A Comparative Test Of New Mean-term Forecasting Models Adapted to Textile Items Sales

Sébastien Thomassey*', Philippe Vroman*, Michel Happiette*, Jean Marie Castelain***

* GEMTEX - Ecole Nationale Supérieure des Arts et Industries Textiles

9, rue de l'Ermitage

59100 Roubaix

FRANCE

** Institut Français du Textile et de l'Habillement

2, rue de la Recherche

59656 Villeneuve d'Ascq

FRANCE

E-mail : sebastien.thomassey@ensait.fr

Abstract: To meet customers' request, to manage their productions and their stocks or to direct their marketing policies, textile companies must improve their supply chain management. This organization requires sales forecasting systems adapted to the uncertain environment of the textile field. The uncertainty is characterized by noisy data and numerous explanatory variables (controlled, available or not) that influence the sales behaviour. This way, some recent researches introduced new forecasting models based on "soft computing". This paper deals with a comparison between new mean-term forecasting models and some traditional ones. The proposed models are hybrid neural model (HNCCX) and hybrid fuzzy model (HFCCX). They use neural networks and fuzzy logic abilities to map the non-linear influences of explanatory variables and consider the seasonality factors. The latter improvement, which is presented here, allows the reduction of models complexity by reducing the number of parameters. These models are also well -adapted to short and discontinuous time series, i.e. when the product sales occur during only some periods a year (major cases of textile items). To evaluate performances, the comparative test has been applied to real data of textile items, selected from an important French ready-to-wear distributor. Extensions are also proposed.

Keywords: Textile-apparel industry, sales forecasting, artificial neural network, fuzzy inference system, production and distribution management, explanatory variables

Sébastien Thomassey received the MSc. degree from the Ecole Nationale Supérieure des Arts et Industries Textiles (ENSAIT) of Roubaix in 1999. He prepared the Ph.D in Automation at the University of Lille 1. He has been engineer searcher at the Institut Français du Textile et de l'Habillement since 1999.

Philippe Vroman received the MSc. degree from the Ecole Nationale Supérieure des Arts et Industries Textiles (ENSAIT) of Roubaix in 1992 and the Ph.D in Automation from the University of Lille 1 in 2000. He has been Assistant Teacher at ENSAIT since 1998.

Michel Happiette received the MSc. degree from the Ecole Nationale Supérieure des Arts et Industries Textiles (ENSAIT) of Roubaix in 1976 and the Ph.D in Automation from the University of Lille 1 in 1983. He has been Assistant Professor at ENSAIT since 1984.

Jean Marie Castelain received the MSc. degree from the Ecole Nationale Supérieure d'Arts et Métiers in 1976 and the Ph.D degree in Mechanical Engineering from the University of Lille 1. Since 1987, he has been Doctor Ability and University Professor. He has been Head of Ecole Nationale Supérieure des Arts et Industries Textiles (ENSAIT) of Roubaix since 1997.

1. Introduction

To make decisions on the design and the driving of any logistic structures, textile managers must rely on efficient and accurate forecasting systems. A suitable forecast of the production, able to predict in due time the sufficient quantity to manufacture, is one of the most important factors for the success of a lean production [1]. Indeed, all the actors of the textile chain can use the Supply Chain Management (SCM) concepts [2], in order to manage and schedule their resources, capacities of stocks according to the customer requirements. But distributors' needs are directly related to the purchases of consumers; therefore all the SCM optimization depends on the forecast quality of the finished articles sales [3].

That supposes to take into account as much information adapted to the forecast lead-time (horizon) as possible. To summarize, the performances of such a forecasting system should meet the following criteria:

- to quickly react to a significant variation of trends and seasonality,
- to identify and smooth purely random events,
- to perform forecasting with short historic sales data,
- to take into account the influences of endogenous variables, linked to the textile product itself (size, color, selling price, merchandizing, promotions, etc.) or controlled by the delivery managers (brand name, distribution channel, sales promotion, advertising campaign, etc.) (Figure 1), and
- to take into account influences of exogenous variables (demographic variables, macroeconomic indicators, climate influences, calendar events, fashion effects, competitors, etc.) (Figure 1).

Different forecasting methods exist: linear/non-linear, adaptive/non-adaptive, explanatory or extrapolative as exponential smoothing models (ex: Holt-Winters), Box&Jenkins model with autoregressive integrated moving average (ARIMA) processes, dynamic regression models with explanatory variables, econometric methods [4], [5], and more recently, artificial neural network (ANN) [6] or fuzzy logic [7] based models. All these models provide disparate results. Indeed, their performances essentially depend on the application area, the forecasting goal, the user experience, and the forecasting horizon.

The major drawback is in that almost all contexts are quite specific and have to use a combination of several forecasting methods [4], [7]. Another difficulty, particularly existing in the textile-garment industrial network, consists of producing forecasts in uncertain environments. All events, which are able to influence the forecasting system, are neither strictly controlled nor identified [8].

After briefly refreshing the background of the distribution management in the textile-garment industry, and the formalization of the forecast issue, the features of some well -known classical models are remembered. Then, we propose the latest developments in hybrid neural (called HNCCX) [9] and hybrid fuzzy (called HFCCX) forecasting models. Many developments from the soft-computing theory have proved their efficiency on the design and development of intelligent systems, even in time series forecasting applications [10]. The aim is to prove the effectiveness of two "soft computing" based models. Neural networks currently take advantage of the learning abilities of neural computation in an uncertain environment, specifically with the introduction of explanatory variables [11]. Fuzzy logic is very interesting in modeling human knowledge [12], [13], [14] and also tolerant with respect to noisy data [14], [15]. The model also includes seasonality factors, which feature the structure of the studied time series. Our model originates from recent developments in artificial neural networks and fuzzy inference system applied to the textile field [16]. The model structure and the learning procedure are adapted to discontinuous, short and noisy time series, as well as a mean-term forecasting horizon. The performance of these models is estimated in a comparative test including the renowned models presented before. This comparative test is carried out with real data of basic textile items, proceeding from a French representative ready-to-wear firm. Finally, after comments and conclusion on the results, possible extensions to the models are suggested.

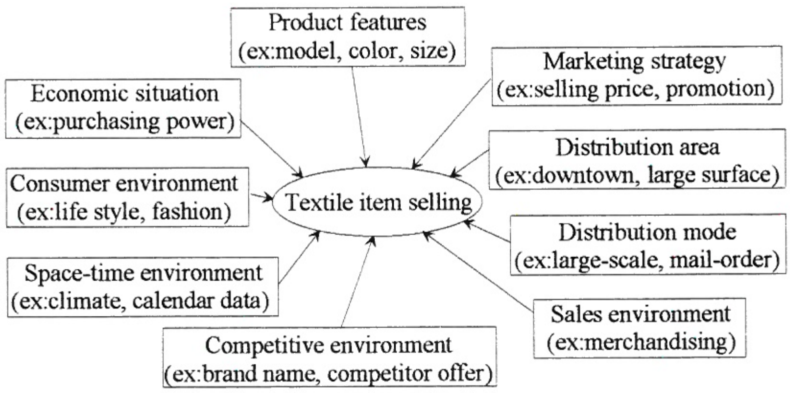

Figure 1. Main Categories of Explanatory Variables

2. Background

First, it seems important to remind of the prediction context and of the features of the forecasting system. The choice of the forecasting model depends indeed on the uncertain environment of the textile-garment field. Then, we argue the choice of our models structure compared to currently used statistical models.

2.1 Textile Background

2.1.1 Forecasting Purpose

New strategies of production management have been introduced for the last two decades. These strategies must face the worldwide competition that has become substantially sharp. They root in a time-based competition and are oriented to just-in-time and synchronous types of manufacturing. These combined management systems reinforce the distribution requirement planning (DRP) method. So, an organization, based on SCM concept, allows driving of the production of item series through the operative chain of textile production-distribution, composed of fibres manufacture to making up while passing through weaving, knitting and dyeing / finishing.

The production flow forwards by a network of firms. The firms represent manufacturing stages and induce proper intermediate inventories. To avoid running out of stock or over-stock, synchronization of the entire network is required. Unfortunately, products are too multiform and the global lead-time (or flow time) is too long to allow appropriate production feedback. A solution would be to introduce a strategic stock, adequately placed on the operative chain, in accordance with the delivery time imposed by customers (in stores) (Figure 2).

With regard to this strategy, textile managers evolve the building of logistic platforms. Downstream, the management method sprung from pull flow systems. Even though upstream, production is based on a push flow system and requires predictive management.

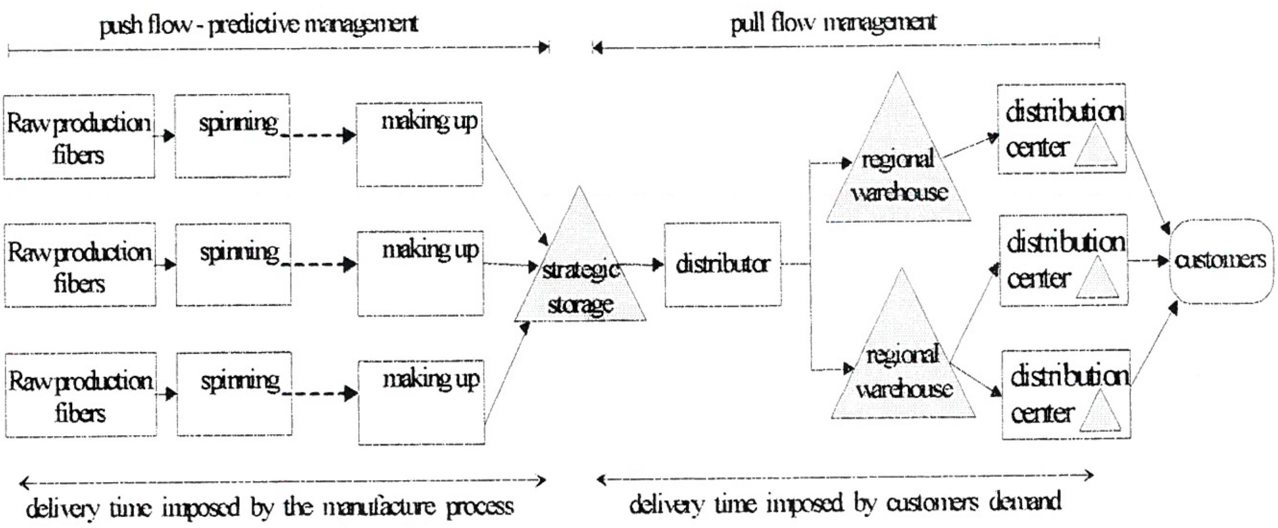

Figure 2. Strategic Stock in the Operative Chain

2.1.2 Sales Factors

The textile field is probably one of the most turbulent and fickle markets. The available sales data are very poor indeed, because of the short life cycle of textile items. They are also noisy, considering the numerous influences of some non-controlled factors (ex: fashion, social behaviors, etc.), and some momentously non-available factors (ex: sales area, local sales promotion, store number). Figure 1 presents the main categories of variables that critically influence the textile item selling.

Some of these variables act on store visitors' frequency, as others will do on customers' purchase decision. They can affect the total quantity of sales and/or only sales shape.

After the enumeration of these explicative variables, some remarks are worth noting:

- the explanatory (i.e. endogenous and exogenous) variables list, for a data series, cannot be exhaustive ;
- the interdependence of these variables does complicate the analysis;
- not all of the explanatory variables which influence the studied series, are always available (for example incomplete promotions histories, competition, ...);
- finally, one of the greatest difficulties lies in the acquisition and interpretation of reliable data (e.g. history of the item sales with opening and closing of distribution centers not recorded) and coherent (e.g. day climate data, monthly sales history).

2.1.3 Forecasting Data Features

One peculiarity of the textile sales analysis is the wide range of products and the significant number of item references. Then, a first preprocessing of data like clustering is imperative. The difficulty is to choose the aggregation level (Figure 3) of the sales forecasts. These levels are determined by the user profile and the manufacturing features of the textile product. This choice must allow the treatment of the explanatory variables influence. Indeed, the higher the aggregation level, the greater the difficulty raised by the interdependencies between endogenous and exogenous variables in modelling the sales. The comparative test herein is carried out with sales data classified according to the model, adding the product sales whatever colour and size. Data clustering is not treated here, for more information see [17].

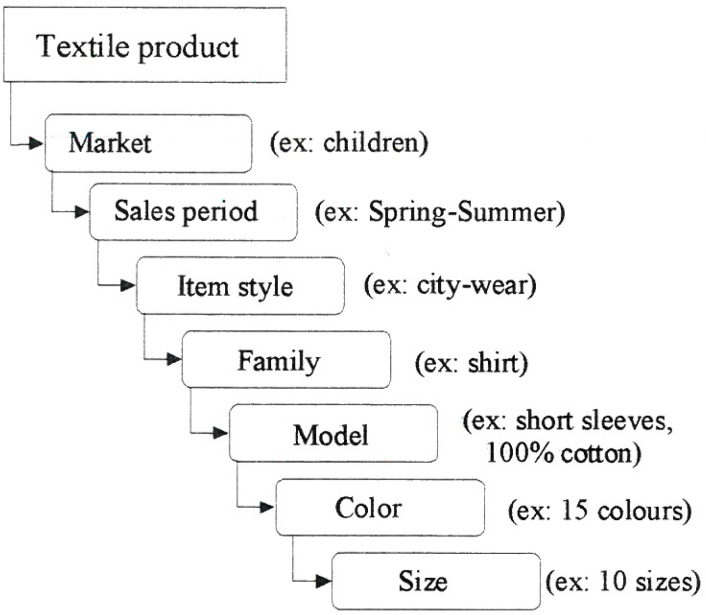

Figure 3. Aggregation Level of Textile Items

The forecasting lead-time required by users also appears to be a textile sales characteristic. Figure 4 presents the production planning of Autumn-Winter textile items, from creation to delivery in textile warehouse. It shows that purchasing managers need to know almost one year before the raw material quantities to order, which corresponds to the forecast total quantity of each textile product range. Production managers also require information about item quantities to manufacture, particularly early in the case of importing goods from far away countries. Moreover, it is necessary to fit in the forecasts after the launch of the textile items, as to let local manufacturers re-stock.

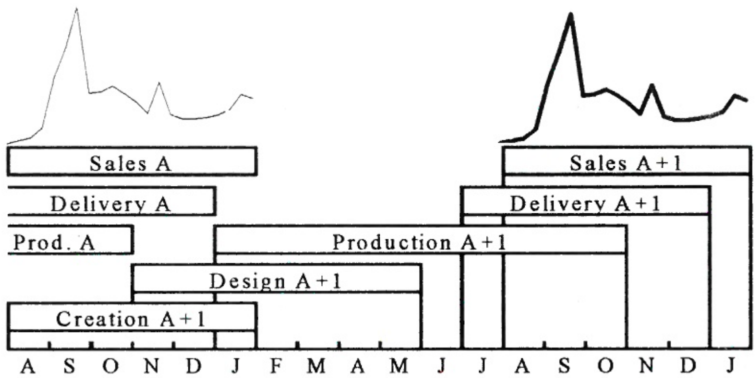

Figure 4. Production Planning of Autumn-Winter Textile Items

Therefore, short-term and mean-term horizons are required. In this paper, focus is on the mean-term forecast, i.e. to return an estimation of the sold quantities and the sales shape during the entire season.

The forecasting period commonly used by marketing managers in textile distribution is week. It practically corresponds to the rhythm of purchasing for consumers

The requirements of the textile managers according to the performance of the forecasting system are essentially based on the accuracy and the confidence interval of the forecasts. The accuracy is expressed by the prediction errors, the difference between forecast and actual sales during the same period (see Section 5.1.1).

2.2 Problem Formalization

A traditional approach to this forecasting modeling is to consider this as a time-series or a temporal signal-processing problem. A significant difference between the proposed models and the classical ones resides in the consideration of a discontinuous time -series. Indeed, it seems that the structure properties of the time -series will vary from the end of the last season to the beginning of the next one. Moreover, to consider continuous time- series means to critically affect the model learning.

As previously explained, the forecasting model must estimate the sales for season A+1 (Figure 4), from historic sales data until season A, the endogenous and exogenous (EE) variables assigned to the same period, and the EE variables known for the next season A+1. The EE variables used are quantitative indicators.

The general form (Figure 5) of the prediction model is then:

$$\hat{X}^{p}_{A+1} = F_{\theta}(X^{p}_{N}, U^{n,p}_{N}, U^{n,p}_{A+1}) = X^{p}_{A+1} + \varepsilon^{p}_{A+1}$$

$N \in [1,A]$ with

$$U^{n,p}_{N} = \begin{bmatrix} u^{1}_{N,1} & \cdots & u^{1}_{N,p} \\ \vdots & \ddots & \vdots \\ u^{n}_{N,1} & \cdots & u^{n}_{N,p} \end{bmatrix}$$

$$X_N^p = [x_{N,1},\ldots,x_{N,p}]$$

$$\varepsilon_{A+1}^p = [\varepsilon_{A+1,1},\ldots,\varepsilon_{A+1,p}]$$

where

p the number of periods a season;

$x_{N,t}$ the real sales value for season N at period t;

$u^i_{N,t}$ the value of the i^{th} EE variable for season N at period t;

$\hat{X}_{A+1}^p$ a p-row vector of sales estimation for season $A+1$ from period 1 to p;

X_N^p represents a p-row vector of real sales data known for the season N, from period 1 to p;

$U_N^{n,p}$ an nxp-matrix of the p values of the n EE variables for season N;

F a non-linear function: $R^{p[A.(n+1)+n]} \rightarrow R^p$;

θ the parameter vector ($\in R^m$) of the model;

ε_t the prediction error at period t ($\varepsilon_t = \hat{x}_t - x_t$).

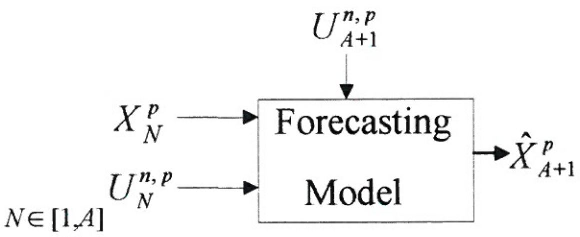

Figure 5. General Form of the Prediction Model

Many models of time series prediction exist (see Section 3). Only non-linear patterns adapted to the treatment of very noisy series are selected. Most of them are based on series statistical analysis [6], [18]. Even if in the last decades, optimization procedures of these models have been developed, no one of these yields better results in an uncertain environment [19].

To improve the results obtained with traditional models, it is imperative to introduce EE variables that characterize the non-linearity of the structure. In addition, the relative complexity of the classical models comfort users to use simpler but more interpretable procedures, with as less intervention of operators as possible, except for judgmental aspects that could help the model learning. Indeed, we consider it is fastidious for the user to significantly improve the learning procedure of the model when identifying statistical model behaviors. However, it could be more rational to introduce his expert judgment through identified market rules, specific to textile items of his distribution firm.

In this paper, our attention has focused on the ability of neural networks to learn the influence of many input variables, even in an uncertain environment, and on the fuzzy theory, to allow the use of interpretable rules by the user, to characterize the influence of input variables.

3. Classical Models for Comparison

Many forecasting models can be listed. Out of these, deterministic models, generally used for chaotic series forecasting, are reserved for marginal applications [20]. More currently used, stochastic models are based on probability and stochastic process theories. The more renowned stochastic models employ mobile average, autoregressive, exponential smoothing, Box&Jenkins or dynamic regression methods. They may consider explanatory variables or they may not. More recently, some works have proposed neural network- and fuzzy logic- based models. Mathematically shown to be at least equivalent to statistic models [21], they nevertheless meet new end-user criteria, for instance: interpretation of the influence of exogenous factors, transparency of the way how the model works, or an easy intervention on the model parameters [10].

Our comparison test has made use of the ARX (AutoRegressive with eXplanatory variables) model linked to the well -known Box&Jenkins (BJ) procedures [4], [22], [23], a dynamic regression model used by the professional software Forecast Pro, which is a market reference, and a Neural AutoRegressive network with eXplanatory variables (NARX) based model, called ENARX, built in precedent works [24], which allows a comparison with neural forecasting models. These three models include explanatory variables, the only condition to improve forecast. Nevertheless, despite the expected performances, the short noisy data involve difficulties for the models to map the rules of the time series. Thus, we also use, for comparison with exponential smoothing models, the simple but efficient Holt-Winters model with multiplicative seasonality (HWS) [4], [23].

4. HNCCX and HFCCX Forecasting Models

As previously shown, classical models are based on a smoothed process, on autoregressive and mobile average process, on dynamic regression process, or a neural multiconnected network. Such models present major drawbacks. The two main disadvantages are the important number of parameters and the low consideration of the seasonality factors, in the case of short time series. The interest of a neural or fuzzy-based model is the ability to map the non-linear influences of the explanatory input variables [25].

The HNCCX and HFCCX models are new alternatives to the forecasting problem stated previously. First, the general form of the models is reminded. These models require a seasonality-based forecasting method to be implemented. Second, the characteristics of HNCCX and HFCCX are summarized.

4.1 General Model

The HNCCX and HFCCX models are Hybrid forecasting models with Neural or Fuzzy estimation of Corrective Coefficients of the eXplanatory variables influence. The two main reasons justifying the building of this model are the singular definition of the problem (see Section 2.1), and the treatment of the influence of the EE variables with structure-limited neural or fuzzy estimators, that consequently allows more numerous input variables. A third, interesting, reason for such a model is that it originates from judgmental treatment of the forecast produced by marketing managers in textile distribution.

The prediction process is divided into three stages:

- sales data are deseasonalized from the influence of EE variables,
- resulting data are used to predict the sales of the next season, with seasonality-based forecasting method,
- sales forecasting are reseasonalized with the influence of EE variables corresponding to the next season.

The general form of the model is shown in Figure 6.

The MH function converts the original sales data to deseasonalized ones X'^p_A, from the corrective coefficients estimated by the MCS function.

$$X'^p_N = \mathrm{MH}(X^p_N, CS^p_N),\ \forall N \in [1, A]$$

The MCS function estimates the influence of EE variables (quantitative indicators) and changes these variables into corrective coefficients of seasonality CS_A^p. It is written:

$$CS_N^p = \mathrm{MCS}(U_N^{n,p}),\ \forall N \in [1, A]$$

The neural (for HNCCX) and the fuzzy (for HFCCX) based MCS models are introduced in Section 4.2 and Section 4.3.

The FM function predicts the sales on season $A+1$ from the deseasonalized sales data known until season A. The method used is one of the simplest forecasting methods: the seasonality mean. Only the seasonality forecast was studied here, the total quantity sold being for our distributor partner less problematic. This basic method was selected in order to show the significant influence of the explanatory variables treatment on the forecast improvement. Besides, its small number of parameters does not require any adjustment as other classical models do. For example, it could be replaced by a traditional ARIMA-based or Holt-Winter method including seasonality treatment.

The obtained predicting data $\hat{X}'^p_{A+1}$ are reseasonalized thanks to the inverse function of MH defined before. Thus

$$\hat{X}^p_{A+1} = \mathrm{MH}^{-1}(\hat{X}'^p_{A+1}, CS^p_{A+1})$$

The coefficients CS^p_{A+1} required by MH^{-1} are simulated with the same MCS function, previously learned on seasons $1,\ldots,A$, from the EE variables U^n_{A+1} of season $A+1$.

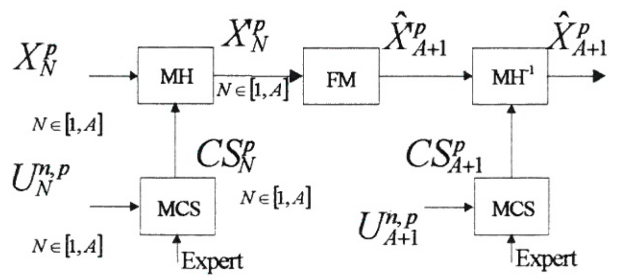

Figure 6. General Form of the HNCCX or HFCCX Forecasting Models

4.2 HNCCX Model

This Section sums up the features of the HNCCX model, the MCS function of which is shown in Figure 7. The neural network used for this function is a Multilayer Feedforward network (MLFF) [26], [27]. The choice is motivated by the network ability of generalization, the large spectrum of updated learning algorithms and particularly by the minimization of the number of parameters proportional to the dimension of the input space [19], [28].

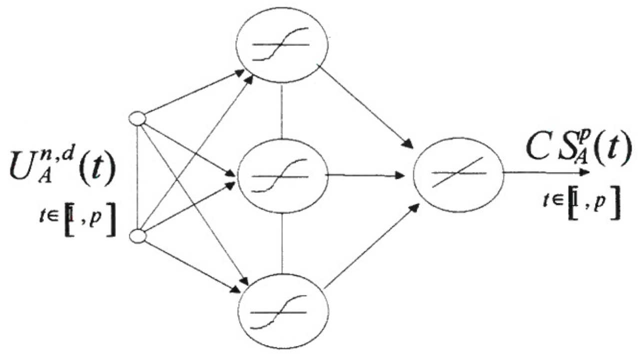

Figure 7. The MCS Function of the HNCCX Model

4.2.1 Model Features

To model the training inputs / output set, a sufficient number of hidden layers and units is necessary. However, a too complex network will not be able to generalize over the simulation set [29], [30]. According to the short historical sales data, the size of the training set (Number of Equations) is limited. Thus the number of parameters (weights and biases), which must be inferior or equal to this number of equations, is restricted.

Activation functions of hidden layer units chosen are sigmoid functions whereas for output units, a linear function is selected.

4.2.2 Learning Procedure

Learning procedure includes an initialization procedure, and minimization, generalization, robustness and weight elimination methods.

The initialization procedure used is based on the Nguyen and Widrow procedure [31], adapted by Demuth and Beale [32].

The learning method chosen is the one step secant algorithm introduced by [33]. This algorithm is presented as a compromise between the BFGS and the conjugate gradient procedures in terms of storage, computation and convergence.

To avoid overfitting, the generalization of the network is first obtained through the selection of a limited structure (number of hidden layers and units), see 4.2.1, and then through the cross validation procedure (also called early stopping) [6].

The robustness method consists in selecting the derivative of the cost functions to the parameters of the model as the influence function.

To simultaneously reduce the complexity of the network during learning, for weight elimination, the statistical stepwise method (SSM) proposed by [34], and experimented in [20] has been used.

4.3 HFCCX Model

This Section summarizes the feature of the HFCCX model, the MCS function of which is shown in Figure 8. The Fuzzy Inference System (FIS) used for this function is a Takagi - Sugeno FIS [35]. This model has the advantage of applying the techniques of training and optimization with more facility than the model of the Mamdani type does[14].

Like for the neural model, the design of fuzzy system requires cleaning and clustering of the training data, scaling of the data, initialization and learning procedures for the rules set selection and tuning of the membership functions parameters.

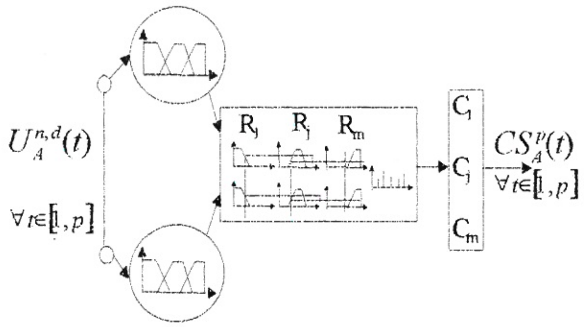

Figure 8. The MCS Function of the HFCCX Model

4.3.1 Model Features

For this analysis, trapezoidal membership functions have been chosen to compute the input degree- of-membership to each fuzzy set.

Output membership functions are singletons. Each rule is associated with an output membership function.

The rules set defined by defect (i.e. without training) consists of all the combinations of rules, which it is possible to form by having all the inputs in each rule.

Fuzzy operators chosen in this study are the product function for the AND operator and the maximum function for the OR operator. The same way, the implication method used is the product function and the aggregation method used just includes all singletons.

4.3.2 Learning Procedure

In general, mainly two approaches can be distinguished: the combination of neural networks (which are learned thanks to such methods as backpropagation) and fuzzy systems [15], [36], [37], [38] or, the use of evolution algorithms such as genetic algorithms[12], [14], [15], [36], [37], [39].

The choice of the optimization method depends on the parameters type to be optimized. Indeed, three types of parameters can be trained in a fuzzy system: input membership, output membership and structure learning. Each category of parameters can be treated by a particular method [36], [40]: linear or non-linear optimization, genetic algorithms...

The method this study makes use of, selects the rules set which gives more precision by genetic algorithms, and tunes, with the Gauss -Newton method, output membership functions of each set of the selected rules.

Each chromosome is thus a chain of m (= rules number) binary numbers that translates rules set. When gene i of a chromosome is equal to " 1 ", then rule i is selected. Conversely, when this gene is equal to " 0 ", the rule is not selected.

From each chromosome, it is thus possible that a fuzzy inference system (FIS) is built, which allows to work out a forecast. Also, the fitness function is proportional to the RMSE criterion (see Section 5.1).

4.4 Evolution of the HFCCX Model

The number of rules for a complete rules set (all inputs are implied in each rule) is equal to: $m = \prod_{i=1}^{n} n_i$, where n_i is the number of membership functions of the input variable i, and n is the number of input variables. By using this complete rules set, when the input number becomes higher than three the number of parameters to be estimated becomes too important, the interpretation of the rules becomes very difficult and some rules can damage system performance [10], [40]. The model structure must also be the smallest possible in order to avoid overfitting problems [41] ; a rules selection becomes then necessary. However, the fitness function just considers the RMSE criterion. Consequently, the rules number stays too significant for an efficient generalization. Thus in order to perform a selection comparable to a neural networks pruning, the number of parameters is penalized in the fitness function. Therefore, the evaluation of the fitness function is proportional to the criterion RMSE + C x SBIC, where C is a constant whose value is function of data used and SBIC is the Scharw's Bayesian Information Criterion (see Section 5.1). RMSE criterion penalizes prediction deviations and SBIC criterion penalizes number of rules. Then, the fitness function allows to improve the accuracy (through the RMSE criterion) with a minimal rules set (through the SBIC criterion).

5. Comparative Test

5.1 Prediction Performance Evaluation

5.1.1 Accuracy

In time series identification and forecasting, a common criterion of the model quality is the root mean square error (RMSE). Given the p-dimension, test set is evaluated on the forecasting season A+1 and the previous quality criteria are then:

$$RMSE_{A+1} = \sqrt{\frac{1}{p}\sum_{t=1}^{t=p}\left(x_{A+1,t} - \hat{x}_{A+1,t}\right)^2}$$

Such criteria penalize important prediction deviations [4].

However, it is also necessary to penalize the number of parameters m of the modeling function, according to the objectives of generalization. Thus to measure the model quality of ARMA or neural networks models, we use the Scharw's Bayesian Information Criterion (SBIC) [4]. As previously said, the smaller the criterion, the better the quality model. They are expressed as follows:

$$SBIC_{A+1} = p.\ln\left(RMSE_{A+1}^{\ 2}\right) + m.\ln(p)$$

with p the number of data considered and m the parameters number of the model.

5.1.2 Robustness and Generalization

The robustness and the generalization of the models are implicitly tested through the learning procedure. The robustness aspect of the forecasting models is motivated by the strongly noisy data used. The generalization of the forecasting models is tested through the comparison of different sales data sets selected from among several articles of diverse colors.

5.2 Sales Data and Explanatory Variables

5.2.1 Sales Data

The choice of sales forecasting data for evaluation is difficult to make. First, and especially in textile industry, no sales data can be considered as representative (see Section 2.1 and Section 2.2). Secondly, such data are not always available and stored at the aggregation level required.

To perform the comparative test, we have chosen a basic textile item, sold during a complete season period. According to the aggregation levels shown in Figure 3, this textile item can be categorized as an Autumn-Winter and man's city- wear item. The chosen item family and model are quantitatively known to be representative for distributors. The test is experimented with 8 sales data series, according to color references. Each data series is composed of 22 periods a season, over 3 years or seasons, including 2 years for the learning.

5.2.2 Explanatory Variables

The input space is built with the EE series that influence the sales prediction. Naturally, the EE variables to be chosen are the most consistent ones. Another requirement is the continuity and the scaling of variables. The explanatory chosen variables are also considered to be very representative, according to marketing managers. Figure 9 shows the three variables used, which are based on the selling price, the holiday periods, and the season sequences. The price is a very significant argument in the purchase decision of the customer and the holidays are generally responsible for a store frequentation rise. The selling price and the holiday periods do not have the same impact on sales according to the season: for example, each year a fall of price before Christmas will not influence the sales as a fall of price in October. That is why the variable based on season sequences is being considered.

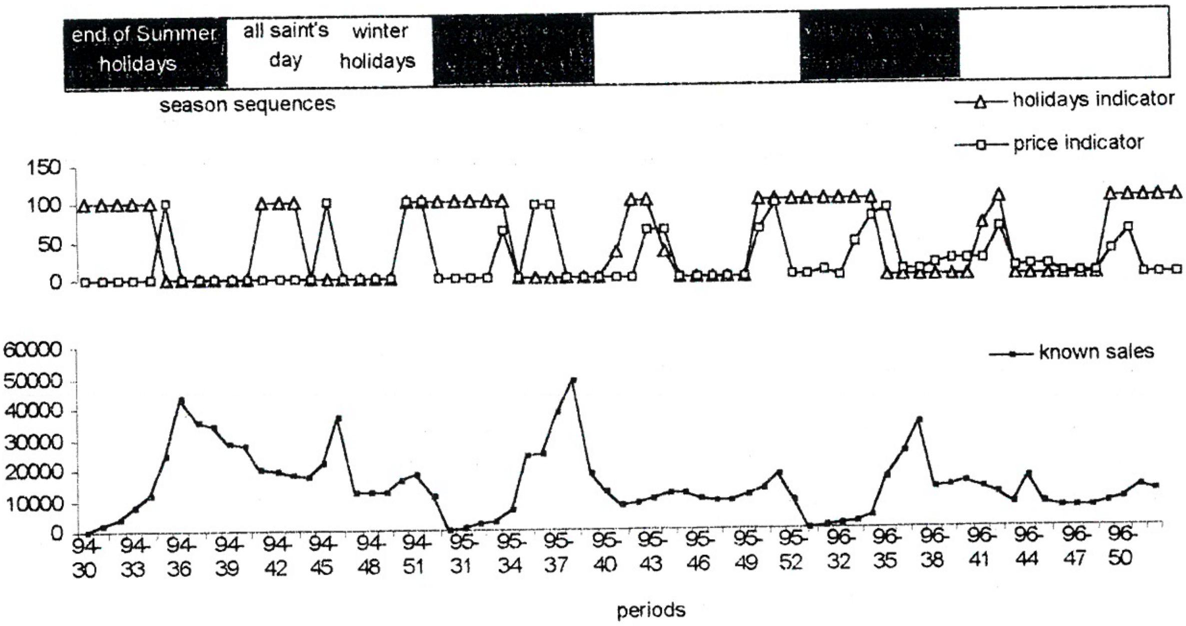

Figure 9. Input Indicators (Price, Season, and Holiday) and Known Sales History

To introduce explanatory variables in the input space of the forecasting models, some modifications need be done. Indeed, as before, these variables must respond to continuity and normalization criteria:

- Selling price. The possessed values correspond to the percentage of stores that practice a decrease of price. In fact, the selling price currently swings between the fixed normal price and a promotion price. The used variable is a price indicator which is proportional to the possessed value.

- Holiday periods. The holiday indicator represents the number of the country areas in holiday time.
- Season sequences. This indicator separates three distinguished sequences in the Autumn-Winter season. These represent the beginning of the term, after the Summer holidays , the All Saints' day period, and the proximity of Winter holidays (see Figure 9).

5.3 Models Tested

The six models used in this comparative test are described in Section 3.

Table 1 shows the features of the different models employed.

Table 1. Features of the Comparison Models

Forecasting model	abbreviation	seasonal model	explanatory variables	forecast horizon	learning methods	parameters	number of parameters	judgmental or expert intervention
Holt-Winter with Seasonality	HWS	yes	no	1 year	Newton based	α, β, γ	3	/
Box&Jenkins	BJ	no	yes	1 year	Newton based	q, r, v(***), w(***)	$p + r + v$	possible choice of model orders
Forecast Pro - Dynamic Regression	DRX-FP	no	yes	1 year	Newton based	regressors	number of regressors	regressor choice (*******)
Neural AutoRegressive network with Explanatory variables based	ENARX	no	yes	1 year	OSS – SSM(******)	$w_{i,j}$, b_i	NP(****)	/
Hybrid Neural Model	HNCCX	yes	yes	1 year	Newton - LM – SSM(*)	$w_{i,j}$, b_i	NP(****)	intuitive expert correction of explanatory variables influence
Hybrid Fuzzy Model	HFCCX	yes	yes	1 year	LM - GA(**)	c_i	m(*****)	intuitive expert correction of explanatory variables influence

(*) Levenberg-Marquardt learning algorithm, Statistical Stepwise Method for weight elimination - (**) Levenberg-Marquardt optimization of membership functions, genetic algorithm for rules selection

(***)number of lagged explanatory variables - (****) - (*****) number of selected rules - (******) One Step secant procedure [Battiti 1992], Statistical Stepwise Method – (*******) the choice is directed by the software

5.4 Comparison Results

This comparison test is carried out based on the following criteria: the RMSE and SBIC criteria, and the model capacity for learning with short time series. Table 2 summarizes the performances of the models, in case of short time series, seasonality, and explanatory variables under an uncertain environment. Figure 10, obtained with the total sales of the textile items, graphically illustrates the main performances of each forecasting model.

As envisaged, the short time histories do not allow classical models (especially DRX-FP and ENARX models) to optimally model the series characteristics. The learning processes of these models need more input data for statistical validation. Besides, the Forecast Pro© software alerts that the data number is too small for a correct usage. To correctly adjust the network to the data set, the structure of ENARX model, with a very restricted number of layers and units, is not sufficient. Although the BJ model yields better results than the two preceding ones, these forecasts remain extremely disturbed. Once again, a more significant training data set would improve the model accuracy.

The HW model does not take into account explanatory variables, but it introduces the seasonality factor. Even if it is difficult to evaluate numerous factors, particularly because of their interdependence, seasonality appears to be one of the most important factors, according to the structure of data series. However, Figure 10 illustrates that the model which ignores the EE variables (i.e. the HWS model), is an undesirable model for our context. Indeed, this model, which considers only history sales, cannot reflect the change of "sales peak" due to different holidays periods or price fall.

Table 2. Models Performances Comparison on A Basic Winter Textile Item

Models	RMSE on total sales	RMSE improvement referring to					
		HWS	BJ	DRX-FP	ENARX	HNCCX	HFCCX
HWS	4873	-	-30.5%	76.9%	-5.1%	-38.0%	-39.8%
BJ	3386	43.9%	-	154.6%	36.6%	-10.7%	-13.3%
DRX-FP	8620	-43.5%	-60.7%	-	-46.3%	-64.9%	-66.0%
ENARX	4625	5.4%	-26.8%	86.4%	-	-34.6%	-36.6%
HNCCX	3023	61.2%	12.0%	185.1%	53.0%	-	-2.9%
HFCCX	2934	66.1%	15.4%	193.7%	57.6%	3.0%	-

Models	SBIC on total sales	SBIC improvement referring to					
		HWS	BJ	DRX-FP	ENARX	HNCCX	HFCCX
HWS	383	-	18.4%	6.6%	75.3%	13.9%	13.5%
BJ	453	-15.6%	-	-10.0%	48.0%	-3.8%	-4.1%
DRX-FP	408	-6.2%	11.1%	-	64.5%	6.9%	6.6%
ENARX	671	-43.0%	-32.4%	-39.2%	-	-35.0%	-35.2%
HNCCX	436	-12.2%	4.0%	-6.4%	53.9%	-	-0.3%
HFCCX	435	-11.9%	4.3%	-6.2%	54.4%	0.3%	-

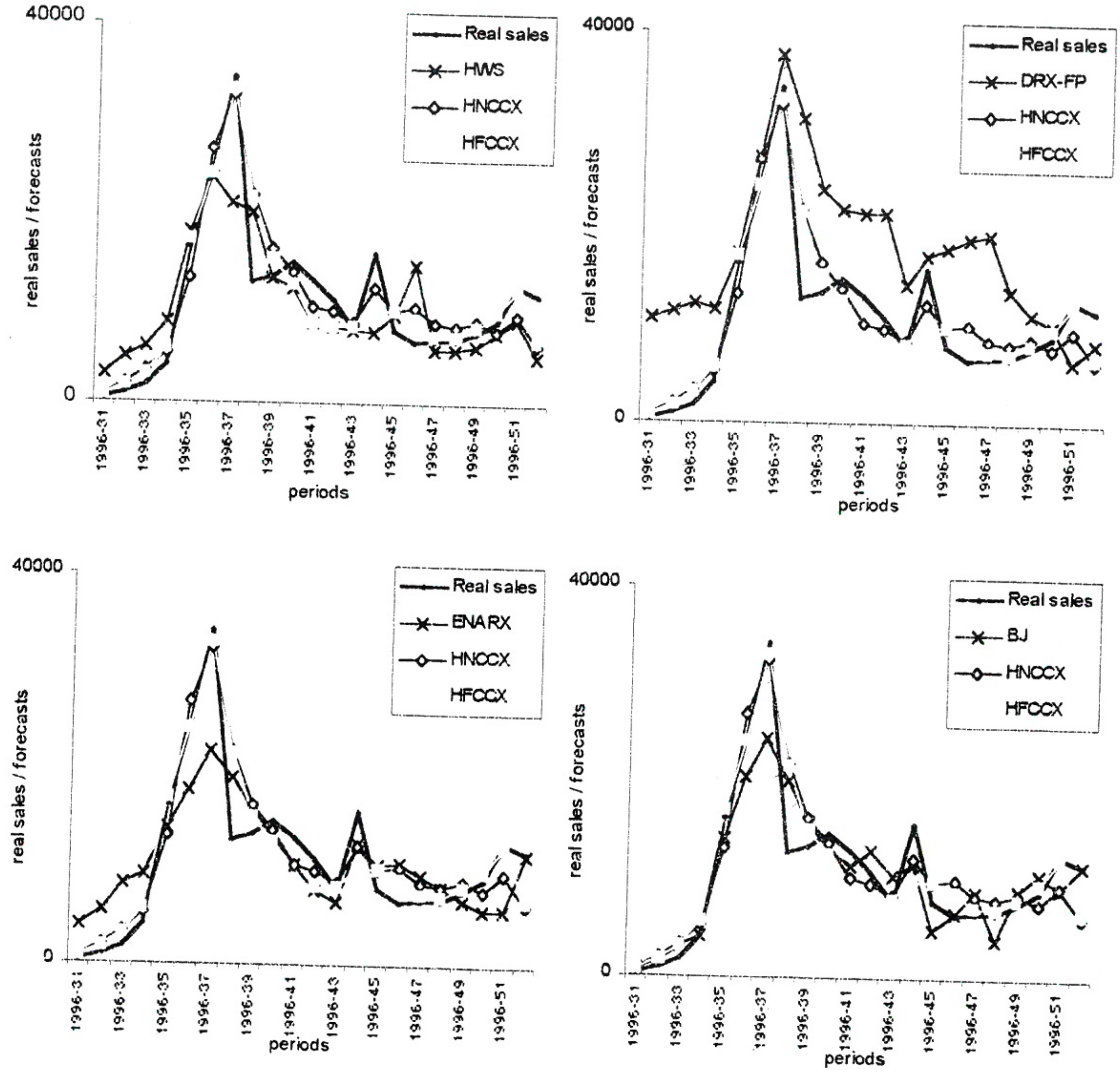

Figure 10. Comparison Between Forecasting Models

As concerns the ability of models that introduce the influence of explanatory variables, to map the nonlinear influences of these variables, Figure 10 shows that, at the end of Summer holidays, HNCCX and HFCCX models will better reply to the sales upsurge than the other models. In the week 44 of the year 1996, the sales have sensibly increased, at a delayed period, with respect to the previous season. The increase should correspond to the combined influence exerted through price decrease, during the All Saints' holidays. At Christmas time, sales keep on increasing until the week 51, due to the Christmas's day and price changes, and then fall, despite the price drop. Figure 10 relates that, on the All Saints' Day problem, only the HNCCX model forecasts a punctual improvement of sales whereas, at Christmas time, the HFCCX model prevails.

This remark is still valid for almost all series. It confirms the features of neural network and fuzzy inference system to map nonlinear relations between inputs and output. It also relates the limits of seasonal-based models, like HWS model that strictly learns the past. Indeed, Figure 9 justifies that the latter model which learned the improvement of sales on the week 46, appeared two years sooner.

Even if the HNCCX and HFCCX models globally give the best results of the comparative test, they have at least one major drawback: the important number of parameters of the neural and fuzzy functions, that requires, for the model learning, a minimum input space. This problem quickly improves with the number of input variables. However, in comparison with an NARX-based classical model [24], the HNCCX model allows the treatment of as many variables with less parameters, and thus with shorter training sets, as possible [9]. The modifications done on the HFCCX model, so that to reduce the structure complexity, decreased the SBIC criterion (and consequently increased the capacity of generalization) while preserving the accuracy of the RMSE criterion at the same value. Besides, the number of parameters keeps smaller than in the case of a commonly used forecasting model only based on a neural network (the ENARX model).

Roughly, the HNCCX and HFCCX models have similar performances. However, their effectiveness sometimes differs over all the season. Fuzzy inference system has the advantage of running through interpretable rules, whereas neural network can be regarded as a "black box". Neural networks show correct learning capacity with no need to define rules; furthermore, they generally employ faster and easier to use learning methods than fuzzy systems do.

6. Conclusion

The aim of this paper has been to demonstrate the advantages brought by such tools as neural networks and fuzzy logic to perform short time series forecasting, with a mean-term horizon, and explanatory variables under an uncertain environment. The hybrid neural model (HNCCX) and the hybrid fuzzy model (HFCCX) have been contrasted on two representative Winter textile items in several colors, with some classical models: the Holt-Winters with seasonality model (HWS), the Box &Jenkins method (BJ), a dynamic regression model (DRX-FP) experimented on the renowned professional Forecast Pro © software and on an NARX-based model (ENARX). All models have been optimized automatically or with the help of human expertise, assisted by statistical diagnostics. Due to the lack of data, the comparison has not been extended to a more significant family of textile items. However, the articles chosen are, according to our distributor partner, typical articles.

Short time series and an uncertain environment mainly account for the relatively good results of the BJ and DRX-FP models, unable to introduce seasonality factors. The HWS model provides better results, even though without the introduction of explanatory variables. Nevertheless, the only way of improving these results is to consider the influence of endogenous and exogenous factors. The ENARX model requires a more significant training data set, so as to use the capacity of its neural network. The best results are obtained with the HNCCX and HFCCX models, due to the ability of neural network and fuzzy logic to map the nonlinear pattern between inputs and output. Despite these results, the negative aspect of the latter models is still their number of parameters, penalized by the SBIC criterion, during the comparative test. However, the number of parameters keeps below that of a forecasting model only based on a neural network. Another problem faced with is the need of an expert correction of the influence of explanatory variables on the learning stage. Indeed, even if this characteristic allows human intervention, the generalization to a large number of items, for example all items of a distributor, seems almost impossible. In spite of this, our models stay convenient for classes of items with similar behavior.

Some possible developments of the model consist in the improvement of learning procedures. An automatic learning of the influence of the explanatory variables becomes essential to an industrial development. The idea is then to build a model that introduces at once the seasonality factors, the ability of neural nets to compute the influence of explanatory variables under uncertain environment or of fuzzy systems to map the forecasters experience, without penalizing the learning process with too many parameters. Such evolutions will be studied in future works.

REFERENCES

1. KINDADE, D. H., CASSILL, N. and WILLIAMSON, N., **The Quick Response Management System: Structure and Components for the Apparel Industry**, JOURNAL OF THE TEXTILE INSTITUTE, Vol. 84, No. 2, 1993, pp. 147-155.

2. LEE, H. L. and SASSER, M. M., **Product Universality and Design for Supply Chain Management**, PRODUCTION PLANNING AND CONTROL, Vol. 6, No. 3, 1995, pp. 270-277.

3. SBOUI, S., RABENASOLO, B., JOLLY-DESODT, A. M. and DEWEALE, N., **Optimisation dynamique de la taille des lots d'approvisionnement : application à la filière textile**, MOSIM '01, Vol. 2, Troyes, 25-27 April 2001, pp. 551-558.

4. BOURBONNAIS, R. and USUNIER, J. C., **Pratique de la prévision des ventes - Conception de systèmes**, EDITIONS ECONOMICA, Paris, 1992.

5. WHEELWRIGHT, S. and MAKRIDAKIS, S., **Forecasting Methods For Management**, 4th edition, WILEY, 1985.

6. PATTERSON, D. W., **Artificial Neural Networks – Theory and Applications**, PRENTICE HALL, 1996.

7. KIM, D. and KIM, C., **Forecasting Time Series With Genetic Fuzzy Predictor Ensemble**, IEEE TRANSACTIONS ON FUZZY SYSTEMS, Vol. 5, No. 4, November 1997, pp. 523-535.

8. DE TONI, A. and MENEGHETTI, A., **The Production Planning Process for A Network of Firms in the Textile-apparel Industry**, INTERNATIONAL JOURNAL OF PRODUCTION ECONOMICS, Vol. 65, 2000, pp. 17-32.

9. VROMAN, P., HAPPIETTE, M. and VASSEUR, C., **A Hybrid Neural Model for Mean-term Sales Forecasting Of Textile Items**, STUDIES IN INFORMATICS AND CONTROL, Vol. 10, No.2, June 2001, pp. 149-167.

10. ZADEH, L. A., **The Roles of Fuzzy Logic and Soft-Computing in the Conception, Design and Deployment of Intelligent Systems**, BT TECHNOLOGY JOURNAL, Vol. 14, No. 4, 1996, pp. 32-36.

11. WEIGEND, A. S. and NEIL, A., **Results of the Time Series Prediction Competition at the Santa Fe Institute**, IEEE International Conference on Neural Networks, San Francisco, CA, 1993.

12. SHIMOJIMA, K., FUKUDA, T. and HASEGAWA, Y., **Self-tuning Fuzzy Modeling with Adaptive Membership Function, Rules, and Hierarchical Structure Based on Genetic Algorithm**, FUZZY SETS AND SYSTEMS 71, 1995, pp. 295-309.

13. ZADEH, L. A., **Soft Computing and Fuzzy Logic**, selected papers, 1994, pp. 48-56.

14. VAN LITH, P. F., BETLEM, B. H. L. and ROFFEL, B., **Fuzzy Clustering, Genetic Algorithms and Neuro-Fuzzy Methods Compared for Hybrid Fuzzy-First Principles Modeling**, ISIAC 2000, June 11-16, 2000.

15. BERSINI, H., **Génération automatique des systèmes de commande floue par les méthodes de gradient et les algorithmes génétiques**, Proceedings of the "Journées Nationales : les applications des systèmes flous", 1992

16. VROMAN, P., RABENASOLO, B., HAPPIETTE, M. and VASSEUR, C., **Optimization of A Neural Network Structure for Textile Sales Forecasting**, IMACS/IEEE CSCC'99 International Multiconference, Athens, Greece, July 1999.

17. BOUSSU, F., HAPPIETTE, M. and RABENASOLO, B., **Sales Partition for Forecasting into Textile Distribution Network**, Proceedings of the IEEE/SMC International Conference on Systems, Man and Cybernetics, Vol. 4, Beijing, China, 1996, pp. 2868-2873.

18. DE GOOIJER, J. G. and KUMAR, K., **Some Recent Development in Non Linear Time Series Modelling, Testing, and Forecasting**, INTERNATIONAL JOURNAL OF FORECASTING 8, 1992, pp. 135-156.

19. SUYKENS, J. A. K. and VANDEWALLE, J., **Nonlinear Modeling: Advanced Black-Box Techniques**, KLUWER ACADEMIC PUBLISHERS, The Netherlands, 1998.

20. MANGEAS, M., **Propriétés statistiques des modèles paramétriques non-linéaires de prévision de séries temporelles: application aux réseaux de neurones à propagation directe**, Thèse de l'Université Paris I, 1997.

21. THIRIA, S., LECHEVALLIER, Y., GASCUEL, O. and CANU, S., **Statistique et méthodes neuronales**, DUNOD, Paris, 1997.

22. BOX, G.E. P. and JENKINS, G. M., **Time Series Analysis - Forecasting and Control**, PRENTICE HALL, 1969.

23. LEWANDOSKY, R., **La prévision à court terme**, DUNOD, Paris, 1983.

24. VROMAN, P., HAPPIETTE, M. and VASSEUR, C., **Design of A Feedforward Neural Network Model for Textile Sales Forecasting**, WAC / ISIAC 2000, 4th World Automation Congress, Hawaii, June 11-16, 2000.

25. YEN, J. and LIAO, J. C., **Using Fuzzy Logic and Genetic Algorithm for Metabolic Modeling**, IEEE, 1996.

26. HORNICK, K., STINCHCOMBE, M. and WHITE, H., **Multilayer Feedforward Networks Are Universal Approximators**, NEURAL NETWORKS 2, 1989, pp. 359-366.

27. LESHNO, M., YA LIN, V., PINKUS, A. and SCHOCKEN, S, **Multilayer Feedforward Networks With Nonpolynomial Activation Function Can Approximate Any Function**, NEURAL NETWORKS 6, 1993, pp. 861-867.

28. THOMAS, P., **Contribution à l'identification de systèmes non linéaires par réseaux de neurones**, Thèse de l'Université Henri Poincaré, Nancy 1, 1997.

29. KARTALOPOULOS, S. V., **Understanding Neural Networks and Fuzzy Logic – Basic Concepts and Application**, IEEE PRESS, 1996.

30. AZOFF, E. M., **Neural Network Time Series Forecasting of Financial Markets**, WILEY FINANCE, England, 1994.

31. NGUYEN, D. and WIDROW, B., **Improving the Learning Speed of 2-layer Neural Networks by Choosing Initial Value of the Adaptive Weights**, Proceedings of International Joint Conference on Neural Networks, Vol. 3, 1990, pp. 21-26.

32. DEMUTH, H. B. and BEALE, M. H., **Neural Network Toolbox User's Guide**, THE MATHWORKS INC., MA, 1994.

33. BATTITI, R., **First and Second Order Methods for Learning: Between Steepest Descent and Newton's Method**, NEURAL COMPUTATION, Vol. 4, No. 2, 1992, pp. 141-166.

34. COTTRELL, M. et al, **Neural Modeling for Time Series: A Statistical Stepwise Method for Weight Elimination**, IEEE TRANSACTIONS ON NEURAL NETWORKS, Vol. 6, No. 6, 1995, 10p.

35. TAKAGI, T. and SUGENO, M., **Fuzzy Identification of Systems and Its Applications**, IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS, Vol. 15, No. 1, 1985, pp. 116-132.

36. OKORNY, M. and CERMAK, P., **Fuzzy Models Identification and Tuning Using the Genetic Algorithms and Neural Network**, EUFIT '97, September 8-11, 1997, pp. 570-573.

37. HARTANI, R., NGUYEN, H.T. and BOUCHON-MEUNIER, B., **Approximation universelle des systèmes flous**, RAIRO – APII - JESA, Vol. 30, No. 5, 1996, pp. 645-663.

38. WU, T. P. and CHEN, S. M., **A New Method for Constructing Membership Functions and Fuzzy Rules from Training Examples**, IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS – PART B: CYBERNETICS, Vol. 29, No. 1, February 1999, pp. 25-40.

39. KLIR, G. J. and YUAN, B., **Fuzzy Sets and Systems, Theory and Applications**, PRENTICE HALL, Englewood Cliffs, NJ, 1995.

40. NELLES, O., FISHER, M. and MULLER, B., **Fuzzy Rules Extraction By A Genetic Algorithm and Constrained Nonlinear Optimization of Membership Functions**, IEEE, 1996.

41. FIORDALISO, A., **A Nonlinear Forecast Combinaison Method Based on Takagi-Sugeno Fuzzy Systems**, INTERNATIONAL JOURNAL OF FORECASTING 14, 1998, pp. 367-379.