Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 36 (2009) 139–154
Expert Systems with Applications
www.elsevier.com/locate/eswa

# A type-2 fuzzy rule-based expert system model for stock price analysis

M.H. Fazel Zarandi [a,*], B. Rezaee [a], I.B. Turksen [b], E. Neshat [a]

[a] Department of Industrial Engineering, Amirkabir University of Technology (Polytechnic of Tehran), P.O. Box 15875-4413, Tehran, Iran
[b] Department of Mechanical and Industrial Engineering, University of Toronto, 5 King College Road, Toronto, ON, Canada M5S2H8

**Abstract**

In this paper, a type-2 fuzzy rule based expert system is developed for stock price analysis. Interval type-2 fuzzy logic system permits us to model rule uncertainties and every membership value of an element is interval itself. The proposed type-2 fuzzy model applies the technical and fundamental indexes as the input variables. This model is tested on stock price prediction of an automotive manufactory in Asia. Through the intensive experimental tests, the model has successfully forecasted the price variation for stocks from different sectors. The results are very encouraging and can be implemented in a real-time trading system for stock price prediction during the trading period.


*Keywords:* Type-2 fuzzy modeling; Fuzzy rule based systems; Forecasting; Stock market

## 1. Introduction

Two common analytical approaches to stock market analysis are fundamental and technical analysis. A fundamental analysis relies on the statistics of the macroeconomics data such as interest rates, money supply, inflationary rates, and foreign exchange rates, as well as the basic financial status of a company. After taking all these factors into account, the analyst can then make a decision to sell or buy a stock. A technical analysis is based on the historical financial time-series data. However, financial time series exhibit quite complicated patterns (for example, trends, abrupt changes, and volatility clustering) and such series are often nonstationary, whereby a variable has no clear tendency to move to a fixed value or a linear trend.

During the last decade, stocks and future traders have come to rely upon various types of intelligent systems to make trading decisions. Lately, artificial neural networks (ANNs) have been applied to this area (Aiken & Bsat, 1999; Chang, Wang, & Yang, 2004; Chi, Chen, & Cheng, 1999; Kimoto & Asakawa, 1990; Lee, 2001; Yao & Poh, 1995; Yoon & Swales, 1991). These models, however, have their own limitations owing to the tremendous noise and complex dimensionality of stock price data and besides, the quantity of data itself and the input variables may also interfere with each other. Therefore, the result may not be convincing.

Other soft computing methods are also applied in the prediction of stock price. These approaches are to use quantitative inputs, like technical indices, and qualitative factors, like political effects, to automate stock market forecasting and trend analysis. Kuo, Chen, and Hwang (2001) used a genetic algorithm base fuzzy neural network to measure the qualitative effects on the stock price. They applied their system to the Taiwan stock market. Aiken and Bsat (1999) used a FNN trained by a genetic algorithm (GA) to forecast three-month US Treasury Bill rates. They concluded that a neural network (NN) can be used to accurately predict these rates. Thammano (1999) used a neuro-fuzzy model to predict future values of Thailand's largest government-owned bank. The inputs of the model are the

* Corresponding author. Tel.: +98 21 641 3034; fax: +98 21 6641 3025.
*E-mail addresses:* zarandi@aut.ac.ir (M.H. Fazel Zarandi), babak_rezaee@aut.ac.ir (B. Rezaee), turksen@mie.utoronto.ca (I.B. Turksen), Neshat_elahe@aut.ac.ir (E. Neshat).

0957-4174/$ - see front matter 
doi:10.1016/j.eswa.2007.09.034

closing prices for the current and prior three months, and the profitability ratios. The output of the model was the stock prices for the following three months. He concluded that the neuro-fuzzy architecture was able to recognize the general characteristics of the stock market faster and more accurately than the basic back propagation algorithm. Also, it could predict investment opportunities during the economic crisis when statistical approaches did not yield satisfactory results. Tansel et al. (1999) compared the ability of linear optimization, ANNs, and GAs to model time series data using the criteria of modeling accuracy, convenience and computational time. They found that linear optimization methods gave the best estimates, although the GAs could provide the same values if the boundaries of the parameters and the resolution are selected appropriately, but that the NNs resulted in the worst estimations. However, they noted that non-linearity could be accommodated by both the GAs and the NNs and that the latter required minimal theoretical background. Baba, Inoue, and Asakawa (2000) used NNs and GAs to construct an intelligent decision support system (DSS) for analyzing the Tokyo Stock Exchange Prices Indexes (TOPIX). The essential feature of their DSS is that it can project the high and low TOPIX values four weeks into the future and suggest buy and sell decisions based on the average projected value and the then-current value of the TOPIX. Kim and Han (2000) used a NN modified by a GA to predict the stock price index. In this instance, the GA was used to reduce the complexity of the feature space, by optimizing the thresholds for feature discretization, and to optimize the connection weights between layers. They concluded that the GA approach outperforms the conventional models.

Abraham, Baikunth, and Mahanti (2001) investigated hybridized soft computing techniques for automated stock market forecasting and trend analysis. They used principal component analysis to preprocess the input data, a NN for one-dayahead stock forecasting, and a neuro-fuzzy system for analyzing the trend of the predicted stock values. Abraham, Philip, and Saratchandran (2003) investigated how the seemingly chaotic behavior of stock markets could be well represented using several connectionist paradigms and soft computing techniques. To demonstrate the proposed technique, they analyzed the 7 year's Nasdaq-100 main index and 4 year's NIFTY index values. They concluded that all the connectionist paradigms considered could represent the stock indices behavior very accurately.

Chang, Liu, and Wang (2006) developed a hybrid model by integrating Self Organization Map (SOM) neural network, genetic algorithms (GA), and fuzzy rule base (FRB) to forecast the future sales of a printed circuit board factory. Chang and Wang (2006) combined fuzzy theory and back-propagation network into a hybrid system, which can be applied in the sales forecasting of printed circuit board (PCB) industries. Chang and Liu (2006) developed Takagi–Sugeno–Kang (TSK) type fuzzy rule based system for stock price prediction. Their TSK fuzzy model applied the technical index as input variables and consequent part is a linear combination of the input variables. The fuzzy rule based model is tested on the Taiwan Electronic Shares from the Taiwan Stock Exchange.

Quite often, the knowledge that is used to construct the rules in a fuzzy logic system (FLS) is uncertain. Three ways in which such rule uncertainty can occur are: (1) the words that are used in antecedents and consequents of rules can mean different things to different people; (2) consequents obtained by polling a group of experts will often be different for the same rule because the experts will not necessarily be in agreement; and (3) noisy training data (Liang & Mendel, 2000). Antecedent or consequent uncertainties translate into uncertain antecedent or consequent membership functions. Type-1 FLSs, whose membership functions are type-1 fuzzy sets, are unable to directly handle rule uncertainties. Type-2 FLSs, the subject of this paper, in which antecedent or consequent membership functions are type-2 fuzzy sets, can handle rule uncertainties.

It should be noted that type-2 fuzzy sets can model and minimize the effects of uncertainties in rule-based fuzzy logic systems. The effects of uncertainties can be minimized by optimizing the parameters of the type-2 fuzzy sets during a training process. The additional parameters of type-2 fuzzy sets over those in type-1 fuzzy sets provide the former with additional design degrees of freedom that make it possible to minimize the effects of uncertainties. Type-2 fuzzy logic is very useful when it is difficult to determine the exact membership functions of fuzzy sets.

The aim of this research is to develop a fuzzy modeling mechanism which is capable of implementing four objectives:

- generating a rule base automatically from numeric data,
- finding the optimal number of rules and fuzzy sets,
- optimizing the parameters of fuzzy membership functions, and
- increasing the robustness of the system.

To achieve these objectives, this paper proposes a fuzzy modeling paradigm by incorporating a fuzzy GK (Gustafson and Kessel, 1979) clustering associated with a proposed cluster validity measure, and a genetic algorithm is interval type-2 fuzzy sets domain.

The rest of the paper is organized as follows: Section 2 reviews the type-2 fuzzy sets and systems and their associated terminologies. In Section 3 the design approach of interval type-2 fuzzy logic system is presented. Section 4 presents the proposed interval type-2 fuzzy system for prediction of single stock price of an automotive manufactory in Asia. Finally, conclusions and comments on further research are appeared in Section 5.

## 2. Type-2 fuzzy logic systems

This section reviews the different aspects of a type-2 fuzzy logic system (FLS) that are needed in this research.

*2.1. Type-2 fuzzy sets (T2FS)*

A type-2 fuzzy set, $\widetilde{A}$ denoted, is characterized by a type-2 membership function, $\mu_{\widetilde{A}}(x,u)$, where, $x \in X$ and $u \in J_x \subseteq [0,1]$ (Mendel & John, 2002):

$$\widetilde{A} = \{((x,u), \mu_{\widetilde{A}}(x,u)) | \ \forall x \in X, \ \forall u \in J_x \subseteq [0,1]\} \quad (1)$$

in which $0 \leqslant \mu_{\widetilde{A}}(x,u) \leqslant 1$. Here, $\widetilde{A}$ can also be expressed as

$$\widetilde{A} = \int_{x\in X} \int_{u \in J_x} \mu_{\widetilde{A}}(x,u)/(x,u) \quad J_x \subseteq [0,1] \quad (2)$$

where, $\iint$ denotes union over all admissible $x$ and $u$.

$J_x$ is called primary membership of $\widetilde{A}$, where $J_x \subseteq [0,1]$ for $x \in X$. The uncertainty in the primary memberships of a type-2 fuzzy set $\widetilde{A}$, consists of a bounded region that is called the footprint of uncertainty (FOU).

When all $\mu_{\widetilde{A}}(x,u) = 1$ in definition of a type-2 fuzzy set $\widetilde{A}$, then we have an interval type-2 fuzzy set (IT2 FS). Although the third dimension of the general T2 FS is no longer needed because it conveys no new information about the IT2 FS, the IT2 FS can still be expressed as a special case of the general T2 FS in (1), as

$$\widetilde{A} = \int_{x\in X} \int_{u \in J_x} 1/(x,u), \quad J_x \subseteq [0,1]. \quad (3)$$

The upper membership function (UMF) and lower membership function (LMF) of $\widetilde{A}$ are two T1 MFs that bound the FOU. The UMF is associated with the upper bound of FOU($\widetilde{A}$) and is denoted $\bar{\mu}_{\widetilde{A}}(x)$ $\forall x \in X$, and the LMF is associated with the lower bound of FOU($\widetilde{A}$) and is denoted $\underline{\mu}_{\widetilde{A}}(x)$ $\forall x \in X$ (Mendel & John, 2002). That is:

$$\begin{aligned} \bar{\mu}_{\widetilde{A}}(x) &\equiv \overline{\mathrm{FOU}(\widetilde{A})} \quad \forall x \in X \\ \underline{\mu}_{\widetilde{A}}(x) &\equiv \underline{\mathrm{FOU}(\widetilde{A})} \quad \forall x \in X \end{aligned} \quad (4)$$

*2.1.1. Operations in type-2 fuzzy sets*

Recall that the membership grades of type-2 sets are type-1 sets; therefore, in order to perform operations like union and intersection on type-2 sets, we need to be able to perform t-conorm and t-norm operations between type-1 sets. This is done by using Zadeh's Extension Principle. This leads to the following definitions (Mendel, John, & Liu, 2006):

(a) The *union* of two T2 FSs, $\widetilde{A}$ and $\widetilde{B}$ ($\mu_{\widetilde{A}}(x) = \int_u f_x(u)/u$ and $\mu_{\widetilde{B}}(x) = \int_w g_x(w)/w$, where $u$, $w \in J_x$), is

$$\begin{aligned} \widetilde{A} \cup \widetilde{B} \Longleftrightarrow \mu_{\widetilde{A}\cup\widetilde{B}}(x) &= \mu_{\widetilde{A}}(x) \coprod \mu_{\widetilde{B}}(x) \\ &= \int_u \int_w (f_x(u) * g_x(w))/(u \vee w) \end{aligned} \quad (5)$$

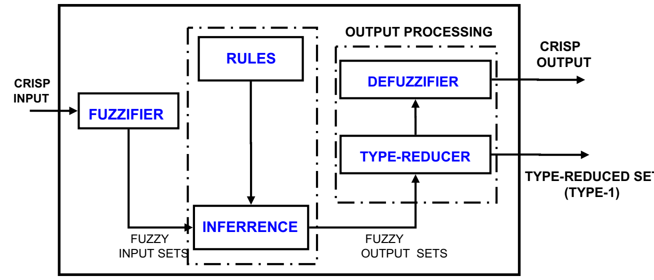



Fig. 1. The structure of type-2 fuzzy logic system (Mendel, 2007).

(b) The *intersection* of two T2 FSs, $\widetilde{A}$ and $\widetilde{B}$, is

$$\begin{aligned} \widetilde{A} \cap \widetilde{B} \Longleftrightarrow \mu_{\widetilde{A}\cap\widetilde{B}}(x) &= \mu_{\widetilde{A}}(x) \prod \mu_{\widetilde{B}}(x) \\ &= \int_u \int_w (f_x(u) * g_x(w))/(u * w) \end{aligned} \quad (6)$$

(c) The *complement* of IT2 FS, $\widetilde{A}$, $\overline{\widetilde{A}}$ is

$$\overline{\widetilde{A}} \Longleftrightarrow \mu_{\overline{\widetilde{A}}}(x) = \neg\mu_{\widetilde{A}}(x) = \int_u f_x(u)/1 - u \quad (7)$$

*2.2. Type-2 fuzzy logic system (FLS)*

The conventional fuzzy rule-base structures employ type-1 fuzzy sets both/either in antecedent and/or consequent parts of the rules. However, recent studies shows that the uncertainty can he captured in a better way by using higher order fuzzy sets, such as type-2 fuzzy sets, which encapsulate more information granules.

Fig. 1 shows the structure of a type-2 fuzzy logic system (FLS). It is very similar to the structure of a type-1 FLS. For a type-1 FLS, the output processing block only contains the defuzzifier. When an input is applied to a type-1 FLS, the inference engine computes the type-1 output set corresponding to each rule. The defuzzifer then computes a crisp output from these rule output sets. For a type-2 FLS, the antecedent and/or consequent sets are type-2, so that each rule output set is type-2. "Extended" versions of type-1 defuzzification methods yield a type-1 set from the type-2 rule output sets. This process is called type-reduction rather than defuzzification, and the resulting type-1 set, the type-reduced set. The defuzzifier in the type-2 FLS can, then, defuzzify the type-reduced set to obtain a crisp output for the type-2 FLS. The fuzzifier maps the crisp input into a fuzzy set. This fuzzy set can, in general, be a type-2 set; however, in this paper, we consider only singleton fuzzification, for which the input fuzzy set has only a single point of nonzero membership (Karnik & Mendel, 1998; Mendel, 2007).

## 3. Designing the type-2 FLS

There are two very different approaches for selecting the parameters of a type-2 FLS (Mendel, 2001). One is the partially dependent approach, where a best possible type-1

FLS is designed first, and then used to initialize the parameters of a type-2 FLS. The other method is a totally independent approach, where all the parameters of the type-2 FLS are tuned from scratch without the aid of an existing type-1 design.

One advantage offered by the partially dependent approach is smart initialization of the parameters of the type-2 FLS. Since the baseline type-1 fuzzy sets impose constraints on the type-2 sets, fewer parameters need to be tuned and the search space for each variable is smaller. Therefore, the computational cost is less than the totally independent approach. So design flexibility is traded for a lower computational burden. Type-2 FLSs designed via the partially dependent approach are able to outperform the corresponding type-1 FLSs (Wu & Tan, 2004), although both the FLSs have the same number of MFs (resolution). However, the type-2 FLS has a larger number of degrees of freedom because the fuzzy set is more complex. The additional mathematical dimension provided by the type-2 fuzzy set enables a type-2 FLS to produce more complex input–output map without the need to increase the resolution.

Our purposed approach is based on partially dependent approach. First, we design a type-1 fuzzy system and then, for increasing the robustness of the system, we create a type-2 fuzzy rule base with uncertain mean and interval secondary membership functions. It uses the same number of fuzzy sets and the same rules as the type-1 FLS. The only difference now is that the antecedent and consequent sets are type-2.

### *3.1. Procedures of development of type-2 fuzzy system*

The procedures of development of the proposed system are as follows:

- Determination of input and output variables of the system.
- Clustering the output space and determination of the number of rules.
- Variable selection.
- Projection of membership values of the output onto the input spaces to generate the membership values of the inputs.
- Tuning the parameters of membership function of input and output variables by using genetic algorithm (GA).
- Transforming type-1 fuzzy rule base to interval type-2 fuzzy rule base.
- Tuning the parameters of interval type-2 membership function of the input and output variables.

#### *3.1.1. Determination of input and output variables of the system*

The first step in system modeling is the identification of input and output variables. This task is usually done by studying the problem domain and by the negotiation with the domain experts. Of course there are an infinite number of possible candidates which should be restricted to certain numbers. In this step, the designers and experts try to determine the most relevant input and output variables.

#### *3.1.2. Clustering the output space and determination of the number of rules*

In this paper, first, the output data is clustered and the primary membership grades of the output clusters are generated. For this purpose, Sugeno and Yasukawa (1993) method is used. We first partition the output space and then obtain the input space clusters by "projecting" the output space partition onto each input variable space, separately.

In order to carry out the process of encoding the output space, we consider one of the most applicable and traceable fuzzy clustering algorithms, i.e., GK clustering.

It is required to obtain a cluster validity criterion in order to determine the optimal number of clusters presented in a data set. In the case of the GK algorithm, there are only a few validation indices found in the literature (Baduska, 1995; Gath & Geva, 1989) with poor performance, and most validation indices proposed for the FCM cannot be applied to the GK clustering directly because they highly depend on the centroid information of the clusters and they do not use the covariance information of the clusters. Most of the validity indices proposed for the FCM (Bezdek, 1975, 1981; Kim, Lee, & Lee, 2003; Razaee, Lelieveldt, & Reiber, 1998) measure intra-cluster compactness and inter cluster separation using cluster centroids. However, interpretation of inter cluster separation of these indices is problematic because such indices quantify cluster separation based on only the distance between cluster centroids. Thus, they are not appropriate for the clusters found by the GK algorithms which are often in the shape of hyper ellipsoids of different orientation and shapes.

Most validity indices focus only on the compactness and the variation of the intra-cluster distance (Fukuyama & Sugeno, 1989; Kwon, 1998; Razaee et al., 1998; Xie & Beni, 1991). Fuzzy hyper volume and density criteria use the hyper volume to assess the density of the resulting clusters measuring mainly compactness of the given fuzzy partition (Baduska, 1998). Some indices, for example Xie and Beni (1991) and Kwon (1998), use the strength of separation between clusters; however, interpretation of these indices is problematic because they quantify cluster separation based only on the distance between cluster centriods. Since the GK clustering involves Mahalanobis distance norm for each cluster, validity indices like Xie and Beni (1991) and Kwon (1998), cannot discriminate the separation of two different pairs of clusters with different clusters and with different orientation. This is shown in the Fig. 2. In this figure even though the pair $(U^{(a)}, V^{(a)})$ provides a better partitioning than the pair $(U^{(b)}, V^{(b)})$, this cannot be reflected properly in $V_{XB}$ and $V_K$ because they calculate the separation between clusters using only centroid distances.

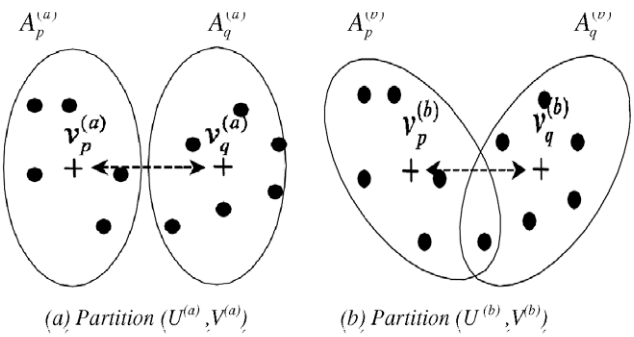



Fig. 2. Two different fuzzy partitions $(U^{(a)}, V^{(a)})$ and $(U^{(b)}, V^{(b)})$ with the same distance between cluster centroids with different orientations for the same data.

*3.1.3. The proposed cluster validity based on fuzzy similarity*

Let $A_p$ and $A_q$ be two fuzzy clusters belonging to a fuzzy partition $(U, V)$ and $c$ be the number of clusters.

**Definition 1.** The relative similarity between two fuzzy sets $A_p$ and $A_q$ at $x_j$ is defined as

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

In (8)

$$f(x_j : A_p \cap A_q) = u_{A_p}(x_j) \wedge u_{A_q}(x_j) \tag{9}$$

where, $\wedge$ is minimum operator. Moreover:

$$f(x_j : A_p - A_q) = \text{Max}(0, u_{A_p}(x_j) - u_{A_q}(x_j)) \tag{10}$$

and

$$f(x_j : A_q - A_p) = \text{Max}(0, u_{A_q}(x_j) - u_{A_p}(x_j)) \tag{11}$$

**Definition 2.** The relative similarity between two fuzzy sets $A_p$ and $A_q$ is defined as

$$S_{\text{rel}}(A_p, A_q) = \sum_{j=1}^{n} S_{\text{rel}}(x_j : A_p, A_q)h(x_j) \tag{12}$$

where,

$$h(x_j) = -\sum_{p=1}^{c} u_{A_p}(x_j)\log(u_{A_p}(x_j)) \tag{13}$$

Here, $h(x_j)$ is the entropy of datum $x_j$ and $u_{A_p}(x_j)$ is the membership value with which $x_j$ belongs to the cluster $A_p$. In (12), $h(x_j)$ measures how vaguely (unclearly) the datum $x_j$ is classified over $c$ different clusters. $h(x_j)$ is introduced to assign a weight for vague data. Vague data are given more weights than clearly classified one. $h(x_j)$ also reflects the dependency of $u_{A_p}(x_j)$ with respect to different $c$ values. This approach makes it possible to focus more on the highly-overlapped data in the computation of the validity index than other indices do.

**Definition 3.** The proposed validity index is as follows:

$$V_{\text{FNT}}(U, V; X) = \frac{2}{c(c-1)} \sum_{p \neq q}^{c} S_{\text{rel}}(A_p, A_q) \tag{14}$$

The optimal number of the clusters is obtained by minimizing $V_{\text{FNT}}$ $(U, V; X)$over the range of $c$ values: $2, \ldots, c_{\max}$.

Thus, $V_{\text{FNT}}$ is defined as the average value of the relative similarity between $c(c-1)/2$ pairs of clusters, where the relative similarity between each cluster pair is defined as the weighted sum of the relative similarity at $x_j$between two clusters in the pair. Hence, the less overlap in a fuzzy partition, and the less vague the data points in that overlap, the lower the value of $V_{\text{FNT}}(U, V; X)$ is resulted.

To demonstrate the effectiveness of the proposed index, we compared it with some other well known indexes on a number of widely used data sets. These indices are as follows:

- Partition coefficient (Bezdek, 1974)

$$V_{\text{PC}}(U) = \frac{\sum_{k=1}^{n}\sum_{i=1}^{c} u_{ik}^2}{n} \tag{15}$$

- Partition entropy (Bezdek, 1975)

$$V_{\text{PE}}(U) = -\frac{1}{n}\left(\sum_{k=1}^{n}\sum_{i=1}^{c}[u_{ik}\log_a(u_{ik})]\right) \tag{16}$$

- Xie and Beni (1991)

$$V_{\text{XB}}(U, V; X) = \frac{\sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^2 \|x_i - v_k\|^2}{n\left(\underbrace{\min}_{i \neq j}\{\|v_i - v_j\|^2\}\right)} \tag{17}$$

- Zahid et al. (1999)

$$V_{\text{SC}}(U, V; X) = \frac{\frac{1}{c}\sum_{i=1}^{c}\|v_i - \bar{v}\|^2}{\sum_{i=1}^{c}(\sum_{k=1}^{n} u_{ik}^m \|x_k - v_i\|^2 / n_i)} - \frac{\sum_{i=1}^{c-1}\sum_{r=1}^{c-i}(\sum_{k=1}^{n}\min(u_{ik}, u_{rk})^2 / n_{ik})}{\sum_{k=1}^{n}\max_i u_{ik}^2 / \sum_{k=1}^{n}\max_i u_{ik}} \tag{18}$$

- Kwon (1998),

$$V_k(U, V; X) = \frac{\sum_{k=1}^{n}\sum_{i=1}^{c} u_{ik}^2 \|x_k - v_i\|^2 + \frac{1}{c}\sum_{i=1}^{c}\|v_i - \bar{v}\|^2}{\underbrace{\min}_{i \neq k}(\|v_k - v_i\|^2)} \tag{19}$$

$V_{\text{XB}}$, $V_{\text{K}}$ and $V_{\text{SC}}$ are modified to accommodate Mahalanobis distance norm instead of Euclidean one in calculating the distance from each data point to the cluster centers. The parameters of the GK algorithm are set as follows: termination criterion $\varepsilon = 10^{-5}$, weighting exponent $m = 2$, $c_{\max} = 12$ and the initial cluster centers are selected by the FCM.

Figs. 3–9 show scatter plots of the seven artificially generated data sets used in the experiments.

Table 1 summarizes the optimal cluster numbers identified by each validity index. For example for Data set 7 all validity indices $V_{\mathrm{PC}}$, $V_{\mathrm{PE}}$, $V_{\mathrm{SC}}$, $V_{\mathrm{XB}}$, and $V_{\mathrm{K}}$ incorrectly have identified the optimal cluster number and only $V_{\mathrm{FNT}}$ identifies it correctly. The optimal number of clusters is obtained by minimizing the proposed validity index. This result indicates that the proposed validity index is more reliable.

### 3.1.4. Variable selection

The phase of input selection in system identification is to find the most dominant input variables which affect the output among a finite number of input candidates. Theoretically, this problem belongs to a more general field of data analysis, i.e., dimension reduction. In the analysis of multivariate data, it is common practice to look for the dimension reduction via linear combinations of the initial variables. Classical techniques, such as principal components (Duda & Hart, 1973), discriminant analysis (Friedman & Rubin, 1996), and canonical correlation (Jain & Dubes, 1980) are examples of this approach. From a practical point of view, another type of dimension reduction is selecting a subset of the variables. The main advantage of

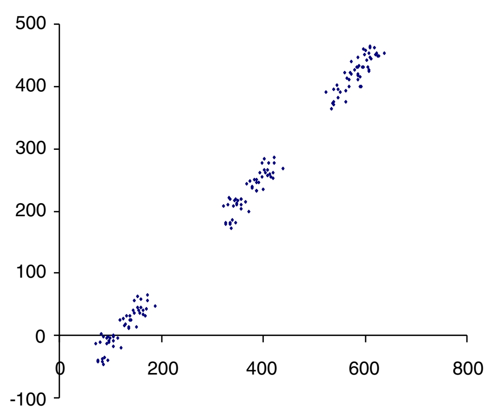


Fig. 3. Data set 1 (optimal cluster number is 3).

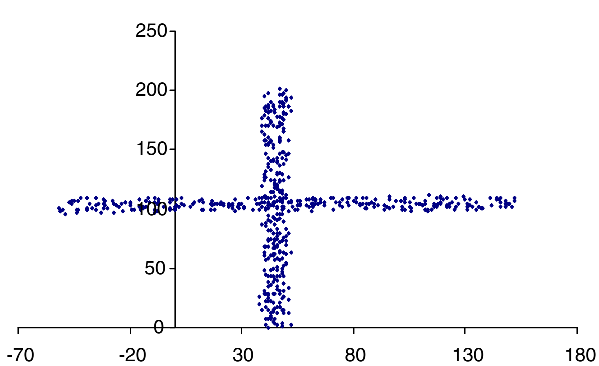


Fig. 4. Data set 2 (optimal cluster number is 2).

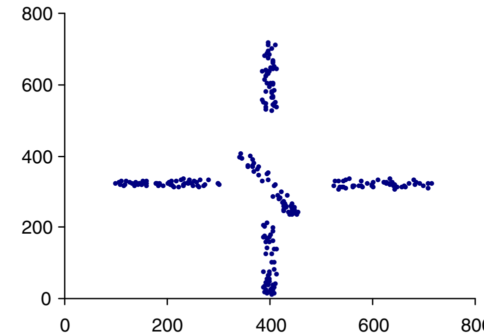


Fig. 5. Data set 3 (optimal cluster number is 5).

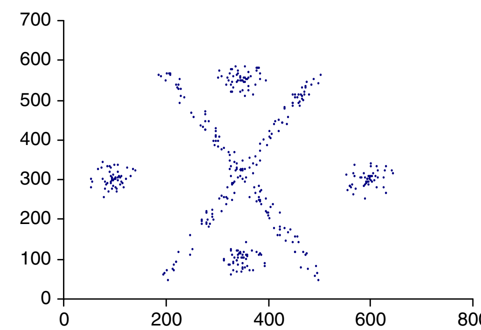


Fig. 6. Data set 4 (optimal cluster number is 6).

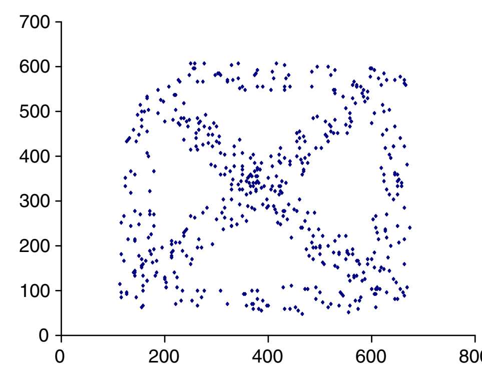


Fig. 7. Data set 5 (optimal cluster number is 6).

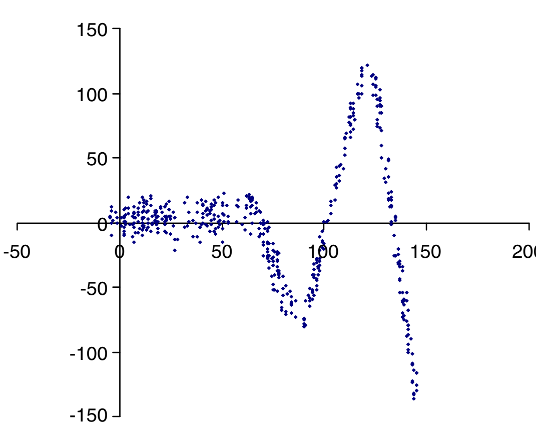
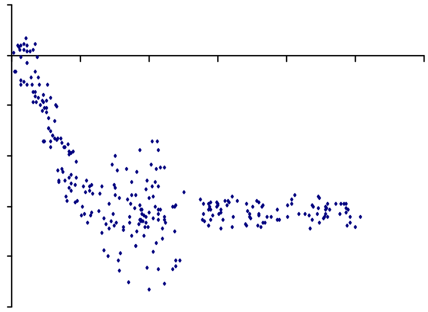


Fig. 8. Data set 6 (optimal cluster number is 4).



this approach is that there is an actual reduction in the number of the measured variables. In this way, we can avoid the interpretational difficulties which could arise in looking at linear combinations of very different kinds of variables. Although it is a common practice to check the weights of variables in a linear combination and to discard those that have "negligible" weights, this is not always easy to do nor are negligible weights always guaranteed.

For variables selection, the Sugeno and Yasukawa (1993) method (the variable selection algorithm) is used. Sugeno and Yasukawa (1993) proposed a combinatorial approach in which al1 possible combinations of input candidates are considered. For each combination, they build two fuzzy models based on two separated sets of data and calculate a performance index called "Regularity Criterion" (RC) based on a method of analyzing two groups of data in an attempt to cause data independence in model formation after that a combination of input variables is chosen which has the minimum value of the performance index.

*3.1.5. Projection of membership functions of output onto input spaces*

After selection of the significant input variables, suitable membership functions should be determined for them. One simple approach is to set the membership grade of each sample input equal to its corresponding output membership grade, obtained from the output data clustering process (Sugeno & Yasukawa, 1993). Therefore, for each output datum, all the corresponding input variables will have the same membership grade. The problem with this technique is that the membership functions assigned in this way are not convex and further approximation is required to shape the convex membership functions. Moreover, there is no reason for the input membership grades to be the same and equal to the output membership grade at each sample point.

With respect to the proposed approach of Fazel Zarandi (1998) that is shown in Fig. 10, first, the ranges in which input variable membership functions that adapt value 1 are determined. Then, the data points are classified, using GK by given $m$ and $c$, which were determined in the pervious stage, as follows:

First determine the interval in which input membership functions adopt value 1 (i.e., $\overline{V_1V_2}$ Fig. 10). Then, the optimum value of $V_1^*$ and $V_2^*$ are determined by searching through the reliable region and analyzing the objective function of classification algorithm, where:

$$J(U^j, X_j) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik}^j)^m \Delta(x_{jk}, \overline{V_1V_2}) \tag{20}$$

In this search, a suitable $\{V_1^*, V_2^*\}$ that minimizes the Eq. (20) is determined, where $n$ is number of data, $c$ is number of clusters (we use $m$ and $c$ that are obtained from output variable clustering stage), $X_j$ and $U^j$ are the $j$th input data set and their partition matrix, respectively, and $\Delta$ is the distance function obtain from Eq. (21):

$$\Delta(x_{jk}, V_1 V_2) = \begin{cases} V_1 - x_{jk} & \text{if} \quad x_{jk} < V_1 \\ 0 & \text{if} \quad V_1 < x_{jk} \leqslant V_2 \\ x_{jk} - V_2 & \text{if} \quad x_{jk} > V_2 \end{cases} \tag{21}$$

$$u_{ik}^j = \frac{1}{\sum_{l=1}^{c} (\Delta_{ik}^j / \Delta_{lk}^j)^{2/(m-1)}}$$

$$V_1' \leqslant V_1, \quad V_2 \leqslant V_2'$$

Now, $U^j$ matrix can be derived from

$$u_{ik}^j = \frac{1}{\sum_{l=1}^{c} (\Delta_{ik}^j / \Delta_{lk}^j)^{2/(m-1)}} \qquad \text{For } 1 \leqslant i \leqslant c, \ 1 \leqslant k \leqslant n \tag{22}$$

*3.1.6. Tuning the parameters of membership function of input and output variables*

In this research we implement genetic algorithms (GAs) for tuning the main parameters of the system. GA is a general-purpose search algorithm that uses principles inspired by natural population genetics to evolve solutions to the problems. It was first proposed by Holland (1975). GAs are theoretically and empirically proven to provide a robust search in complex spaces, thereby offering a valid approach to problems requiring efficient and effective searches (Goldberg, 1989; Sakawa, 2002).

Fig. 11 contains the flow chart of a basic GA. First, a chromosome population is randomly generated. Each chromosome encodes a candidate solution of the optimization problem. The fitness of all individuals with respect to the optimization task is then evaluated by a scalar objective function (fitness function). According to Darwin's principle, highly fit individuals are more likely to be selected to reproduce offspring. Genetic operators such as crossover and mutation are applied to the parents in order to produce a new generation of candidate solutions. As a result of this evolutionary cycle of selection, crossover and mutation, more and more suitable solutions to the optimization problem emerge within the population.

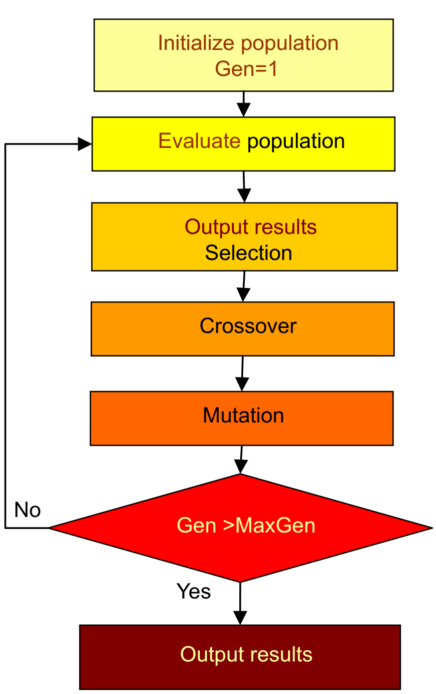


Fig. 11. The flow chart of a basic GA.

In GA, first a population of chromosomes is formed. Each chromosome represents a possible solution to the problem. The population will undergo operations similar to genetic evolution, namely reproduction, crossover, and mutation. In many researches, GAs have been successfully used to tune and design the membership functions (MFs) and the rules of type-1 FLSs. GAs optimize these parameters representing the MFs and the rules of type-1 FLSs in the direction having better performance.

For applying genetic algorithms, the feature parameters of a type-1 FLS have to be encoded into a form of a chromosome. In encoding schemes for a type-1 FLS, a type-1 MF is represented as a mean and a standard deviation (*std*) in a Gaussian.

Many GA-based FLS designing processes have been used to represent the rule table as genes. In other way, they described a rule to a order set of fuzzy sets for each variables. In this paper, we use in latter approach. Each type-1 rule can be represented as shown in (23). Using the above encoding scheme for a type-1 MFs, each MF is encoded together with each rule:

$$R^l : \text{IF } x_1 \text{ is } F_1^l \text{ and } x_2 \text{ is } F_2^l \text{ and } \ldots \text{ and } x_p \text{ is } F_p^l,$$
$$\text{THEN } y \text{ is } G^{l''} \iff (m_{x_1}^l, \sigma_{x_1}^l)(m_{x_2}^l, \sigma_{x_2}^l)\ldots(m_{x_p}^l, \sigma_{x_p}^l)(m_y^l, \sigma_y^l) \tag{23}$$

For encoded type-1 FLSs, genetic operations such as crossover and mutation, are performed to evaluate and optimize chromosomes. These crossover and mutation operators are closely related to the encoding scheme of MFs and rules. So, according to encoding scheme, effective genetic operators have to be proposed and used. In this paper, we only consider the type-1 FLSs with the fixed number of rules and a complete rule set. Therefore, the length of chromosomes is the same and the meaning of each gene according to its position is also the same. We easily use simple one-point or two-point crossover. The mutation plays a role in fine tuning near local points rather for introducing radically different chromosomes into the population.

In GA, evaluation is performed to select and rank chromosomes in a population. Fitness function for evaluation is one of the most important factors to determine the performance of solutions and to control the speed of evolution.

In the genetic based FLSs, it is not so easy to evaluate the performances of the FLSs. In most cases, the results of executions are evaluated as a performance measure after simulating a designed FLS for given situations or for certain time interval. Therefore, the fitness function uses the difference or errors between desired and simulated outputs.

Table 2
Variables of the system

| Variable name | Variable description | Variable |
|---|---|---|
| Demand index (DI) | It become from combination of volume and price. If purchase pressure is more than sell pressure DI is positive and vice versa | – |
| Moving average divergence convergence (MACD) | It uses two different exponential smoothing moving average lines and helps us determining 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) |
| Relative strength index (RSI) | It measures increase or decrease in close price for a specific time period | $\mathrm{RS} = \frac{\sum \text{X days' up closes}}{\sum \text{X days' down closes}}$ $\mathrm{RSI} = \frac{100}{1+\mathrm{Rs}}$ |
| 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 sum of single period close prices dividend on numbers of periods | $\mathrm{MA}(n) = \frac{\sum_{i=1}^{n} \text{close prices of } i \text{ th period}}{n}$ |
| $R$-squared ($R^2$) | It shows trend power | – |
| Linear regression slop (LRS) | It shows trend general movement | – |
| Average true range (ATR) | It measures price fluctuation | $\mathrm{TR}_i = \mathrm{MAX}\left\{\begin{array}{l}(\text{day}_i\text{Max Price-day}_i \text{ Min Price}),\\(\text{day}_i\text{Max Price-day}_{i-1} \text{ Min Price}),\\(\lvert\text{day}_{i-1} \text{ close Price-day}_i \text{ Min Price}\rvert)\end{array}\right\}$ $\mathrm{ATR} = \frac{\sum_{i=1}^{n} \mathrm{TR}_i}{n}$ often $n = 1$ |
| Price channel (top) | It is maximum of price in last four week period | Max{close price of last four week} |
| Price channel (bottom) | It is minimum of price in last four week period | Min{close price of last four week} |
| Price per earning per share ($P/E$) | It shows the payback period or its inversion shows the stock rate of return | $P/E = \frac{\text{Price}}{\text{EPS}}$ $\mathrm{EPS} = \frac{\text{Firm Earning}}{\text{Number of shares}}$ |
| Volume (vol) | It shows how much money transact in trading | – |
| Open price (open) | It shows the open price which is most of the time is equal to last day close price | – |
| Range (R) | It is the range of price in one specific day | Ranges = Maximum Price − Minimum Price |
| Changes | It is the difference between today close and the last day | $\text{Changes}_i = \text{day}_i \text{ Price} - \text{day}_{i-1} \text{ Price}$ |
| Cose price (close) | It shows the close price of the day | – |

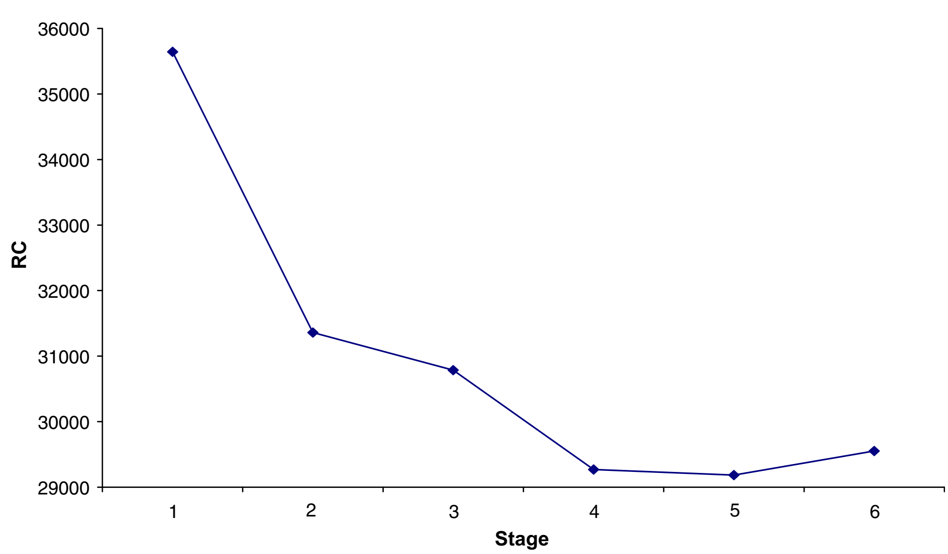
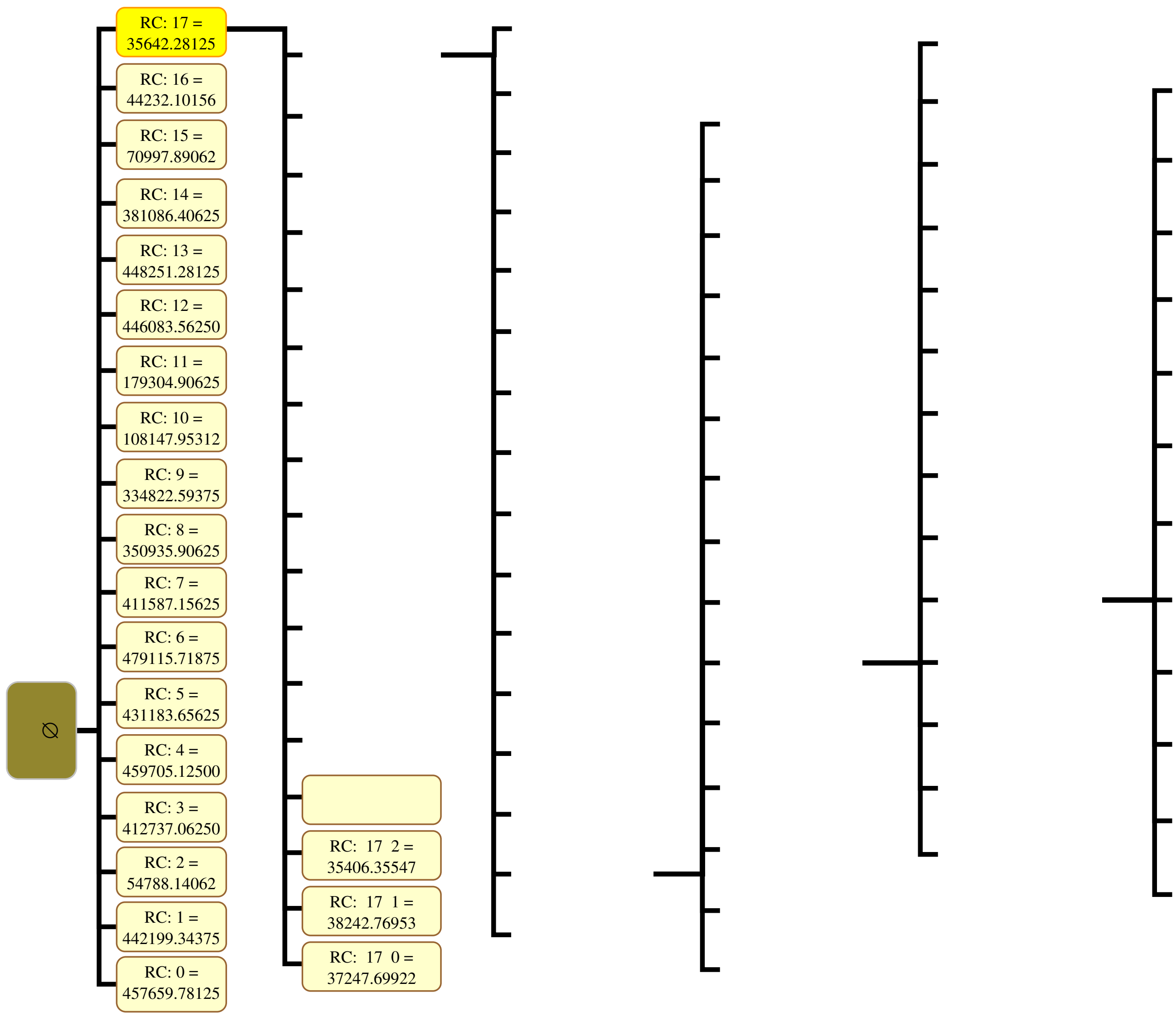



Fig. 12. Behavior of RC in automotive manufactory model.


RC: 17 = 35642.28125
RC: 16 = 44232.10156
RC: 15 = 70997.89062
RC: 14 = 381086.40625
RC: 13 = 448251.28125
RC: 12 = 446083.56250
RC: 11 = 179304.90625
RC: 10 = 108147.95312
RC: 9 = 334822.59375
RC: 8 = 350935.90625
RC: 7 = 411587.15625
RC: 6 = 479115.71875
RC: 5 = 431183.65625
RC: 4 = 459705.12500
RC: 3 = 412737.06250
RC: 2 = 54788.14062
RC: 1 = 442199.34375
RC: 0 = 457659.78125
RC: 17 2 = 35406.35547
RC: 17 1 = 38242.76953
RC: 17 0 = 37247.69922

*3.1.7. Transformation type-1 to interval type-2 membership functions*

For transforming type-1 fuzzy set to an interval type-2 fuzzy set with uncertain mean, we consider the case of a Gaussian primary MF having a fixed standard deviation $\sigma_k^l$ and an uncertain mean that takes on values in $[m_{k1}^l, m_{k2}^l]$ (Mendel, 2000), i.e.,

$$u_k^l(x_k) = \exp\left[-\frac{1}{2}\left(\frac{x_k - m_k^l}{\sigma_k^l}\right)\right], \quad m_k^l \in \left[m_{kl}^l, m_{k2}^l\right] \tag{24}$$

where, $k = 1, \ldots, p$; $p$ is number of antecedents; $l = 1, \ldots, M$; and $M$ is number of rules. The upper MF is

$$\bar{u}_k^l(x_k) = \begin{cases} N\left(m_{k1}^l, \sigma_k^l, x_k\right), & x_k < m_{k1}^l \\ 1, & m_{k1}^l \leqslant x_k \leqslant m_{k2}^l \\ N\left(m_{k2}^l, \sigma_k^l, x_k\right), & x_k > m_{k2}^l \end{cases} \tag{25}$$

where

$$N\left(m_{k1}^l, \sigma_k^l, x_k\right) \cong \exp\left(-\frac{1}{2}\left(\frac{x_k - m_{k1}^l}{\sigma_k^l}\right)^2\right) \tag{26}$$

The lower MF is

$$\underline{u}_k^l(x_k) = \begin{cases} N(m_{k2}^l, \sigma_k^l, x_k), & x_k \leqslant \frac{m_{k1}^l + m_{k2}^l}{2} \\ N(m_{k1}^l, \sigma_k^l, x_k), & x_k > \frac{m_{k1}^l + m_{k2}^l}{2} \end{cases} \tag{27}$$

*3.1.8. Tuning the parameters of interval type-2 membership functions*

Given an input–output training pair $(x^{(t)}, y^{(t)})$, $x^{(t)} \in R^p$ and $y^{(t)} \in R$, we now wish to design an interval singleton type-2 FLS with output $f(x^t)$ such that the following error function is minimized:

$$e(t) = \frac{1}{2}[f(x^{(t)}) - y^{(t)}]^2 \quad t = 1, \ldots, N \tag{28}$$

Based on the analysis by Liang and Mendel (2000), only the upper and lower MFs and the two end-points of the centroid of the consequent set determine $f(x^{(t)})$. So, we want to tune the upper and lower MFs and the consequent parameters $Y^i = [y_l^i, y_r^i]$. Since an interval type-2 FLS can be characterized by two FBF expansions that generate the points $y_r$ and $y_l$, respectively, we can focus on tuning the parameters of just these two type-1 FLSs. They are all tuned using a steepest descent method (Liang & Mendel, 2000).

## 4. Implementation of the proposed model in stock price forecasting

In this section, we present a type-2 fuzzy model for data analysis of stock price of an automotive manufacturing in Asia. The input and output candidates of the system are shown in Table 2. In this table, close price is the output variables.

The data of the above automotive manufacturing's stock price is modeled into a multiple-input–single-output (MISO) system.

The steps of the development of the MISO model are as follows:

1. For variable selection, the Sugeno and Yasukawa (1993) method (the variable selection algorithm) is used. In this case, we begin with a fuzzy model with one input. We generate 18 models: one model for one particular input then RC of each model calculated and selected one model to minimize RC from among the one input models. Next, we fix the one input selected above and add another input to our fuzzy model from among the remaining 17 candidates. At this stage our fuzzy model has two inputs. The second input is selected in the first step, according to the value of RC. The above process is continued until the value of RC increases. The result is shown in Figs. 12 and 13. As it is shown, *open price* is selected at the first step, *Price Channel* (*bottom*) at the second step, *Changes* at the third step, *DI+* at the fourth step and *DI-* at the fifth step. At the sixth step, all of the values of RC for the sixth input are bigger than the minimal RC at the fifth step. So the search is terminated at this stage. Using this method, the variables *Open price*, *Price Channel* (*bottom*), *Changes*, *DI+*, *DI−*, are selected.
2. GK algorithm is implemented to cluster the stock price. Then the proposed cluster validity index based on similarity measure ($V_{\mathrm{FNT}}$) is implemented to determine the most suitable number of clusters or rules ($c$). Here the initial cluster centers were selected by the FCM. As shown in Fig. 14, the best number of clusters based on this cluster validity index is obtained by the minimum value of the index. This result is 8 clusters. So, the type-1 system contains of 8 rules.
3. The output space is projected onto the input spaces to select the most critical inputs and the membership functions of input and output are assigned by estimation, using genetic algorithm. It is assumed that inputs and output membership functions are Gaussian.
4. For generating interval type-2 fuzzy rule bases that the antecedent and consequent sets are interval type-2 sets, a Gaussian primary MF is implemented with uncertain mean and fixed standard deviation.

In this research 365 data points have been selected where, 265 data points are used for generating rules and the rest for testing the model. After clustering, we find the optimal number of the clusters to be 8. The number of rules is 8, too Fig. 15.

### 4.1. Proposed type-2 fuzzy model

We create an interval type-2 FLS from the type-1 FLS. Similar to the type-1 FLS the interval type-2 FLS uses singleton fuzzification, product t-norm, product inference, and center-of-sets type-reduction. It also uses the same number of fuzzy sets and the same rules as the type-1

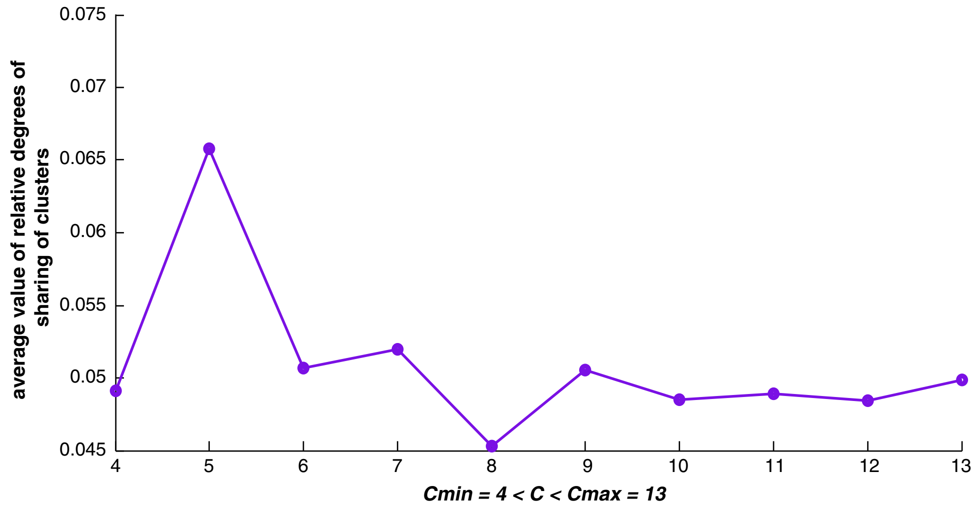



Fig. 14. Behaviour of FNT in automotive manufactory model.

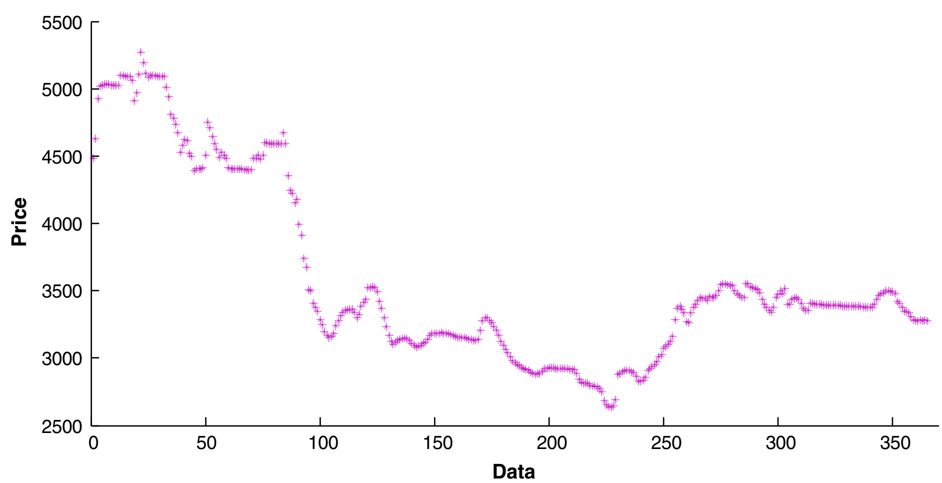



Fig. 15. Automotive manufactory data set.

FLS. The only difference now is that the antecedent and consequent sets are type-2 which has a fixed standard deviation and an uncertain mean that takes on values in an interval, i.e., (Mendel, 2000).

In Fig. 16 the interval type-2 rule base of automotive manufactory is shown. As shown in this figure, there are five inputs (*Open Price*, *Price Channel* (*bottom*), *Changes*, *DI+*, *and DI−*), one output (***close***) and eight rules.

Fig. 17 shows the comparison between observed data and the output of the proposed model. Also it can be observed that the type reduction output which is lower and upper bound of the output of the model and the crisp output is the average of them.

For validation of system, we compare the result of the proposed type-2 fuzzy model with the results of the following models:

*A. Multiple regressions*

We have used the regression analysis with "Minitab". The regression equation is:

$$\begin{aligned} y = {} & -0.000005 - 0.000000x_1 - 0.000000x_2 + 0.000010x_3 \\ & + 0.000000x_4 + 0.000003x_5 - 0.000000x_6 + 0.000000x_7 \\ & + 0.000001x_8 + 0.000000x_9 - 0.000000x_{10} + 0.000000x_{11} \\ & - 0.000000x_{12} + 0.000000x_{13} + 1.00x_{14} + 0.000000x_{15} \\ & - 0.000000x_{16} + 0.000000x_{17} + 1.00x_{18} \end{aligned}$$

After we have found a linear method for showing the relationship between these 19 parameters we must also look at the residual plots to examine the formula. As it is observed in the Fig. 18, the linear formula for this relationship is a good estimation.

*B. Type-1 fuzzy model*

In type-1 fuzzy model we use Sugeno software and obtain a fuzzy model with five rules, five inputs and one output. The inputs are *Open Price*, *Price Channel* (*bottom*), *Changes*, *DI+* and *DI−* and the output is *Close Price*. We use Mamdani-style inference, min–max, sum–product

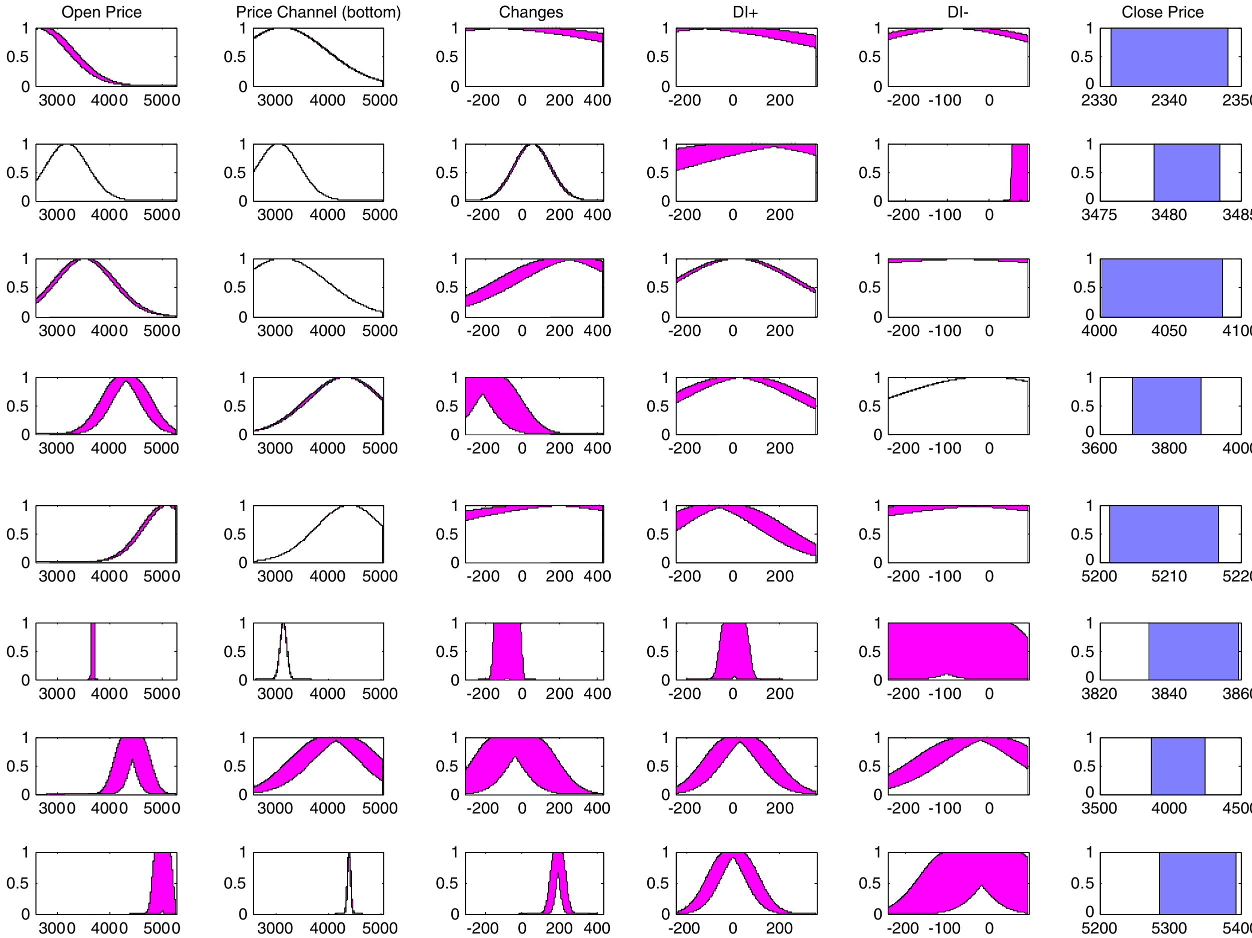



Fig. 16. Interval type-2 rule base of automotive manufactory.

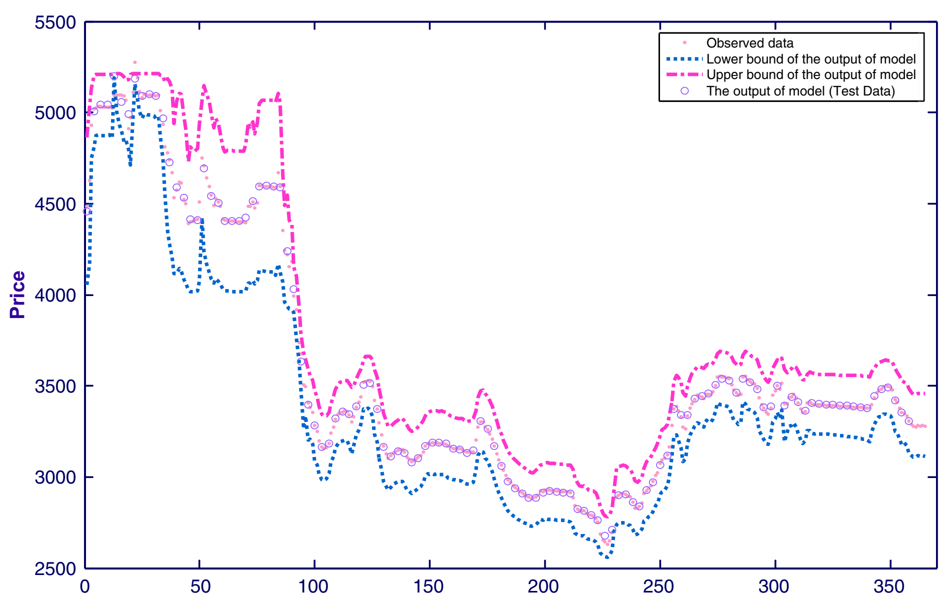



Fig. 17. Comparison between observed data and the output of the proposed model.

operators and some defuzzification methods such as *centroid*, *bisector*, *mom* (middle of maximum), *som* (smallest of maximum), *lom* (largest of maximum), or Custom, for a custom operation.

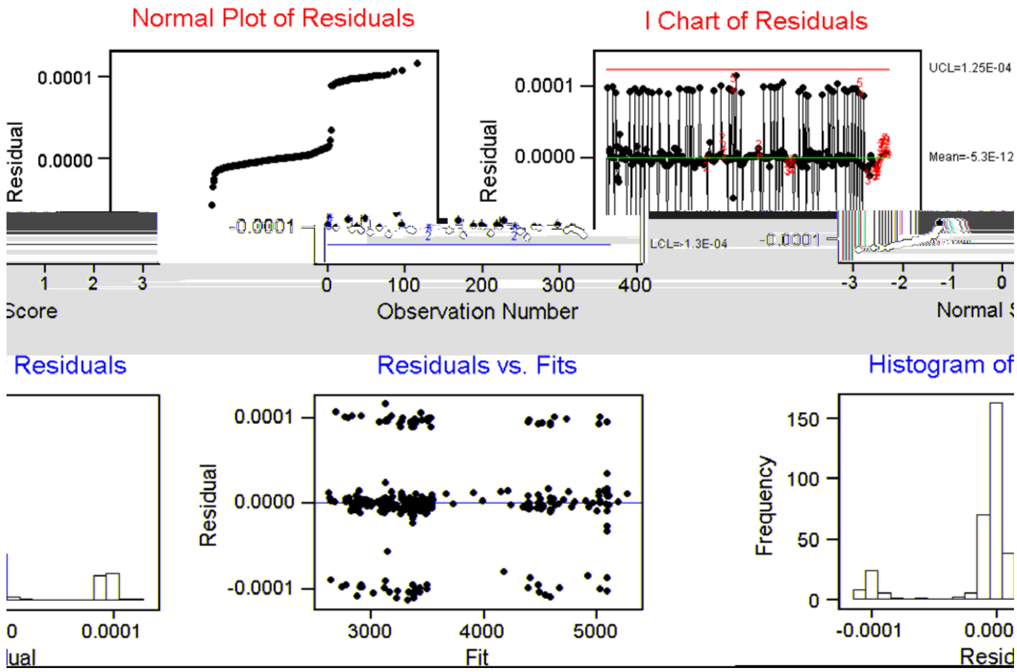



Fig. 18. The residual plots.

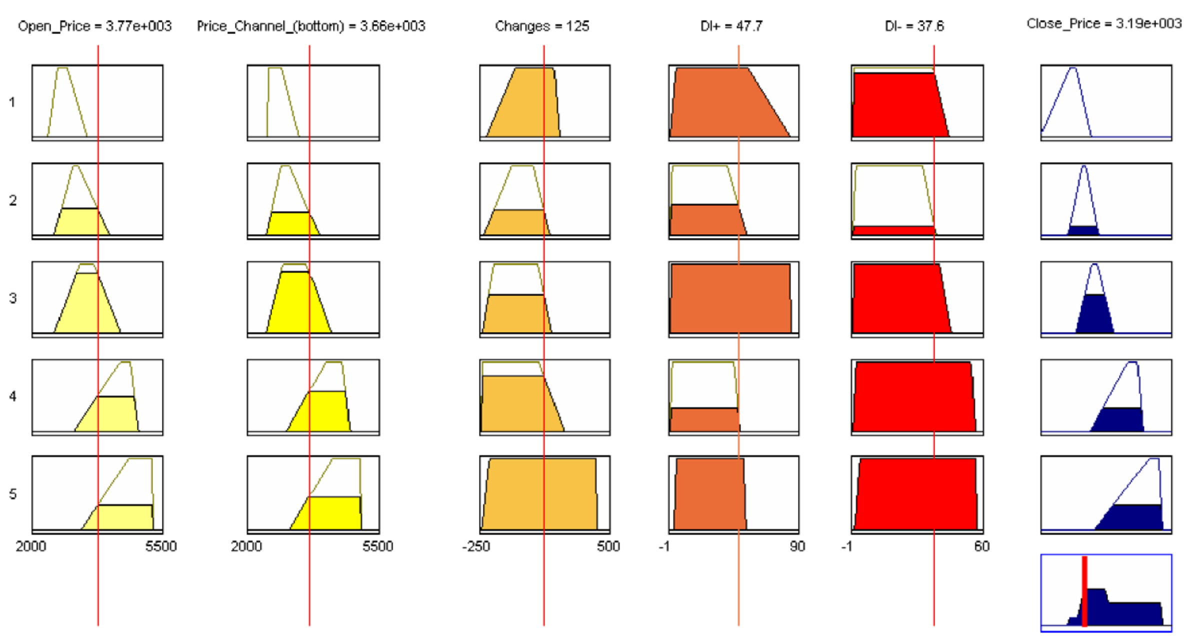



Fig. 19. Type-1 FLS for automotive manufactory data.

The best result of this system is obtained by min–max operators and *som* (smallest of maximum) defuzzification method. The error of this model is 153.13. This rule based system in is demonstrated Fig. 19.

The comparison of the proposed type-2 fuzzy model with that of Sugeno and Yasukawa (1993) and multiple regression approach is shown in Table 3. The Table shows that while the *root mean square error* (RMSE) of multiple

Table 3
Comparing between models

| | Root mean square error (RMSE) | Number of rules |
|---|---|---|
| Multiple regression | 309.17 | – |
| Sugeno–Yasukawa's model | 153.13 | 5 |
| Proposed model | 14.21 | 8 |

regression is 309.17 and that of Sugeno and Yasukawa's model is 153.13, the RMSE of our model is 14.21, which is rather improved. where

$$\mathrm{RMSE} = 2\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_i^*)^2} \qquad (29)$$

## 5. Conclusion

In this paper, interval type-2 fuzzy system for stock price prediction was presented. After investigating the system domain, the inputs and output of the system were determined.

Indirect approach is used to fuzzy system modelling by implementing the proposed cluster validity index for determining the number of rules in fuzzy clustering approach. Then Sugeno and Yasukawa method was used to select the most important variables for the rule-base fuzzy logic system. Next, the output membership values were projected onto the input spaces to generate the membership values of input variables, and the membership functions of inputs and output were tuned by genetic algorithm. Then, the type-1 method has been implemented for inference. After that, for increasing the robustness of the system, we transformed the type-1 fuzzy rule base in to an interval type-2 fuzzy rule base. For generating interval type-2 fuzzy rule base the Gaussian primary MF with uncertain mean and fixed standard deviation was used; then interval type-2 inference engine, product t-norm, sum s-norm and the center-of-set type reduction was used after that Center-of-Set defuzzification was done to get the result.

The proposed system shows its superiority with respect to robustness, flexibility and error minimization. The system may be used by financial institutes to price financial derivatives. Also it can be used by speculators for more profitable trading in stock market.

## References


Abraham, A., Baikunth, N., & Mahanti, P. K. (2001). Hybrid intelligent systems for stock market analysis. *Lecture Notes in Computer Science, 2074*, 337–345.

Abraham, A., Philip, N. S., & Saratchandran, P. (2003). Modeling chaotic behavior of stock indices using intelligent paradigms. *Neural, Parallel and Scientific Computations, 11*, 143–160.

Aiken, M., & Bsat, M. (1999). Forecasting market trends with neural networks. *Information Systems Management, 16*(4), 42–48.

Baba, N., Inoue, N., & Asakawa, H. (2000). Utilization of neural networks and GAs for constructing reliable decision support systems to deal stocks. In *IEEE–INNS–ENNS international joint conference on neural networks (IJCNN'00) 5* (pp. 5111–5116).

Baduska, R. (1995). New approach to constructing fuzzy relational models from data, In *Proceedings of European control conference, Rome, Italy* (pp. 1207–1212).

Baduska, R. (1998). *Fuzzy modeling for control*. Kluwer Academic Publishers.

Bezdek, J. C. (1974). Cluster validity with fuzzy set. *Journal of Cybernetics, 3*(3), 58–72.

Bezdek, J. C. (1975). Mathematical models for systematic and taxonomy, In *G. Estabrook (Ed.), Proceedings of the eighth international conference on numerical taxonomy, Freeman, San Franscisco, CA* (pp. 143–166).

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. NY, USA: Plenum Press.

Chang, P. C., & Liu, C. H. (2006). A TSK type fuzzy rule based system for stock price prediction, *Expert Systems with Applications*, On-line available.

Chang, P. C., Liu, C. H., & Wang, Y. W. (2006). A hybrid model by clustering and evolving fuzzy rules for sale forecasting in printed circuit board industry. *Decision Support Systems, 42*(3), 1254–1269.

Chang, P. C., & Wang, Y. W. (2006). Fuzzy delphi and back-propagation model for sales forecasting in PCB Industry. *Expert Systems with Applications, 30*(4), 715–726.

Chang, P. C., Wang, Y. W., & Yang, W. N. (2004). An investigation of the hybrid forecasting models for stock price variation in Taiwan. *Journal of the Chinese Institute of Industrial Engineering, 21*(4), 358–368.

Chi, S. C., Chen, H. P., & Cheng, C. H. (1999). A forecasting approach for stock index future using Grey theory and neural networks. In *IEEE international joint conference on neural networks* (pp. 3850–3855).

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. Wiley and Sons Inc.

Fazel Zarandi, M. H. (1998). Aggregate system analysis for prediction of tardiness and mixed zones of continuous casting with fuzzy methodology. PhD thesis.

Friedman, H. P., & Rubin, J. (1996). On some invariant criteria for grouping data. *Journal of the American Statistical Association, 62*, 1159–1178.

Fukuyama, Y., & Sugeno, M. (1989). A new methodof choosing the number of clusters for the fuzzy c-means method. In *Proceedings of the fifth fuzzy systems symposium* (pp. 247–250).

Gath, I., & Geva, A. B. (1989). Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*(7), 773–781.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning*. Reading, MA: Addison-Wesley.

Gustafson, D., & Kessel, W. (1979). Fuzzy clustering with a fuzzy covariance matrix. In *Proceedings IEEE CDC* (pp. 761–766). San Diego, CA, USA.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. MI: University of Michigan Press.

Jain, A. K., & Dubes, R. C. (1980). *Algorithms for clustering data*. Englewood Cliffs, NJ: Prentice-Hall.

Karnik, N. N., & Mendel, J. M. (1998). Introduction to type-2 fuzzy logic systems. In *Proceedings of the 1998 IEEE fuzzy conference, May 1998* (pp. 915–920).

Mendel, J. M. (2001). *Uncertain rule-based fuzzy logic systems: introduction and new directions*. Upper-Saddle River, NJ: Prentice-Hall.

Mendel, J. M. (2007). Advances in type-2 fuzzy sets and systems. *Information Sciences, 177*(1), 84–110.

Mendel, J. M., & John, R. I. (2002). Type-2 fuzzy sets made simple. *IEEE Transaction on Fuzzy Systems, 10*, 117–127.

Mendel, J. M., John, R. I., & Liu, F. (2006). Interval type-2 fuzzy logic systems made simple. *IEEE Transactions on Fuzzy Systems, 14*(6), 808–821.

Razaee, M. R., Lelieveldt, B. P. F., & Reiber, J. H. C. (1998). A new cluster validity index for the fuzzy c-means. *Pattern Recognition Letters, 19*, 237–246.

Sakawa, M. (2002). *Genetic algorithms and fuzzy multiobjective optimization*. Boston: Kluwer Academic.

Sugeno, M., & Yasukawa, T. (1993). A fuzzy logic based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems, 1*, 7–31.

Tansel, I. N., Yang, S. Y., Venkataraman, G., Sasirathsiri, A., Bao, W. Y., & Mahendrakar, N. (1999). Modeling time series data by using neural netwroks and genetic algorithms. In C. H. Dagli, A. L. Buczak, J. Ghosh, M. J. Embrechts, & O. Ersoy (Eds.), *Smart engineering system design: neural networks, fuzzy logic, evolutionary programming, data mining, and complex systems*. *Proceedings of the Artificial Neural Networks in Engineering Conference (ANNIE '99)* (pp. 1055–1060). New York: ASME Press.

Thammano, A. (1999). Neuro-fuzzy model for stock market prediction. In C. H. Dagli, A. L. Buczak, J. Ghosh, M. J. Embrechts, & O. Ersoy (Eds.), *Smart engineering system design: neural networks, fuzzy logic, evolutionary programming, data mining, and complex systems*. *Proceedings of the artificial neural networks in engineering conference (ANNIE '99)* (pp. 587–591). New York: ASME Press.

Wu, D., & Tan, W. W. (2004). A type-2 fuzzy logic controller for the liquid-level process. In *Proceedings of IEEE international conference on fuzzy systems, Budapest, Hungary, July 2004* (pp. 953–958).

Xie, X. L., & Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*(8), 841–847.

Yao, J., & Poh, H. L. (1995). Forecasting the KLSE index using neural networks. *IEEE International Conference on Neural Networks, 2*, 1012–1017.

Yoon, Y., & Swales, J. (1991). Prediction stock price performance: A neural network approach. *Proceedings of twenty-fourth annual hawaii international conference on system science*, 156–162.

Zahid, N., Limouri, M., & Essaid, A. (1999). A new cluster-validity for fuzzy clustering. *Pattern Recognition, 32*, 1089–1097.