An improved sales forecasting approach by the integration of genetic fuzzy systems and data clustering: Case study of...

Article *in* Expert Systems with Applications · August 2011

DOI: 10.1016/j.eswa.2011.01.132 · Source: DBLP

Esmaeil Hadavandi
Amirkabir University of Technology

Hassan Shavandi
Sharif University of Technology

Expert Systems with Applications 38 (2011) 9392–9399

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

An improved sales forecasting approach by the integration of genetic fuzzy systems and data clustering: Case study of printed circuit board

Esmaeil Hadavandi [a], Hassan Shavandi [a,*], Arash Ghanbari [b]

[a] Department of Industrial Engineering, Sharif University of Technology, P.O. Box 11365-9466, Tehran, Iran
[b] Department of Industrial Engineering, College of Engineering, University of Tehran, P.O. Box 11155-4563, Tehran, Iran

ARTICLE INFO

Keywords:
Sales forecasting
Genetic fuzzy system
Data clustering
Printed circuit boards
Hybrid intelligence approach

ABSTRACT

Success in forecasting and analyzing sales for given goods or services can mean the difference between profit and loss for an accounting period and, ultimately, the success or failure of the business itself. Therefore, reliable prediction of sales becomes a very important task. This article presents a novel sales forecasting approach by the integration of genetic fuzzy systems (GFS) and data clustering to construct a sales forecasting expert system. At first, all records of data are categorized into k clusters by using the K-means model. Then, all clusters will be fed into independent GFS models with the ability of rule base extraction and data base tuning. In order to evaluate our K-means genetic fuzzy system (KGFS) we apply it on a printed circuit board (PCB) sales forecasting problem which has been used as the case in different studies. We compare the performance of an extracted expert system with previous sales forecasting methods using mean absolute percentage error (MAPE) and root mean square error (RMSE). Experimental results show that the proposed approach outperforms the other previous approaches.

1. Introduction and literature review

Sales is the most direct measure of the outcome of marketing efforts and so sales forecasting models play a significant role in marketing planning. In the very competitive and dynamic environment that most businesses face, forecasting is a beneficial tool and an indispensable strategy for business survival. Reliable prediction of sales can improve the outcome of business, because it allows functional areas, such as production, sales and marketing, and finance, to effectively develop programs to grow the company. Examples of these programs include sales planning, budgeting, and promotion and advertising plan.

Artificial intelligence techniques such as, artificial neural networks (ANNs), fuzzy logic, and genetic algorithms (GAs) are popular techniques to solve the complex engineering and optimization problems (Konar, 2005). These techniques can be combined together in various ways to form hybrid techniques. Examples of such hybrids in the literature are a combination of neural networks and fuzzy systems in marketing (Kuo & Chen, 2004; Kwong, Wong, & Cha, 2009; Li, 2000) and also in stock price prediction (Atsalakis & Valavanis, 2009).

Hybrid models have more flexibility and can be used to estimate the non-linear relationship, without the limits of traditional models such as Time Series models. Therefore, more and more researchers tend to use hybrid forecasting models to deal with forecasting problems.

Aiken and Bsat (1999) used a feed forward neural networks trained by a genetic algorithm to forecast three-month US Treasury Bill rates. They concluded that hybrid model can be used to accurately predict these rates. Chang and Liu (2008) used a Takagi–Sugeno–Kang (TSK) type fuzzy rule based system (FRBS) for stock price prediction. They used simulated annealing (SA) for training the best parameters of fuzzy systems. They found that the forecasted results from TSK fuzzy rule base model were much better than those of back propagation network (BPN) or multiple regressions. Efendigil, Önüt, and Kahraman (2009) proposed a forecasting mechanism which was modeled by artificial intelligence approaches including the comparison of both ANNs and adaptive network-based fuzzy inference systems (ANFIS) to forecast the fuzzy demand with incomplete information.

Since most sales data are non-linear in relation and are subject to noise, researchers tend to use hybrid artificial intelligence forecasting models to deal with sales forecasting problem. Kuo and Xue (1999) utilized fuzzy neural network (FNN) in order to grasp the experts' knowledge for sales forecasting. The proposed system was also compared with the other methods, single ANN and ARMA and outperformed them.

Chang and Lai (2005) proposed a hybrid system to combine the self-organizing map (SOM) of neural network with a case-based reasoning (CBR) method, for sales forecasting of new released

* Corresponding author. Tel.: +98 21 6616 5711; fax: +98 21 6602 2702.
E-mail addresses: es.hadavandi@gmail.com (E. Hadavandi), shavandi@sharif.edu (H. Shavandi), arashghanbari@yahoo.com, arghanbari@ut.ac.ir (A. Ghanbari).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.01.132

Table 1
History of PCB sales forecasting.

Authors	Year	Case study	Methods	Accuracy	
				MAPE	RMSE
Chang, Wang and Tsai	2005	PCB industry	GA + ANN	3.13	NA
Chang and Wang	2006	PCB industry	Fuzzy logic + ANN	3.09	NA
Chang, Liu and Wang	2006	PCB industry	SOM + ANN + GA + FRBS	2.16	21,346
Chang, Wang and Liu	2007	PCB industry	WEFuNN	2.11	24,909
Chang, Liu and Lai	2008	PCB industry	FCBR	4.82	43,385
Chang, Liu and Fan	2009	PCB industry	K-means clustering + FNN	2.19	20,287

books. The result of the prediction of SOM/CBR was compared with the results of K/CBR, which is divided by K-means, and traditional CBR. They found out that the SOM/CBR is more accurate and efficient when being applied to the forecast of the data than K/CBR or traditional CBR.

Thomassey, Happiette, and Castelain (2005) proposed a hybrid sales forecasting model based on fuzzy logic, neural networks and evolutionary procedures, permitting the processing of uncertain data. The experimental results showed that the performance of the hybrid model was superior to other models such as ARMAX. Thomassey and Happiette (2007) also proposed a hybrid sales forecasting model, based on clustering and Probabilistic Neural Network in textile industry and they concluded the proposed model enabled an accurate estimation of the behavior of future sales.

One of the most popular approaches is the hybridization between fuzzy logic and GAs leading to genetic fuzzy systems (GFSs) (Cordón, Herrera, Hoffmann, & Magdalena, 2001). A GFS is basically a fuzzy system augmented by a learning process based on evolutionary computation, which includes genetic algorithms and other evolutionary algorithms (EAs) (Eiben & Smith, 2003).

In recent years some articles have been published in the favor of using GFS in marketing area (Casillas & Martínez-López, 2009a, 2009b; Martínez-López & Casillas, 2009; Orriols-Puig, Casillas, & Martínez-López, 2009). They have all obtained satisfactory results and concluded that using GFSs is very promising for marketing analysis.

There are several researches which have used GFSs for forecasting problems (Damousis & Dokopoulos, 2001; Shim, Seong, Ko, & So, 1999), but there is no research in the literature that uses a GFS with the ability of extracting the whole knowledge base of the fuzzy system for sales forecasting problems.

This article presents a hybrid artificial intelligence (AI) methodology for sales forecasting to extract useful patterns of information with a descriptive rule induction approach based on genetic fuzzy systems. Our method combines a data clustering technique and a GFS type (Cordón & Herrera, 1997) to learn the whole knowledge base of the system. We test the capability of the proposed method by applying it on a sales forecasting case study called "*printed circuit board (PCB) industry in Taiwan*" which has been frequently used by the other authors as a case.

1.1. PCB sales forecasting

Printed circuit board industry in Taiwan occupies a significant portion of Taiwan's manufacturing. There are several studies in the literature which have considered PCB sales forecasting as the case study. In the following we'll present a brief review of them.

Chang, Wang, and Tsai (2005) developed an Evolving Neural Network (ENN) forecasting model by integrating GAs and ANNs, Along with trend and seasonal factors considered by Winter's model. In this research effective economical factors are chosen by the Grey Relation Analysis (GRA). The experimental results show that the performance of ENN is superior to traditional statistical models and Back Propagation Network. Chang and Wang (2006) proposed a hybrid model by integrating fuzzy logic and ANN to construct a fuzzy back-propagation network (FBPN) for sales forecasting. Parameters which were chosen as inputs to the FBPN were no longer considered as of equal importance, but some sales managers and production control experts were requested to express their opinions about the importance of each input parameter in predicting the sales with linguistic terms, which could be converted into pre-specified fuzzy numbers. Experimental results indicated that the FBPN approach outperforms other traditional methods such as Grey Forecasting, Multiple Regression Analysis and back propagation networks in mean absolute percentage error (MAPE) measures.

Chang, Liu, and Wang (2006) developed a hybrid model by integrating SOM, ANNs, GAs and FRBS. They used a simple GA to determine the near-optimal number of fuzzy terms for each variable. They found that performance of the model was superior to previous methods that proposed for PCB sales forecasting such as Chang et al. (2005). Chang, Wang, and Liu (2007) used GRA to select a combination of key factors influencing sales and then developed a weighted evolving fuzzy neural network (WEFuNN) model for PCB sales forecasting. They concluded that this hybrid model is superior to previous hybrid models.

Chang and Liu (2008) developed a multi-stage framework for combining cased-based reasoning and fuzzy multicriteria decision making as a decision support model. The experimental results showed that performance of the fuzzy case based reasoning (FCBR) model is superior to traditional statistical models and BPN. Chang, Liu, and Fan (2009) used a K-means clustering and fuzzy neural network (FNN) to forecast the future sales of PCB. They used stepwise regression to select many possible variables which may influence PCB monthly sales amount, then used K-means for clustering data in different clusters to be fed into independent FNN models. The forecasted results were compared with other forecasting models such as BPN, ANFIS and FNN. Results showed that the proposed approach outperforms them.

Table 1 summarizes various methods developed dealing with PBC sales forecasting.

2. K-means genetic fuzzy system

As mentioned before, this article presents a hybrid artificial intelligence method called K-means genetic fuzzy system (KGFS) to construct an expert system (ES) for sales forecasting problems. We apply a combination of K-means data clustering, GA and fuzzy logic approach to construct the ES. At first we apply K-means clustering technique to cluster our raw data. Then data in different clusters divided by K-means will be fed into independent genetic fuzzy systems.

The framework of KGFS is shown in Fig. 1, and the details of each stage are described below.

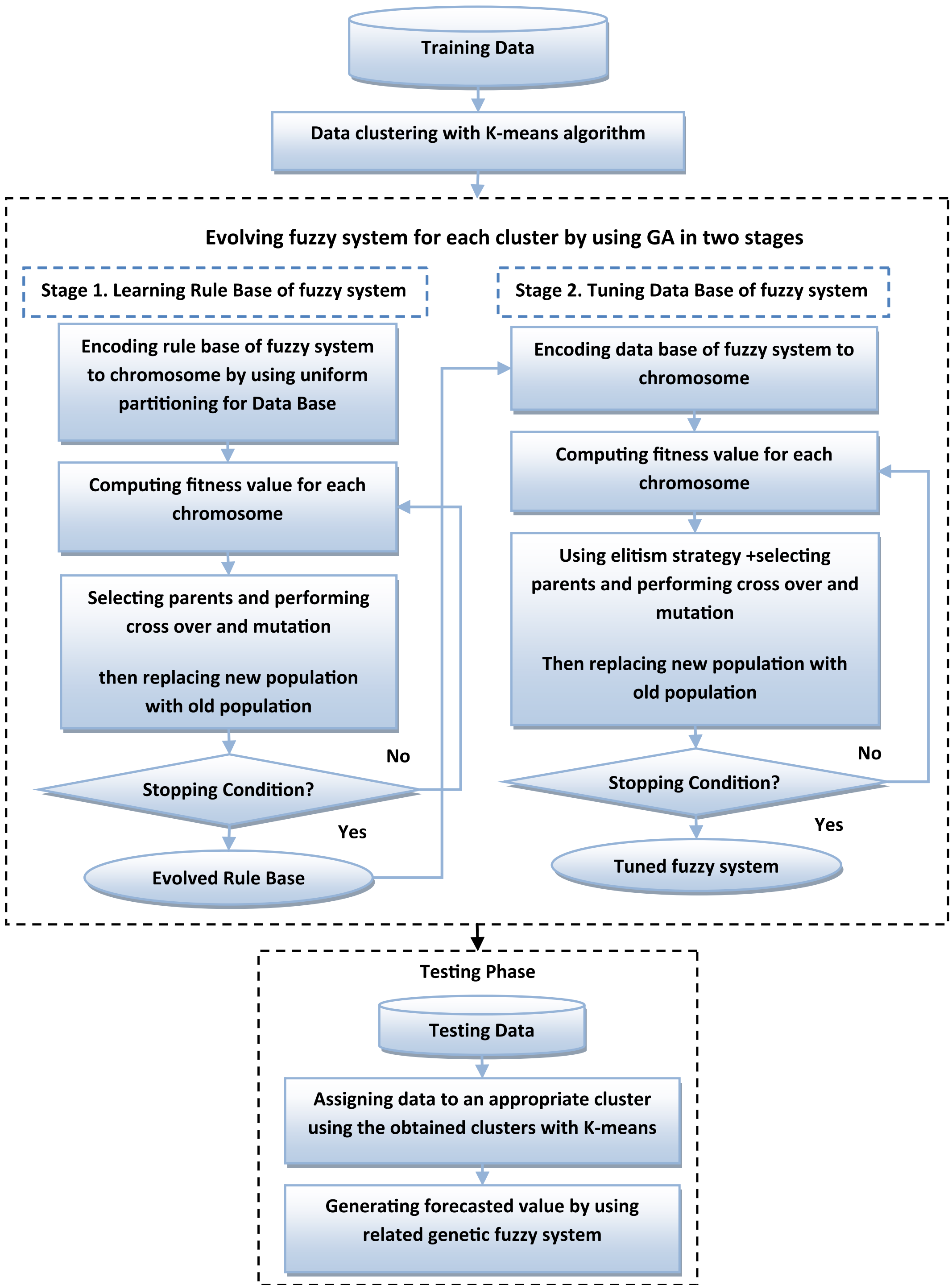

Fig. 1. Framework of KGFS.

2.1. K-means clustering analysis

The K-means method (Anderberg, 1973) is a technique employed for partitioning a set of objects into k groups such that each group is homogeneous with respect to certain attributes based on the specific criterion. We use K-means data clustering to reduce the effect of data noise and have a better forecasted accuracy. The procedures of the K-means method can be summarized as follows:

Step 1: Randomly select k initial cluster centroids, where k is the number of the clusters.
Step 2: Assign each object to a cluster which it is the closest based on the distance between the object and the cluster mean.
Step 3: Calculate the new mean for each cluster and reassign the objects to the clusters.
Step 4: Stop if the criterion converges. Otherwise go back to Step 2.

The Tanagra 1.4 was used for applying K-means clustering in this research (Rakotomalala, 2005).

2.2. Developing a genetic fuzzy system

Nowadays fuzzy rule-based systems have been successfully applied to a wide range of real-world problems from different areas. In order to design an intelligent system of this kind for a concrete application, several tasks have to be performed. One of the most important and difficult ones is to derive an appropriate knowledge base (KB) about the problem. The KB stores the available knowledge in the form of fuzzy linguistic IF–THEN rules. It is composed of the rule base (RB), constituted by the collection of rules in their symbolic forms, and the data base (DB), which contains the linguistic term sets and the membership functions defining their meanings (Casillas, Cordón, Herrera, & Villar, 2004).

The difficulty presented by human experts to express their knowledge in the form of fuzzy rules has made researchers develop automatic techniques to perform this task. In this sense, a large number of methods has been proposed to automatically generate fuzzy rules from numerical data. Usually, they use complex rule generation mechanisms such as neural networks (Nauck, Klawonn, & Kruse, 1997) or genetic algorithms (Cordón et al., 2001).

GAs have been demonstrated to be a powerful tool for automating the definition of the KB, since adaptive control, learning, and self-organization may be considered in a lot of cases as optimization or search processes. In particular, the application to the design, learning, and tuning of KBs has produced quite promising results. These approaches can be given the general name of genetic fuzzy systems (Cordon & Herrera, 1995).

The GFS type which we use in this article consists of two general stages; stage 1 derives rule base of FRBS and stage 2 tunes data base of FRBS. In the following we'll describe theses two stages.

2.2.1. Genetic derivation of the rule base for FRBS

A previously defined DB constituted by uniform fuzzy partitions with triangular membership functions crossing at height 0.5 is considered. The number of linguistic terms forming each one of them can be specified by the GFS designer, and then Pittsburgh approach is used for learning RB. Each chromosome encodes a whole fuzzy rule set and the derived RB is the best individual of the last population. Pittsburgh approach can be decomposed in the following steps.

Step 1 – Coding mechanism.
Many GFSs employ the decision table proposed by (Thrift, 1991) as the common classical representation for the RB of an FRBS. A fuzzy decision table represents a special case of a crisp relation (the ordinary type of relations we are familiar with) defined over the collections of fuzzy sets corresponding to the input and output variables. A chromosome is obtained from the decision table by going row-wise and coding each output fuzzy set as an integer number with start from 1 to number of output variable linguistic terms. Fuzzy decision table for an FRBS with two inputs (X_1, X_2) and one output (Y) variable, with three fuzzy sets ($A_{11}, A_{12}, A_{13}, A_{21}, A_{22}, A_{23}$) related to each input variable and four fuzzy sets (B_1, B_2, B_3, B_4) related to the output variable and Applying this code to the fuzzy decision table represented in Fig. 2.
Step 2 – Generating the initial population.
Initial chromosomes (N_{pop}) are randomly generated; while the alleles are in the set $\{1, 2, \ldots, N_B\}$ (N_B is the number of output variables' linguistic terms). All consequent labels have the same probability to be assigned to each gene.
Step 3 – Calculating the fitness values.
As regards the fitness function, it is based on an application-specific measure usually employed in the design of GFSs, the mean squared error (MSE) over a training data set, which is represented by the following expression:

$$\mathrm{MSE}(C_j) = \frac{1}{N}\sum_{i=1}^{N}(Y_i - P_i)^2$$

where Y_i is the actual value and P_i is the output value of ith training data obtained from the FRBS using the RB coded in jth chromosome (C_j) and N is the number of training data.
Step 4 – Generate ($N_{pop} - 1$) chromosome using the genetic operations.
We use a binary tournament selection scheme for the selection procedure. In binary tournament selection, two members of the population are selected at random and their fitness compared and the best one according to fitness value will be chosen to reproduce. Also we use one-point crossover and uniform mutation for genetic operators.

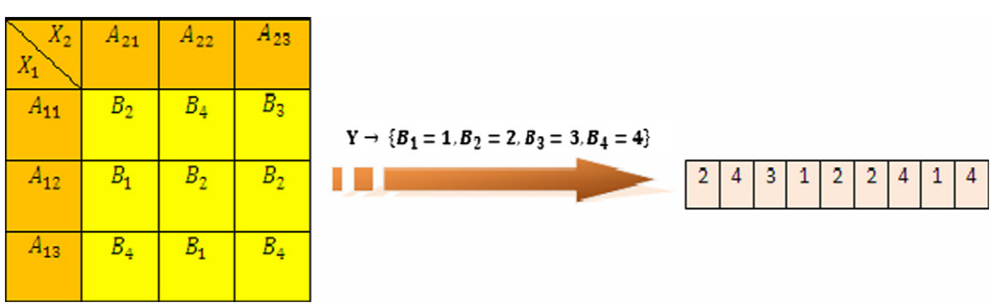

Fig. 2. Coding decision table as a chromosome.

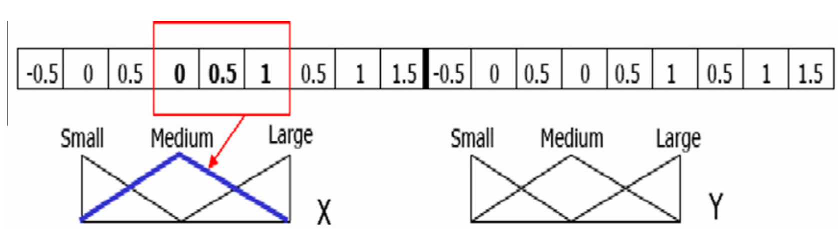

Fig. 3. Coding data base as chromosomes.

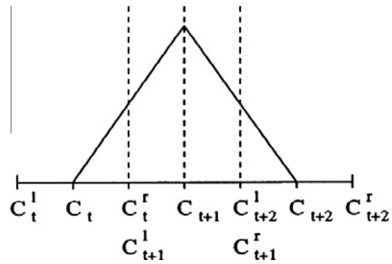

Fig. 4. Membership function and interval of performance for the tuning process.

Table 2
Suitable features of KGFS.

GFS-optimum feature	Cluster 1	Cluster 2	Cluster 3
Number of labels	2	2	2
Population size	95	75	80
Number of generations	30	3460[a]	50
Cross probability	0.6	0.75	0.85
Mutation probability	0.08	0.04	0.04

[a] Variable ranges in cluster 2 were much more larger than the other clusters, so searching in such variable space and learning the patterns needs more generations.

Step 5 – Add the best rule set in the current population to the newly generated $(N_{pop} - 1)$ chromosome to form the next population.
Step 6 – If the number of generations equals the maximum generation number, then stop; otherwise go to *step 3*.

Table 3
KGFS forecasted values vs. actual values of PCB sales data.

Month	Actual values	KGFS forecasts
2003/1	649,066	645649.2
2003/2	466,750	462041.4
2003/3	633,615	636362.1
2003/4	693,946	701704.2
2003/5	785,838	799244.6
2003/6	679,312	678026.7
2003/7	723,914	730172.5
2003/8	757,490	755321.4
2003/9	836,846	848193.6
2003/10	833,012	852101.9
2003/11	860,892	849898.4
2003/12	912,182	852563.3

2.2.2. Genetic tuning of data base.

After generation of rule base, we utilize the genetic tuning process that was proposed by Cordón and Herrera (1997). This tuning process slightly adjusts the shape of the membership functions of a preliminary DB defined. This approach can be decomposed in the following steps.

Step 1 – Coding data base as Chromosomes.
Each chromosome encodes a different DB definition. We use triangular membership functions for input and output variables' linguistic terms. Each triangular membership function is encoded by 3 real values, a primary fuzzy partition is represented as an array composed of 3*N* real values, with *N* being the number of linguistic terms for each variable (we take the same number of linguistic terms for each input and output

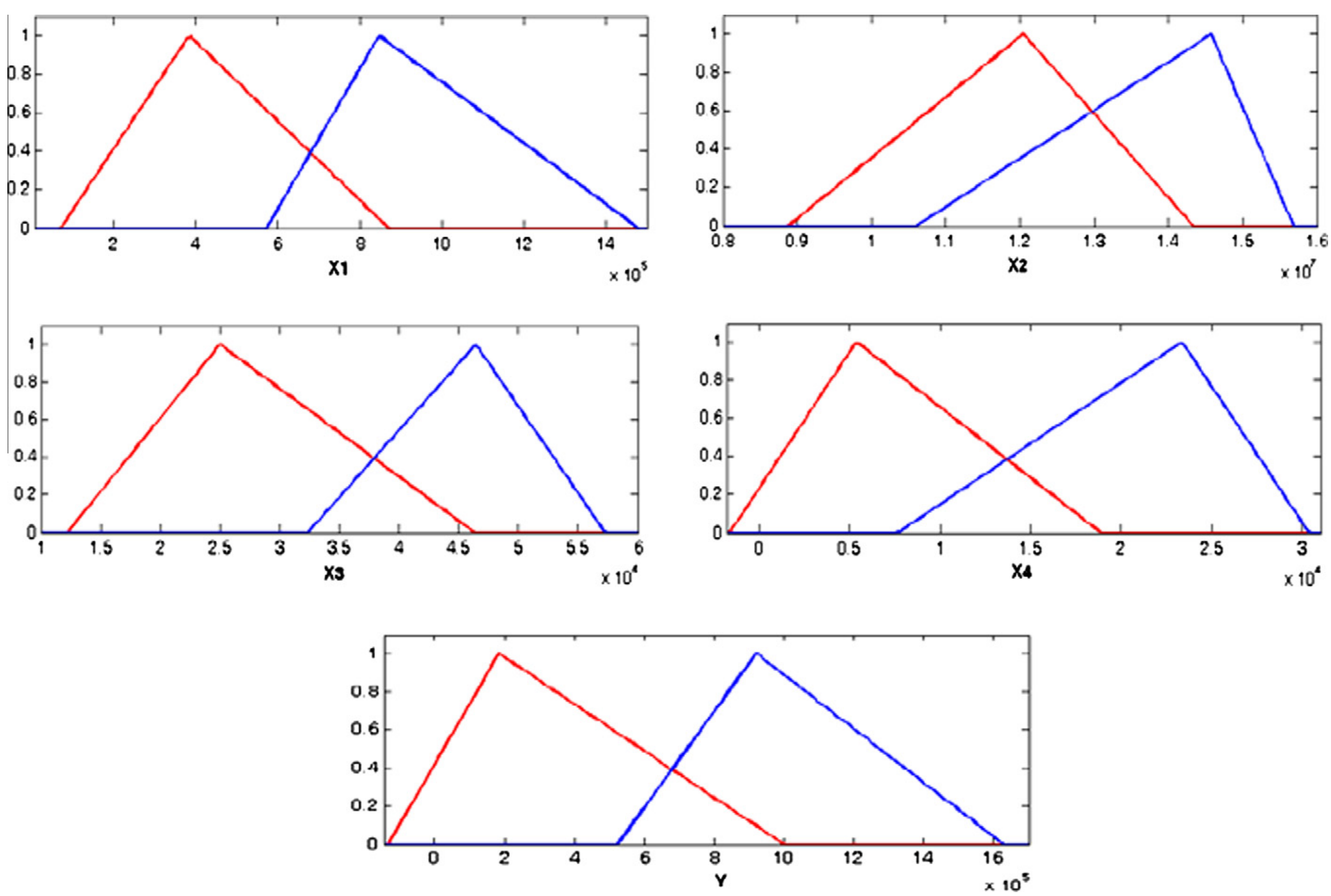

Fig. 5. The tuned membership functions of input and output variables for cluster 1's GFS.

linguistic variable). The complete DB for a problem, in which Mlinguistic variables are involved, is encoded into a fixed-length real coded chromosome C_j built by joining the partial representations of each one of variable fuzzy partitions as is shown in the following:

$$C_i^j = (a_{i1}^j, b_{i1}^j, c_{i1}^j, a_{i2}^j, b_{i2}^j, c_{i2}^j, \ldots, a_{iN}^j, b_{iN}^j, c_{iN}^j) \quad (1)$$

$$C_j = C_1^j C_2^j C_3^j \cdots C_{M-1}^j C_M^j \quad (2)$$

A sample coded data base with one input variable as well as one output variable is shown in Fig. 3. Each variable is defined by a fuzzy linguistic term such as small, medium, and large.

Step 2 – Generating the initial population.

The initial population (N_{pop}) is created by using the initial DB definition. The first chromosome (C_1) is encoded directly from initial DB definition. The remaining individuals ($N_{pop} - 1$) are generated by associating an interval of performance, $[c_h^l, c_h^r]$ to every gene c_h in C_1, $h = 1, 2, \ldots, 3NM$. Each interval of performance will be the interval of adjustment for the corresponding variable, $c_h \in [c_h^l, c_h^r]$.

If $t \bmod 3 = 1$, then c_t is the left value of the support of a triangular fuzzy number. The triangular fuzzy number is defined by the three parameters (c_t, c_{t+1}, c_{t+2}) and the intervals of performance are as follows:

$$c_t \in [c_t^l, c_t^r] = \left[c_t - \frac{c_{t+1} - c_t}{2}, c_t + \frac{c_{t+1} - c_t}{2}\right]$$
$$c_{t+1} \in [c_{t+1}^l, c_{t+1}^r] = \left[c_{t+1} - \frac{c_{t+1} - c_t}{2}, c_{t+1} + \frac{c_{t+2} - c_{t+1}}{2}\right]$$
$$c_{t+2} \in [c_{t+2}^l, c_{t+2}^r] = \left[c_{t+2} - \frac{c_{t+2} - c_{t+1}}{2}, c_{t+2} + \frac{c_{t+3} - c_{t+2}}{2}\right]$$

Fig. 4 shows these intervals.

Step 3 – Fitness value function.

We use MSE over a training data set as fitness function. This fitness function is applied on the chromosomes considering the tuned membership functions and the rule base extracted in the previous phase.

Step 4 – Selection and elitism.

The best 10% of the population is copied without changes in the elitism set. Elitism set ensurs that the best chromosomes will not be destroyed during crossover and mutation. The selection process is then implemented. We use binary tournament selection scheme to select chromosomes for mating pool. The size of the mating pool equals 90% of the population size.

Step 5 – Genetic operators.

We use *BLX*-0.1 crossover (Eshelman & Schaffer, 1993) and uniform mutation in the proposed genetic tuning process.

Step 6 – Replacement.

The current population is replaced by the newly generated offsprings, which forms the next generation by integrating the elitism set.

Step 7 – Stopping criteria.

If the number of generations equals the maximum generation number, then stop; otherwise go to *step 3*.

3. Experimental results

In this section we develop the proposed KGFS using monthly PCB sales data from 1999/1 to 2003/12. Actually, we use the first 4 years (48 data) for training the model and the last year (12 data) for testing the model. We use the same variables used in

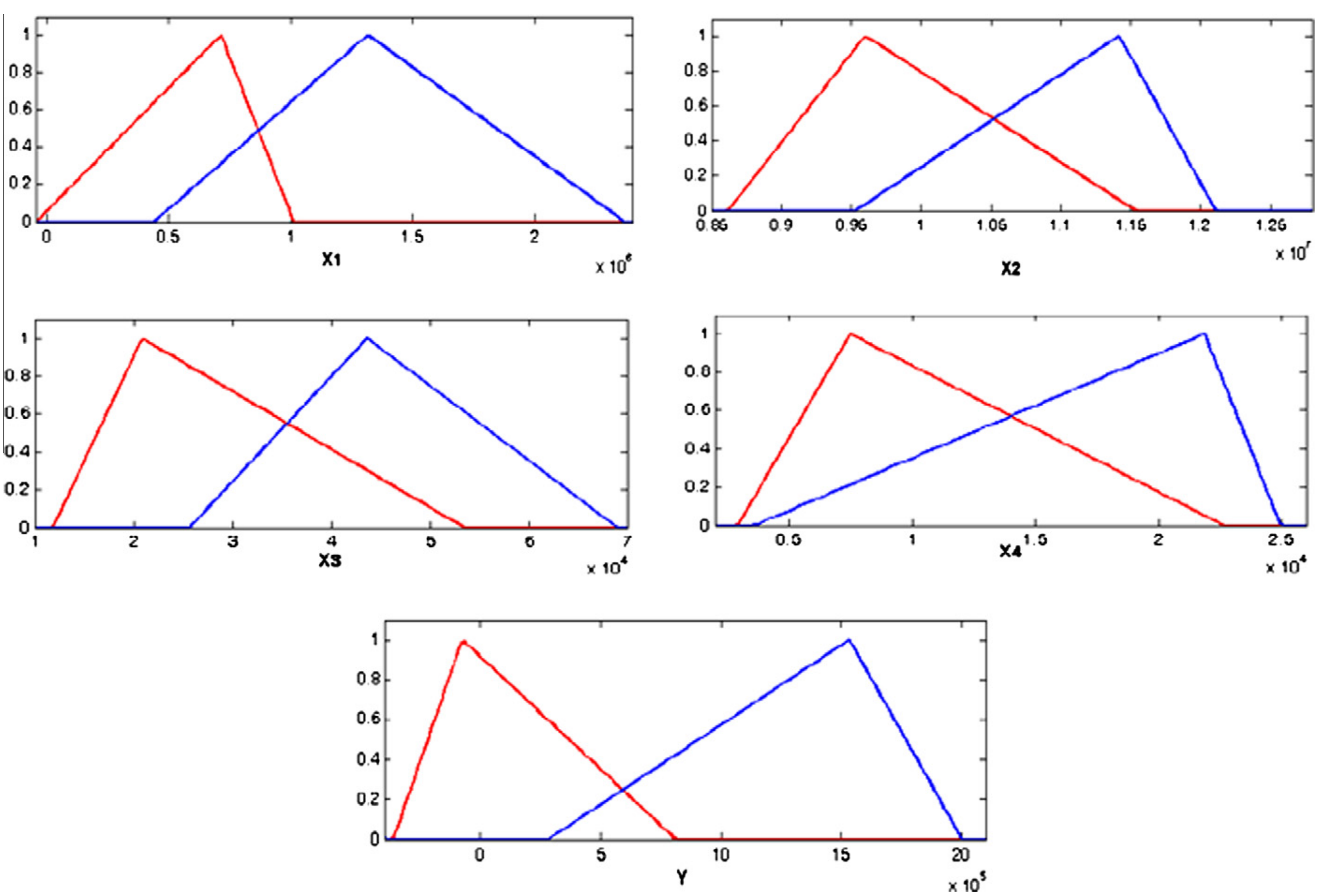

Fig. 6. The tuned membership functions of input and output variables for cluster 2's GFS.

mentioned PCB sales forecasting studies. The influential variables are as follows:

X_1 : Preprocessed historical data that addresses the effects of trend, seasonality and random noise;
X_2 : Consumer price index;
X_3 : Liquid crystal element demand;
X_4 : Total production value of PCB;
Y : Actual historical monthly PCB sales Data.

where X_i ($i = 1, 2, 3, 4$) represent input values and Y is the output value. For more information on model variables refer to Chang, Liu, and Fan (2009).

3.1. Constructing KGFS sales forecasting ES

In the first stage, all records of data are inputted into the K-means model and three different clusters are generated (we obtained data based on some experiments that in our case the suitable number of clusters (k) is 3), where each cluster contains a portion of train and test data. In the second stage we build a GFS for each cluster using related training data. Finally, sales forecasting will be done by means of each cluster's test data. The suitable features of GFSs for each cluster after examination of different values are shown in Table 2.

The proposed KGFS approach was applied to forecasting the sales data of the PCB case and the results are summarized in Table 3.

The tuned membership functions of input and output variables for GFSs of cluster 1, cluster 2 and cluster 3 are shown in Figs. 5–7, respectively.

3.2. Comparisons of KGFS model with other previous models

For the purpose of evaluating KGFS forecasting accuracy, we will compare outputs of this method with other methods mentioned in Section 1.1. We perform this task by two evaluation statistics called mean absolute percentage error (MAPE) and root mean square error (RMSE), which are calculated as follows:

$$\mathrm{MAPE} = 100 \times \frac{1}{N}\sum_{i=1}^{N}\frac{|Y_i - P_i|}{Y_i} \quad (3)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - P_i)^2} \quad (4)$$

where Y_i is actual value and P_i is the forecasted value of ith test data obtained from KGFS model and N is the number of training data. Summary of KGFS evaluations in comparison with the other methods are shown in Table 4.

Regarding Table 4, our proposed approach has improved the forecasting accuracy of sales. Namely, KGFS has made 31% improvement in the best obtained MAPE from previous studies and also 4.3% improvement in the best obtained RMSE by them. So, KGFS outperforms the rest of methods due to MAPE and RMSE evaluations and this shows that it can be considered as a promising alternative for sales forecasting problems.

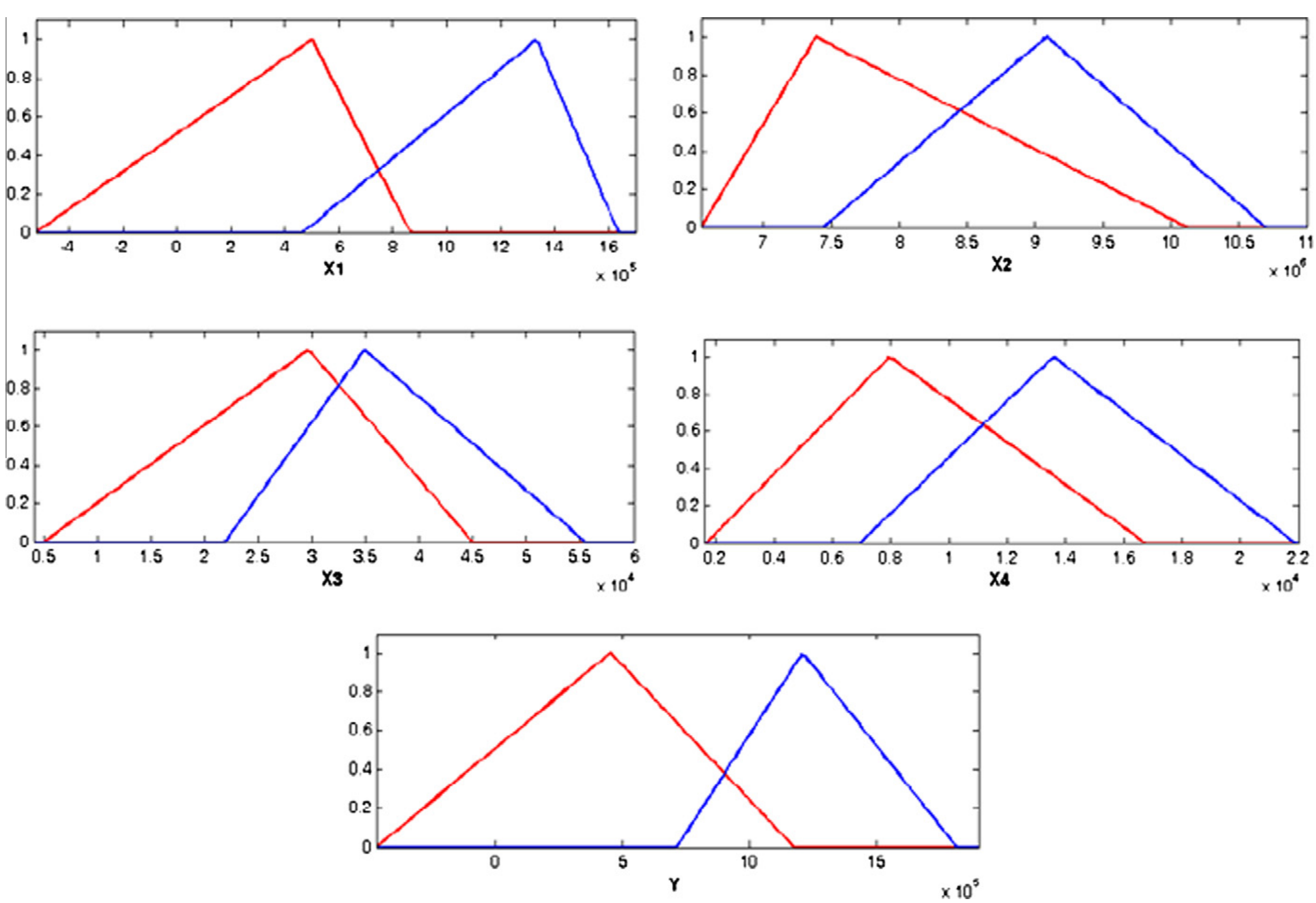

Fig. 7. The tuned membership functions of input and output variables for cluster 3's GFS.

Table 4
KGFS evaluation vs. other methods.

Method	MAPE	RMSE
Chang and Lai (2005)	3.13	NA
Chang and Wang (2006)	3.09	NA
Chang et al. (2006)	2.16	21,346
Chang et al. (2007)	2.11	24,909
Chang and Liu (2008)	4.82	43,385
Chang et al. (2009)	2.19	20,287
KGFS (the proposed method)	1.46	19,354

4. Conclusion

This paper presents a novel approach based on genetic fuzzy systems and K-means clustering (KGFS) in building a sales forecasting expert system. The KGFS approach has the following novel features:

- It reduces effects of noisy data by means of clustering data set into *k* different clusters.
- GAs have been demonstrated to be a powerful tool for automating the definition of the fuzzy rule based systems. KGFS uses a genetic algorithm for extracting rule base of the fuzzy expert system.
- For the purpose of accuracy improvement, it tunes the data base of the expert system using a unique genetic algorithm.
- The marketing manager may perceive too much uncertainty when taking the decisions. With this purpose, our method successfully hybridizes fuzzy logic with genetic algorithms and data clustering to extract useful and descriptive information patterns from sales data.
- Due to the proposed model, we can visualize the sales level as a continuous function of its input parameters and as a result we can answer "what-if" questions about sales and making plans for a coming period.
- We can be assured of working with optimum solutions, expressed in an easy, semantically understandable way of reasoning of the human being.

We applied KGFS on printed circuit board sales problem which had been used in different papers as the case study. Results show that forecasting accuracy of KGFS outperforms the previous approaches regarding MAPE and RMSE evaluations, and KGFS can be used as a suitable forecasting tool in sales forecasting problems.

References

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

Anderberg, M. R. (1973). *Cluster analysis for applications*. New York: Academic Press.

Atsalakis, G. S., & Valavanis, K. P. (2009). Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert Systems with Applications, 36*, 10696–10707.

Casillas, J., Cordón, O., Herrera, F., & Villar, P. (2004). A hybrid learning process for the knowledge base of a fuzzy rule-based system. In *Proceedings of the 2004 international conference on information processing and management of uncertainty in knowledge-based systems, Perugia, Italy* (pp. 2189–2196).

Casillas, F., & Martínez-López, F. (2009a). Mining uncertain data with multiobjective genetic fuzzy systems to be applied in consumer behaviour modelling. *Expert Systems with Applications, 36*(2), 1645–1659.

Casillas, J., & Martínez-López, F. (2009b). A knowledge discovery method based on genetic-fuzzy systems for obtaining consumer behaviour patterns. An empirical application to a web-based trust model. *International journal of Management and Decision Making, 10*, 402–428.

Chang, P.-C., & Lai, C.-Y. (2005). A hybrid system combining self-organizing maps with case-based reasoning in wholesaler's new-release book forecasting. *Expert Systems with Applications, 29*(1), 183–192.

Chang, P.-C., & Liu, C.-H. (2008). A TSK type fuzzy rule based system for stock price prediction. *Expert Systems with Applications, 34*, 135–144.

Chang, P.-C., Liu, C.-H., & Fan, C. Y. (2009). Data clustering and fuzzy neural network for sales forecasting: A case study in printed circuit board industry. *Knowledge-Based Systems, 22*(5), 344–355.

Chang, P.-C., Liu, C.-H., & Wang, Y.-W. (2006). A hybrid model by clustering and evolving fuzzy rules for sales decision supports 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).

Chang, P.-C., Wang, Y.-W., & Liu, C.-H. (2007). The development of a weighted evolving fuzzy neural network for PCB sales forecasting. *Expert Systems with Applications, 32*(1), 86–96.

Chang, P.-C., Wang, Y.-W., & Tsai, C.-Y. (2005). Evolving neural network for printed circuit board sales forecasting. *Expert Systems with Applications, 29*, 83–92.

Cordon, O., & Herrera, F. (1995). A general study on genetic fuzzy systems. In J. Periaux, G. Winter, M. Galen, & P. Cuesta (Eds.), *Genetic algorithms in engineering and computer science* (pp. 33–57). Wiley.

Cordón, O., & Herrera, F. (1997). A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples. *International Journal of Approximate Reasoning, 17*(4), 369–407.

Cordón, O., Herrera, F., Hoffmann, F., & Magdalena, L. (2001). *Genetic fuzzy systems: Evolutionary tuning and learning of fuzzy knowledge bases*. Singapore: World Scientific.

Damousis, I., & Dokopoulos, P. (2001). A fuzzy expert system for the forecasting of wind speed and power generation in wind farms. In *International conference on power industry computer applications PICA, Sydney, Australia* (pp. 63–69).

Efendigil, T., Önüt, S., & Kahraman, C. (2009). A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis. *Expert Systems with Applications, 36*(3), 6697–6707.

Eiben, A., & Smith, J. (2003). *Introduction to evolutionary computation*. Berlin: Springer.

Eshelman, L., & Schaffer, J. (1993). Real-coded genetic algorithms and interval-schemata. In L. D. Whitley (Ed.), *Foundations of genetic algorithms* (Vol. 2). San Mateo, CA: Morgan Kaufmann Publishers.

Konar, A. (2005). *Computational intelligence: Principles techniques*. Berlin: Springer.

Kuo, R. J., & Chen, J. A. (2004). A decision support system for order selection in electronic commerce based on fuzzy neural network supported by real-coded genetic algorithm. *Expert Systems with Applications, 26*(2), 141–154.

Kuo, R., & Xue, K. (1999). Fuzzy neural networks with application to sales forecasting. *Fuzzy Sets and Systems, 108*, 123–143.

Kwong, C., Wong, T., & Cha, K. (2009). A methodology of generating customer satisfaction models for new product development using a neuro-fuzzy approach. *Expert Systems with Applications, 36*(8), 11262–11270.

Li, S. (2000). The development of a hybrid intelligent system for developing marketing strategy. *Decision Support Systems, 27*, 395–409.

Martínez-López, F., & Casillas, J. (2009). Marketing intelligent systems for consumer behaviour modelling by a descriptive induction approach based on genetic fuzzy. *Industrial Marketing Management.*

Nauck, D., Klawonn, F., & Kruse, R. (1997). *Foundations of neuro-fuzzy systems*. New York: Wiley.

Orriols-Puig, A., Casillas, J., & Martínez-López, F. (2009). Unsupervised learning of fuzzy association rules for consumer behavior modeling. *Mathware and Soft Computing, 16*(1), 29–43.

Rakotomalala, R. (2005). TANAGRA: A free software for research and academic purposes. In *EGC*'2005, RNTI-E-3 (Vol. 2, pp. 697–702).

Shim, M., Seong, S., Ko, B., & So, M. (1999). Application of evolutionary computation at LG electronics. In *IEEE international conference on fuzzy systems FUZZ-IEEE*'99, Seoul, South Korea (pp. 1802–1806).

Thomassey, S., & Happiette, M. (2007). A neural clustering and classification system for sales forecasting of new apparel items. *Applied Soft Computing, 7*, 1177–1187.

Thomassey, S., Happiette, M., & Castelain, J.-M. (2005). A global forecasting support system adapted to textile distribution. *International Journal of Production Economics, 96*, 81–95.

Thrift, P. (1991). Fuzzy logic synthesis with genetic algorithms. In *Fourth international conference on genetic algorithms (ICGA*'91), San Diego, USA (pp. 509–513).