008-0219

USING THE FUZZY INFERENCE SYSTEM AS A PROPOSED METHOD TO DEVELOP OPERATIONS STRATEGIES IN THE STOCK MARKET

Fabio Augusto L. Cichini - UNESP - São Paulo State University

Av. Eng. Luiz Edmundo Carrijo Coube, 14-01 – Bauru – SP – Brazil

E-mail: fabiocichini@bol.com.br - Phone: 55 (14) 31036122

Rosani de Castro - UNESP - São Paulo State University

Av. Eng. Luiz Edmundo Carrijo Coube, 14-01 – Bauru – SP – Brazil

E-mail: rosani@feb.unesp.br - Phone: 55 (14) 31036122

Rogerio A. Flauzino UNESP - São Paulo State University

Av. Eng. Luiz Edmundo Carrijo Coube, 14-01 – Bauru – SP – Brazil

E-mail: raflauzino@feb.unesp.br - Phone: 55 (14) 31036122

POMS 19th Annual Conference

La Jolla, California, U.S.A.

May 9 to May 12, 2008

Abstract

In Brazil, managing investment funds through artificial intelligence-based systems is an activity still in its developing stages, however, in 2006 the managed volume reached 0,8% of the 800 billion reais negotiated by the fund industry in Brazil (BALTAZAR, 2006).

Considering the growing demand for this type of computing tool, a system based on the Fuzzy Logic has been developed to propose a method to develop operation strategies in the stock market, with the possibility of improving investment options. Generally speaking, in applications involving Fuzzy modeling and systems, it is convenient to use a cost function expressing the deviation between the desired results and those supplied by the system and, by means of a parametric adjustment, we obtain rules that minimize the cost function. The obtained results picture the efficiency of the adjustment algorithm developed and motivates future investigations.

Keywords: Stock market, Fuzzy inference system, quotation estimates, operations strategies

1. Introduction

According to Ross (1995), the role of the financial market is to promote the encounter between two converging interests: Acquiring resources and saving resources. The financial market is very important for today's society, because "it allows a permanent adjustment of the financing needs and capability of the economic agents – companies, financial institutions, individuals, Government and public groups" (SAINT-GEOURS, 1998).

According to Bovespa (São Paulo Stock Exchange-2000), Brazilian economy is going through several changes where organized production forms have developed and competition and productivity have increased. The new productive paradigm requires a capital increase. Due to this fact and also to the lack of old financing mechanisms, besides the stabilization and return of economical growth, companies have become more interested in open capital. Some companies that until recently did not consider the capital market as a financing source, have started to change their thinking and attitude. To further reinforce this paradigm, new companies without some of the traditional bad habits have appeared, and see the capital market as an alternative to obtain resource. In general, the prejudice that prevented companies from searching for financing in the capital market has been replaced by an open mindedness that appreciates its importance as a financing source and sees conditions for sustainable growth in it.

However, the changes in the capital market development, although favorable, are not enough. There are many aspects today that halt this development. Many of them do not have a short term solution. The tax structure is counterproductive and, besides, it burdens the variable income operations with the CPMF (tax on financial trade); the lack of management structure in the Comissão dos Valores Mobiliários (securities and exchange commission), the preferential shares stock. All these factors imply in the lack of quality papers in the market, of well structured and solid companies with good growth perspectives (BOVESPA, 2000).

Despite some factors that still stem the effective development of the Brazilian capital market, a new type of computer tool appears in the world market, and it has strongly changed the way we work. Dealing with challenges expressed in mathematical terms, the computer is indeed an excellent tool. The investment funds management through computers has reached 1.5 trillion dollars, 7% of the total negotiated in the world. They are computer systems based on artificial intelligence which, besides executing financial assets estimations, also buy and sell them. They are called Quantitative Systems and, according to a research based on seventy American investment funds, it has been shown that they present better results than their human competitors. Management by these systems is in its first steps of development in Brazil, but in 2006 the volume thus managed doubled, reaching 0.8% of the 800 billion Reais traded by the funds industry in the country (BALTAZAT, 2006).

Based on this information, a computer system was developed using the historical data made available by Bovespa on the stock in its trade portfolio, statistics tools and an artificial intelligence technique called Fuzzy Logic, with the objective of estimating the value of stocks at Bovespa and thus obtaining the best investment options.

2. Stock market analysis tools

According to Noronha (2007), there are basically two methods for analyzing stock value in the capital market: fundamental analysis and technical analysis. Fundamental analysis takes into consideration the study of the causes of market movement. The fundamental school considers studies originated in economic-financial analysis of companies, their management patterns and the growth perspective in their segment. To perform this analysis, the behavior of the micro and macro segments of economy is observed. For Chaves (2004), the objective of fundamental analysis is to calculate the asset's intrinsic value, which is compared to the market and classified as satisfactory, evaluated or sub-evaluated. Next step is to decide and define which asset shall be negotiated.

On the other hand, technical analysis takes into account the market movement. The technical school uses information from data generated by prices and business volume oscillations, statistic charts and indicators, focusing on the historical market behavior to determine its current or future situation. The main objective of the technical analysis is to know when is the best moment to buy or sell an asset. For Murphy (1986), technical analysis is based on three principles: all the information on the agents and their expectations, as well as the psychological, political and economic fundaments are included in the market price; prices change according to tendencies; and the price behavior in the past repeats itself in the future. According to Noronha (2007), technical analysis does not consider the reason why prices change, but the reflexes in prices caused by their agents' interaction. For technical analysis, even if we know the fundamental factors affecting the price of a product or a company's stock, such as weather, strikes, political decisions and demand factors, we will not have all the factors needed to understand the changing in prices, because it is not those factors that affect them, but the way the market agents react to those factors. For technical analysis, it is the market that should be studied, because it is the only place where offer and demand, the mass psychology with its fears, beliefs, and expectations meet (BORGES, 2007).

As the fuzzy systems used here are built from historical data of the stocks, this study is based on technical analysis for, from a data set available in the market, it aims to estimate which will be the values assumed by the studied stock, thus obtaining the advantage of knowing the right moment to buy a certain asset and the best moment to sell it, getting the highest profit in the operation.

3. Fundaments and use of the Fuzzy Systems

According to the work of Shaw and Simões (1999), there still is a great disparity between the creative ability of a person to solve problems and the possibilities that computers provide for such solutions. This is due to the fact that humans reason in an uncertain, variable way, while computers work by precise, binary reasoning. A system that allows the machines to reason in the same imprecise way as human beings is called an intelligent system. Hence the term artificial intelligence, representing the study of how people solve problems and how machines can use this behavior to solve problems.

Fuzzy logic is a tool that uses the human way of reasoning to work in a control system. According to Yen and Langari (1999), fuzzy logic can be considered from two points of view. The first, in a strict sense, sees it as a logical system that generalizes the binary logical system. The second, in a comprehensive sense, sees fuzzy logic as all the theory and technology that employ fuzzy sets, classes without well-defined boundaries.

In the study of Gomide and Gudwin (1994), fuzzy logic can be accepted as the best way to represent human reasoning, which is partial and approximated. Fuzzy modeling and control are tools that aim to deal strictly with qualitative information.

Yen and Langary (1999), considering the basic concepts of fuzzy logic, explain: fuzzy sets; linguistic variables; distribution and possibilities; fuzzy rules of condition-action.

A fuzzy set is a set with a gradual boundary. The fuzzy set generalizes the classical set, because it allows a degree of partial membership of each element in relation to the set. For a classical set, an element belongs or does not belong to a set. On the other hand, in a fuzzy set, an element has a certain degree of membership in relation to the said set. This degree is defined in the interval [0,1], where 0 means total exclusion of an element in relation to the set, 1 means total inclusion of the element and any value between these two represents a partial membership.

Linguistic variables are the composition of a symbolic value and a numeric value. A linguistic variable assumes values in a set of linguistic terms to express the concepts and knowledge of human communication. For example, a ‘weight’ linguistic variable may have the terms {Low, Medium, High}. To assign the terms a numeric meaning, to each one of them we associate a fuzzy set, defined from a common speech universe, in this case represented by the linguistic variable 'Weight'.

Distribution of possibilities is the result of the association of a fuzzy set to a linguistic variable, to make known the membership function that controls this association, that is, which membership values each element of the linguistic variable will assume in relation to the fuzzy set.

Fuzzy rules of condition-action associate a set of conditions that describe parts of input data to an output action that will modify the system to the required conditions. Also called fuzzy rules, they are linguistic propositions given in the system input variables. For example, it may be that for a “High weight” the rule associates the control action “Cut down fat ingestion”.

After some basic definitions of fuzzy logic, we will deal more specifically with the use of this technique for complex system implementation.

According to Gomide and Gudwin (1994), the first step for the traditional implementation of a process is the derivation of the mathematical model that gives the detailed description of the process. Such proceeding requires a detailed knowledge of the process, something practically unfeasible for complex systems.

Kruse et al. (1994) says that a method of simplifying complex systems is to use an imprecision and uncertainty logic during the modeling phase. With fuzzy control systems, the attention is turned to decreasing complexity, achieved by the intelligent use of imperfect information. The quality of a model should not be measured by the precision degree of isolated data, but by the consideration of criteria such as accuracy, integrity, adequacy, efficiency and convenience in use. Therefore, it is not a surprise that a model presenting reduction of the system complexity by the inclusion of imperfect information can be better.

According to Hiemstra (1994), an adequate support system for stock investment decisions gives a precise prediction model that must define and store knowledge to offer a clear and interactive prediction. Fuzzy logic serves the purpose, because fuzzy systems are approximate models of Input-Output functions that do not require a mathematical model of how the Outputs depend on the Inputs.

According to Bilobrevec et al. (2004), fuzzy inference systems are applicable to problems that do not yet have methodology that is adjusted to resolution and control. Simplicity and low cost are extremely favorable factors to the fuzzy systems.

Considering this information, this study on stock value estimation was based on fuzzy logic, because the object in question represents a complex system of which the ruling mathematic system is unknown.

4. Multilayer Fuzzy System Model

Fuzzy inference systems can be understood as systems that use the concepts and operations defined by the fuzzy set theory, using the fuzzy inference process to perform its operational functions. Basically, these operational functions contain the fuzzification of the system inputs, the inference of the associated rules, the aggregation of the rules and the posterior defuzzificaton of the aggregation result, that is, the output or outputs of the fuzzy system.

We can observe from this that fuzzy inference systems have distinct and clearly defined functions. Therefore, to understand such systems, they can be represented by a multilayer model.

Considering the operational functions performed by fuzzy inference systems, a three-layer model is convenient to represent them. Thus, the fuzzy inference system proposed in this thesis can be given by the sequential composition of the input layer, the fuzzy rules inference layer and the output layer.

The input layer functions as connection of the input variables (coming from the outside environment) to the fuzzy inference system, as well as their fuzzyfication through the membership functions associated to each input.

In the fuzzy rules inference layer, or simply inference layer, the fuzzyficated input variables combine with each other, according to defined rules, using the operations defined by the fuzzy set theory as support.

The results of each individual rule are aggregated following a specific method, thus forming the output fuzzy set. This set is then defuzzyficated, resulting in the fuzzy inference system output. The aggregation process and the defuzzyfication process of the output fuzzy set are executed by the output layer. It is important to observe that the output layer, besides performing both processes described above, is also responsible for storing the membership functions of the output variables.

The three layers involved in the proposed multilayer fuzzy system are presented with more details in the next section.

4.1. Input Layer

As presented before, the fuzzy inference system input layer connects the inputs from the real world to the fuzzy system, and performs the fuzzyfication of these inputs according to membership functions associated to the fuzzy system.

The system inputs' fuzzyfication determines the membership degree of each input in relation to the fuzzy sets associated to each input variable. As many fuzzy sets as necessary can be associated to each fuzzy system input variable. Therefore, given a one-input fuzzy system, and to this input associated N membership functions, that is, N fuzzy sets defining the said input, the input layer output will be a column vector with N elements representing the input membership degrees in relation to these fuzzy sets.

If this one-input fuzzy system input is defined by scaler x, the fuzzy system input layer output will be vector $\boldsymbol{Y}$, that is:

$$\boldsymbol{Y}(x) = \boldsymbol{p}(x) = \begin{bmatrix} p_1(x) \\ p_2(x) \\ \vdots \\ p_N(x) \end{bmatrix} \tag{1}$$

where $p_k(.)$ represents the membership function defined for the x input related to the k-th fuzzy set associated to this input.

The generalization of the concept of input layer for a fuzzy system with m input variables is done if each of these fuzzy system inputs is modeled with a sublayer of the input layer. Therefore, the output vector of the input layer $\boldsymbol{Y(x)}$ can be represented as in (2), that is:

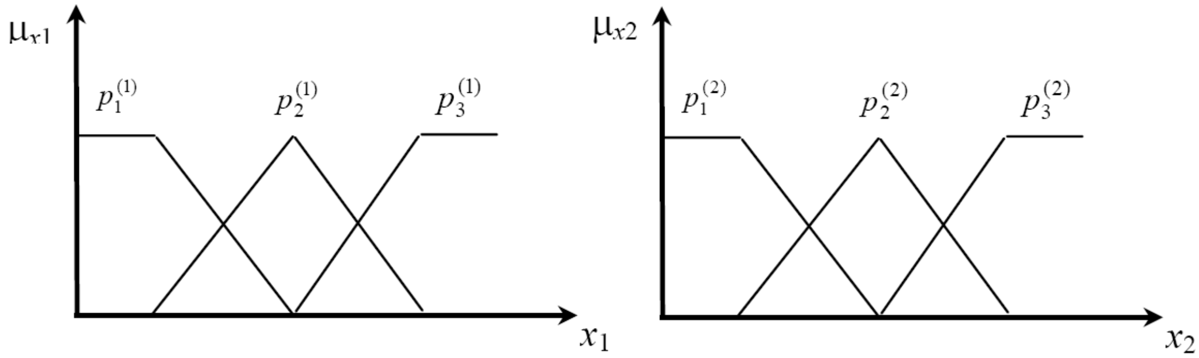
$$\boldsymbol{Y}(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{Y}^{(1)} \\ \boldsymbol{Y}^{(2)} \\ \vdots \\ \boldsymbol{Y}^{(\boldsymbol{m})} \end{bmatrix} = \begin{bmatrix} \boldsymbol{p}^{(1)}(x_1) \\ \boldsymbol{p}^{(2)}(x_2) \\ \vdots \\ \boldsymbol{p}^{(\boldsymbol{m})}(x_m) \end{bmatrix} \tag{2}$$

where x_i is the i-th input of the fuzzy system, $\boldsymbol{p}^{(k)}(.)$ is the k-th vector of fuzzy inference function associated to the x_k and $\boldsymbol{Y}^{(k)}$ input. Each sublayer has its own fuzzy sets defined by the fuzzy inference functions vector $\boldsymbol{p}^{(k)}(.)$.

Therefore, vector $\boldsymbol{Y(x)}$ given in (1) is defined as the vector concatenation of vectors $\boldsymbol{Y}^{(1)}$, $\boldsymbol{Y}^{(2)}$, ..., $\boldsymbol{Y}^{(m)}$. For a system with two input variables, the expression (2) would be:

$$\boldsymbol{Y}(\boldsymbol{x}) = [\underbrace{p_1^{(1)}(x_1)\ \ p_2^{(1)}(x_1)\ \ \cdots\ \ p_N^{(1)}(x_1)}_{\boldsymbol{p}^{(1)}(x_1)}\ \ |\ \ \underbrace{p_1^{(2)}(x_2)\ \ p_2^{(2)}(x_2)\ \ \cdots\ \ p_N^{(2)}(x_2)}_{\boldsymbol{p}^{(2)}(x_2)}]^T \tag{3}$$

As an example, Picture 1 shows the fuzzy sets associated to a two-input system, x_1 and x_2, with three fuzzy sets in each one.



Picture 1 – Representation of the multilayer fuzzy system inputs

As mentioned previously, the membership functions define the fuzzy sets associated to each input. In the input layer, it is the membership functions that determine the membership degree of each variable in relation to the fuzzy set defined by them.

Membership functions can be defined through any function. The only requirement of such functions is that they be defined in the closed domain [0,1]. However, it is convenient to define the membership functions in a simple way, proper to the respective computational implementation, for higher processing speed and rational memory use.

4.2. Inference Layer

The inference layer of a fuzzy system processes the fuzzy inference rules defined for the system. The inference layer also provides a knowledge basis on the process in question. The inference rules are processed together, in the same way as the sublayers of the input layer.

In this context, the definition of this set of rules is fundamental for the accurate functioning of the fuzzy inference system. There are several methods to extract the fuzzy rules from the adjustment data set.

Weighing the inference rules is an adequate way of representing the most important rules in the fuzzy system, or even allowing conflicting rules to co-relate without loss of their verbal completion.

Thus the *i*-th fuzzy rule is represented as in (4):

$$R_i(\boldsymbol{Y}(\boldsymbol{x})) = w_i r_i(\boldsymbol{Y}(\boldsymbol{x})) \tag{4}$$

where $R_i(.)$ is the function representing the weighed value of the *i*-th fuzzy rule, w_i is the weighing factor of the *i*-th fuzzy rule and $r_i(.)$ represents the fuzzy value of the *i*-th fuzzy rule.

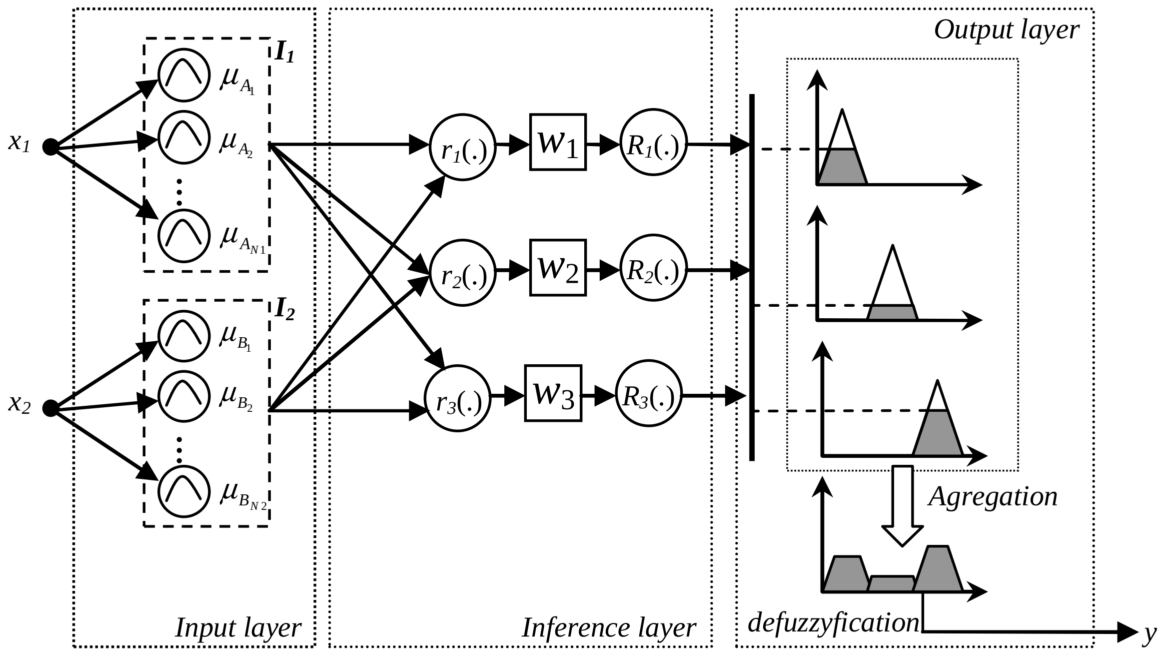
4.3. Output Layer

The fuzzy inference system output layer aggregates the inference rules and defuzzyficates the fuzzy set generated by this aggregation.

In the fuzzy inference system project, the choice of both the aggregation method and the defuzyyfication method is a very important decision. The aggregation method of the fuzzy inference rules must be such that the fuzzy set resulting from the aggregation can adequately infer the knowledge specified by its fuzzy rules set. Analogically, the chosen method for defuzzyfication must express, in a non-fuzzy value, the fuzzy set resulting from the fuzzy aggregation.

Besides these operational aspects, the aggregation and defuzzyfication methods must comply with the computational performance requirements, to reduce the computational struggle necessary for the fuzzy system processing. This study uses as aggregation method the 'maximum' operator and as defuzzyfication method the 'area center' (PEDRYCZ & GOMIDE, 1998), producing one only defuzzyficated system output.

Here, the fuzzy system output layer is also adjusted, and is represented by fuzzy sets of the output layer. The adjustment of this layer happens in a similar way to the input layer adjustment. Picture 2 presents a chart of the multilayer fuzzy model for a system with two inputs and three activated inference rules.



Picture 2 – Multilayer fuzzy system

4.4. Fuzzy Inference Systems Structural Adjustments

Generally speaking, in applications involving identification of fuzzy systems and modeling, it is convenient to use cost functions expressing the deviation between the wished results and the results given by the fuzzy system. Thus, this function can be expressed in terms of the average quadratic error between the fuzzy system output and the wished results, which are given in form of a training set, similar to the artificial neural networks with supervised training. Besides the average quadratic error, some data regularization indicators can be added to the cost function to improve the system response to noise in the training data.

Even though it creates a new task in the creation process of the fuzzy inference system, the definition of the cost function associated to the system, together with the techniques outlined in this study, save work during the membership functions tuning stages and creation of the fuzzy inference rules.

The creation algorithm of the rule base used here is based on the optimization algorithm *Hill Clibing* (ROGER JANG *et al*, 1997). Therefore, the developed process aims to perform a search in a determined valid space. Formally, given a search space S and a feasible set F, $F \subseteq$

S, we find r^* such as the expression (5) is true, that is:

$$\xi(r^*) \le \xi(r)\ ; \quad \forall r \in F \tag{5}$$

where $\xi(.)$ represents the cost function associated to the fuzzy system, r is a determined rule base belonging to the set of all the valid rule bases and r^* represents the rule base giving the smallest value to the cost function in relation to the feasible region F.

However, the search for a rule base that minimizes the cost function ξ is very complex, since the universe F representing all the feasible rule bases has innumerable possibilities. Besides, the surface defined by ξ is made of points of minimums and maximums, thus compromising the search for a solution which satisfies the global minimum condition.

Therefore, it is convenient to adopt search strategies that operate locally. The expression (5) can then be changed as follows:

$$\xi(r^{(1)}) \le \xi(r^{(2)})\ ; \ \forall\ r^{(2)} \in N(r^{1}) \tag{6}$$

where:

$$N(r^{(1)}) = \{r^{(2)} \in F : dist(r^{(1)}, r^{(2)}) \le \varepsilon\} \tag{7}$$

where $dist(r^{(1)}, r^{(2)})$ is a function that determines the distance between $r^{(1)}$ and $r^{(2)}$, and ε is a constant that defines the neighboring function radius $N(.)$.

Therefore, due to the complexity of the search process of a base rule that satisfies condition (6), the structural optimization algorithm proposed here operates in two different phases. In the first phase, several rule bases are chosen at random. The base rule offering the lowest cost is accepted. This way, the search space will be restricted to a limited area defined by the set $N(r^{(1)})$, where $r^{(1)}$ represents the rule base chosen that offered the lowest cost compared to the others.

Taking as starting point the rule base $r^{(1)}$ chosen in the first phase, in the second phase the local adjustment is done through the insertion of little variations in the original rule base. If the inserted variation causes a cost reduction, such variation is accepted and the process goes on until no significant improvement is observed in the system by a certain number of iterations.

4.5. Fuzzy Inference Systems Parametrical Adjustments

This section presents a method of fuzzy inference system tuning based on error backpropagation algorithms and optimization techniques. This method minimizes a cost function associated to the fuzzy inference system.

The formalization of a fuzzy inference system in a multilayer system, as presented in Section 4.5, can be justified not only by the distinct operational division of each of these layers, but also by the presence in each of them of distinct free parameters. As previously shown, in the input layer we have the parameters of the fuzzy sets membership functions of the system input space; the rule weighing is in the fuzzy inference layer; the parameters of the fuzzy sets membership functions of the output space are in the output layer.

Considering that the adjustment set $\{\boldsymbol{x}, y\}$ is fixed during all the adjustment process, then the cost function responsible for the tuning of the free parameters of the fuzzing mapping g: $\boldsymbol{x} \rightarrow y$ can be written as follows:

$$\xi_g = \xi\left(\boldsymbol{h}^{(1)}, \boldsymbol{h}^{(2)}, \boldsymbol{h}^{(3)}\right) \tag{8}$$

where $\boldsymbol{h}^{(1)}$, $\boldsymbol{h}^{(2)}$ and $\boldsymbol{h}^{(3)}$ represent respectively the vectors of the input membership functions parameters, the inference rule weighing vector and the output membership function parameter.

Therefore, after the minimization of the cost function ξ_g, the several free parameters incorporated in vectors $\boldsymbol{h}^{(1)}$, $\boldsymbol{h}^{(2)}$ and $\boldsymbol{h}^{(3)}$ will correspond to the values that minimize the respective cost function.

Consequently, the problem of parametric adjustment of a fuzzy inference system can be interpreted as an unrestrained optimization problem that can be solved by any available method. This study used the descending gradient method for the cost function minimization and the calculation of the gradient was estimated using the difference equation method.

To elucidate this adjustment process, take a fuzzy system with one variable with two Gaussian membership functions in the input layer; three rules in the inference layer; and one variable with two Gaussian membership functions in the output layer. Thus, the vectors with parameters $\boldsymbol{h}^{(1)}$, $\boldsymbol{h}^{(2)}$ and $\boldsymbol{h}^{(3)}$ will be:

$$\boldsymbol{h}^{(1)} = [\mu_1^e \quad \sigma_1^e \quad \vdots \quad \mu_2^e \quad \sigma_2^e]^T \tag{9}$$

$$\boldsymbol{h}^{(2)} = [w_1 \quad w_2 \quad w_3]^T \tag{10}$$

$$\boldsymbol{h}^{(3)} = [\mu_1^s \quad \sigma_1^s \quad \vdots \quad \mu_2^s \quad \sigma_2^s]^T \tag{11}$$

where μ_i^e and σ_i^e are respectively the average and the variance of the *i*-th input membership function; w_1, w_2 and w_3 are the fuzzy rule weighing factors; μ_i^s and σ_i^s are respectively the average and the variance of the *i*-th output membership function.

In case of layer by layer adjustment, simply apply the descendent gradient method to each one of them. For the second layer, the following expression represents the rule weighing factors:

$$\boldsymbol{h}^{(2)}(k) = \boldsymbol{h}^{(2)}(k-1) + \eta \cdot \nabla\xi(\boldsymbol{h}^{(2)}(k-1)) \tag{12}$$

where η is the learning rate. Here, the gradient vector $\nabla\xi(\boldsymbol{h}^{(\cdot)})$ can be generally given by:

$$\nabla\xi(\boldsymbol{h}^{(\cdot)}) = \left[\frac{\partial\xi}{\partial h_1^{(\cdot)}} \quad \frac{\partial\xi}{\partial h_2^{(\cdot)}} \quad \cdots \quad \frac{\partial\xi}{\partial h_n^{(\cdot)}} \right]^T \tag{13}$$

where $h_i^{(\cdot)}$ represents the *i*-th element of the parameter vector $\boldsymbol{h}^{(\cdot)}$ and *n* is the number of elements of this vector. In this case, as the analytical expression of the gradient vector is difficult to obtain, we use its numeric approximation through the difference equation method (DENNIS & SCHNABEL, 1983), that is:

$$\frac{\partial\xi}{\partial h_i^{(\cdot)}} \cong \frac{\xi(\boldsymbol{h}^{(\cdot)} + \rho \cdot \boldsymbol{e}_i) - \xi(\boldsymbol{h}^{(\cdot)})}{\rho} \tag{14}$$

where ρ is a very small value and $\boldsymbol{e}_i = [0\ 0 \ldots 1 \ldots 0]^T$, where the element 1 is in the position *i*.

The cost function ξ_g in this case is the average quadratic error function in relation to the tuning standards *p*, that is:

$$\xi_g = \frac{1}{p}\sum_{k=1}^{p}\xi_{g(k)} \tag{15}$$

where $\xi_{g(k)}$ is the quadratic error in relation to the *k*-th training standard, that is:

$$\xi_{g(k)} = (d_k - y_k)^2 \tag{16}$$

where d_k is the wished output value in relation to the *k*-th training standard and y_k is the output given by the fuzzy system also in relation to the *k*-th training standard.

5. Results of Computational Simulations

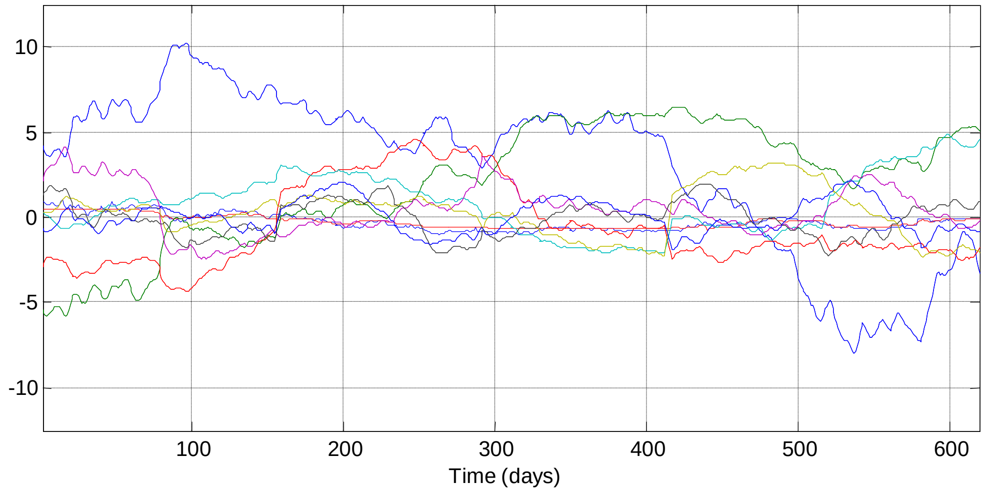
Computational simulations were carried out based on historical values of quotations by Bovespa in 625 days of financial trade. Of all the stocks negotiated by the said market 47 were chosen. The stocks used in this study were chosen considering the stocks traded on all days of the study horizon, so we used the stock with business in all 625 days of the study.

Therefore, the objective of the fuzzy system developed was, with data obtained from stocks quotations, to estimate the stocks behavior for future values, based on the temporal behavior of the whole set of stocks under observation. The past horizon used for the estimation was five days, making 235 input variables.

The great number of input variables for the estimation of the stock temporal behavior can make the direct application of estimation methods impossible, including the fuzzy inference system. There was then the need to pre-process the data to reduce their number. In this context, the technique of analysis by principal components deserves attention. This technique investigates the input variables spaces to obtain their respective principal components. The original variables can be obtained from these components by simple linear combination. In other words, the analysis technique by principal components gives a transformation matrix. This matrix reduces the dimension of the input variables used in modeling the studied system.

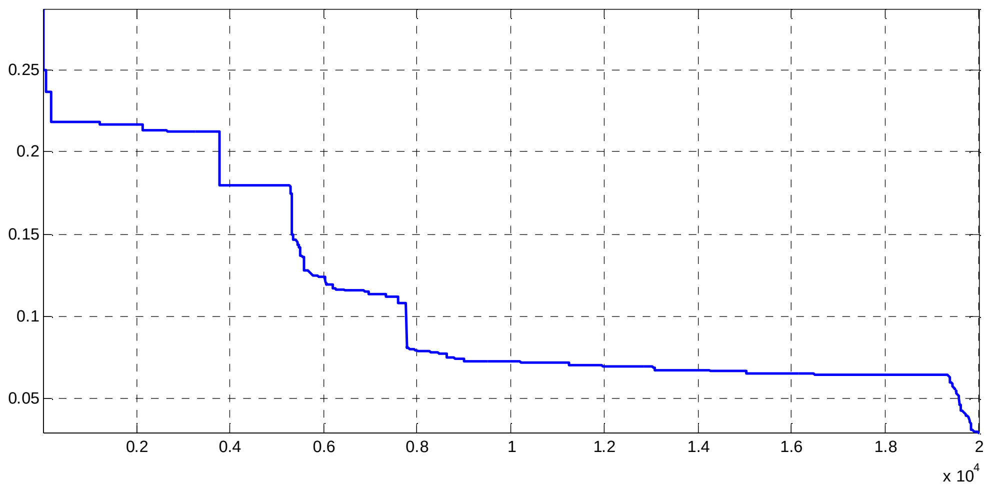
The application of the analysis technique by principal components in this study reduced the variable number from 235 to only eight. Thus, by linear combination it is possible to rebuild the original 235 variables from these eight variables. The principal components are graphically represented in terms of their temporal behavior in Picture 3.

Each of the 47 stocks used in the study was estimated by fuzzy inference systems adjusted according to the algorithms presented in Section 3. Each of the fuzzy inference systems was constituted by eight inputs, that is, one for each principal component. Each of the inputs, denoted in terms of fuzzy logic from speech universes, had a total of three membership functions. The rule base used consisted of 10 fuzzy rules. The system output, like the input, had three membership functions.



Picture 3 – Temporal behavior of the principal components used as estimation variables of future stock quotation

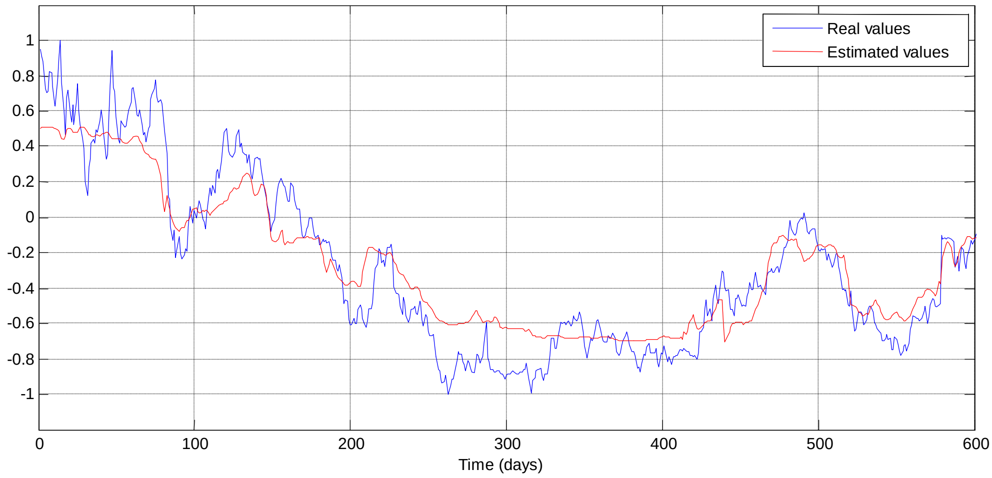
To illustrate how the proposed algorithm is able to adjust the fuzzy inference systems, Picture 4 shows how the average quadratic error was reduced during the adjustment algorithms iterations.



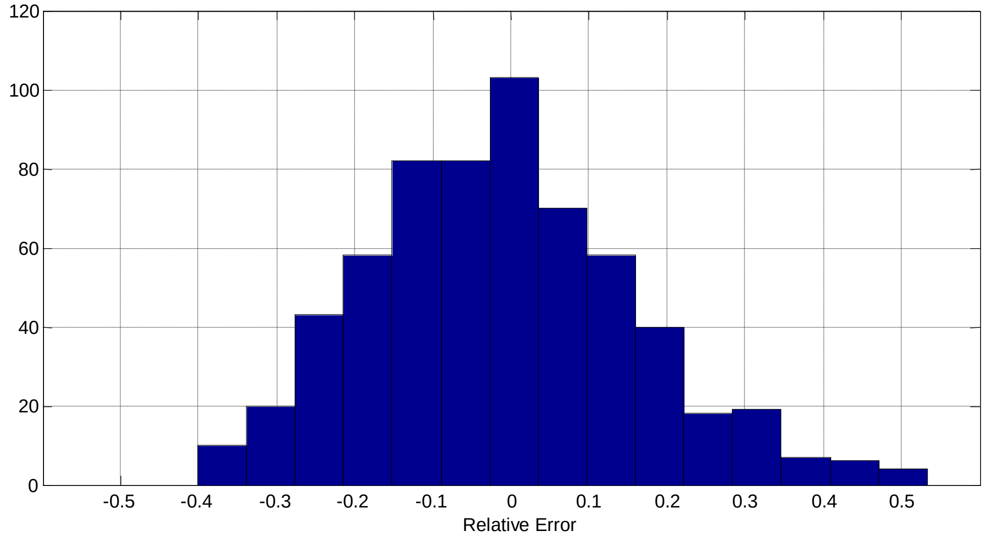
Picture 4 – Average quadratic error during the adjustment process of the fuzzy inference system

In Picture 4 we can see the sharp reduction of the average quadratic error during the adjustment algorithm iterations, reflecting the proposed technique efficiency. Complementing the presentation of the obtained results, the graph in Picture 5 shows the wished values compared to the estimated data for the stock code Bovespa EBTP4.

In Picture 5 we can verify the model's efficiency in estimating the data for the investigated stock. Complementing the presentation of the obtained results for the investigated stock, Picture 6 shows the histogram of the estimation relative error.



Picture 5 – Real values and estimated values for stock code Bovespa EBTP4



Picture 6 – Histogram of the relative percent error for estimation of future values of stock code Bovespa EBTP4

In Picture 6 we can verify the great proximity of the relative error distribution with the Gaussian distribution. The average error was -0.0123 with standard deviation of 0.1710. The fact that the estimation error distribution has a Gaussian behavior leads to the conclusion that the presented approach deserves more investigation with the possibility of even better results than the ones here demonstrated.

6. Conclusions

The objective of this study was to present the developments made for the estimation of stock future values from a horizon of past values. The horizon used for the estimation was 5 days and a total of 47 stocks were elected as estimation variables. The reduction of the input space was obtained by the analysis technique of principal components and fuzzy inference systems were used for the aimed estimation.

The results obtained show the efficiency of the developed algorithm adjustment and motivate future investigations to improve the studies summarized in this article.

References

BALTAZAR, A. P. *Uma cena em extinção.*Veja, São Paulo, v.41, n.1978, p.100-104, 2006.

BILOBROVEC, M.; MARÇAL, R. F. M.; KOVALESKI, J. L. *Implementação de um sistema de controle inteligente utilizando a Lógica Fuzzy*. In:SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO, 11., 2004, Bauru. Anais... Bauru: UNESP,2004. p.1-7.

BORGES, R.; *Curso de análise técnica: Introdução.* Disponível em: <http://www.ricardoborges.com/iniciando_bovespa.htm>. Acesso em 23 abr. 2007

BOVESPA. *O mercado de capitais: Desafios e oportunidades para o mercado de capitais brasileiro.* São Paulo:MB associados, 2000. Disponível em: <http://www.bovespa.com.br/pdf/mercado_capitais_desafios.pdf>. Acesso em 02 mai.2007.

CHAVES, D. A. T. *Análise técnica e fundamentalista: divergências, similaridades e complementariedades.* 2004.119 f. Monografia (Graduação em Administração) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2004.

GOMIDE, F. A. C.; GUDWIN, R. R. *Modelagem, controle, sistemas e lógica fuzzy.* SBA Controle & Automação, Campinas, v.4, n.3, p.97-115, 1994.

HIEMSTRA, Y. *A Stock Market Forecasting Support System Based on Fuzzy Logic.* In: HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES.27.,1994,Maui. Proceedings...Maui: IEEE Computer Society Press, 1994. p. 281-287.

KRUSE, R.; GEBHARDT, J.; KLAWONN, F. *Foundations of Fuzzy Systems.*Chichester: John Wiley & Sons Ltd, 1994. 265 p.

MURPHY, J. *Technical Analysis of the Future Markets.* New York: Institute of Finance, 1986.

NORONHA, M.; *Curso básico de análise gráfica.* Disponível em: <http://www.timing.com.br/download/AG01.pdf>. Acesso em: 25 abr. 2007.

ROSS, S. A.; WESTERFIELD, R. W.; JAFFE, J. F. *Administração financeira: Corporate finance.* São Paulo: Atlas, 1995.

SAINT-GEOURS, J. *Os mercados financeiros*. São Paulo: Ática, 1998.

SHAW, I. S.; SIMÕES, M. G. *Controle e Modelagem Fuzzy*. São Paulo: Edgard Blücher, FAPESP, 1999.

YEN, J.; LANGARI. R. *Fuzzy Logic: Intelligence, Control, and Information.*New Jersey: Prentice Hall, 1999. 548 p.