Accurate, Transparent, and Compact Fuzzy Models for Function Approximation and Dynamic Modeling through Multi-objective Evolutionary Optimization

Fernando Jiménez[1], Antonio F. Gómez-Skarmeta[1], Hans Roubos[2], and Robert Babuška[2]

[1] Dept. Informática, Inteligencia Artificial y Electrónica
University of Murcia, Spain
{fernan,skarmeta}@dif.um.es
[2] Control Engineering Laboratory
Delft University of Technology, the Netherlands
{J.A.Roubos,R.Babuska}@ITS.TUDelft.NL

Abstract. Evolutionary algorithms to design fuzzy rules from data for systems modeling have received much attention in recent literature. Many approaches are able to find highly accurate fuzzy models. However, these models often contain many rules and are not transparent. Therefore, we propose several objectives dealing with transparency and compactness besides the standard accuracy objective. These objectives are used to find multiple Pareto-optimal solutions with a multi-objective evolutionary algorithm in a single run. Attractive models with respect to compactness, transparency and accuracy are the result.

Keywords: Takagi-Sugeno fuzzy model, Pareto optimality, multi-objective evolutionary algorithm.

1 Introduction

This paper deals with fuzzy model parameter estimation and structure selection. In fuzzy model identification, we can, in general, take into account three criteria to be optimized: compactness, transparency and accuracy. Different measures for these criteria are proposed here. Compactness is related to the size of the model, i.e. the number of rules, the number of fuzzy sets and the number of inputs for each rule. Transparency is related to linguistic interpretability [1,2] and locality of the rules. Often one is interested in the local behavior of the global nonlinear model. Such information can be obtained by constraining the model-structure during identification. Transparency and model interpretability for data-based fuzzy models received a lot of interest in recent literature [3,4,5, 6].

Evolutionary Algorithms (EA) [7,8] have been recognized as appropriate techniques for multi-objective optimization because they perform a search for multiple solutions in parallel [9,10,11]. EAs have been applied to learn both the

E. Zitzler et al. (Eds.): EMO 2001, LNCS 1993, pp. 653–667, 2001.

antecedent and consequent part of fuzzy rules, and models with both fixed and varying number of rules have been considered [12,13]. Also, EAs have been combined with other techniques like fuzzy clustering [14,15,16] and neural networks [17,18]. This has resulted in many complex algorithms and, as recognized in [1] and [2], often the transparency and compactness of the resulting rule base is not considered to be of importance. In such cases, the fuzzy model becomes a black-box, and one can question the rationale for applying fuzzy modeling instead of other techniques like, e.g., neural networks. If the fuzzy model or a neural network is handled as a black-box model it will typically store the information in a distributed manner among the neurons or fuzzy sets and their associated connectivity [19].

Most evolutionary approaches to multi-objective fuzzy modeling consist of multiple EAs, usually designed to achieve a single task each, which are applied sequentially to obtain a final solution. In these cases each EA optimizes the problem attending to one criterion separately which is an impediment for the global search. Simultaneous optimization of all criteria is more appropriate. Other approaches are based on classical multi-objective techniques in which multiple objectives are aggregated into a single function to be optimized [16]. In this way a single EA obtains a single compromise solution. Current evolutionary approaches for multi-objective optimization consist of a single multi-objective EA, based on the Pareto optimality notion, in which all objectives are optimized simultaneously to find multiple non-dominated solutions in a single run of the EA. These approaches can also be considered from the fuzzy modeling perspective [20]. The advantage of the classical approach is that no further interaction with the decision maker is required, however it is often difficult to define a good aggregation function. If the final solution cannot be accepted, new runs of the EA may be required until a satisfying solution is found. The advantages of the pareto approach are that no aggregation function has to be defined, and the decision maker can choose the most appropriate solution according to the current decision environment at the end of the EA run. Moreover, if the decision environment changes, it is not always necessary to run the EA again. Another solution may be chosen out of the family of non-dominated solutions that has already been obtained.

In this paper we propose a single multi-objective EA to find, with a low necessity for human intervention, multiple non-dominated solutions for fuzzy modeling problems. In section 2, fuzzy modeling and the criteria taken into account, are discussed. The main components of the multi-objective EA are described in section 4. Section 5 proposes several optimization models for fuzzy modeling and a decision making strategy. In section 6, experiments with the EA for a test problem are shown and compared with results in literature. Section 6 concludes the paper and indicates lines for future research.

2 Fuzzy Model Identification

2.1 Fuzzy Model Structure

We consider rule-based models of the Takagi-Sugeno (TS) type [21] which are especially suitable for the approximation of dynamic systems. The rule consequents are often taken to be linear functions of the inputs:

$$R_i : \textbf{If } x_1 \text{ is } A_{i1} \textbf{ and } \ldots x_n \text{ is } A_{in} \textbf{ then} \tag{1}$$
$$\hat{y}_i = \zeta_{i1}x_1 + \ldots, \zeta_{in}x_n + \zeta_{i(n+1)}, \; i = 1, \ldots, M$$

Here $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ is the input vector, $\hat{y}_i$ is the output of the ith rule, A_{ij} $(j = 1, \ldots, n)$ are fuzzy sets defined in the antecedent space by membership functions $\mu_{A_{ij}} : \mathrm{R} \to [0, 1]$, $\zeta_{ij} \in \mathrm{R}$ $(j = 1, \ldots, n+1)$ are the consequent parameters, and M is the number of rules. The total output of the model is computed by aggregating the individual contributions of the rules:

$$\hat{y} = \sum_{i=1}^{M} p_i(\mathbf{x})\hat{y}_i \tag{2}$$

where $p_i(\mathbf{x})$ is the normalized firing strength of the ith rule:

$$p_i(\mathbf{x}) = \frac{\prod_{j=1}^{n} \mu_{A_{ij}}(x_j)}{\sum_{i=1}^{M} \prod_{j=1}^{n} \mu_{A_{ij}}(x_j)} \tag{3}$$

We apply the frequently used trapezoidal membership functions to describe the fuzzy sets A_{ij} in the rule antecedents:

$$\mu_{A_{ij}}(x) = \max\left(0, \min\left(\frac{x - a_{ij}}{b_{ij} - a_{ij}}, 1, \frac{c_{ij} - x}{c_{ij} - b_{ij}}\right)\right) \tag{4}$$

2.2 Multi-objective Identification

Identification of fuzzy models from data requires the presence of multiple criteria in the search process. In multi-objective optimization, the set of solutions is composed of all those elements of the search space for which the corresponding objective vector cannot be improved in any dimension without degradation in another dimension. These solutions are called *non-dominated* or *Pareto-optimal.* Given two decision vectors $\mathbf{a}$ and $\mathbf{b}$ in a universe U, $\mathbf{a}$ is said to *dominate* $\mathbf{b}$ if $f_i(a) \leq f_i(b)$, for all objective functions f_i, and $f_j(a) < f_j(b)$, for at least one objective function f_j, for minimization. A decision vector $\mathbf{a} \in U$ is said to be *Pareto-optimal* if no other decision vector dominates $\mathbf{a}$.

The Pareto-optimality concept should be integrated within a decision process in order to select a suitable compromise solution from all non-dominated alternatives. In a decision process, the decision maker expresses preferences which should be taken into account to identify preferable non-domination solutions. In this way, *preference articulation* implicitly defines a *utility function* which discriminates between candidate solutions. Approaches based on weights, goals and priorities have been used more often. Moreover, preference articulation can be achieved in different ways depending on how the computation and the decision processes are combined in the search for compromise solutions. Three broad classes can be identified, *a priori*, *a posteriori*, and *progressive* articulation of preferences.

2.3 Rule Set Simplification Techniques

Automated approached to fuzzy modeling often introduce redundancy in terms of several similar fuzzy sets that describe almost the same region in the domain of some variable. According to some similarity measure, two or more similar fuzzy sets can be merged to create a new fuzzy set representative for the merged sets [22]. This new fuzzy set substitutes the ones merged in the rule base. The merging process is repeated until fuzzy sets for each model variable cannot be merged, i.e., they are not similar. This simplification may results in several identical rules, which are removed from the rule set.

We consider the following similarity measure between two fuzzy sets A and B:

$$S(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{5}$$

If $S(A,B) > \theta_S$ (we use $\theta_S = 0.6$) then fuzzy sets A and B are merged in a new fuzzy set C as follows:

$$\begin{aligned} a_C &= min\{a_A, a_B\} \\ b_C &= \alpha b_A + (1-\alpha) b_B \\ c_C &= \alpha c_A + (1-\alpha) c_B \\ d_C &= max\{d_A, d_B\} \end{aligned} \tag{6}$$

where $\alpha \in [0,1]$ determine the influence of A and B on the new fuzzy set C.

3 Criteria for Fuzzy Modeling

We consider three main criteria to search for an acceptable fuzzy model: (i) accuracy, (ii) transparency, and (iii) compactness. It is necessary to define quantitative measures for these criteria by means of appropriate objective functions which define the complete fuzzy model identification.

The accuracy of a model can be measured with the *mean squared error*:

$$MSE = \frac{1}{K} \sum_{k=1}^{K} (y_k - \hat{y}_k)^2 \tag{7}$$

where y_k is the true output and $\hat{y}_k$ is the model output for the kth input vector, respectively, and K is the number of data samples.

Many measures are possible for the second criterion, transparency. Nevertheless, in this paper we only consider one of most significant, *similarity*, as a first starting point. The similarity S among distinct fuzzy sets in each variable of the fuzzy model can be expressed as follows:

$$S = \max_{\substack{i,\,j,\,k \\ A_{ij} \neq B_{ik}}} S(A_{ij}, B_{ik}),\ i = 1, \dots, n,\ j = 1, \dots, M,\ k = 1, \dots, M \tag{8}$$

This is an aggregated similarity measure for the fuzzy rule-based model with the objective to minimize the maximum similarity between the fuzzy sets in each input domain.

Finally, measures for the third criterion, the compactness, are the number of rules M and the number of different fuzzy sets L of the fuzzy model. We assume that models with a small number of rules and fuzzy sets are compact.

In summary, we have considered three criteria for fuzzy modeling, and we have defined the following measures for these criteria:

Criteria	Measures
Accuracy	MSE
Transparency	S
Compactness	M, L

4 Multi-objective Evolutionary Algorithm

The main characteristics of the Multi-Objective Evolutionary Algorithm are the following:

1. The proposed algorithm is a Pareto-based multi-objective EA for fuzzy modeling, i.e., it has been designed to find, in a single run, multiple non-dominated solutions according to the Pareto decision strategy. There is no dependence between the objective functions and the design of the EA, thus, any objective function can easily be incorporated. Without loss of generality, the EA minimizes all objective functions.
2. Constraints with respect to the fuzzy model structure are satisfied by incorporating specific knowledge about the problem. The initialization procedure and variation operators always generate individuals that satisfy these constraints.
3. The EA has a variable-length, real-coded representation. Each individual of a population contains a variable number of rules between 1 and max, where max is defined by a decision maker. Fuzzy numbers in the antecedents and the parameters in the consequent are coded by floating-point numbers.
4. The initial population is generated randomly with a uniform distribution within the boundaries of the search space, defined by the learning data and model constraints.

5. The EA search for among simplified rule sets, i.e, all individuals in the population has been previously simplified (after initialization and variation), which is an added ad hoc technique for transparency and compactness. So, all individuals in the population have a similarity S between 0 and 0.6.
6. Chromosome selection and replacement are achieved by means of a variant of the preselection scheme. This technique is, implicitly, a niche formation technique and an elitist strategy. Moreover, an explicit niche formation technique has been added to maintain diversity respect to the number of rules of the individuals. Survival of individuals is always based on the Pareto concept.
7. The EAs variation operators affect at the individuals at different levels: (i) the rule set level, (ii) the rule level, and (iii) the parameter level.

4.1 Representation of Solutions and Constraint Satisfaction

An individual I for this problem is a rule set of M rules as follows:

$$\begin{array}{l} R_1: \; A_{11} \; \ldots \; A_{1n} \quad \zeta_{11} \; \ldots \; \zeta_{1n} \; \zeta_{1(n+1)} \\ \ldots \\ R_M: A_{M1} \; \ldots \; A_{Mn} \; \zeta_{M1} \; \ldots \; \zeta_{Mn} \; \zeta_{M(n+1)} \end{array}$$

The constraints on the domain of the variables for a fuzzy model come given by the semantic of a fuzzy number. Thus, a fuzzy number A_{ij} ($i = 1, \ldots, M$, $j = 1, \ldots, n$) can be represented by means of four real values $a_{ij}, b_{ij}, c_{ij}, d_{ij} \in [l_j, u_j]$, with $a_{ij} \leq b_{ij} \leq c_{ij} \leq d_{ij}$. The consequent parameters are also real values constrained by a domain, i.e. $\zeta_{ij} \in [l, u]$ ($i = 1, \ldots, M$, $j = 1, \ldots, n+1$). Other constraint are related with the number of rules M of the model, which can be defined between a lower number 1 and a upper number max fixed by the decision maker.

In the following sections we describe easy initialization and variation procedures to generate random individuals which satisfy these constraints.

4.2 Initial Population

Initial population is completely random, except that the number of individuals with M rules, for all $M \in [1, max]$, should be between $minNS$ and $maxNS$ to ensure diversity respect to the number of rules, where $minNS$ and $maxNS$, with $0 \leq minNS \leq \frac{PS}{max} \leq maxNS \leq PS$ (PS is the population size), are the minimum and maximum niche size respectively (see next subsection).

To generate an individual with M rules, the procedure is as follows: for each trapezoidal fuzzy number A_{ij} ($i = 1, \ldots, M$, $j = 1, \ldots, n$), four random real values from $[l_j, u_j]$ are generated and sorted to satisfy the constraints $a_{ij} \leq b_{ij} \leq c_{ij} \leq d_{ij}$. Parameters ζ_{ij} ($i = 1, \ldots, M$, $j = 1, \ldots, n+1$) are real values generated at random from $[l, u]$. After, the individual is simplified according to the procedure described in a previous section.

4.3 Selection and Generational Replacement

We use a variant of the preselection scheme [8] which has been one of the results of previous works for general constrained multi-objective optimization problems by EA [23].

In each iteration of the EA, two individuals are picked at random from the population. These individuals are crossed $nChildren$ times and children mutated producing $2 \cdot nChildren$ offspring. After, the best of the first offspring replaces the first parent, and the best of the second offspring replaces to the second parent only if:

- the offspring is better than the parent, and
- the number of rules of the offspring is equal to the number of rules of the parent, or the niche count of the parent is greater than $minNS$ and the niche count of the offspring is smaller than $maxNS$

An individual I is better than another individual J if I dominates J. The best individual of a collection is any individual I such that there is no other individual J which dominates I. The niche count of an individual I is the number of individuals in the population with the same number of rules as I.

Note that the preselection scheme is an implicit niche formation technique to maintain diversity in the populations because an offspring replaces an individual similar to itself (one of their parents). Implicit niche formation techniques are more appropriate for fuzzy modeling than explicit techniques, such as sharing function, which can provoke an excessive computational time. However, we need an additional mechanism for diversity with respect to the number of rules of the individuals in the population. One of the reasons is that the number of rules is an integer parameter and the variation operators can generate individuals with quite different numbers of rules of the parents. The preselection scheme is not effective in such a case. The added explicit niche formation technique ensures that the number of individuals with M rules, for all $M \in [1, max]$, is greater or equal to $minNS$ and smaller or equal to $maxNS$. Moreover, the preselection scheme is also an elitist strategy because the best individual in the population is replaced only by a better one.

4.4 Variation Operators

As already said, an individual is a set of M rules. A rule is a collection of n fuzzy numbers (antecedent) plus $n+1$ real parameters (consequent), and a fuzzy number is composed of four real numbers. In order to achieve an appropriate exploitation and exploration of the potential solutions in the search space, variation operators working in the different levels of the individuals are necessary. In this way, we consider three levels of variation operators: rule set level, rule level, and parameter level. After a sequence of crossovers and mutations, the offspring are simplified according to the rule set simplication procedure as described previously.

Five crossover and four mutation operators are used in the EA. In the following, $\alpha \in [0, 1]$ is a random number from a uniform distribution.

Rule Set Level Variation Operators

- **Crossover 1**: Given two parents $I_1 = (R_1^1 \ldots R_{M_1}^1)$ and $I_2 = (R_1^2 \ldots R_{M_2}^2)$, this operator exchanges information about the number of rules of the parents and information about the rules of the parents, but no rule is internally crossed. Two children are produced: $I_3 = (R_1^1 \ldots R_a^1 R_1^2 \ldots R_b^2)$ and $I_4 = (R_{a+1}^1 \ldots R_{M_1}^1 R_{b+1}^2 \ldots R_{M_2}^2)$, where $a = round(\alpha \cdot M_1 + (1-\alpha) \cdot M_2)$ and $b = round((1-\alpha) \cdot M_1 + \alpha \cdot M_2)$. The number of rules of the children is between M_1 and M_2.
- **Crossover 2**: This operator increases the number of rules of the two children as follows: the first child contains all M_1 rules of the first parent and $\min\{\max - M_1, M_2\}$ rules of the second parent; the second child contains all M_2 rules of the second parent and $\min\{\max - M_2, M_1\}$ rules of the first parent.
- **Mutation 1**: This operator deletes or adds, both with equal probability, one rule in the rule set. For deletion, one rule is randomly deleted from the rule set. For rule-addition, one rule is randomly generated, according to the initialization procedure described, and added to the rule set.

Rule Level Variation Operators

- **Crossover 3**: Given two parents $I_1 = (R_1^1 \ldots R_i^1 \ldots R_{M_1}^1)$ and $I_2 = (R_1^2 \ldots R_j^2 \ldots R_{M_2}^2)$, this operator produces two children $I_3 = (R_1^1 \ldots R_i^3 \ldots R_{M_1}^1)$ and $I_4 = (R_1^2 \ldots R_j^4 \ldots R_{M_2}^2)$, with $R_i^3 = \alpha R_i^1 + (1-\alpha) R_j^2$ and $R_j^4 = \alpha R_j^2 + (1-\alpha) R_i^1$, where i, j are random indexes from $[1, M_1]$ and $[1, M_2]$ respectively.
- **Crossover 4**: Given two parents $I_1 = (R_1^1 \ldots R_i^1 \ldots R_{M_1}^1)$ and $I_2 = (R_1^2 \ldots R_j^2 \ldots R_{M_2}^2)$, this operator produce two children $I_3 = (R_1^1 \ldots R_i^3 \ldots R_{M_1}^1)$ and $I_4 = (R_1^2 \ldots R_j^4 \ldots R_{M_2}^2)$, where R_i^3 and R_j^4 are obtained with the *uniform* crossover.
- **Mutation 2**: This operator removes a randomly chosen rule and inserts a new one which is randomly generated by the rule-initialization procedure.

Parameter Level Variation Operators

- **Crossover 5**: Given two parents, and one rule of each parent randomly chosen, this operator crosses the fuzzy numbers corresponding to a random input variable or the consequent parameters. The crossover is arithmetic.
- **Mutation 3**: This operator mutates a random fuzzy number or the consequent of a random rule. The new fuzzy number or consequent is generated at random.
- **Mutation 4**: This operator changes the value of one of the antecedent fuzzy sets a, b, c or d of a random fuzzy number, or a parameter of the consequent ζ, of a randomly chosen rule. The new value of the parameter is generated at random within the constraints by a non-uniform mutation.

5 Optimization Models and Decision Making

After preliminary experiments in which we have checked different optimization models, the following remarks can be maded:

1. The minimization of the number of rules M of the individuals has negative influence on the evolution of the algorithm. The reason is than this parameter is not an independent variable to optimize, as the amount of information in the population decreases when the average number of rules is low, which is not good for exploration. Then, we do not minimize the number of rules during the optimization, but we will take it into account at the end of the run, in a posteriori articulation of preferences applied to the last population.
2. It is very important to note that a very transparent model will be not accepted by a decision maker if the model is not accurate. In most fuzzy modeling problems, excessively low values for similarity hamper accuracy, for which these models are normally rejected. Alternative decision strategies, as *goal programming*, enable us to reduce the domain of the objective functions according to the preferences of a decision maker. Then, we can impose a goal g_S for similarity, which stop minimization of the similarity in solutions for which goal g_S has been reached.
3. The measure L (number of different fuzzy sets) is considerably reduced by the rule set simplification technique. So, we do not define an explicit objective function to minimize L.

According to the previous remarks, we finally consider the two following optimization models:

Optimization Model 1:

$$\begin{array}{l} Minimize\ f_1 = MSE \\ Minimize\ f_2 = S \end{array} \tag{9}$$

Optimization Model 2:

$$\begin{array}{l} Minimize\ f_1 = MSE \\ Minimize\ f_2 = max\{g_S, S\} \end{array} \tag{10}$$

At the end of the run, we consider the following a posteriori articulation of preferences applied to the last population to obtain the final compromise solution:

1. Identify the set $X^* = \{x_1^*, \ldots, x_p^*\}$ of non-dominated solutions according to:

$$\begin{array}{l} Minimize\ f_1 = MSE \\ Minimize\ f_2 = S \\ Minimize\ f_3 = M \end{array} \tag{11}$$

2. Choose from X^* the most accurate solution x_i^*; remove x_i^* from X^*;
3. If solution x_i^* is not accurate enough or there is no solution in the set X^* then STOP (no solution satisfies);

4. If solution x_i^* is not transparent or compact enough then go to step 2;
5. Show the solution x_i^* as output.

Computer aided inspection shown in Figure 3 can help in decisions for steps 2 and 3.

6 Experiments and Results

Consider the 2^{nd} order nonlinear plant studied by Wang and Yen in [24,25]:

$$y(k) = g(y(k-1), y(k-2)) + u(k) \tag{12}$$

with

$$g(y(k-1), y(k-2)) = \frac{y(k-1)y(k-2)(y(k-1)-0.5)}{1+y^2(k-1)y^2(k-2)} \tag{13}$$

The goal is to approximate the nonlinear component $g(y(k-1), y(k-2))$ of the plant with a fuzzy model. As in [24], 400 simulated data points were generated from the plant model (12). Starting from the equilibrium state $(0,0)$, 200 samples of identification data were obtained with a random input signal $u(k)$ uniformly distributed in $[-1.5, 1.5]$, followed by 200 samples of evaluation data obtained using a sinusoidal input signal $u(k) = \sin(2\pi k/25)$. The resulting signals and the real surface are shown in Figure 1.

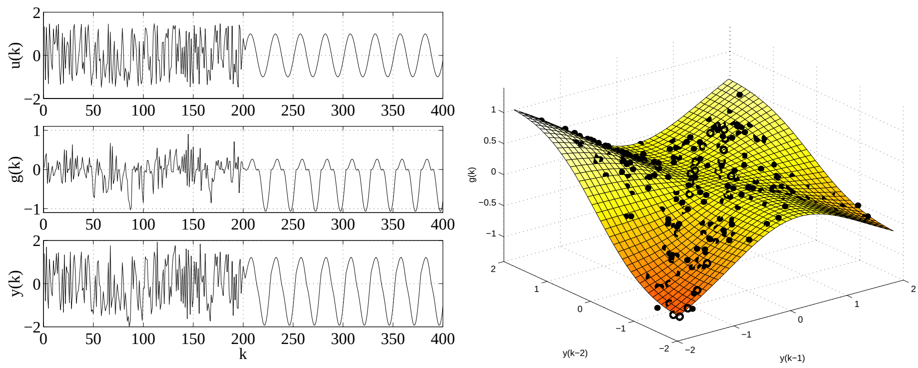


Fig. 1. *Left*: Input $u(k)$, unforced system $g(k)$, and output $y(k)$ of the plant in (12). *Right*: Real surface.

The following values for the parameters of the EA were used in the simulations: population size 100, crossover probability 0.8, mutation probability 0.4, number of children for the preselection scheme 10, minimum number of individuals for each number of rules 5, and maximum number of individuals for each number of rules 20. All crossover and mutation operators are applied with the same probability. The EA stops when the solutions satisfy the decisor maker.

We show results obtained with the EA by using the optimization models (9) ($max = 5$) and (10) ($max = 5$, $g_S = 0.25$). Figure 2 shows the non-dominated solutions in the last population according to (11) for both optimization models (9) and (10). One can appreciate the effectiveness of the preselection technique and the added explicit niche formation technique to maintain diversity in the populations. The main differences between the optimization models (9) and (10) are that with (9) the EA obtains more diversity but the fuzzy models are less accurate. Goal-based model have the disadvantage that it is necessary to choose, a priori, a good goal for the problem, although this value is representative of the maximum degree of overlapping of the fuzzy sets allowed by a decisor.

According to the described decision process, we finally choose a compromise solution showed in Figure 3 by means of different graphics for the obtained model. Figure 3(a) shows the local model. The surface generated by the model is shown in Figure 3(b), fuzzy sets for each variable are showed in Figure 3(c), and finally, the identification and validation results as well as the prediction error are shown in Figure 3(d).

We compared our results, with those obtained by the four different approaches proposed in [25] and [26]. The best results obtained for in each case are summarized in Table 1, with an indication of the number of rules, number of different fuzzy sets, consequent type, and obtained MSE for training and evaluation data. In [25], the low MSE on the training data is in contrast with the MSE for the evaluation data which indicates overtraining. The solution in [26] is similar to the solutions in this paper with respect to the accuracy, transparency and compactness, but hybrid techniques (initial fuzzy clustering and a sequence of specific genetic algorithms) were required in [26]. Solutions in this paper are obtained with a single EA and they have been chosen among different alternatives, which is an advantage for an appropriate decision process.

Table 1. Fuzzy models for the dynamic plant. All models are of the Takagi-Sugeno type.

Ref.	No. of rules	No. of sets	Consequent	MSE train	MSE eval
[25]	36 rules (initial)	12 (B-splines)	Linear	$1.9 \cdot 10^{-6}$	$2.9 \cdot 10^{-3}$
	24 rules (optimized)	-	Linear	$2.0 \cdot 10^{-6}$	$6.4 \cdot 10^{-4}$
[26]	7 rules (initial)	14 (triangular)	Linear	$1.8 \cdot 10^{-3}$	$1.0 \cdot 10^{-3}$
	5 rules (optimized)	5 (triangular)	Linear	$5.0 \cdot 10^{-4}$	$4.2 \cdot 10^{-4}$
This paper[1]	5 rules	5 (trapezoidal)	Linear	$2.0 \cdot 10^{-3}$	$1.3 \cdot 10^{-3}$
This paper[2]	5 rules	6 (trapezoidal)	Linear	$5.9 \cdot 10^{-4}$	$8.8 \cdot 10^{-4}$

[1] Solution corresponds to the solution marked with * in Figure 2(left), and Figure 3(left)

[2] Solution corresponds to the solution marked with * in Figure 2(right), and Figure 3(right)

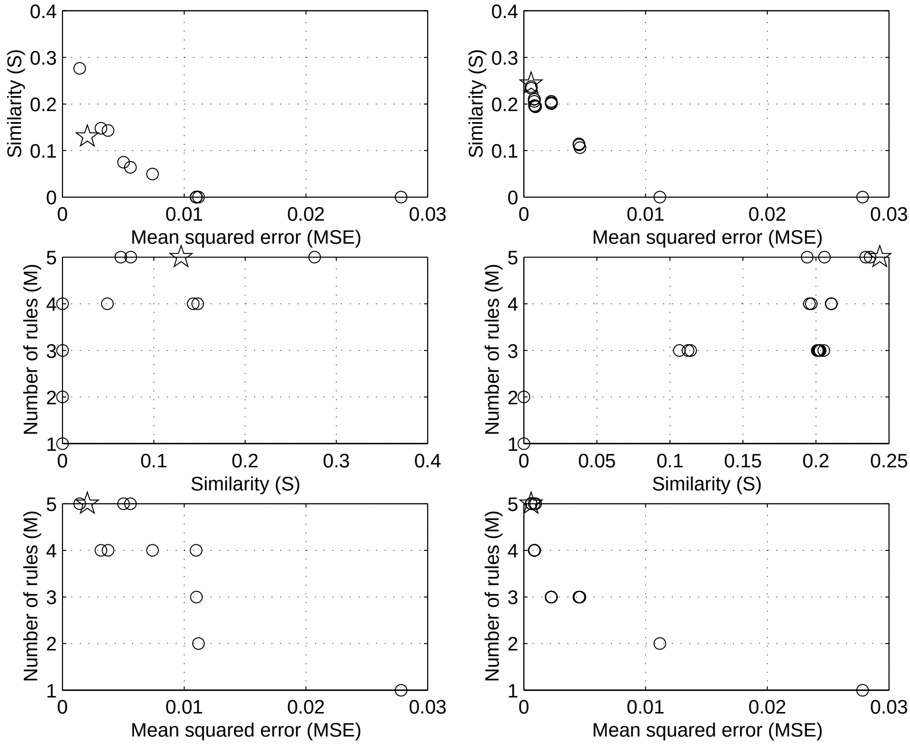
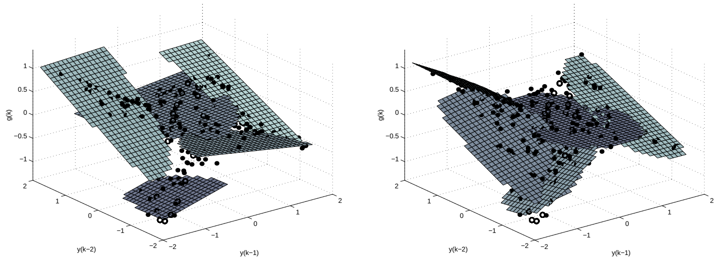


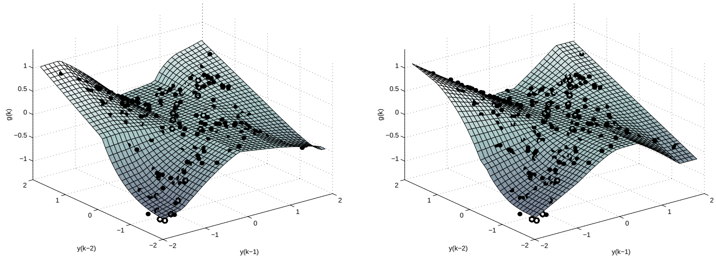
Fig. 2. *Left:* Non-dominated solutions according to (11) obtained with the Pareto-based multi-objective EA by using the optimization model (9). *Right:* Non-dominated solutions according to (11) obtained with the Pareto-based multi-objective EA by using the optimization model (10). Solution marked with * is the final compromise solution.

7 Conclusions and Future Research

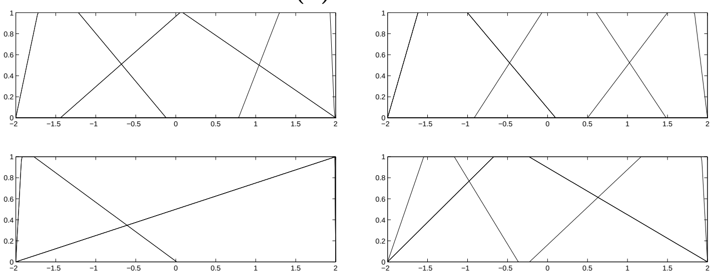
This paper remarks some initial results in the combination of Pareto-based multi-objective evolutionary algorithms and fuzzy modeling. Criteria such as accuracy, transparency and compactness have been taken into account in the optimization process. Some of these criteria have been partially incorporated into the EA by means of ad hoc techniques, such as rule set simplification techniques. An implicit niche formation technique (preselection) in combination with other explicit techniques with low computational costs have been used to maintain diversity. These niche formation techniques are appropriate in fuzzy modeling if excessive amount of data are required. Excessive computational times would result if sharing function were used. Elitism is also implemented by means of the preselection technique. A goal based approach has been proposed to help to obtain more accurate fuzzy models. Results obtained are good in comparison with other more complex techniques reported in literature, with the advantage that the proposed



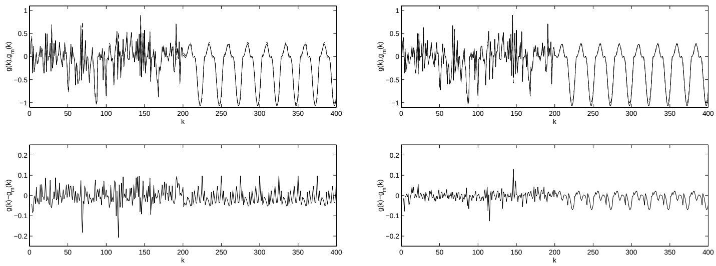
(a) Local model



(b) Surface



(c) Fuzzy sets



(d) Identification and validation, and error prediction

Fig. 3. Accurate, transparent and compact fuzzy models for the plant model (12). *Left:* Final compromise solution obtained with the Pareto-based multi-objective EA by using the optimization model (9) and a posteriori decision process. *Right:* Final compromise solution obtained with the Pareto-based multi-objective EA by using the optimization model (10) and a posteriori decision process.

technique identifies a set of alternative solutions. We also proposed an easy decision process with a posteriori articulation of preferences to choose finally a compromise solution.

One of the main differences between the proposed EA and other approaches for fuzzy modeling is the reduced complexity because we use a single EA for generating, tuning and simplification processes. Moreover, human intervention is only required at the end of the run to choose one of the multiple non-dominated solutions found by the EA.

In our future works we will consider other and more complex fuzzy modeling test problems in order to check the robustness of the EA, other measures to optimize transparency, e.g., similarity in the consequent domain instead or together with of the antecedent domain, scalability of the algorithm, and applications in the real word by means of research projects.

References

1. J. Valente de Oliveira. Semantic constraints for membership function optimization. *IEEE Transactions on Fuzzy Systems*, 19(1):128–138, 1999.
2. M. Setnes, R. Babuška, and H. B. Verbruggen. Rule-based modeling: Precision and transparency. *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications & Reviews*, 28:165–169, 1998.
3. H. Pomares, I. Rojas, J. Ortega, J. Gonzalez, and A. Prieto. A systematic approach to a self-generating fuzzy rule-table for function approximation. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, 30(3):431–447, 2000.
4. Y. Jin. Fuzzy modeling of high-dimensional systems. *IEEE Transactions on Fuzzy Systems: Complexity Reduction and Interpretability Improvement.*, 8:212–221, 2000.
5. O. Cordón and F. Herrera. A proposal for improving the accuracy of linguistic modeling. *IEEE Transactions on Fuzzy Systems*, 8(3):335–344, 2000.
6. T.A. Johansen, R. Shorten, and R. Murray-Smith. On the interpretation and identification of dynamic Takagi-Sugeno fuzzy models. *IEEE Transactions on Fuzzy Systems*, 8(3):297–313, 2000.
7. J. Biethahn and V. Nissen (eds). *Evolutionary Algorithms in Management Applications.* Springer-Verlag Berlin Heidelberg, 1995.
8. D.E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, 1989.
9. C.M. Fonseca and P.J. Fleming. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary Computation*, 3(1):1–16, 1995.
10. J. Horn and N. Nafpliotis. Multiobjective optimization using the niched pareto genetic algorithm. IlliEAL Report No. 93005, July 1993.
11. N. Srinivas and K. Deb. Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation*, 2(3):221–248, 1995.
12. C.-H. Wang, T.-P. Hong, and S.-S. Tseng. Integrating fuzzy knowledge by genetic algorithms. *Fuzzy Sets and Systems*, 2(4):138–149, 1998.
13. H. Ishibuchi, T. Nakashima, and T. Murata. Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, 29(5):601–618, 1999.

14. L.O. Hall, I.B. Özyurt, and J.C. Bezdek. Clustering with genetically optimized approach. *IEEE Transactions on Evolutionary Computing*, 3(2):103–112, 1999.
15. H.-S. Hwang. Control strategy for optimal compromise between trip time and energy consumption in a high-speed railway. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 28(6):791–802, 1998.
16. A.F. Gómez-Skarmeta and F. Jiménez. Fuzzy modeling with hibrid systems. *Fuzzy Sets and Systems*, 104:199–208, 1999.
17. I. Jagielska, C. Matthews, and T. Whitfort. An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems. *Neurocomputing*, 24:37–54, 1999.
18. M. Russo. FuGeNeSys – a fuzzy genetic neural system for fuzzy modeling. *IEEE Transactions on Fuzzy Systems*, 6(3):373–388, 1998.
19. S. Mitra and Y. Hayashi. Neuro-fuzzy rule generation: Survey in soft computing framework. *IEEE Transactions on Neural Networks*, 11(3):748–768, 2000.
20. A.F. Gómez-Skarmeta, F. Jiménez, and J. Ibánez. Pareto-optimality in fuzzy modelling. In *EUFIT'98*, pages 694–700, Aachen, Alemania, 1998.
21. T. Takagi and M. Sugeno. Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics*, 15:116–132, 1985.
22. M. Setnes, R. Babuška, U. Kaymak, and H.R. van Nauta Lemke. Simlilarity measures in fuzzy rule simplification. *IEEE Transaction on Systems, Man and Cybernetics, Part B: Cybernetics*, 28(3):376–386, 1999.
23. F. Jiménez, J.L. Verdegay, and A.F. Gómez-Skarmeta. Evolutionary techniques for constrained multiobjective optimization problems. In *Genetic and Evolutionary Computation Conference (GECCO-99), Workshop on Multi-Criterion Optimization Using Evolutionary Methods*, pages 115–116, Orlando, Florida, USA, 1999.
24. L. Wang and J. Yen. Extracting fuzzy rules for system modeling using a hybrid of genetic algorithms and Kalman filter. *Fuzzy Sets and Systems*, 101:353–362, 1999.
25. J. Yen and L. Wang. Application of statistical information criteria for optimal fuzzy model construction. *IEEE Transactions on Fuzzy Systems*, 6(3):362–371, 1998.
26. H. Roubos and M. Setnes. Compact fuzzy models through complexity reduction and evolutionary optimization. In *proceedings 9th IEEE Conference on Fuzzy System*, pages 762–767, San Antonio, USA, May 7-10, 2000.