Econometric Reviews, 30(6):609–631, 2011

ISSN: 0747-4938 print/1532-4168 online
DOI: 10.1080/07474938.2011.553569

FUZZY AUTOREGRESSIVE RULES: TOWARDS LINGUISTIC TIME SERIES MODELING

José Luis Aznarte[1], Jesús Alcalá-Fdez[2], Antonio Arauzo[3], and José Manuel Benítez[2]

[1]*Department of Artificial Intelligence, Universidad Nacional de Educación a Distancia, Madrid, Spain*
[2]*Department of Computer Science and Artificial Intelligence, Universidad de Granada, Spain*
[3]*Department of Rural Engineering, Universidad de Córdoba, Spain*

□ *Fuzzy rule–based models, a key element in soft computing (SC), have arisen as an alternative for time series analysis and modeling. One difference with preexisting models is their interpretability in terms of human language. Their interactions with other components have also contributed to a huge development in their identification and estimation procedures. In this article, we present fuzzy rule–based models, their links with some regime-switching autoregressive models, and how the use of soft computing concepts can help the practitioner to solve and gain a deeper insight into a given problem. An example on a realized volatility series is presented to show the forecasting abilities of a fuzzy rule–based model.*

Keywords Fuzzy models; Regime-switching models; Soft computing; Time series; Volatility.

JEL Classification C45; C53.

1. INTRODUCTION

Research concerning the combination of intelligent computing technologies was started by Zadeh (1994), who proposed the term *soft computing* (SC), defining it as a "collection of methodologies that aim to exploit the tolerance for imprecision and uncertainty to achieve tractability, robustness, and low solution cost." According to Zadeh, the principal constituents of SC are fuzzy logic, artificial neural networks, and probabilistic reasoning, with the last one subsuming evolutionary computation.

Address correspondence to Dr. José Luis Aznarte, Departmento de Inteligencia Artificial, E.T.S. de Ingeniería Informática, Universidad Nacional de Educación a Distancia, C/Juan del Rosal, 16, Ciudad Universitaria, Madrid 28040, Spain; E-mail: jlaznart@dia.uned.es

One of the most successful parts of soft computing is fuzzy logic, which has seen an enormous spread thanks, amongst other reasons, to its industrial applications. More precisely, fuzzy rule–based models (FRBMs), due to their modeling power, linguistic interpretation, and universal approximation capabilities arise as one of the most popular elements of SC (Dubois and Prade, 1980; Jang et al., 1997; Klir and Yuan, 1995; Nauck et al., 1997).

In the framework of time series analysis, FRBMs as well as artificial neural networks have been used. Both paradigms yield promising results, but the later has been more easily accepted by statistic researchers and economists. Asymptotic theory and hypothesis testing were developed for neural networks Teräsvirta et al. (2006a), and it was not until recently that the same can be said about FRBMs (cf. Aznarte et al., 2010).

Considering with Aznarte et al. (2007) that FRBMs can be seen as a generalization of some popular autoregressive models, ranging from the linear model to nonlinear ones as the smooth transition autoregressive (STAR) model Teräsvirta (1994) or the neuro coefficient smooth transition (NCSTAR) model (Medeiros and Veiga, 2005), one might be interested in studying what can be learned from the SC paradigm with respect to these regime switching autoregressive models.

In this article, we introduce fuzzy logic, fuzzy rules, and the FRBMs in Sections 2 and 3. In Section 5 we introduce a recently developed bio-inspired algorithm to identify and estimate FRBMs, whereas in Section 4 we analyze the use of FRBMs in the time series analysis problem and relate them to other nonlinear autoregressive models. Finally, in Section 6 we show an application of the model proposed in Section 5 to a specific realized volatility series, and we gather our conclusions in Section 7.

2. FUZZY SETS

In everyday conversations, we use many words and expressions that we understand but are intrinsically vague, from "the weather is cold" to "the euro is getting strong." Fuzzy sets were proposed to deal in a mathematical way with such vague words. Fuzzy sets, and the models derived from them, can handle vague concepts as "a set of *tall* people" or "apartments that are *close* to a subway station." In these sentences, the words "tall" and "close" give ambiguous ideas that are nonetheless meaningful to us. Classic set theory does not handle this type of vague expressions and it forces us to define terms in an exact manner as "a set of people taller than 190 cm" or "apartments within 500 m from a subway station." Measuring a person's height will tell us if he or she belongs to the former set. These conventional sets, defined in an exact way, are called "crisp sets" in fuzzy set theory.

Economics and econometrics are far from being free from imprecision or uncertainty. In the process of reducing to numbers and mathematical concepts some economy related quantities and magnitudes, we have to deal with a wealth of vague terms (confidence, fear, instability, risk...) which are meaningful for us but difficult to encode in computer language. For example, a set of stocks with *small volatility* or countries with *high unemployment rates* are fuzzy sets that need some transformation in order to be handled by classic set theory.

In crisp set theory, the characteristic function χ_A of a set on the universe X is defined as $\chi_A : X \to \{0, 1\}$ with

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in X, \\ 0 & \text{if } x \notin X. \end{cases} \tag{1}$$

This function indicates that if the element x belongs to A, χ_A is 1, and it does not belong to A, χ_A is 0. Characteristic functions are rarely used for the application of crisp sets. However, when we extend this idea to fuzzy sets, the role of characteristic functions becomes significant, and they are called *membership functions*.

Membership functions can be better understood through the following example. Let us think of the group of companies included in the Dow Jones Index. This group is our universe of discourse, X, and we can easily define crisp subsets in it: pharmaceutical companies, companies which operate in more than one market, companies with benefits last year. All these subsets have an implicit characteristic equation as the one defined in expression (1). There are operations that we can perform on these sets, including union, intersection, and complement. There are also laws studied by set theory such as the commutative law, the associative law, the distributive law, and so on, that are applicable to the operations on these sets. We can represent these sets and their operations graphically using Venn diagrams.

Now the problem becomes more difficult if we need to define the set of members of X which have high volatility at a given time. Once we are able to measure volatility (Andersen and Bollerslev, 1998), the idea of "volatile" (understood as *having a greater volatility than expected*) may change from person to person, from company to company, and depending on the situation. It is impractical to divide companies into a "volatile" group and a "non volatile" group, and representing these sets using Venn diagrams poses a difficulty. The degree of "volatile" can vary from "little volatile" to "extremely volatile" and hence is necessary to express this degree, for example, by assigning a number between 0 and 1. Then the degree 1 would mean that the company completely belongs to the set of "volatile companies," whereas 0 would denote that the company does not belong to the set (and thus does not present a greater volatility than expected).

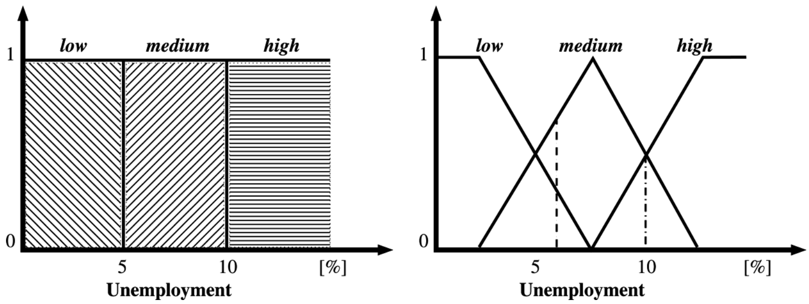

FIGURE 1 Characteristic and membership functions for crisp (left) and fuzzy (right) sets defined on the unemployment rate of European countries.

Membership functions of fuzzy sets helps us define such degree of belongingness, and, as opposed to crisp sets' characteristic functions, they are defined over a continuous range, $\mu_A : X \to [0, 1]$. The closer the value of $\mu_A(x)$ to 1, the higher the degree of membership of the element x in fuzzy set A. If $\mu_A(x) = 1$, the element completely belongs to the fuzzy set, whereas if $\mu_A(x) = 0$, the element does not belong to A at all. Consequently, crisp sets are a special case of fuzzy sets.

Using Fig. 1 we can compare membership functions and characteristic functions to demonstrate the features of fuzzy sets. We can define several sets of European countries considering their unemployment rate[1]. The figure shows, on the left hand, the characteristic functions for "low," "medium," and "high," and on the right hand the corresponding membership functions.

Take the case of Slovenia, which in July 2009 had an unemployment rate of 6.0%. Under the crisp set classification shown in the left part of Fig. 1, Slovenia belongs to the "medium" unemployment set. On the other hand, consider France, which had an unemployment rate of 9.8%. This country is also classified as having a "medium" unemployment rate in the crisp division. Now, in the fuzzy sets defined in the right part of the figure, things become more familiar to us: Slovenia still belongs to the "medium" unemployment set, but now with a membership degree of about 0.7, and it also belongs to the "low" height set with a degree of 0.3 (dashed line). Concerning France, it has a membership degree of 0.5 in the "medium" unemployment set, the same value as in the "high" unemployment set (dash-dot line). If we were to express their unemployment rate linguistically, Slovenia would have something like a low-medium unemployment rate, and France would have a moderately high unemployment rate.

[1]Source: Eurostat Statistical Office (01/09/2009), http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=0&language=en&pcode=teilm020.

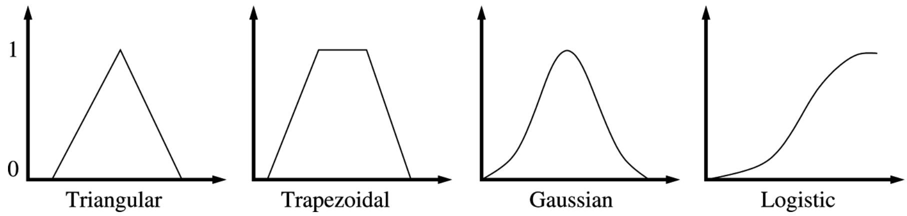

FIGURE 2 Popular choices for the membership functions.

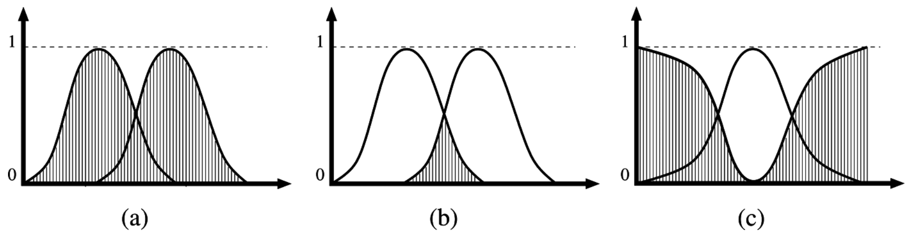

FIGURE 3 Operations on fuzzy sets: (a) union, (b) intersection, and (c) complement.

There are several popular choices for the shape of a membership function. Apart from the triangular fuzzy sets shown in Fig. 1, there are other common representations: trapezoidal, Gaussian, logistic, etc. They are usually chosen attending to interpretability and computational complexity considerations. Figure 2 shows some of them.

As was mentioned before, fuzzy sets were developed as an extension of classic sets. In particular, the operations on sets are also defined as extensions of crisp ones. Union is modeled through t-norms in the fuzzy set theory, intersection (t-conorm), and complement of fuzzy sets are defined as shown graphically in Fig. 3. It is easy to prove that their crisp counterparts are special cases of these. There are different possible choices for each operator, leading to related set of operators. This results in greater flexibility allowing for different interpretations and applications.

3. FUZZY RULES AND INFERENCE

The classical model for representation and processing of knowledge is Aristotelian logic, which was extended into propositional logic and predicate logic. The core logical construct used in predicate logic is the *production rule*. These rules are defined as

IF a set of conditions is satisfied, THEN it is possible to infer
a set of conclusions.

The first part of this rule is known as *antecedent* or premise, whereas the second is known as *consequent* or conclusion. If the clauses included in the antecedent and consequent of a rule are expressed in terms of fuzzy sets, it is called *fuzzy rule*, and the associated type of reasoning is known as *Approximate Reasoning*.

In classical logic terms, when we want to represent that an object x satisfies property P, that is, x belongs to the set of objects satisfying P, we write "x is P." If the property is fuzzy, then the set P is fuzzy and the proposition "x is P" is called *fuzzy proposition*. Using the logic connectives (negation, conjunction, disjunction, and implication) extended for fuzzy sets, we can connect fuzzy propositions.

If we replace the propositions in a classical production rule by fuzzy propositions, we get a *fuzzy rule*:

$$\text{IF } x_1 \text{ IS } A_1 \cap \cdots \cap x_n \text{ IS } A_n, \quad \text{THEN } y_1 \text{ IS } B_1 \cup \cdots \cup \ y_m \text{ IS } B_m, \tag{2}$$

where $\cap$ and $\cup$ are fuzzy conjunction (t-norm) and disjunction (t-conorm), respectively.

Amongst the most important properties of fuzzy rules we have:

a) Uncertainty representation. The most essential representation of information used by human beings is language. It is plagued of vagueness and uncertainty. Fuzzy sets represent perfectly this kind of uncertainty. Fuzzy rules represent knowledge with those concepts.
b) Compact information representation model. With just one fuzzy rule, we can express all the information contained in a set of classical rules.
c) Local character. The information described in a single fuzzy rule usually affects only a local zone of the complete domain of the problem. Its interactions with other descriptive elements of the problem is restricted to those which are in its neighborhood. This eases construction and interpretation of fuzzy rules.

As a way to combine fuzzy propositions, the *compositional rule of inference* was introduced by Zadeh (1973). It extends the classical logic "modus ponens" into fuzzy logic, and allowed for the construction of fuzzy inference systems, or FRBMs.

3.1. FRBMs

A model using fuzzy rules (and the compositional rule of inference) is a FRBM. These models can be used to describe unknown or complex systems. In the Control field they have proved to be very appropriate, where they are usually called *fuzzy controllers*. Its application to control problems with hard or impossible mathematical solutions has been a

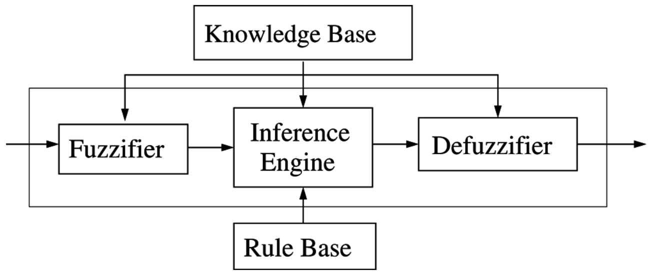

FIGURE 4 Structure of an FRBM.

milestone for the acceptance and fast expansion of the techniques based in fuzzy set theory.

FRBMs map spaces of dimension n into spaces of dimension m. They are formed by two main components, *knowledge base* and *inference engine*. The knowledge base is further subdivided into *data base* and *rule base*. When inputs and outputs are real values, two additional components are necessary: *fuzzifier* and *defuzzifier*. A schematic view of the system is displayed in Fig. 4.

The knowledge base includes the rule base and the data base. The first one includes all the rules capturing the knowledge of the system. The definitions of the membership functions for the labels of the rules are kept in the data base. The inference engine computes the fuzzy output from the fuzzy inputs by applying a fuzzy implication function. It also aggregates the outputs of the different applicable rules producing a single fuzzy output value. The fuzzifier transforms the real input values into fuzzy values, usually in uni-valued fuzzy sets or singletons. Finally, the defuzzifier condenses in a single real number the inferred fuzzy output.

In short, the fuzzy inference process is as follows: Given an input, the *firing strength* of each rule is obtained. This firing strength comes from the matching degree of each component of the input with the corresponding proposition in the fuzzy rule's antecedent. When the inputs are real values and the fuzzification process transforms them into singleton fuzzy sets, this matching degree is just the membership degree of each real input value to the fuzzy set defined by each rule proposition. Then these degrees are aggregated according to the connective operators linking the propositions. Usually, these are conjunctive operators, so the firing strength of the rule will be the value of the function modeling the conjunction (minimum, product, ...) applied on all the coupling degrees. The fuzzy output of the rule will be its consequent, in which the fuzzy sets will have a maximum membership degree equal to the firing strength of the rule. Lastby in the case of systems with real output, the fuzzy output is transformed into real values by applying the correspondent defuzzification method.

FRBMs have some interesting properties, amongst which its "Universal Approximation" capability stands out. In Castro (1995), Castro and Delgado (1996), and Kosko (1994) several wide classes of fuzzy controllers are proved to be universal approximators. This result entails FRBMs as right tools for approximating other systems even when the data associated to them are not fuzzy.

We can distinguish two main applications of FRBMs:

1. *Linguistic description* of systems where there is uncertainty or vagueness in the inputs and/or outputs. This also includes the cases in which precision is not important. What is fundamental is to obtain a description of the inner relationships amongst inputs and outputs of an unknown system in terms of words and expressions of the natural language.
2. *System Approximation.* Exploiting its universal approximation property, FRBMs are an alternative to classical mathematical models in unknown system identification. Its strengths are clear when dealing with complex systems, where its simplicity and ease of use wake them preferable to other methods.

For a deeper discussion on FRBMs, see for example Yager and Filev (1994).

3.1.1. *Mamdani's FRBM*

In 1975, the first practical application of an FRBM was shown (Mamdani and Assilian, 1975). It was a simple FRBM devoted to the control of a steam engine. This model had a fuzzifier, a rule base, a data base, and a defuzzifier.

The rules used by this model are of the form:

$$\begin{aligned} R_i : \mathrm{IF} x_1 \text{ IS } A_1^i \text{ AND } \ldots \text{ AND } x_n \text{ IS } A_1^n, \\ \text{THEN } y_i : \text{ IS } B_1^i \text{ AND } \ldots \text{ AND } y_m \text{ IS } B_1^m, \end{aligned} \tag{3}$$

For example, if we were to control an imaginary risk alarm system, we could have a rule lke

IF volatility IS *LOW* AND mean IS *MEDIUM* THEN risk is *LOW*.

The respective values of LOW, MEDIUM, and LOW would be encoded as fuzzy labels wiht predefined membership functions μ_{LOW}(volatility), μ_{MEDIUM}(mean), and μ_{LOW}(risk).

Their simplicity and ease of interpretation later granted these type of rules the category of standard, being one of the most common in practice.

The inference process uses the minimum as conjunction operator and as implication function. Hence, given an input $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, the firing strength of the rule is

$$\gamma_i = \min\left(\mu_{A_1^i}(x_1), \ldots, A_n^i(x_n)\right). \tag{4}$$

The fuzzy outputs of the rules are B_1^i, $B_2^i, \ldots, B_n^i$, cut at the level γ_i. The aggregation operator is modeled as a disjunction, usually the *maximum* operator. As defuzzification interface it is common to use the center of gravity or the maximum average.

As stated above, this is the older and simpler FRBM, which also contributed to its success. Moreover, its simplicity makes it really easy and inexpensive to be implemented in hardware, and this favored its expansion to many applications, ranging from small home appliances to big engineering devices as container cranes or automatic train drivers.

3.1.2. *Takagi-Sugeno-Kang's FRBM*

In 1985, (Sugeno and Nishida, 1985; Sugeno and Kang, 1988; Takagi and Sugeno, 1985) proposed a different FRBM which was much more effective in approximation tasks. Its rules, usually known as TSK rules, were of the type

$$\begin{aligned} &R_i : \text{IF} x_1 \text{ IS } A_1^i \text{ AND } x_2 \text{ IS } A_1^2 \text{ AND } \ldots \text{ AND } x_n \text{ IS } A_n^i, \\ &\quad \text{THEN } y = p_i(x_1, x_2, \ldots, x_n), \end{aligned} \tag{5}$$

being $p_i(x_1, x_2, \ldots, x_n)$ a linear function. That is, the output adopts a purely functional shape.

The firing degree of the rules is obtained in a way similar to the one used in Mamdani's FRBM. The only difference is that the conjunction is modeled through product instead of minimum:

$$\gamma_i = \prod\left(A_1^i(x_1), A_2^i(x_2), \ldots, A_n^i(x_n)\right). \tag{6}$$

The final output of the model is given by

$$y = \frac{\sum_{i=1}^r \gamma_i p_i(x_1, x_2, \ldots, x_n)}{\sum_{i=1}^r \gamma_i}. \tag{7}$$

3.1.3. *Additive Fuzzy Models*

Additive Fuzzy Models (AFM), proposed by Kosko (1994) are characterized by a different way of performing the inference.

Instead of using the maximum as an aggregation operator or computing the weighted average to obtain the final output of the model, in AFM, this final aggregation is modeled through a weighted sum. In the case of TSK-type rules, (5), the output takes the form

$$y = \sum_{i=1}^{r} w_i p_i(x_1, x_2, \ldots, x_n). \tag{8}$$

These are the main classes of FRBMs. When a linguistic goal is being faced, the Mamdani type of system is the choice. However, if the objective is to obtain a highly accurate system, TSK or additive fuzzy system are the most effective way. This will be the case when applying FRBMs to the modeling and forecasting of time series. The structure of the systems are simple enough to easily grasp their behavior, especially due to the localness of their components. Notwithstanding their structure is flexible enough to render a good enough shape in the approximation of nonlinear systems. And certainly for the successful application of a given kind of models, in addition to their concept, identification and estimation procedures are essential. A specially powerful algorithm for this task is described in next Section 5. In the next section, we will see how these models are used to face the time series forecasting problem, and the relations that link them to standard nonlinear autoregressive models.

4. FRBMs FOR TIME SERIES MODELING AND FORECASTING

As it has been stressed above, when confronted to function approximation problems, the preferred FRBM type is the one using linear functions in the consequents of the rules, known as TSK model. In the framework of time series modeling and forecasting, these models use rules of the form

$$\begin{aligned} &\text{IF } y_{t-1} \text{ IS } A_1 \text{ AND } \ldots \text{ AND } y_{t-p} \text{ IS } A_p \\ &\quad \text{THEN } y_t^* = b_0 + b_1 y_{t-1} + \cdots + b_p y_{t-p} + \varepsilon_t. \end{aligned} \tag{9}$$

In this rule, the variables y_{t-h} are lagged values of the time series, $\{y_t\}$ (although it is also possible to include exogenous variables). In Aznarte et al. (2007) these rules were considered as local autoregressive (AR) models. Indeed, the antecedent of these fuzzy rules expresses the region of the input space where the AR encoded in its consequent is applicable. That article suggested the inclusion of the random shock term, ε_t leaving us with what is reasonable to call *fuzzy autoregressive rules.*

Let us consider the following AR process (Aznarte et al., 2007):

$$y_t = 1.6 - 0.04 y_{t-1} + 0.06 y_{t-2} + \varepsilon_t, \tag{10}$$

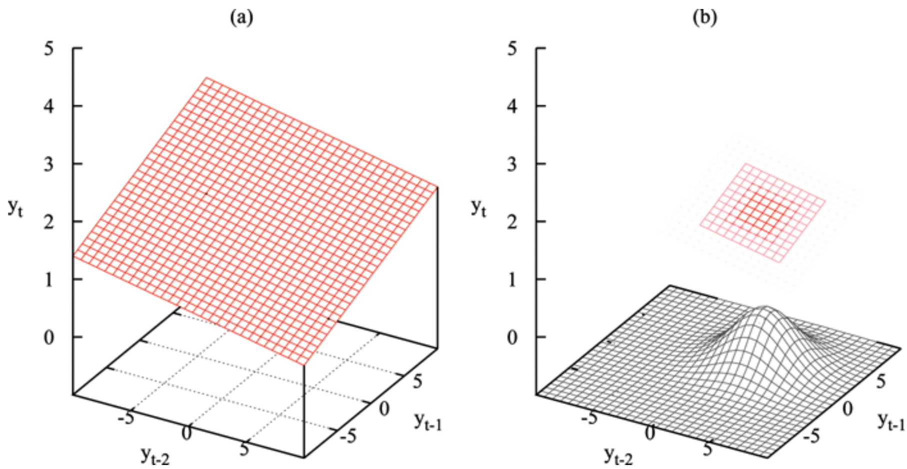

FIGURE 5 (a) Plane defined by the AR(2) model (10). (b) Graphical representation of the fuzzy rule which makes use of the AR(2) model.

which can be seen as a definition of the relationship between the output variable y_t and the input variables, y_{t-1} and y_{t-2}. The deterministic part of this relationship can be displayed graphically as shown in Fig. 5 (a).

We might build a fuzzy autoregressive rule whose consequence is exactly the aforementioned AR model:

$$\text{IF } y_{t-2} \text{ IS } A_1 \text{ AND } y_{t-1} \text{ IS } A_2, \text{ THEN } y_t^* = 1.6 - 0.04y_{t-1} + 0.06y_{t-2} + \varepsilon_t. \tag{11}$$

Considering Gaussian-type membership functions, and ignoring multiplicative constants, the expression for μ_{A_i} is

$$\mu_{A_i}(x) = \exp\left(\frac{-(x-\mu_i)^2}{2\sigma_i^2}\right), \quad i = 1, 2. \tag{12}$$

A graphical representation of this fuzzy rule is shown in Fig. 5 (b), in which $\mu_1 = \mu_2 = 2.5$ and $\sigma_1 = \sigma_2 = 2.0$.

From the graphical representation shown in Fig. 5, it is fairly clear that the application of the fuzzy rule amounts to the application of the AR(2) model in the state-space subset defined by the belongingness of the input variables to the membership functions defined for its antecedent. It must also be noted that this state-space subset is a fuzzy subset, and hence its borders are not crisp.

4.1. Relations with Nonlinear AR Models

Based on the concept of fuzzy autoregressive rules, we can show that there are strong links between the nonlinear autoregressive models and FRBM (Aznarte et al., 2007).

It is rather clear that there is some parallelism between the two aforementioned families of models. At a high level, models from both sides are composed of a set of elements (AR – fuzzy rules) which happen to be closely related, as stated above. On a lower level, both families rely on building a hyper-surface on the state-space which tries to model the relationship between the lagged variables of a time series. Moreover, both define this hyper-surface as the composition of hyper-planes which apply only in certain parts of the state-space.

Amongst nonlinear autoregressive models we find the logistic Smooth Transition Autoregressive (STAR) model, first proposed by Teräsvirta (1994). The original formulation for this model was

$$y_t = \mathbf{b}_0\mathbf{x}_t + \mathbf{b}_1\mathbf{x}_t f(s_t; \gamma, c) + \varepsilon_t, \tag{13}$$

where the smoothing function f is the logistic function and the transition variable is called s_t (and is usually a lagged endogenous variable: $s_t = y_{t-d}$ for a certain integer $d > 0$). This model nests the threshold autoregressive model (TAR) (Tong, 1990) when $\gamma_i \to \infty$.

An immediate extension of model (13) is the multiple regime STAR, given by

$$y_t = \mathbf{b}_0\mathbf{x}_t + \sum_{i=1}^{k} \mathbf{b}_i\mathbf{x}_t f(s_t; \gamma_i, c_i) + \varepsilon_t. \tag{14}$$

In this specification, the STAR model is composed of a linear term plus a set of locally applied AR models, which, in turn, can be seen as fuzzy AR rules with logistic membership function and a single term in the antecedent. Hence, the STAR model is a special case of a fuzzy rule–based model. Needless to say, the TAR model is also a special case of an FRBM, where the membership function is not "fuzzy" anymore but a crisp step function.

This equivalence can be clearly seen in Fig. 6. On the left (a) side of this figure, we see two linear models together with a transition/membership function. On the right (b) side we see the surface generated by the STAR model or the fuzzy rule–based model whose functional expression is

$$\begin{aligned} y_t = &\left(1.7 - 0.03y_{t-1} + 0.02y_{t-2}\right) + \left(1.6 + 0.01y_{t-1} + 0.06y_{t-2}\right) \\ &\times f(y_{t-1}; 2.5, 1.75) + \varepsilon_t. \end{aligned} \tag{15}$$

The special case of an FRBM which derives from a STAR model has some limitations concerning the type of surfaces it can generate and concerning its linguistic power: it allows for only one fuzzy variable (the

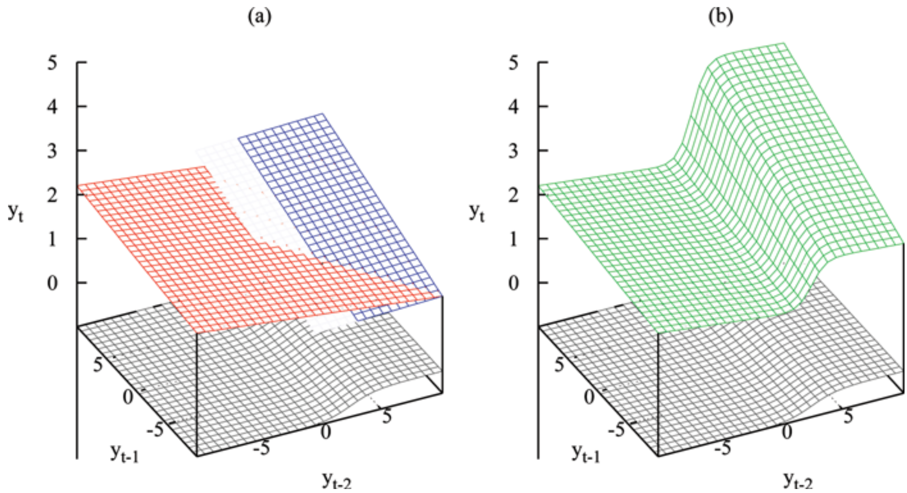

FIGURE 6 (a) Two AR models and a transition/membership function. (b) The corresponding STAR model (or fuzzy rule–based model).

one defined over s_t). Hence the associated fuzzy rules are of a simplified type. The rule base of the previous example would be composed of two rules:

$$\begin{array}{ll} \text{IF } y_{t-1} \text{ IS } A \text{ THEN} & y_t^* = 1.7 - 0.03y_{t-1} + 0.02y_{t-2} + \varepsilon_t \\ \text{In any case,} & y_t^* = 1.6 + 0.01y_{t-1} + 0.06y_{t-2} + \varepsilon_t. \end{array} \tag{16}$$

In the first rule, A has a membership function $\mu_A(y_{t-1}) = f(y_{t-1}, 2.5, 1.75)$ whose linguistic interpretation could be *higher than about* 2.5. The second rule is called a *default rule*.

In a real world situation, it is usual to find auto-correlation in more than one lagged values of the series. In that case, we might think of a slightly more complicated model, allowing for changes of regimes in the domain of another transition variable, for example y_{t-2}. The knowledge encoded in such a model could be expressed, for example, by the following rule base:

$$\begin{array}{ll} \text{IF } y_{t-1} \text{ IS } A \text{ AND } y_{t-2} \text{ IS } B, \text{ THEN} & y_t^* = 1.7 - 0.03y_{t-1} + 0.02y_{t-2} + \varepsilon_t \\ \text{In any case,} & y_t^* = 1.6 + 0.01y_{t-1} + 0.06y_{t-2} + \varepsilon_t. \end{array} \tag{17}$$

If we defined B as *higher than about* 0, we could have $\mu_B(y_{t-2}) = f(y_{t-2}, 0, 1.0)$. In that case, our model generates the surface shown in the right (b) side of Fig. 7.

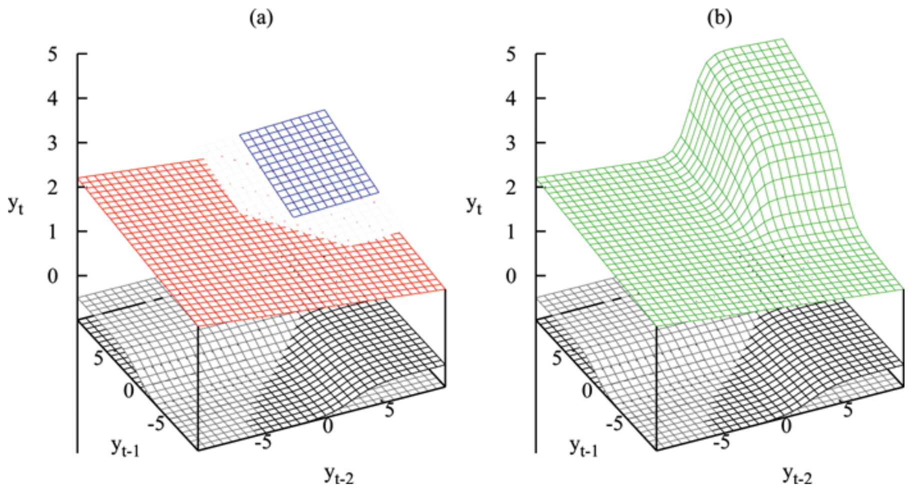

FIGURE 7 (a) Two AR models and a transition/membership function. (b) The corresponding STAR model (or fuzzy rule–based model).

Using the product to model implication and conjunction of fuzzy logic terms, this model is written as

$$y_t = (1.6 + 0.01y_{t-1} + 0.06y_{t-2}) + (1.6 + 0.01y_{t-1} + 0.06y_{t-2}) \times \prod_{i=1}^{2} f(y_{t-i}, c_i, \gamma_i), \tag{18}$$

where $c_1 = 2.5$, $\gamma_1 = 1.75$, $c_2 = 0.0$, $\gamma_2 = 1.0$.

With the addition of more nonlinear terms or fuzzy rules, this extended STAR (or FRBM) adds a great deal of flexibility and modeling power to the standard STAR, and it has been proved to be a member of the class of *universal approximators*.

In fact, in the literature we can find a related extended STAR model which is based on the merge of soft computing (SC) with standard statistic techniques. This model, called Neuro-Coefficient STAR (NCSTAR) (Medeiros and Veiga, 2005), uses an artificial neural network to dynamically adapt the coefficients of a linear model. It is noted as

$$y_t = \mathbf{b}_0\mathbf{x}_t + \sum_{i=1}^{k} \mathbf{b}_i\mathbf{x}_t f(\boldsymbol{\omega}_i\mathbf{z}_t) + \varepsilon_t, \tag{19}$$

where $\mathbf{z}_t$ is a $q \times 1$ vector of transition variables and $\boldsymbol{\omega}_i = [\omega_{1i}, \ldots, \omega_{qi}]'$ are real parameters. By comparing this model with the Additive FRBM, described in Section 3.1.3, p. 6, it is easy to prove that this model is functionally equivalent to an FRBM (Aznarte, 2008).

5. ADVANCES IN THE IDENTIFICATION AND ESTIMATION OF FRBMs

Since the development of FRBMs, it is clear that good identification and estimation procedures for such models were crucial towards their applicability and success. Hundreds of proposals have arisen in this relatively short period, and because of the relations that link FRBMs to the family of nonlinear autoregressive models, which will be mentioned in Section 4, they are also of interest in the econometrics area.

As stated previously, fuzzy logic and all the ideas derived from it are an integral component of a wider knowledge body known as SC, which was defined as a "collection of methodologies" (Zadeh, 1994). The hybridization of those methodologies has proved to be a very fruitful entity. This is true for example in the estimation and identification processes of an FRBM, where the most successful proposals were those that gathered the representation power of the fuzzy approach with the main strengths of other SC elements, as we show below.

Basically, the identification of an FRBM implies finding a good set of fuzzy rules. This task can be divided into finding the antecedents of such rules and fixing the corresponding consequents. Both tasks are strongly conditioned by an important consideration: there is a trade-off between the interpretability of the FRBM and its accuracy. A model with more interpretable rules usually tends to be less precise in its outputs, and the other way around, the more accurate an FRBM is, the less interpretable it usually is. Hence, this is a key criterion in building a rule base. For a review on the interpretability–accuracy balance, see Casillas et al. (2003) and Casillas and Herrera (2003).

For example, if one is interested in obtaining a very clear linguistic interpretation of a problem at hand, one should choose Mamdani type rules, whose consequents are expressed in linguistic (fuzzy) terms. If the accuracy of the final model is more important, TSK type rules should be chosen, as their linear consequents are more appropriate for approximative tasks and they allow to obtain a more accurate model.

At the same time, considering the antecedents, there are two main approaches. The global semantics–based approach considers a global collection of fuzzy sets for all the fuzzy rules. Hence the meaning for each label is exactly the same in every rule. On the other hand, there is the local semantics–based approach, in which each fuzzy rule has its own local fuzzy sets associated, which entails for a higher accuracy paid by a decrease in their interpretability.

All of the nonlinear autoregressive models, seen as FRBM, are on one extreme of this interpretability vs. accuracy compromise line: they are strongly biased towards accuracy. This is because they use a local semantics-based approach and TSK-type consequents. However, through the use of a conveniently designed estimation procedure you can strive for accuracy while preserving most of the interpretability.

Next, we will briefly show a successful proposal for identification and estimation of accurate FRBMs, which could be applicable to nonlinear autoregressive models after the equivalence results shown in Section 4.

5.1. LEL-TSK: Local Evolutionary Learning of TSK Rules

In Alcalá et al. (2007), the use of Mamdani fuzzy rules as local prototypes was proposed to obtain accurate TSK rules. An algorithm composed of two main stages is developed following the so-called MOGUL paradigm (Cordón et al., 1999) (*a Methodology to Obtain Genetic fuzzy rule–based systems Under the iterative rule Learning approach*). MOGUL is based on evolutionary algorithms—specifically genetic algorithms—to automatically design FRBMs by identifying and estimating the fuzzy rule base while keeping a generic structure. The efficiency of the algorithm is greatly improved by functionally decomposing the search space into two parts, which are explored sequentially.

Two key properties of a rule should be taken into account when building them: the quality of rule with respect to the accuracy and also the interactions among them. The local identification of prototypes induces competition amongst rules, considering only the quality of the approximation performed by each rule. Then to increase the generalization ability of the model, interaction among rules is very important. In the MOGUL approach, a post-processing stage is adopted for this purpose.

The procedure is built upon an wide number of previous proposals resulting in an effective but rather sophisticated algorithm. Hence a detailed exposition is out of the scope of this article. Refer to Alcalá et al. (2007) for a complete description of this algorithm. A brief description of this proposal is presented below.

5.1.1. *Local Process for Identifying Prototypes*

To obtain a set of local Mamdani fuzzy rules which can be used as prototypes, a method described in Cordón and Herrera (2001) that is based on local covering measures to induce competition among rules can be used, considering the *completeness* and *consistency* properties (Cordón and Herrera, 1999).

As a starting point, a grid-like partition of the input space is performed by evenly distributing the transition (or membership) functions over the range of each variable. Over the resulting rules, a fitness function is defined taking into account their local adequacy to the available data series. This fitness function is basically concerned with the degree of agreement between the rule and positive examples and the possible negative examples. An iterative process goes on building new prototypes while there are uncovered data.

To obtain the linear consequents, once the set of local fuzzy prototypes is obtained, the existing local linear input–output relation is computed using the data located in each input subspace by means of another variant of evolutionary strategy which minimizes the Mean Square Error (MSE).

It is important to note that, in the original proposal for this method, triangular membership functions were used. There is no loss of generality, as deriving the method for any of the membership functions mentioned in Section 2 is straightforward.

After all prototypes have been selected, they are refined—tuned—through the evolutionary strategy defined in Cordón and Herrera (2001). This tuning is performed within the local area corresponding for each prototype.

This concludes the first stage of the process. The result is an FRBM whose accuracy can be further improved without seriously affecting the interpretability. So far, the identification of the FRBM has been accomplished, along with a pre-estimation phase. An effective estimation step is still pending.

5.1.2. *Post-Processing Stage*

Once this first version of the FRBM has been obtained, two different processes are considered to improve the performance of the system: the *genetic simplification process* and the *genetic tuning process*:

1. *Genetic simplification process*: This process, described in Cordón and Herrera (2001), is based on a standard binary-coded genetic algorithm and also considers the completeness property. Its aim is to select the subset of rules whose interaction leads to best results amongst the rules generated in the previous stage.
2. *Genetic tuning process*: This method is an adaptation of Cordón and Herrera (1999). It further adjusts the parameters of the FRBM by a combination of an evolutionary strategy and a genetic algorithm.

An implementation of the algorithm is publicly available within the KEEL Project at (Alcalá-Fdez et al., 2009) repository `http://www.keel.es`.

5.2. Specifying Traditional Nonlinear Models Using SC

These equivalence links shown in Section 4 are not just a curious fact without further consequences. On the one hand, it allows to establish the statistical properties and a sound statistical framework for FRBMs, which is expected to improve the way we understand and use these models. On the other hand, being able to directly apply the methods and techniques coming from SC into the field of nonlinear autoregressive models is an interesting possibility.

For example, we could apply the LEL-TSK technique described in Section 5.1 to identify and estimate STAR or NCSTAR models. In order to do so, first we need to select the kind of membership—or transition—function according to the kind of desired regime-switching model. Then we have to process the available data through the first stage of the LEL-TSK algorithm: local process for identifying prototypes. As a result of this stage, we will have a first set of fuzzy rules. This set of rules conveys the number of identified regimes, which is given by the cardinal of the set. The rule set also includes the transition functions, which are the membership function in the antecedent of each rule, and the AR model for each regime, which is the consequent of the rule.

The second stage of the LEL-TSK procedure will simplify and fine-tune the obtained regime-switching model, discarding possible irrelevant or redundant regimes and reestimating the coefficients of the AR regimes.

In the end, we will obtain a regime-switching model with the same structure as originally defined. Notwithstanding, the model will quite possibly have a higher accuracy because of the superior performance of evolutionary computation methods when addressing highly nonlinear optimization problems.

Next section offers, as a preliminar example of the benefits of the FRBM approach, a comparison of the results of the application of the models described in this article to a problem which might be of econometric interest.

6. APPLICATION EXAMPLE

The predictability of some common financial time series, as stock prices or level of indexes, is a controversial issue which has been questioned in the past. These series tend to behave as random-walk processes which, in theory, makes its prediction impossible. On the other hand, financial time series are usually subject to regime shifting, which means that their statistical properties do not remain constant through time. For example, some series do react to business cycle or to slow-down periods.

Furthermore, a large amount of random day-to-day variations are usually found in these series, which makes them still more unpredictable. Examples of events which influence the behavior of a financial series are the announcements of firm specific news, fiscal measures, employment reports, as well as political events. The effect of this on a specific series is to increase its noisy nature, making it difficult to distinguish good prediction algorithms from bad ones, since even a naive predictor can produce good results. See Hellström and Holmström (1998) for a detailed discussion.

This inherent difficulty of the economic series, together with the expected immediate benefit of good forecasts, made them one of the preferred application fields for SC researchers. There is a great number of papers which deal with modeling and prediction of financial time series under the various disciplines included in SC, i.e., neural networks Kaastra and Boyd (1996); Zhang et al. (1998), bioinspired algorithms Chen (2002); Mahfoud and Mani (1996), support vector machines Cao and Tay (2001); Huang et al. (2005); Tay and Cao (2001) etc.

One of the most prolific amongst these disciplines have been fuzzy-related models, where a great effort has been devoted to financial series forecasting. Be it in the form of hybrid models (neuro-fuzzy Ang and Quek (2006); Lin et al. (2003); Rast (1997), bioinspired Deboeck (1994); Ju et al. (1997)) or variants of fuzzy rule–based models de Andrés Sánchez and Terce no Gómez (2003); Dourra and Siy (2002); Serguieva and Hunter (2004); Wang (2002); Yu (2005), the list of references is long.

6.1. Financial Application: IBM Volatility Series

This experiment dealt with realized volatility of the IBM Dow Jones Industrial Average index stocks. The raw intraday data are constituted of tick-by-tick quotes extracted from the NYSE Trade and Quote (TAQ) database. The period of the analysis starts in January 2, 1995, and ends in December 30, 2005. In order to estimate the daily realized volatility, we followed the procedure proposed by McAleer and Medeiros (2008), leaving the last year (2005) for testing and using the rest to identify and estimate the models. In order to stationarize the series, a diff-log type transformation was also employed.

With the purpose of assessing the benefits of the fuzzy rule–based approach with respect to classic and state-of-the-art statistical and econometric methods, we used four models. Besides the FRBM model described in Section 5.1, we built and estimated a linear AR model, a STAR model (14), and an NCSTAR model (19). These models were built using the implementation found in the publicly available GNU's `R` package `tsDyn` di Narzo et al. (2009), which follows the hypothesis testing iterative identification procedure proposed in Teräsvirta (1994) and Teräsvirta et al. (2006b). In all cases, a common lagged structure was used, with $(y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4})$ as input variables. An initial p-value of 0.95 was selected, which resulted in 2 regimes both for the STAR and NCSTAR models (a full specification of these models is shown in the Appendix).

Table 1 shows the results obtained by the four models. The root mean squared error (RMSE) and the mean absolute errors (MAE) are the standard error measures, while the normalized mean squared error

TABLE 1 Root mean squared error, mean absolute error, and normalized mean squared error values obtained in training and test by the four considered models

Model	RMSE		MAE		nMSE	
	In-sample	Test	In-sample	Test	In-sample	Test
AR	0.53421	0.53070	0.41012	0.41824	0.79038	0.78002
STAR	0.53250	0.53345	0.40944	0.42142	0.78533	0.78812
NCSTAR	0.53145	0.52767	0.40772	0.41538	0.78224	0.77115
FRBM	0.53114	0.44939	0.40880	0.34046	0.76123	0.76764

(nMSE) is computed as the mean squared error divided by the variance of the data.

The columns marked "in-sample" show the results obtained when the models were applied to forecast the data used during identification and estimation, whereas the "test" columns show the results corresponding to the testing dataset (data from 2005).

As we can see, all the three nonlinear models behaved similarly when doing in-sample forecasting, outperforming the linear AR model. This confirms that the complexity of the series is very difficult to be captured by a simple linear model. The greater generalization ability of the fuzzy rule–based model is evident when considering the results obtained in the testing dataset: up to around 24% of improvement against the STAR model in mean average error.

In order to determine if results are statistically significant, we applied the Diebold and Mariano test Diebold and Mariano (1995) to compare the forecasts of AR, STAR, and NCSTAR against the FRBM. For each of these models, the *p*-values obtained were 0.972, 0.999, and 0.997, respectively, which allows us to conclude that the forecasts are statistically different.

These empirical results clearly confirm the usefulness of the FRBM approach in terms of efficiency, and hence justify its consideration as an accurate model for time series modeling and forecasting.

7. CONCLUSIONS

This article deals with one of the core elements of SC, the fuzzy rule–based model (FRBM). To present it, we briefly explored fuzzy logic and fuzzy sets, and offered a quick review of the state of the art in FRBM identification.

In Section 4, we have shown how FRBMs relate to existent nonlinear autoregressive models when dealing with time series modeling and forecasting. It has been proved that an FRBM can be considered as a generalization of several standard statistical models including the smooth transition autoregressive model (STAR), and this allows for an exchange of ideas and methods from one field to another.

Section 5 deals with a recent development in the identification and estimation of FRBM. The LEL-TSK procedure presented in this section uses bio-inspired algorithms to determine the structure and the parameters of fuzzy models, and it is can be also used in the framework of other related nonlinear autoagressive models.

As an example of the benefits that can be obtained from the exchange, between both areas, we modeled the IBM realized volatility series in Section 6.1 through an FRBM using LEL-TSK. The results were compared against those obtained with three common autoregressive models (NCSTAR, STAR, and AR), showing that the FRBM significantly outperforms them.

This fact suggests that some of the proposals included in SC could be used in the framework of the autoregressive nonlinear time series analysis, especially the algorithms for identification and estimation of FRBMs. Further experiments may allow us to adapt these proposals to other complex models whose identification and estimation remain difficult. Interpretability issues of autoregressive nonlinear models in terms of fuzzy logic are also to be studied.

A. PARAMETERS OF THE MODELS

A.1. STAR Model

$$\begin{aligned} y_t = & -0.0299 - 0.6024 y_{t-1} - 0.4140 y_{t-2} - 0.3231 y_{t-3} - 0.1862 y_{t-4} \\ & + (-0.0585 + 0.1880 y_{t-1} + 0.1153 y_{t-2} + 0.1089 y_{t-3} + 0.0781 y_{t-4}) \\ & \times f(y_{t-1}; 40.0838, -0.1466) + \varepsilon_t. \end{aligned} \tag{20}$$

A.2. NCSTAR Model

$$\begin{aligned} y_t = & \, 4.91 + 0.08 y_{t-1} + 0.53 y_{t-2} + 0.83 y_{t-3} + 1.40 y_{t-4} \\ & + (9.77 + 0.65 y_{t-1} + 0.41 y_{t-2} + 0.23 y_{t-3} - 0.19 y_{t-4}) \\ & \times f(\mathbf{x}_t; \boldsymbol{\omega}) + \varepsilon_t, \end{aligned} \tag{21}$$

with $\boldsymbol{\omega} = [\gamma, \mathbf{c}] = [-0.02, (0.40, 0.46, 0.52, 0.62)]$.

ACKNOWLEDGMENTS

The research included in this work is part of the first author's Ph.D. dissertation at the Department of Computer Science and Artificial Intelligence of the University of Granada. It benefited form the comments of an editor and two referees, which indeed helped in improving the first manuscript. It has been partially supported by the Ministerio de Ciencia e Innovación of the Spanish Government through Research Grant Ref. TIN2009-14575.

REFERENCES

Alcalá, R., Alcalá-Fdez, J., Casillas, J., Cordón, O., Herrera, F. (2007). Local identification of prototypes for genetic learning of accurate TSK fuzzy rule-based systems. *International Journal of Intelligent Systems* 22:909–941.

Alcalá-Fdez, J., Sánchez, L., García, S., del Jesus, M. J., Ventura, S., Garrell, J. M., Otero, J., Romero, C., Bacardit, J., Rivas, V. M., Fernández, J. C., Herrera, F. (2009). KEEL: A software tool to assess evolutionary algorithms to data mining problems. *Soft Computing* 13:307–318.

Andersen, T. G., Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39:885–905.

de Andrés Sánchez, J., Terce no Gómez, A. (2003). Estimating a term structure of interest rates for fuzzy financial pricing by using fuzzy regression methods. *Fuzzy Sets and Systems* 139:313–331.

Ang, K. K., Quek, C. (2006). Stock trading using RSPOP: A novel rough Set-Based Neuro-Fuzzy approach. *Neural Networks, IEEE Transactions on* 17:1301–1315.

Aznarte, J. L. (2008). Modelling time series through fuzzy rule based models: A statistical approach. PhD thesis, Universidad de Granada.

Aznarte, J. L., José Manuel Benítez, Juan Luis Castro (2007). Smooth transition autoregressive models and fuzzy rule-based systems: Functional equivalence and consequences. *Fuzzy Sets and Systems* 158:2734–2745.

Aznarte, J. L., Medeiros, M. C., Benítez, J. M. (2010). Linearity testing for fuzzy rule based models. *Fuzzy Sets and Systems* 161:1836–1851.

Cao, L., Tay, F. E. H. (2001). Financial forecasting using support vector machines. *Neural Computing & Applications* 10:184–192.

Casillas, J., Cordón, O., Herrera, F., Magdalena, L. (2003). *Accuracy Improvements in Linguistic Fuzzy Modeling*. New York: Springer.

Casillas, J., Herrera, F. (2003). *Interpretability Issues in Fuzzy Modeling*. Berlin: Springer.

Castro, J. L. (1995). Fuzzy logic controllers are universal approximators. *IEEE Trans. Systems, Man and Cybernetics* 25:629–635.

Castro, J. L., Delgado, M. (1996). Fuzzy systems with defuzzification are universal approximators. *IEEE Trans. Systems, Man and Cybernetics* 26:149–152.

Chen, S. H. (2002). *Genetic Algorithms and Genetic Programming in Computational Finance*. Norwell, MA, USA: Kluwer Academic Pub.

Cordón, O., del Jesus, M. J., Herrera, F., Lozano, M. (1999). Mogul: A methodology to obtain genetic fuzzy rule-based systems under the iterative rule learning approach. *International Journal of Intelligent Systems* 14:1123–1153.

Cordón, O., Herrera, F. (1999). A two-stage evolutionary process for designing tsk fuzzy rule-based systems. *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics* 29:703–715.

Cordón, O., Herrera, F. (2001). Hybridizing genetic algorithms with sharing scheme and evolution strategies for designing approximate fuzzy rule-based systems. *Fuzzy Sets and Systems* 118:235–255.

Deboeck, G. (1994). *Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets*. Hoboken, NJ, USA: John Wiley & Sons Inc.

Diebold, F. X., Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13:253–263.

Dourra, H., Siy, P. (2002). Investment using technical analysis and fuzzy logic. *Fuzzy Sets and Systems* 127:221–240.

Dubois, D., Prade, H. (1980). *Fuzzy Sets and Systems*. New York: Academic Press.

Hellström, T., Holmström, K. (1998). Predicting the stock market. Technical report, Center of Mathematical Modeling (CMM), Mälardalen University.

Huang, W., Nakamori, Y., Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers and Operations Research* 32:2513–2522.

Jang, J.-S. R., Sun, C.-T., Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence.* Upper Saddle River, NJ, USA: Prentice Hall.

Ju, Y. J., Kim, C. E., Shim, J. C. (1997). Genetic-based fuzzy models: Interest rate forecasting problem. *Computers & Industrial Engineering* 33:561–564.

Kaastra, I., Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing* 10:215–236.

Klir, G. J., Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic.* Upper Saddle River, NJ, USA: Prentice Hall.

Kosko, B. (1994). Fuzzy systems as universal approximators. *IEEE Trans. Computers* 43:1324–1333.

Lin, J. W., Hwang, M. I., Becker, J. D. (2003). A fuzzy neural network for assessing the risk of fraudulent financial reporting. *Managerial Auditing Journal* 18:657–665.

Mahfoud, S., Mani, G. (1996). Financial forecasting using genetic algorithms. *Applied Artificial Intelligence* 10:543–565.

Mamdani, E. H., Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *Intern. J. Man-Machines Studies* 7:1–13.

McAleer, M., Medeiros, M. C. (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics* 147:104–119.

Medeiros, M. C., Veiga, A. (2005). A exible coeficient smooth transition time series model. *IEEE Transactions on Neural Networks* 16:97–113.

di Narzo, A. F., Stigler, M., Aznarte, M. J. L. (2009) *tsDyn: Time series analysis based on dynamical systems theory*, R package version 0.7.

Nauck, D., Klawonn, F., Kruse, R. (1997). *Foundations of Neuro-Fuzzy Systems.* Chichester, UK: Wiley and Sons.

Rast, M. (1997). Forecasting financial time series with fuzzy neural networks. In: *Intelligent Processing Systems, 1997. ICIPS '97. 1997 IEEE International Conference on*, Vol. 1. pp. 432–434.

Serguieva, A., Hunter, J. (2004). Fuzzy interval methods in investment risk appraisal. *Fuzzy Sets and Systems* 142:443–466.

Sugeno, M., Kang, G. T. (1988). Structure identification of fuzzy model. *Fuzzy Sets and Systems* 28:15–33.

Sugeno, M., Nishida, M. (1985). Fuzzy control of model car. *Fuzzy Sets and Systems* 16:103–113.

Takagi, T., Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Systems, Man and Cybernetics* 15:116–132.

Tay, F. E. H., Cao, L. (2001). Application of support vector machines in financial time series forecasting. *Omega* 29:309–317.

Teräsvirta, T. (1994). Specification, estimation and evaluation of smooth transition autoregresive models. *Journal of the American Statistical Association* 89:208–218.

Teräsvirta, T., Medeiros, M. C., Rech, G. (2006a). Building neural network models for time series: A statistical approach. *Journal of Forecasting* 25:49–75.

Teräsvirta, T., Medeiros, M. C., Rech, G. (2006b). Building neural network models for time series: A statistical approach. *Journal of Forecasting* 25:49–75.

Tong, H. (1990). *Non-Linear Time Series: A Dynamical System Approach.* Oxford, UK: Oxford University Press.

Wang, Y. F. (2002). Predicting stock price using fuzzy grey prediction system. *Expert Systems with Applications* 22:33–38.

Yager, R. R., Filev, D. P. (1994). *Essentials of Fuzzy Modeling and Control.* New York, USA: John Wiley & Sons, Inc.

Yu, H. K. (2005). Weighted fuzzy time series models for TAIEX forecasting. *Physica A: Statistical Mechanics and Its Applications* 349:609–624.

Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Systems, Man and Cybernetics* 1:28–44.

Zadeh, L. A. (1994). Soft computing and fuzzy logic. *IEEE Software* 11:48–56.

Zhang, G., Eddy Patuwo, B., Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting* 14:35–62.