5. Hypotheses to be tested |
| Analysis | <ol style="list-style-type: none"> 1. Data collection and processing 2. Computation of test statistics 3. Interpretation of results for the experimenter |

2. Problem description and experimental design

2.1 Problem description

For experimental study, the ANFIS is used to predict points of the price time series which are historical data of the stock price of an automotive part-making manufacturing share from an Asian stock exchange. The candidate variables of the system in the form of input and output are shown in Table 2. The rational combination of 27 technical and fundamental variables is used for prediction. In this table, closing price is the output variable. Prediction in stock market is a hot topic for research.

The modern school's view in stock market is dynamic systems and chaotic behavior of stock prices. From this standpoint, stock price movements have very complex and nonlinear correlations with some variables, which require advanced mathematical modeling. One of the challenges of modern capital market analysis is to develop theories that are capable of explaining these complex movements in asset prices and returns. The study of stock market has led financial economists to apply statistical techniques from chaos theory for analyzing stock market data. Based on these new techniques, recent empirical studies document nonlinearities in stock market data. The neural network model is also appropriate for capturing all the nonlinear dynamic relationships in stock market.

Table 2
Candidate inputs and output variables of the system

| Variable name | Variable description | variable |
|--|--|--|
| Date | It is the day of trading | - |
| Volume | It shows how transaction are made in trading | - |
| Value | It shows the values of transaction | - |
| Number of trading | It shows the number of transactions | - |
| Open price | It shows the opening price | - |
| Maximum price | It shows the maximum price of stock in trading day | - |
| Minimum price | It shows the minimum price of stock in trading day | - |
| Typical price (TypP) | It is the average of maximum, minimum and closing price | Average {min price, max price, close price} |
| Price change (Pchange) | It is the difference between today's closing and the last day price | $Change_i = day_i \text{ Price} - day_{i-1} \text{ Price}$ |
| Number of buyers | It shows the number of buyers | - |
| Moving average convergence divergence (MACD) | It uses two different exponential smoothing moving average lines and helps us determine the price trends | MACD = exponential smoothing average 12 days (weeks) - exponential smoothing average 26 days (weeks) |
| Moving average MACD (MA-MACD) | It uses exponential smoothing moving average line in contrast to MACD | MA-MACD = exponential smoothing average 9 days (weeks) |
| Positive directional movement index (DI+) | It shows power of up moving trend | If $DI+ > DI-$ then you should buy else you should sell |
| Negative directional movement index (DI-) | It shows power of down moving trend | |
| Moving average (MA) | It is the sum of single period closing prices dividend on numbers of periods | $MA = \frac{\sum_{i=1}^n \text{close prices of } i\text{th period}}{n}$ |
| Price channel (top) | It is the maximum of price in the last four week periods | Max {close price of last four week} |
| Price channel (bottom) | It is the minimum of price in the last four week periods | Min {close price of last four week} |
| Range (R) | It is the range of price in one specific day | Range = Maximum Price - Minimum Price |
| Price per earning per share (P/E) | It shows the payback period or its inversion shows the stock rate of return | |
| Price and volume trend (PVT) | It relates price and volume in the stock market | $PVT = PVT_{prev} + volume \times \frac{\text{close}_{tod} - \text{close}_{prev}}{\text{close}_{prev}}$ |
| Stochastic oscillator | It compares a stock's closing price to its price range over a given period of time | Stochastic Oscillator = $100 \times \frac{\text{close} - \text{min}}{\text{max} - \text{min}}$ |
| Accumulation/Distribution index (AccDist) | adds or subtracts each day's volume in proportion to where the closing price is between the day's high and low | $AccDist = AccDist_{prev} + volume * CLV$ $CLV = \frac{(\text{close} - \text{min}) - (\text{max} - \text{close})}{\text{max} - \text{min}}$ |
| Momentum (Mom) | It shows the difference between today's closing price and the closing price of N previous days | Momentum = $\text{close}_{today} - \text{close}_{N \text{ previous days}}$ |
| Rate of change (ROC) | It shows the difference between today's closing price and the closing price of N previous days as a fraction | $ROC = \frac{\text{close}_{today} - \text{close}_{N \text{ days ago}}}{\text{close}_{N \text{ days ago}}}$ |
| Close price | It is the closing price of the trading day | - |

In the study performed with ANFIS, we run a full factorial experiment with different levels of value for each of the factors. We use the six important factors reported by Zanchettin et al. (2005). More specially, for initial rule base construction, the Sugeno and Yasukawa (1993) approach, subtractive

clustering approach, and Emami et al. (1999) fuzzy modeling algorithm are used. For testing the effect of different cluster validity index (CVI), three CVI proposed by Fukuyama and Sugeno (1989), Kwon (1998), and Fazel Zarandi et al. (2009) are considered.

2.2 Experimental design

Three different numbers of data points are used to test the effects of various factors. For each case, half of the data points are used for training phase of the ANFIS and the rest of them are used for testing. To show whether initial clustering algorithm is significant or not, we consider random initial clustering and agglomerative hierarchical clustering (AHC) algorithm. Besides, three clustering algorithms include fuzzy c-means (FCM), Gustafson-Kessel (GK), and fuzzy noise rejection data partitioning algorithm proposed by Melek et al. (2005) are tested. Table 3 demonstrates the full list of controlled factors in ANFIS experiment configuration. The description of the above models, algorithms, and techniques are as follows:

Table 3
ANFIS experiment configuration

| No. | Factor | Levels | | |
|-----|--------------------------------------|-----------------|----------------|----------------------|
| 1 | Input MF number | 15 | 20 | 27 |
| 2 | Output MF shape | linear | constant | - |
| 3 | Initial rule base construction model | Sugeno-Yasukawa | Sub-clustering | Emami et al. |
| 4 | Training Epochs | 10 | 30 | 50 |
| 5 | Degree of fuzziness | 2 | 2.5 | 3 |
| 6 | Cluster Validity Index | Fukuyama-Sugeno | Kwon | Fazel-Neshat-Turksen |
| 7 | Number of data points | 400 | 700 | 1000 |
| 8 | Initial clustering algorithm | Random | AHC | - |
| 9 | Clustering Algorithm | FCM | GK | Melek et al. |

2.2.1 Initial rule based construction approaches

For testing the effects of the initial rule based construction subtractive clustering (Sugeno & Yasukawa, 1993; Chiu, 1994) and the fuzzy modeling algorithms proposed by Emami et al. (1999) are used. Sugeno and Yasukawa (1993) proposed an algorithm to extract the fuzzy *if-then* rules from the historical data with four main steps. The first three steps are part of the *structure identification* stage and the final step is the *fuzzy reasoning* stage. The first step is fuzzy clustering of the output variable. The second one is to determine the most relevant input variables with a myopic neighborhood search algorithm. The third step is construction the antecedent part of the fuzzy rules. This is achieved by projecting the output membership degrees into the already selected significant input variables. Finally, the fourth step is the fuzzy inference as stated earlier. Emami et al. (1999) proposed the same approach with some revisions. For clustering the output space, they used agglomerative hierarchical hard clustering for initial prototypes. Also for the purpose of input membership assignment, first, they performed fuzzy line clustering for input membership functions, and then eliminated ineffective input candidates. If we do not have a clear idea about how many clusters there should be for a given set of data, *Subtractive clustering* (Chiu, 1994), is a fast, one-pass algorithm for estimating the number of clusters and the cluster centers in a set of data. The cluster estimates can be used to initialize iterative optimization-based clustering methods and model identification methods like ANFIS.

2.2.2 Cluster validity indexes

In order to use fuzzy clustering algorithms, initially, we have to choose the optimum number of clusters. For this purpose, three different cluster validity indexes are tested. First, the following criterion proposed by Fukuyama and Sugeno (1989) is implemented:

$$S(c) = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^m \left(\|x_k - v_i\|^2 - \|v_i - \bar{x}\|^2 \right),$$

where,

- n : number of data to be clustered,
- c : number of clusters,
- x_k : k th data, usually a vector,
- \bar{x} : average of data,
- v_i : vector expressing the center of i th cluster,
- $\|\cdot\|$: norm,
- μ_{ik} : grade of k th data belonging to i th cluster, and
- m : degree of fuzziness.

Second, a validity index, which is proposed by Kim et al. (2004) and modified by Fazel Zarandi et al. (2009), is used. In other words, we minimize:

$$V_{FNT}(U, V; X) = \frac{2}{c(c-1)} \sum_{p \neq q}^c S_{rel}(A_p, A_q). \quad (5)$$

The optimal number of the clusters is obtained by minimizing $V_{FNT}(U, V; X)$ over the range of c values: $2, \dots, c_{max}$; where $S_{rel}(A_p, A_q)$ is the relative similarity between two fuzzy sets A_p and A_q and is defined as:

$$S_{rel}(A_p, A_q) = \sum_{i=1}^n S_{rel}(x_j : A_p, A_q) h(x_j), \quad (6)$$

where,

$$h(x_j) = -\sum_{p=1}^c u_{A_p}(x_j) \log(u_{A_p}(x_j)), \quad (7)$$

and $S_{rel}(x_j : A_p, A_q)$ is the relative similarity between two fuzzy sets A_p and A_q at x_j which is defined as,

$$S_{rel}(x_j : A_p, A_q) = \frac{f(x_j : A_p \cap A_q)}{f(x_j : A_p \cap A_q) + f(x_j : A_p - A_q) + f(x_j : A_q - A_p)}. \quad (8)$$

Here, $h(x_j)$ is the entropy of datum x_j and $u_{A_p}(x_j)$ is the membership value where x_j belongs to the cluster A_p . Finally, we use Kwon (1998) validity index as a third CVI test for our experiments. The index is as follow,

$$V_k(U, V; X) = \frac{\sum_{k=1}^n \sum_{i=1}^c u_{ik}^2 \|x_k - v_i\|^2 + \frac{1}{c} \sum_{i=1}^c \|v_i - \bar{v}\|^2}{\min \left(\|v_k - v_i\|^2 \right)}. \quad (9)$$

1.2.1. Initial clustering algorithms

For the choice of the initial clustering, an agglomerative hierarchical clustering algorithm (AHC) and random algorithm are tested. In random algorithm, we just create the initial membership function,

randomly. The AHC algorithm puts each of the n data vectors in an individual cluster. Then, by defining a matrix of dissimilarities $D = [d_{ij}]$, the AHC merges two or more of these clusters, moving to a higher level of data partition. The process is repeated to form a sequence of nested clustering in which the number of clusters decreases gradually until the minimum required number of clusters c is reached (Melek et al., 2005). In specific terms, we calculate the $(c \times N)$ matrix of dissimilarities $D = [d_{ij}]$ as the following Euclidean-based distance,

$$d_{ij} = d(X_i, X_j) = \sqrt{\frac{2n_i n_j}{n_i + n_j}} \|\mathbf{v}_{hi} - \mathbf{v}_{hj}\| \quad (10)$$

where \mathbf{v}_{hi} and \mathbf{v}_{hj} are mean vectors of the hard clusters X_i and X_j , respectively, and n_i (n_j) is the number of data in the hard cluster X_i (X_j).

3. Statistical results

Nine control factors (variables) were considered for ANFIS experimental design, where seven of them have three levels and the rest have two levels, resulting in 8748 combinations. Each one of the levels combinations of control factors was replicated five times, totalizing 43740 analyses. Response variable is set to root mean square error (RMSE). Table 4 gives the ANFIS variance analysis. The ANOVA table contains the sources of variation, degrees of freedom, sum of squares, mean square, F-ratio test statistics, and the corresponding significance levels. Note that the input MF number, output MF shape, initial rule based construction approach, and cluster validity index factors present the greatest statistical relevance, because the higher F-ratio value or the smaller probability means the grater important and relevance of the corresponding factor.

Therefore, based on our results, these factors have the most important effects on the performance of ANFIS when the proposed method is used to the prediction of stock price problem. The initial rule based construction approach, corresponding to $\approx 50\%$ of the system variance, the input MF number, corresponding to $\approx 39\%$ of the system variance, cluster validity index, corresponding to $\approx 4\%$ of the system variance, and output MF shape, corresponding to $\approx 3\%$ of the system variance.

Through the variance analysis of the factorial experiment, 6 factors were seen to be significant at the 5 percent level: input MF number, output MF shape, initial rule base construction approach, degree of fuzziness, training epochs, and cluster validity index. In addition, four interactions among some factors are significant: input MF number and output MF shape, input MF number and initial rule based construction approach, output MF shape and training epochs, initial rule based construction approach and initial clustering algorithm.

In Table 4, the probability for the number of data points variable is 0.057. Hence, despite our expectation, three different controlled data point numbers have no significant influence on the performance of the ANFIS. This is an interesting result because it implies that in our problem, stock price prediction, the number of data points is not a significant source of variance in the prediction performance of ANFIS. Therefore, even few numbers of data points can be fed to ANFIS and the results might be acceptable.

As the probability value for the initial clustering approach controlled variable is greater than 5 percent, this factor is not statistically significant. Also small participation of the clustering algorithm in the performance of ANFIS in this problem can be easily interpreted from its low F-test value, 2.56, in Table 4. Since the interactions are significant, one should be very cautious in interpreting the main effects. A significant interaction here means that, for example, the effect of initial rule based construction approach on ANFIS performance (RMSE) at one degree of fuzziness is different from its effect at the other degree of fuzziness. The same statement holds for other significant interactions.

Table 4
ANOVA table for the prediction of stock price using ANFIS

| Source | df | SS | MS | F | Probability |
|----------------------------------|----|----------|----------|----------|-------------|
| Main Factors | | | | | |
| 1 Input MF number | 2 | 1876.728 | 938.3638 | 138.6667 | 0.0000000 |
| 2 Output MF shape | 1 | 67.85092 | 67.85092 | 10.02667 | 0.0000535 |
| 3 Initial rule base construction | 2 | 2418.091 | 1209.046 | 178.6667 | 0.0000000 |
| 4 Training Epochs | 2 | 51.97092 | 25.98546 | 3.84 | 0.0221153 |
| 5 Degree of fuzziness | 2 | 44.75273 | 22.37637 | 3.306667 | 0.0374221 |
| 6 Cluster Validity Index | 2 | 188.6833 | 94.34165 | 13.94133 | 0.0000013 |
| 7 Number of data points | 2 | 39.03593 | 19.51797 | 2.884267 | 0.0568045 |
| 8 Initial clustering algorithm | 1 | 18.76728 | 18.76728 | 2.773333 | 0.0633919 |
| 9 Clustering algorithm | 2 | 34.64728 | 17.32364 | 2.56 | 0.0782940 |
| Significant Interactions | | | | | |
| 1×2 interaction | 2 | 69.29456 | 34.64728 | 5.12 | 0.0062857 |
| 1×3 interaction | 4 | 128.1949 | 32.04873 | 4.736 | 0.0091617 |
| 2×4 interaction | 4 | 86.6182 | 21.65455 | 3.2 | 0.0415788 |
| 3×8 interaction | 4 | 95.28001 | 23.82 | 3.52 | 0.0303178 |

These can be seen graphically by plotting the corresponding treatment RMSEs as shown in Fig. 2. Note that the lines in the figures are not parallel.

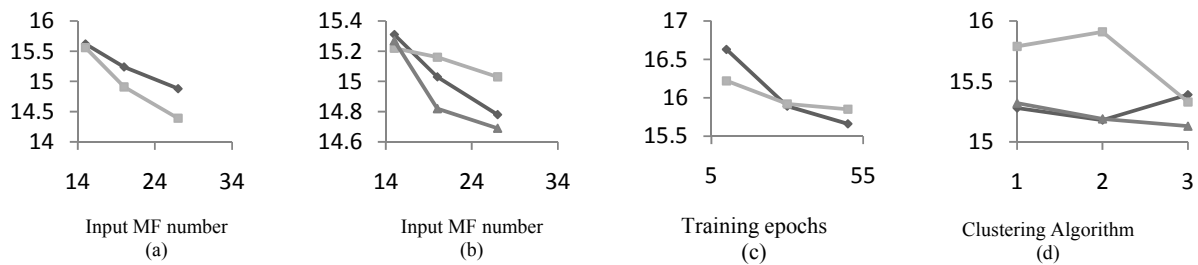


Fig. 2. ANFIS factors' interactions plot: (a) input MF number versus output MF shape, (b) input MF number versus initial rule based construction approach, (c) output MF shape versus training epochs, and (d) initial rule based construction approach and initial clustering algorithm, The vertical axis for all figures are RMSE

4. Conclusion

Starting with the popularity and prevalent use of the neuro-fuzzy systems, specially ANFIS, we explained the inherent difficulty of its designing and parameter setting process. To tackle this problem, one suitable approach is to perform design of experiment technique to identify the most statistically significant factors on the performance of the ANFIS. Since ANFIS is widely used in prediction problems, we applied it for the stock price prediction problem. Historical time series of an automotive part-making company available in Tehran stock exchange market were used for experimental purposes. The results of experiments showed that the most relevant parameters for ANFIS in stock price prediction problem are input membership function number, output membership function shape, initial rule based construction approach and cluster validity index. Moreover, four significant interactions among factors were identified in this experiment. The experiment brings a valuable insight into ANFIS designing process and can reduce difficulties in designing of ANFIS, reducing the search space, and the complexity of the systems' tuning. In automatic optimization techniques, this result can be mapped in the cost functions, for the adaptability to consider first the optimization of parameters with greatest influence in the ANFIS behavior and performance. For

future works, one can completely accomplish this approach in other proposed neuro-fuzzy systems and also can use different popular application problems.

References

- Babuška, R., & Verbruggen, H. (2003). Neuro-fuzzy methods for nonlinear system identification. *Annual reviews in control* 27, 73-85.
- Buragohain, M., & Mahanta, C. (2008). A novel approach for ANFIS modelling based on full factorial design. *Applied Soft Computing*, 8, 609–625.
- Chen, M. S. (1999). A Comparative Study of Learning Methods in Tuning Parameters of Fuzzy Membership Functions. In *Proceedings of IEEE SMC '99 Conference*, Tokyo, Japan, 40-44.
- Chiu, S. (1994). Fuzzy Model Identification Based on Cluster Estimation. *Journal of Intelligent & Fuzzy Systems* 2(3).
- Echanobe, J., Campo, I. D., & Bosque, G. (2008). An adaptive neuro-fuzzy system for efficient implementations. *Information Sciences*, 178, 2150–2162.
- Emami, M. R., Turksen I. B., & Goldenberg, A. A. (1999). A unified parameterized formulation of reasoning in fuzzy modeling and control. *Fuzzy Sets and Systems* 108, 59–81.
- Fazel Zarandi, M. H., Rezaee, B., Turksen, I. B., & Neshat, E. (2009). A type-2 fuzzy rule-based expert system model for stock price analysis. *Expert Systems with Applications* 36, 139–154.
- Fukuyama, Y., & Sugeno, M. (1989). A new method of choosing the number of clusters for fuzzy c-means method. In *Proceeding of 5th fuzzy system symposium* (in Japanese), 247-250.
- Hicks, C. R. (1999). *Fundamental Concepts in the Design of Experiments*. Oxford University Press: New York.
- Jang, J-S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference systems. *IEEE Transactions on Systems, Man & Cybernetics*, 23(3), 665–685.
- Jang, J-S. R., & Mizutani, E. (1996). Levenberg-Marquardt Method for ANFIS Learning. In *Proceedings of Biennial Conference of the North American Fuzzy Information Processing Society*, 87 -91.
- Kim, Y. I, Kim, D. W., Lee, D., & Lee, K. H. (2004). A cluster validation index for GK cluster analysis based on relative degree of sharing. *Information Sciences*, 168, 225-242.
- Kosko, B. (1994). Fuzzy systems as universal approximators. *IEEE Transactions on Computers*, 43(11), 1329-1333.
- Kwon, S. H. (1998). Cluster validity index for fuzzy clustering. *Electronics Letters* 34(22), 2176–2177.
- Mascioli, F.M., Varazi, G.M, & Martinelli, G. (1997). Constructive Algorithm for Neuro-Fuzzy Networks. In *Proceedings of the Sixth IEEE International Conference Fuzzy Systems*, Spain, 1, 459- 464.
- Melek, W. W., Goldenberg, A. A., & Emami, M. R. (2005). A fuzzy noise-rejection data partitioning algorithm. *International journal of approximate reasoning* 38, 1-17.
- Panella, M., & Gallo, A. S. (2005). An Input-Output Clustering Approach to the Synthesis of ANFIS Networks. *IEEE Transactions on Fuzzy Systems*, 13(1), 69-81.
- Pedrycz, W., & Reformat, M. (2003). Evolutionary Fuzzy Modeling. *IEEE Transactions on Fuzzy Systems* 11(5), 652-665.
- Riverol, C., & Di Sanctis, C. (2009). A fuzzy filter for improving the quality of the signal in adaptive-network-based fuzzy inference systems (ANFIS). *Applied Soft Computing*, 9(1), 305–307.
- Shoorehdeli, M. A., Teshnehlab, M., Sedigh, A. K, & Khanesar, M. A. (2009). Identification using ANFIS with intelligent hybrid stable learning algorithm approaches and stability analysis of training methods. *Applied Soft Computing*, 9(2), 833–850.
- Sugeno, M., & Yasukawa, T. (1993). A fuzzy logic based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems*, 1, 7–31.
- Tang, A. M, Quek, C., & Ng, G. S. (2005). GA-TSKfnn: Parameters tuning of fuzzy neural network using genetic algorithms. *Expert Systems with Applications* 29, 769-781.
- Teshnehlab, M., Aliyari Shoorehdeli, M., & Sedigh, A.K. (2008). Novel hybrid learning for tuning ANFIS parameters as an identifiers using fuzzy PSO. In *Proceedings of IEEE International Conference on Networking, Sensing and Control*, Sanya, China, 111-116.
- Tsoukalas, L. H., & Uhrig, R. E. (1997). *Fuzzy and Neural Approaches in Engineering*. John Wiley & Sons: New York.
- Zanchettin, C., Minku, F. L, & Ludermir, T. B. (2005). Design of Experiments in Neuro-Fuzzy Systems. In *Proceedings of the Fifth International Conference on Hybrid Intelligent Systems (HIS'05)*, 6, doi: 10.1109/ICHIS.2005.34.