An improved approach based on fuzzy clustering and Back-Propagation Neural Networks with adaptive learning rate for sales forecasting: Case study of PCB industry

Attariuas Hicham[1], , Bouhorma Mohamed[2], El Fallahi Abdellah[3]

[1,2 2]Laboratory LIST, Group Research in Computing and Telecommunications (ERIT) FST Tangier, BP : 416, Old Airport Road, Morocco

[3]Laboratory LITT, Group Research in Computing and Logistic (ERIT) ENSA Tetouan, Mhannech II, BP : 2121, Morocco

Abstract

This paper describes new hybrid sales forecasting system based on fuzzy clustering and Back-propagation (BP) Neural Networks with adaptive learning rate **(FCBPN).**The proposed approach is composed of three stages: (1) Winter's Exponential Smoothing method will be utilized to take the trend effect into consideration; (2) utilizing Fuzzy C-Means clustering method (Used in an clusters memberships fuzzy system (**CMFS**)), the clusters membership levels of each normalized data records will be extracted; (3) Each cluster will be fed into parallel BP networks with a learning rate adapted as the level of cluster membership of training data records. Compared to many researches which use Hard clustering, we employ fuzzy clustering which permits each data record to belong to each cluster to a certain degree, which allows the clusters to be larger which consequently increases the accuracy of the proposed forecasting system . Printed Circuit Board (PCB) will be used as a case study to evaluate the precision of our proposed architecture. Experimental results show that the proposed model outperforms the previous and traditional approaches. Therefore, it is a very promising solution for industrial forecasting.

Keywords: *Sales forecasting, fuzzy clustering, fuzzy system, Printed circuit boards, back propagation network, Hybrid intelligence approach.*

1. Introduction and related research

Reliable prediction of sales becomes a vital task of business decision making. Companies that use accurate sales forecasting system earn important benefits. Sales forecasting is both necessary and difficult. It is necessary because it is the star-ting point of many tools for managing the business: production schedules, finance, marketing plans and budgeting, and promotion and advertising plan. It is difficult because it is out of reach regardless of the quality of the methods adopted to predict the future with certainty. The parameters which intervenient are numerous, complex and often unquantifiable.

Recently, the combined intelligence technique using artificial neural networks (ANNs), fuzzy logic, Particle Swarm Optimization (PSO), and genetic algorithms (GAs) has been demonstrated to be an innovative forecasting approach. Since most sales data are non-linear in relation and complex, many studies tend to apply Hybrid models to time-series forecasting. Kuo and Chen (2004)[20] use a combination of neural networks and fuzzy systems to effectively deal with the marketing problem.

Many researchers conclude that the application of BPN is an effective method as a forecasting system, and can also be used to find the key factors for enterprisers to improve their logistics management level. Zhang, Wang and Chang (2009) [28] utilized Back Propagation neural networks (BPN) in order to forecast safety stock. Zhang, Haifeng and Huang (2010)[29] used BPN for Sales Forecasting Based on ERP System. They found out that BPN can be used as an ac-curate sales forecasting system.

The rate of convergence of the traditional back-propagation networks is very slow because it's dependent upon the choice of value of the learning rate parameter. However, the experimental results (2009 [25]) showed that the use of an adaptive learning rate parameter during the training process can lead to much better results than the traditional neural net-work model (BPN).

Many papers indicate that the system which uses the hybridization of fuzzy logic and neural networks can more accurately perform than the conventional statistical method and single ANN. Kuo and Xue (1999) [21] proposed a fuzzy neural network (FNN) as a model for sales forecasting. They utilized fuzzy logic to extract the expert's fuzzy knowledge. Toly Chen (2003) [27] used a model for wafer fab prediction based on a fuzzy back propagation network (FBPN). The proposed system is constructed to incorporate production control expert judgments in enhancing the performance of an existing crisp back propagation network. The results showed the performance of the FBPN was better than that of the BPN.

Efendigil, Önü, and Kahraman (2009) [16] utilized a forecasting system based on artificial neural networks ANNs and adaptive network based fuzzy inference systems (ANFIS) to predict the fuzzy demand with incomplete information.

A Hybrid Intelligent Clustering Forecasting System was proposed by Kyong and Han (2001)[22]. It was based on Change Point Detection and Artificial Neural Networks. The basic concept of proposed model is to obtain significant intervals by change point detection. They found out that the proposed models are more accurate and convergent than the traditional neural network model (BPN).

Recently, some researchers have shown that the use of the hybridization between fuzzy logic and GAs leading to genetic fuzzy systems (GFSs) (Cordón, Herrera, Hoffmann, & Magdalena (2001) [13]) has more accurate and efficient results than the traditional intelligent systems. Casillas, & MartínezLópez (2009) [24], MartínezLópez & Casillas (2009) [23], utilized GFS in various case Management. They have all obtained good results.

This article proposes a new hybrid system based on fuzzy clustering and Back-propagation Neural Networks with adaptive learning rate (FCBPN) for sales forecasting in Printed Circuit Board (PCB) industry, which has been frequently used by the other authors as a case study.

1.1 PCB sales forecasting

Due to the important role of Printed Circuit Board (PCB) industry in Taiwan's economy, there are several studies in the literature which have considered PCB sales forecasting as the case study (table 1).

Chang, Wang and Tsai (2005)[3] used Back Propagation neural networks (BPN) trained by a genetic algorithm (ENN) to estimate demand production of printed circuit board (PCB). The experimental results show that the performance of ENN is greater than BPN.

Chang and Wang (2006) [6] used a fuzzy backpropagation network (FBPN) for sales forecasting. The opinions of sales managers about the importance of each input, were converted into prespecified fuzzy numbers to be integrated into a proposed system. They concluded that FBPN approach outperforms other traditional methods such as Grey Forecasting, Multiple Regression Analysis and back propagation networks.

Chang, Liu, and Wang (2006)[7] proposed a fusion of SOM, ANNs, GAs and FRBS for PCB sales forecasting. They found that performance of the model was superior to previous methods that proposed for PCB sales forecasting.

Chang, Wang and Liu (2007) [10]developed a weighted evolving fuzzy neural network (WEFuNN) model for PCB sales forecasting. The proposed model was based on combination of sales key factors selected using GRA. The experimental results that this hybrid system is better than previous hybrid models.

Chang and Liu (2008)[4] developed a hybrid model based on fusion of cased-based reasoning (CBR) and fuzzy multicriteria decision making. The experimental results showed that performance of the fuzzy cased based reasoning (FCBR) model is superior to traditional statistical models and BPN.

Chang, Liu and Fan (2009) [5] developed a K-means clustering and fuzzy neural network (FNN) to estimate the future sales of PCB. They used K-means for clustering data in different clusters to be fed into independent FNN models. The experimental results show that the proposed approach outperforms other traditional forecasting models, such as, BPN, ANFIS and FNN.

Hadavandi, Shavandi and Ghanbari (2011) [18] proposed a novel sales forecasting approach by the integration of genetic fuzzy systems (GFS) and data clustering to construct a sales forecasting expert system. They use GFS to extract the whole knowledge base of the fuzzy system for sales forecasting problems. Experimental results show that the proposed approach outperforms the other previous approaches.

The rest of the article is organized as follows: Section 2 describes the proposed model which called fuzzy clustering and Back-propagation Neural Networks with adaptive learning rate (FCBPN). Experiment results and analysis are discussed in Section 3.Finally, in Section 4, conclusions are presented.

2. Development of the FCBPN model

The proposed architecture consists of three stages as shown in Figure 1: (1) the output of Winter's method will taken as an input ($\mathbf{K_4}$) on FCBPN system to remove the trend effect; (2) all normalized records of data are categorized into **K** clusters by using the fuzzy C-means model; (3) the fuzzy distances from all records data ($\mathbf{x_i}$) to different cluster centers ($\mathbf{c_j}$) founded by fuzzy C-means (membership cluster)will be introduced into a parallel BP networks with a learning rate adapted according to the level of cluster membership of each record of training data set.

2.1. Data preprocessing stage

Historical data of an electronic company in Taiwan is used to choose the key variables (K_1, K_2, K_3) (see table 2) that are to be considered in the forecasted model. Monthly sales amount of Printed Circuit Board (PCB) is considered as a case of the forecasting model which has been used as the case in different studies.

2.1.1. Winter's Exponential Smoothing

In order to take the effects of seasonality and trend into consideration, winter's Exponential Smoothing is used to preliminarily forecast the quantity of PCB production. For time serial data, Winter's Exponential Smoothing is used to preprocess all the historical data and use them to predict the production demand (see figure 2), which will be entered into the proposed hybrid model as input variable (**K_4**)(see table 2).Similar to the previous researches, we assume $\alpha = 0.1$, $\beta = 0.1$ and $\gamma = 0.9$.

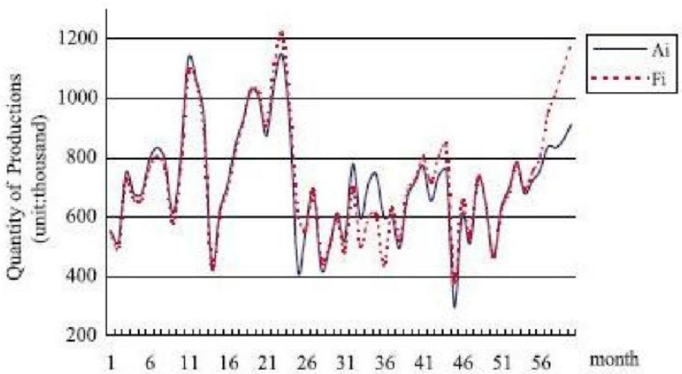

FIGURE 2: Comparison of forecasting Results of Winter's Exponential Smoothing to The Real Number.

2.2. Extract membership levels to each cluster (CMFS)

Using Fuzzy C-Means clustering method (utilized in an adapted fuzzy system (CMFS)), the clusters centers of the normalized data records will be founded, and consequently, we can extract the clusters membership levels of each normalized data records.

2.2.1. Data normalization

The input values (**K1, K2, K3, K4**) will be ranged in the interval [0.1 , 0.9] to meet property of neural networks. The

$$X_i = 0.1 + 0.8 * (K_i - min(K_i))/(max(K_i) - min(K_i)). \quad (1)$$

$$N_i = 0.1 + 0.8 * (K_i - min(K_i))/(max(K_i) - min(K_i)). \quad (1)$$

Where **K_i** presents a key variable, **N_i** presents normalized input (see table 2), max (**K_i**) and min (**K_i**) represent maximum and minimum of the key variables, respectively.

Input	Description
K_1	Consumer price index
N_1	Normalized Consumer price index
K_2	Liquid crystal element demand
N_2	Normalized Liquid crystal element demand
K_3	PCB total production value
N_3	Normalized PCB total production value
K_4	Preprocessed historical data (WES)
N_4	Normalized preprocessed historical data (WES)
Y^0	Actual historical monthly PCB sales Data
Y	Normalized Actual historical monthly PCB sales Data

TABLE 2: Description of input forecasting model.

2.2.2. Fuzzy c-means clustering

In hard clustering, data is divided into distinct clusters, where each data element belongs to exactly one cluster. In Fuzzy c-means (FCM) (developed by Dunn 1973 [14] and improved by Bezdek 1981 [1]), data elements can belong to more than one cluster, and associated with each element is a set of membership levels. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^m \|x_i - c_j\|^2 \qquad , 1 \le m < \infty.$$

Where u_{ij} is the degree of membership of x_i in the cluster **j** , x_i is the i^{th} of measured data and c_j is the center of the j^{th} cluster. The algorithm is composed of the following steps :

Step 1: Initialize randomly the degrees of membership matrix $U = [u_{ij}]$, $U^{(0)}$

Step 2: Calculate the centroid for each cluster $C(k) = [c_j]$ with $U(k)$:

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m . x_i}{\sum_{i=1}^{N} u_{ij}^m}$$

Step 3: For each point, update its coefficients of being in the clusters **($U(k)$,$U(k+1)$)** :

$$u_{ij} = \frac{1}{\sum_{i=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{2/(m-1)}}$$

Step 4: If $\| U^{(k+1)} - U^{(k)} \| < \varepsilon$, $0 < \varepsilon < 1$. then STOP; otherwise return to step 2.

This procedure converges to a local minimum or a saddle point of J_m. According to Bezdek [1], the appreciated parameter combination of two factors (m and **ε**) is **m** = 2 and **ε**= 0.5

Authors	Year	Methods	accuracy	
			Mape	RMSE
Chang, Wang and Tsai	2005	GA + ANN	3,13	NA
Chang, and Wang	2006	Fuzzy logic + ANN	3,09	NA
Chang, Liu and Wang	2006	SOM+ANN+GA+FRBS	2,16	21,346
Chang, Wang and Liu	2007	WEFuNN	2,11	24,909
Chang, Liu and Lai	2008	FCBR	4,82	43,385
Chang, Liu and Fan	2009	K-means clustering+FNN	2,19	20,287
Hadavandi and Ghanbari	2011	K-means clustering+KGFS	1,46	19,354

TABLE 1: Summarizes various methods developed for PBC sales forecasting.

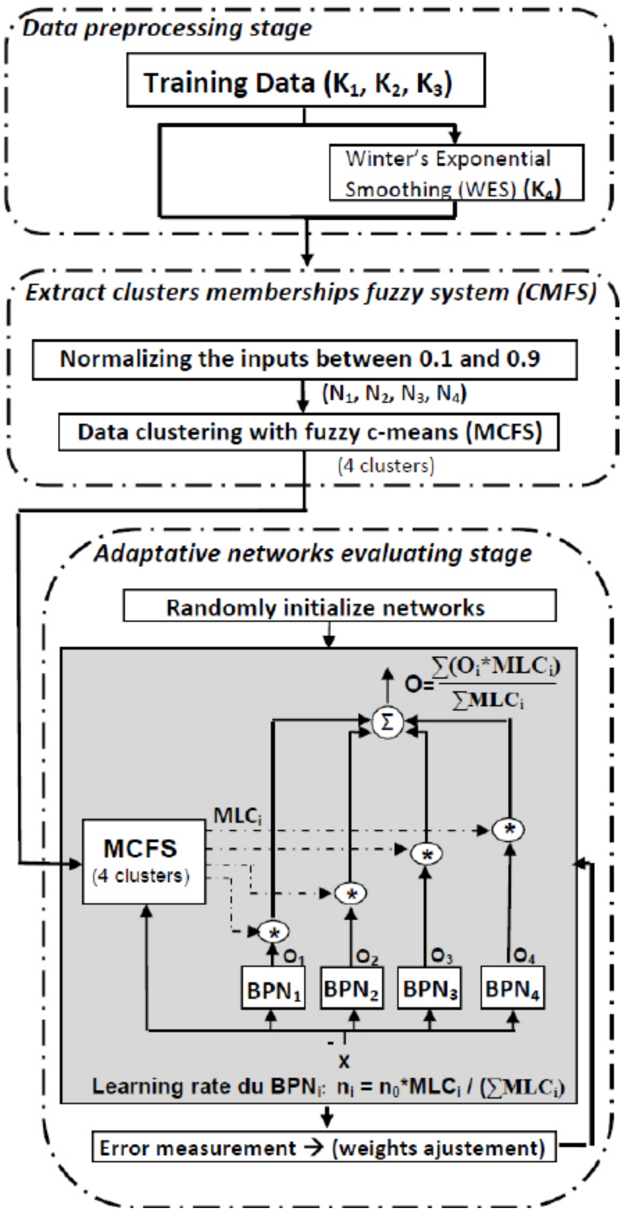

FIGURE 1: The architecture of FCBPN model.

Using fuzzy c-means, Table 3 shows that the use of four clusters is the best among all different clustering numbers.

Clustering groups	fuzzy c-means total distance
Clustering 2 groups	23.7904
Clustering 3 groups	20.2777
Clustering 4 groups	18.1477

TABLE 3: Comparison of different clustering algorithms in total distance.

2.2.3. The degree of Membership levels (**MLC$_k$**)

In this stage, we will use the sigmoid function (figure 3) to improve the precision of results and to accelerate the training process of neural networks. Then, the advanced fuzzy distance between records data (**X$_i$**) and a cluster center (**c$_k$**) (**AFD$_k$**) will be presented as follow:

$$AFD_k(X_i) = sigmf(\|C_k - X_i\|, [50, 0.5])$$

$$sigmf(x,[a,c]) = \frac{1}{1+e^{-a(x-c)}}$$

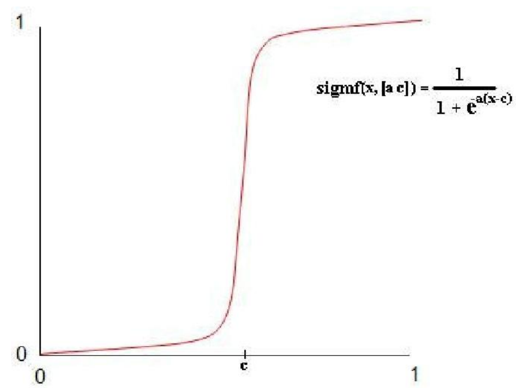

FIGURE 3: Sigmoid function, a = 50 and c = 0,5.

The membership levels of belonging of a record **X$_i$** to **k^{ith}** cluster (**MLC$_k$ (X$_i$)**) is related inversely to the distance from records data **X$_i$** to the cluster center **c$_k$ (AFD$_k$ (X$_i$))**:

$$MLC_k(X_i) = \frac{1}{AFD_k(X_i)}$$

The clusters memberships' fuzzy system (CMFS) return the memberships level of belonging of data record X to each clusters:

$$CMFS(X) = (MLC_1(X), MLC_2(X), MLC_3(X), MLC_4(X))$$

Thus, we can construct a new training sample (X_i , $MLC_1(X_i)$, $MLC_2(X_i)$, $MLC_3(X_i)$, $MLC_4(X_i)$) for the adaptive neural networks evaluating (Figure **1**).

2.3. Adaptive neural networks evaluating stage

The artificial neural networks (ANNs) concept is originated from the biological science (neurons in an organism). Its components are connected according to some pattern of connectivity, associated with different weights. The weight of a neural connection is updated by learning. The ANNs possess the ability to identify nonlinear patterns by learning from the data set. The back-propagation (BP) training algorithms are probably the most popular ones. The structure of BP neural networks consists of an input layer, a hidden layer, as well as an output layer. Each layer contains **I** ; **J** and **L** nodes denoted. The $\boldsymbol{w_{ij}}$ is denoted as numerical weights between input and hidden layers and so is $\boldsymbol{w_{jl}}$ between hidden and output layers as shown in figure 4.

In this stage, we propose an adaptive neural networks evaluating system which consists of four neural networks. Each cluster **K** is associated with the $\boldsymbol{K^{ith}}$ BP network. For each cluster, the training sample will be fed into a parallel Back Propagation networks (BPN) with a learning rate adapted according to the level of clusters membership ($\boldsymbol{MLC_k}$) of each records of training data set. The structure of the proposed system is shown in figure **1**.

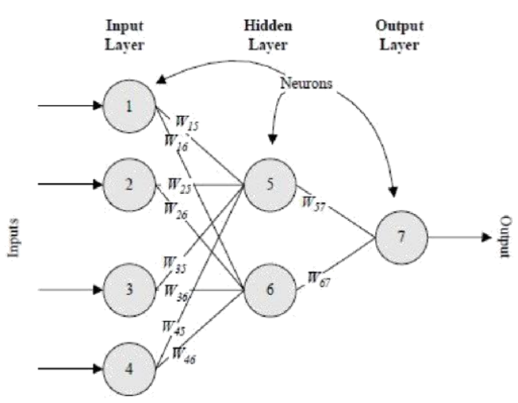

FIGURE 4: The structure of back-propagation neural network

The Adaptive neural networks learning algorithm is composed of two procedures: (a) a feed forward step and (b) a back-propagation weight training step. These two separate procedures will be explained in details as follows:

Step 1- All BP networks are initialized with the same random weights.

Step 2 - Feed forward.

For each $\boldsymbol{BPN_k}$ (associate to the $\boldsymbol{K^{th}}$ cluster), we assume that each input factor in the input layer is denoted by $\boldsymbol{x_i}$. $\boldsymbol{y_j^k}$ and $\boldsymbol{o^k_l}$ represent the output in the hidden layer and the output layer, respectively. And $\boldsymbol{y_i^k}$ and $\boldsymbol{o^k_l}$ can be expressed as follows:

$$Y_j^k = f(X_j) = f(w_{oj}^k + \sum_{i=0}^{I} w_{ij}^k x_i)$$

and

$$o_l^k = f(Y_l^k) = f(w_{ol}^k + \sum_{j=1}^{J} w_{jl}^k y_j^k)$$

where the $\boldsymbol{w_{oj}}^k$ and $\boldsymbol{w_{ol}}^k$ are the bias weights for setting threshold values, $\boldsymbol{f}$ is the activation function used in both hidden and output layers and $\boldsymbol{X_j^k}$ and $\boldsymbol{Y_l^k}$ are the temporarily computing results before applying activation function $\boldsymbol{f}$. In this study, a sigmoid function (or logistic function) is selected as the activation function. Therefore, the actual outputs $\boldsymbol{y_j^k}$ and $\boldsymbol{o^k_l}$ in hidden and output layers, respectively, can also be written as :

$$y_j^k = f(X_j^k) = \frac{1}{1+e^{-X_j^k}}$$

and

$$o_l^k = f(Y_l^k) = \frac{1}{1+e^{-Y_l^k}}$$

The activation function f introduces the nonlinear effect to the network and maps the result of computation to a domain (0, 1). In our case, the sigmoid function is used as the activation function.

$$f(t) = \frac{1}{1+e^{-t}} \qquad , f' = f(1-f)$$

The globate output of the adaptive neural networks is calculated as :

$$o_l = \frac{\sum_{k=1}^{4}(MLC_k(X_j) \times o_l^k)}{\sum_{k=1}^{4} MLC_k(X_j)}$$

As shown above, the effect of the output $\boldsymbol{o}^k{}_l$ on the global output $\boldsymbol{o}_l$ is both strongly and positively related to the membership level ($\boldsymbol{MLC_k}$) of data record X_i to $\boldsymbol{k}^{ith}$ cluster.

Step 3- Back-propagation weight training.

The error function is defined as:

$$E = \frac{1}{2}\sum_{l=1}^{L} e_l^2 = \frac{1}{2}\sum_{l=1}^{L}(t_l - o_l)^2$$

Where $\boldsymbol{t_k}$ is a predefined network output (or desired output or target value) and $\boldsymbol{e_k}$ is the error in each output node. The goal is to minimize $\mathbf{E}$ so that the weight in each link is accordingly adjusted and the final output can match the desired output. The learning speed can be improved by introducing the momentum term. Usually, falls in the range [0, 1]. For the iteration n and for $\boldsymbol{BPN_k}$ (associated to $\boldsymbol{k}^{th}$ cluster), the adaptive learning rate in $\boldsymbol{BPN_k}$ and the variation of weights $\boldsymbol{\Delta w_k}$ can be expressed as

$$\eta_k = \frac{MLC_k(X_j)}{\sum_{k=1}^{4} MLC_k(X_j)} \times \eta$$

$$\Delta w_k(n+1) = \eta_k \times \Delta w_k(n) + \alpha \times \frac{\delta E}{\delta w_k(n)}$$

As shown above, we can conclude that the variation of the $\boldsymbol{BPN_k}$ network weights ($\boldsymbol{w_{oj}}^k$ and $\boldsymbol{w_{ol}}^k$) are more important as long as the the membership level ($\boldsymbol{MLC_k}$) of data record $\boldsymbol{X_j}$ to $\boldsymbol{k}^{th}$ cluster is high. If the value of membership level ($\boldsymbol{MLC_k}$) of data record $\boldsymbol{X_j}$ to $\boldsymbol{k}^{ith}$ cluster is close to zero then the changes in $\boldsymbol{BPN_k}$ network weights are very minimal.

The configuration of the proposed BPN is established as follows:

– Number of neurons in the input layer: I =4

– Number of neurons in the output layer: L = 1

– Single hidden layer

– Number of neurons in the hidden layer: J =2

– Network-learning rule: delta rule

– Transformation function: sigmoid function

– learning rate: =0.1

– Momentum constant: = 0.02

– learning times : 20000

3. Experiment results and analysis

3.1. Constructing FCBPN System

The data test used in this study was collected from sales forecasting case study, called printed circuit board (PCB) industry in Taiwan. The total number of training samples was collected from January 1999 to December 2002 while the total number of testing samples was from January 2003 to December 2003.The proposed FCBPN system was applied as case to forecast the sales data of the PCB. The results are presented in table 5.

Month	Actual values	Forecasted values
2003/1	649,066	657,749
2003/2	466,750	493,585
2003/3	633,615	635,837
2003/4	693,946	674,867
2003/5	785,838	747,22
2003/6	679,312	693,531
2003/7	723,914	720,687
2003/8	757,490	754,198
2003/9	836,846	830,237
2003/10	833,012	852,37
2003/11	860,892	876,213
2003/12	912,182	893,217

TABLE 5: The forecasted results by FCBPN method.

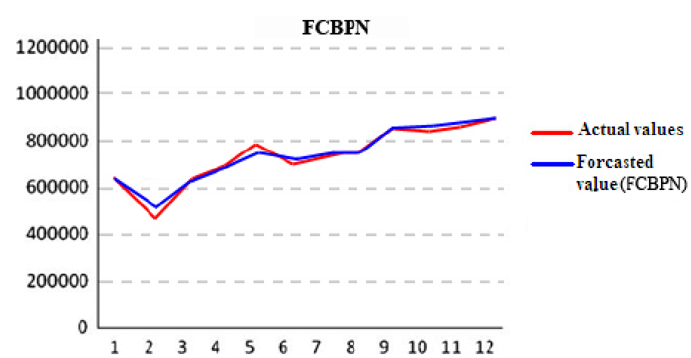

FIGURE 5: The MAPE of FCBPN .

3.2. Comparisons of FCBPN model with other previous models

Experimental comparison of outputs of FCBPN with other methods shows that the proposed model outperforms the previous approaches (tables 5-10). We apply two different performance measures called mean absolute percentage error (MAPE) and root mean square error (RMSE), to compare the FCBPN model with the previous

Authors	Year	Methods	accuracy	
			Mape	RMSE
Chang, Wang and Tsai	2005	GA + ANN	3,13	NA
Chang, and Wang	2006	Fuzzy logic + ANN	3,09	NA
Chang, Liu and Wang	2006	SOM+ANN+GA+FRBS	2,16	21,346
Chang, Wang and Liu	2007	WEFuNN	2,11	24,909
Chang, Liu and Lai	2008	FCBR	4,82	43,385
Chang, Liu and Fan	2009	K-means clustering+FNN	2,19	20,287
Hadavandi and Shavandi	2011	K-means clustering+KGFS	1,46	19,354
Attariuas and Fellahi	2012	Fuzzy clustering+FCBPN	1,97	18,009

TABLE 5: History of PCB sales forecasting.

methods: KGFS, KFNN, FNN, WES, BPN and RBFNN.

$$MAPE = 100 \times \frac{1}{N}\sum_{t=1}^{N}\frac{|Y_t - P_t|}{Y_t}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(Y_t - P_t)^2}$$

where, $\boldsymbol{P_t}$ is the expected value for period $\boldsymbol{t}$, $\boldsymbol{Y_t}$ is the actual value for period t and N is the number of periods.

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

TABLE 6: KGFS forecasted values vs actual values of PCB sales data.

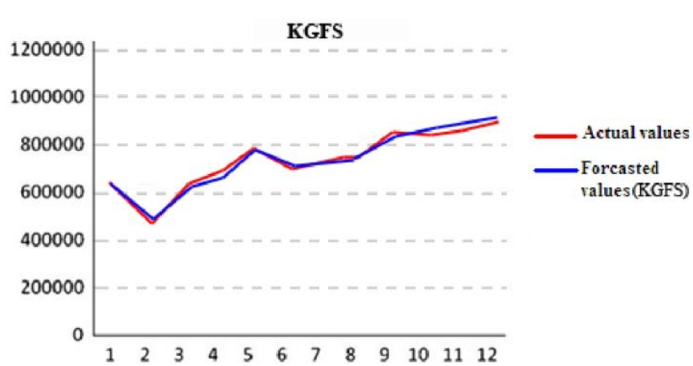

FIGURE 6: The MAPE of KGFS.

Month	Actual values	KGFS forecasts
2003/1	649,066	584,901.9
2003/2	466,750	483,872.3
2003/3	633,615	713,874.6
2003/4	693,946	711,356.1
2003/5	785,838	769,881.6
2003/6	679,312	684,634.5
2003/7	723,914	721,192.4
2003/8	757,490	770,609
2003/9	836,846	817,423.4
2003/10	833,012	851,827
2003/11	860,892	884,484.1
2003/12	912,182	912,129.1

TABLE 7: The forecasted results by FNN methods.

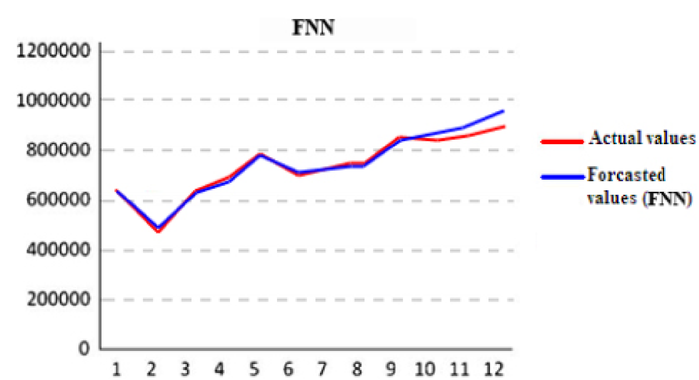

FIGURE 7: The forecasted results by FNN methods.

Month	Actual values	Forecasted values
2003/1	649,066	649,066
2003/2	466,750	466,750
2003/3	633,615	633,615
2003/4	693,946	693,946
2003/5	785,838	785,838
2003/6	679,312	679,312
2003/7	723,914	723,914
2003/8	757,490	757,490
2003/9	836,846	836,846
2003/10	833,012	833,012
2003/11	860,892	860,892
2003/12	912,182	912,182

Table 8: The forecasted results by Winter's method.

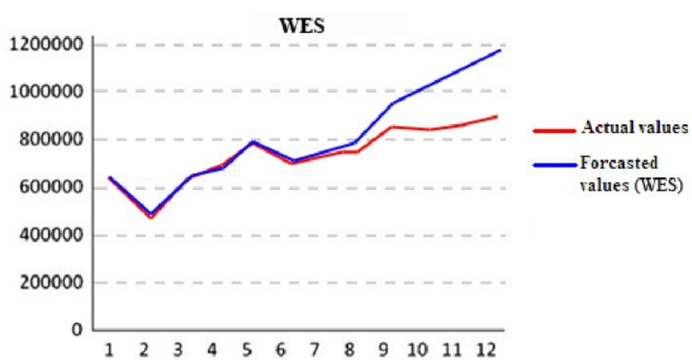

Figure 8 :The MAPE of WES

Month	Actual values	Forecasted values
2003/1	649,066	622,402.3
2003/2	466,750	456,226
2003/3	633,615	618,346
2003/4	693,946	669,445.5
2003/5	785,838	795,971.6
2003/6	679,312	682,646.4
2003/7	723,914	741,996.5
2003/8	757,490	789,756.8
2003/9	836,846	945,738.1
2003/10	833,012	1,006,899
2003/11	860,892	1,077,823
2003/12	912,182	1,141,621

Table 9 :The MAPE of BPN

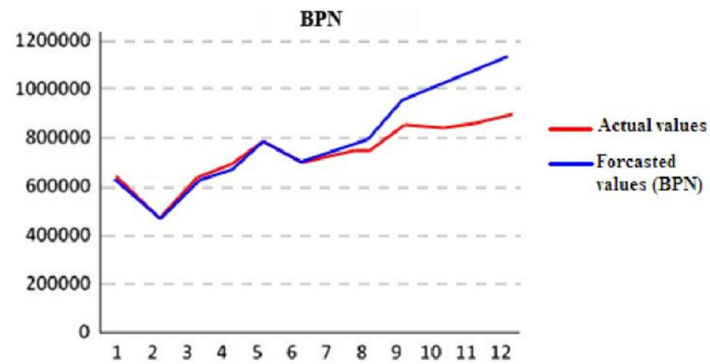

FIGURE 9: The MAPE of BPN.

Month	Actual values	Forecasted values
2003/1	649,066	574,952
2003/2	466,750	465,762
2003/3	633,615	634,449
2003/4	693,946	693,809
2003/5	785,838	785,878
2003/6	679,312	678,838
2003/7	723,914	723,553
2003/8	757,490	759,976
2003/9	836,846	875,283
2003/10	833,012	800,874
2003/11	860,892	860,366
2003/12	912,182	905,347

TABLE 10: The forecasted results by RBFNN method.

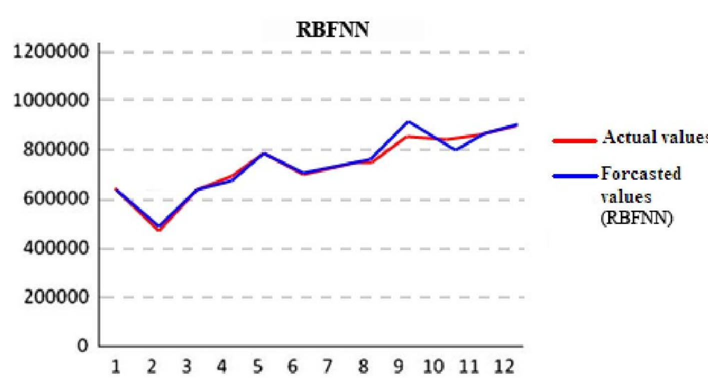

FIGURE 10: The MAPE of RBFNN.

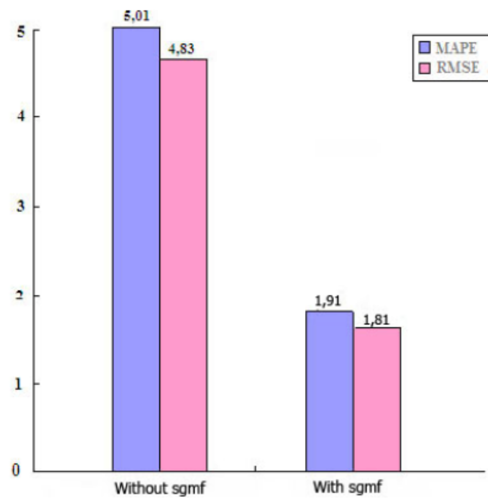

FIGURE 11: The performance improvement of FCBPN after using sigmoid fonction.

As shown in figure 11, the use of CMFS enhanced by the sigmoid function in the proposed acupuncture (FCBPN) has better precision results than the use of the fuzzy c-means clustering in the test stage. FCBPN has made 1.97 as MAPE evaluation and 18009 as RMSE evaluation. Therefore, the forecasting accuracy of FCBPN outperforms the previous approaches regarding MAPE and RMSE evaluations which are summarized in Table 4.

4. Conclusion

This article proposes a new hybrid system based on fuzzy clustering and Back-propagation Neural Networks with adaptive learning rate (FCBPN) for sales forecasting.

The experimental results of the proposed approach show that the effectiveness of the FCBPN outperforms the previous and traditional approaches : WES, BPN, RBFNN, KFNN, FNN, WES, BPN and KGFS. Furthermore, it also demonstrates that our modeling approach (FCBPN) has properties, such as, fast convergence, high precision, robust and accurate forecasting techniques.

Compared to previous researches which tend to use the classical hard clustering methods (K-means clustering) to divide data set into subgroups in order to reduce the noise and form more homogeneous clusters (Chang, 2009 [5]), the advantage of our proposed system (FCBPN) is that it uses a fuzzy clustering (fuzzy c-means clustering) which permits each data record to belong to each cluster to a certain degree, which allows the clusters to be larger which consequently increases the accuracy of forecasting system results.

We applied FCBPN for sales forecasting in Printed Circuit Board (PCB) as a case study. The results demonstrated the effectiveness and superiority of the FCBPN compared to the previous approaches regarding MAPE and RMSE evaluations. Other academic researchers and industrial practitioners may find these contributions interesting.

References

[1] Bezdek J. C. (1981) : "Pattern Recognition with Fuzzy Objective Function Algoritms", Plenum Press, New York.

[2] 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).

[3] Chang, P.-C, & Lai, C.-Y. (2005). A hybrid system combining self-organizing maps with case-based reasoning in wholesalerbs new-release book forecasting. Expert Systems with Applications, 29(1),183 192.

[4] 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.

[5] 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.

[6] 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.

[7] 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).

[8] 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.

[9] 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.

[10] Chang, P.-C , Yen-Wen Wang, Chen-Hao Liu(2007). The development of a weighted evolving fuzzy neural network for PCB sales forecasting. Expert Systems with Applications, 32(1),86 96.

[11] Chang PeiChann , Wang Di-di , Zhou Changle (2011). A novel model by evolving partially connected neural network for stock price trend forecasting. Expert Systems with Applications, Volume 39 Issue 1, January, 2012 .

[12] 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),369407.

[13] Cordón, O., Herrera, F., Hoffmann, F., & Magdalena, L.

(2001). Genetic fuzzy systems : Evolutionary tuning and learning of fuzzy knowledge bases. Singapore : World Scientific.

[14] Dunn J. C. (1973) : "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters", Journal of Cybernetics 3 : 32 57.

[15] Eberhart, R. C., and Kennedy, J. (1995). A new optimizer using particle swarm theory. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 39 43. Piscataway, NJ : IEEE Service Center.

[16] Efendigil, T., Önü, 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.

[17] Esmin, A (2007).Generating Fuzzy Rules from Examples Using the Particle Swarm Optimization Algorithm. Hybrid Intelligent Systems, 2007. HIS 2007. 7th International Conference IEEE.

[18] Hadavandi Esmaeil,Shavandi Hassan, & Ghanbari Arash (2011),An improved sales forecasting approach by the integration of genetic fuzzy systems and data clustering : Case study of printed circuit board.Expert Systems with Applications 38 (2011) 9392 9399.

[19] Goldberg, D. A. (1989). Genetic algorithms in search, optimization, and machine learning. Reading, MA : Addison-Wesley.

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

[21] Kuo .R.J & Xue. K.C (1999).Fuzzy neural networks with application to sales forecasting.Fuzzy Sets and Systems, 108,123 143.

[22] Kyong Joo Oh, Ingoo Han (2001),An intelligent clustering forecasting system based on change-point detection and artificial neural networks :application to financial economics.System Sciences, 2001. Proceedings of the 34th Annual Hawaii

[23] 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.

[24] 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.

[25] Saeid Iranmanesh, M. Amin Mahdavi (2009). A Differential Adaptive Learning Rate Method for Back-Propagation Neural Networks.World Academy of Science, Engineering and Technology 50 2009.

[26] Sivanandam.S. N, Visalakshi P : Dynamic task scheduling with load balancing using parallel orthogonal particle swarm optimisation. IJBIC 1(4) : 276 286 (2009)

[27] Toly Chen (2003).A fuzzy back propagation network for output time prediction in a wafer fab.Applied Soft Computing 2/3F (2003),211 222.

[28] Zhang, Wang & Chang (2009) . A Model on Forecasting Safety Stock of ERP Based on BP Neural Network.Proceedings of the 2008 IEEE ICMIT

[29] Zhang, Haifeng and Huang(2010).Sales Forecasting Based on ERP System through BP Neural Networks.ISICA'10 Proceedings of the 5th international conference on Advances in computation and intelligence

Attariuas Hicham received the computer engineer degree in 2009 from **ENSAIS** national school of computer science and systems analysis in Rabat, Morocco. Currently, he is a PhD Student in Computer Science. Current research interests: fuzzy system, intelligence system, bac-propagation network, genetic intelligent system.

Bouhorma Mohamed received the the PhD degree in Telecommunications and Computer Engineering. He is a Professor of Telecommunications and Computer Engineering in Abdelmalek Essaadi University. He has been a member of the Organizing and the Scientific Committees of several symposia and conferences dealing with Intelligent system, Mobile Networks, Telecommunications technologies.

El Fallahi Abdellah . received the the PhD degree in neural systems in 2008 from Valencia University, Spain. He is Professor in the logistics and transport department at the National School of applied sciences.His teaching is devoted to the logistics and transport, Integer and Linear Programming in Mathematics and heuristics .His research interest focuses on the development of meta-heuristics for hard optimization problems.