Fuzzy Modeling with Multi-Objective Neuro-Evolutionary Algorithms

F. Jiménez[1], G. Sánchez[1], A.F. Gómez-Skarmeta[1]
H. Roubos[2], R. Babuška[2]

[1]Dept. Ingeniería de la Información y las Comunicaciones
Universidad de Murcia, 30071 Espinardo, Murcia, Spain
[2]Control Engineering Laboratory, Delft University of Technology, the Netherlands
E-mail: fernan@dif.um.es

Abstract— **Interpretability aspects of fuzzy models have received quite some attention in recent years and may be obtained by using transparent rule-structures and well characterized fuzzy membership functions. Moreover, model compactness is important for the interpretability and is related to the number of rules and fuzzy sets. Besides these two criteria, the model accuracy should always be taken into account. In this way, several criteria appear in fuzzy modeling and then multi-objective evolutionary algorithms are suitable to capture several non-dominated solutions in a single run of the algorithm. For fuzzy modeling, we describe a multi-objective neuro-evolutionary algorithm that considers all three objectives. The algorithm applies an accuracy criterium and a transparency criterium, based on fuzzy set similarity, while compactness is achieved by a specific technique, incorporated ad hoc within the evolutionary alrorithm. Results are shown for an approximation problem studied before by other authors.**

Keywords— **Fuzzy Modeling, Multi-objective Evolutionary Algorithm, Neural Networks, Interpretability.**

I. INTRODUCTION

In recent years, fuzzy modeling, as a complement to conventional modeling techniques, has become an active research topic and has found successful applications in many areas. Evolutionary Algorithms (EAs) [4] have been applied to learn both the antecedent and consequent part of fuzzy rules, and models with both fixed and varying number of rules have been considered [25], [9]. Also, EAs have been combined with other techniques like fuzzy clustering [7], [5] and neural networks [10], [18]. This has resulted in many complex algorithms and, as recognized in [24] and [19], often the transparency and compactness of the resulting rule base are not considered to be of importance; accuracy aspect prevails and interpretability aspects are partly ignored. In such cases, the fuzzy model becomes a black-box, and one can question the rationale for applying fuzzy modeling instead of other techniques.

Transparency and model interpretability for data-based fuzzy models received quite some interest in recent literature [16], [13], [2], [14]. Based on this literature, we introduce three important criteria to be optimized in fuzzy model identification: compactness, transparency and accuracy. 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, while transparency is related to linguistic interpretability of the fuzzy sets and locality of the rules [15], [19]. Often one is interested in the local behavior of the global nonlinear model. Such information may be obtained by constraining the model structure during identification or by using these criteria by multi-objective optimization techniques, like multi-objective EAs [1], [11]. Most evolutionary approaches to multi-objective fuzzy modeling consist of multiple EAs, usually designed to achieve a single task each, which are then applied sequentially. In these cases, each EA optimizes the model attending to a single criterion separately, which is an impediment for global search. Simultaneous optimization of all criteria is more appropriate. Therefore, others used approaches based on classical multi-objective techniques in which multiple objectives are aggregated into a single function to be optimized [5], [17]. In this way, the EA obtains a compromise solution that consists of the weighted criteria.

Another promising method to handle multi-criteria optimization problems is the 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 [3], [6], [21]. The solution for a such a multi-objective optimization problem is a set of so-called Pareto optimal solutions. A Pareto-based multi-objective EA incorporates the Pareto concept to identify multiple solutions through a single run of the algorithm. This practically leaves obsolete the classical tendency to aggregate the different objectives using a weight vector or similar approach to obtain a single function which is then optimized. In that case, the method often requires several executions of the EA with different weights, in order to identify one Pareto solution in each run. In this aspect, multi-objective optimization, based on the Pareto-optimality concept, distinguishes itself from related optimization methods, like gradient techniques, simulated annealing or neural networks.

Pareto-based multi-objective evolutionary approaches can also be considered from the fuzzy modeling perspective [8], [12]. The advantage of the the classical optimization approach with aggregated objectives, a single solution is obtained without further interaction with the decision maker. However, it may often be difficult to define a good aggregation function. Then, if the solution cannot be accepted, new runs of the EA are required until a satisfying solution is found. The advantages of the Pareto approach are that no aggregation function need to be defined, and that 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 multi-objective neuro-EA to find multiple non-dominated solutions for fuzzy modeling problems. In section 2, the fuzzy model is defined and in section 3 and 4, techniques to improve transparency and compactness of rule set and training are approached respectively. The criteria taken into account (accuracy, tranparency and compactness) are discussed in section 5. The main components of the multi-objective neuro-EA are described in section 6. Section 7 shows the results obtained for a test problem. Section 8 concludes the paper and indicates lines for future research.

II. FUZZY MODEL IDENTIFICATION

We consider rule-based models of the Takagi-Sugeno (TS) type [23]. The rule consequents are taken to be linear functions of the inputs:

$$\begin{array}{l} R_i : \textbf{If } x_1 \textit{ is } A_{i1} \textbf{ and} \ldots x_n \textit{ is } A_{in} \textbf{ then} \\ \hat{y}_i = \zeta_{i1}x_1 + \ldots + \zeta_{in}x_n + \zeta_{i(n+1)} \end{array} \quad (1)$$

where $i = 1, \ldots, M$, and $\mathbf{x} = (x_1, \ldots, x_n)$ is the input vector, $\hat{y}_i$ is the output of the ith rule, A_{ij} are fuzzy sets defined in the antecedent space by membership functions $\mu_i : \Re \rightarrow [0,1]$, $\zeta_{ij} \in \Re$ 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 each rule:

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

where $p_i(\mathbf{x})$ $i = 1, \ldots, M$ is the normalized firing strength of the ith rule:

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

where $i = 1, \ldots, M$.

Each fuzzy set A_{ij} is described by an asymmetric gaussian membership function.

$$\mu_{A_{ij}}(x_j) = \begin{cases} \exp\left(-\frac{(c_{ij}-x_j)^2}{2\sigma_{lij}^2}\right) & if\ x_j < c_{ij} \\ \exp\left(-\frac{(x_j-c_{ij})^2}{2\sigma_{rij}^2}\right) & if\ x_j \geq c_{ij} \end{cases} \quad (4)$$

where $i = 1, \ldots, M$ and $j = 1, \ldots, n$.

This fuzzy model is defined by a radial basis function neural network. The number of neurons in the hidden layer of an RBF neural network is equal to the number of rules in the fuzzy model. The firing strength of the ith neuron in the hidden layer matches the firing strength of the ith rule in the fuzzy model. We apply an asymmetric gaussian membership function defined by three parameters, the center c, left variance σ_l and rigth variance σ_r. Therefore, each neuron in the hidden layer has these three parameters that define its firing strength value. The neurons in the output layer perform the computations for the first order linear function described in the consequents of the fuzzy model, therefore, the ith neuron of the output layer has the parameters ζ_i that correspond to the linear function defined in the ith rule of the fuzzy model.

III. A TECHNIQUE TO IMPROVE THE TRANSPARENCY AND COMPACTNESS OF THE FUZZY RULE SETS

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. This new fuzzy set substitutes the ones merged in the rule base. On the other hand, if there are two fuzzy sets which are similar, but not very similar, the best approach is to split the fuzzy sets, so that their similarity improves. The merging-splitting process is repeated until fuzzy sets for each model variable cannot be merged, i.e., they are not similar. This process 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:

$$S(A,B) = \max\left\{\frac{A \cap B}{A}, \frac{A \cap B}{B}\right\} \quad (5)$$

If $S(A,B) > \theta_1$ (we use $\theta_1 = 0.9$), fuzzy sets A and B are merged in a new fuzzy set C as follows:

$$c_C = \alpha c_A + (1-\alpha)c_B \quad (6)$$

$$\sigma_{lC} = \frac{c_C - \min\{c_A - n_\sigma\sigma_{lA}, c_B - n_\sigma\sigma_{lB}\}}{n_\sigma} \quad (7)$$

$$\sigma_{rC} = \frac{\max\{c_A + n_\sigma\sigma_{rA}, c_B + n_\sigma\sigma_{rB}\} - c_C}{n_\sigma} \quad (8)$$

where $n_\sigma = 3$ and $\alpha \in [0,1]$ determines the influence of A and B on the new fuzzy set C:

$$\alpha = \frac{\sigma_{rA} + \sigma_{lA}}{\sigma_{rA} + \sigma_{lA} + \sigma_{rB} + \sigma_{lB}} \quad (9)$$

If $\theta_2 < S(A,B) < \theta_1$ (we use $\theta_2 = 0.6$), fuzzy sets A and B are split as follows:

$$\sigma_C^l = \sigma_C^l(1-\beta) \quad (10)$$

$$\sigma_C^r = \sigma_C^r(1-\beta) \quad (11)$$

where $\beta \in [0,1]$ denotes the amount of spliting between A and B; in our algorithm, we use $\beta = 0.1$.

IV. TRAINING OF THE RBF NEURAL NETWORKS

The RBF neural networks associated with the fuzzy models can be trained with a gradient method to obtain more accuracy. However, in order to mantain the transparency and compactness of the fuzzy sets, only the consequent parameters are trained. The training algorithm incrementally updates the parameters based on the currently presented training pattern. The network parameters are updated by applying the gradient descent method to an error function. The error function for the th training pattern is given by the MSE function defined in (15). The update rule is:

$$\zeta_{ij}^{new} = \zeta_{ij}^{old} - \eta \frac{\partial MSE}{\partial \zeta_{ij}} \tag{12}$$

where $i = 1, \ldots, M$, $j = 1, \ldots, n+1$ and η is the learning rate. This rule is applied a number of epochs. Our algorithm use a value of $\eta = 0.01$ and a number of 10 epochs. The detailed derivation of $\Delta\zeta_{ij} = \frac{\partial MSE}{\partial \zeta_{ij}}$ is the following:

$$\Delta\zeta_{ij} = -(\hat{y}_k - y_k)p_i(\mathbf{x})x_j \tag{13}$$

$$\Delta\zeta_{i(n+1)} = -(\hat{y}_k - y_k)p_i(\mathbf{x}) \tag{14}$$

where $i = 1, \ldots, M$, $j = 1, \ldots, n$ and $p_i(\mathbf{x})$ is the firing value for the ith rule defined as in equation 3.

V. CRITERIA FOR FUZZY MODELING

Identification of fuzzy models from data requires de 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 the universe U, $\mathbf{a}$ is said to *dominate* $\mathbf{b}$ if $f_i(\mathbf{a}) \leq f_i(\mathbf{b})$, for all objective functions f_i and $f_j(\mathbf{a}) < f_j(\mathbf{b})$, 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 $\mathbf{a}$.

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

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

$$MSE = \frac{1}{K}\sum_{k=1}^{K}(y_k - \hat{y}_k)^2 \tag{15}$$

where y_k is the actual output and $\hat{y}_k$ is the desired 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 the 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:

$$S = \max_{\substack{i,j,k \\ A_{ij} \neq B_{kj}}} S(A_{ij}, B_{kj}) \tag{16}$$

where $i = 1, \ldots, M$, $j = 1, \ldots, n$ and $k = 1, \ldots, M$.

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.

VI. MULTI-OBJECTIVE NEURO-EVOLUTIONARY ALGORITHM

The main characteristics of the Multi-Objective Neuro-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 EAs, thus, any objective function can easily be incorporated. Without loss of generality, the EA minimizes all objective functions.
2. 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.
3. 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.
4. 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.
5. The EA searchs for among rule sets treated with the technique described in section III and trained as defined in section IV, i.e, all individuals in the population have been treated with technique III to mantain transparency and trained (after initialization and variation), which is an added ad hoc technique for transparency, compactness and accuracy. 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.

A. Representation of solutions

An individual for this problem is a rule set of M rules defined by the weights of the RBF neural network. With n input variables, we have for each individual the following parameters:

• centers c_{ij}, left variances σ_{lij} and right variances σ_{rij}, $i = 1, \ldots, M$, $j = 1, \ldots, n$

• coefficients for the linear function of the consequent ζ_{ij}, $i = 1, \ldots, M$, $j = 1, \ldots, n+1$

B. Initial population

The initial population is generated randomly. The number of individuals with M rules, for all $M \in [1, max]$, must be between $minNS$ and $maxNS$ to ensure diversity respect to the number of rules, where $minNS$ and $maxNS$, with $0 \leq minNS \leq PS/max \leq maxNS \leq PS$ (PS is the population size), are the minimum and maximum niche size respectively (see next section).

To generate an individual with M rules, the procedure is as follows: for each fuzzy number A_{ij} ($i = 1, \ldots, M$, $j = 1, \ldots, n$), three random real values from $[l_j, u_j]$ are generated and sorted to satisfy the constraints $c_{lij} \leq c_{ij} \leq c_{rij}$ ($\sigma_{lij} = (c_{ij} - c_{lij})/n_\sigma$ and $\sigma_{rij} = (c_{rij} - c_{ij})/n_\sigma$). 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.

C. 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 three real numbers. In order to achieve an appropiate 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.

C.1 Rule set level variation operators

• **Rule Set Crossover**: This operator interchanges rules. Given two parents $I_1 = (R_1^1 \ldots R_{M_1}^1)$ and $I_2 = (R_1^2 \ldots R_{M_2}^2)$, 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 $a = round(\alpha \cdot M_1)$ and $b = round((1-\alpha) \cdot M_2)$. The number of rules of the children is between M_1 and M_2.

• **Rule Set Increase Crossover**: 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.

• **Rule Set Mutation**: 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.

C.2 Rule Level Variaton Operators

• **Rule Arithmetic Croossover**: Performs arithmetic crossover of two random rules. 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.

• **Rule Uniform Crossover**: Performs uniform crossover of two random rules. 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.

C.3 Parameter level variation operators

• **Arithmetic Crossover**: Given two parents, and one rule of each parent randomly chosen, this operator performs an arithmetic cross of the fuzzy numbers corresponding to a random input variable or the consequent parameters.

• **Non-Uniform Mutation**: This operator changes the value of one of the antecedent fuzzy sets 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 given by a non-uniform mutation.

• **Uniform Mutation**: Similar to former, but within the constraints given by an uniform mutation.

• **Small Mutation**: Similar to former, but within the constraints given by an small mutation. The small mutation produced an small change in the individual and it is suitable for fine tuning of the real parameters.

D. Selection and generational replacement

In each iteration, the neuro-EA executes the following steps:

1. Two individuals are picked at random from the population.
2. These individuals are crossed and mutated to produce two offspring.
3. Performs technique III in the offspring.
4. Performs the training in the offspring.
5. The first offspring replaces the first parent and the second offspring replaces the second parent only if:

• the offspring is betther 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 niche count of an individual I is the number of individuals in the population with the same number of rules as I. The preselection scheme is an implicit niche formation technique to maintain diversity in the population because an offspring replaces an individual similar to itself (one of their parents). Implicit niche formation techniques are more appropiate for fuzzy modeling than explicit techniques, such as sharing function, which can provoke an excessive computational time. However, we need and additional mechanism for diversity with respect to the number of rules of the individuals in the population. 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.

E. Optimization model

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

1. Instead of minimizing of the number of rules M we have decided to search for rules sets with a number of rules within an interval $[1, max]$ where a decision maker can feel comfortable. The explicit niche formation technique ensures the EA always contains a minimum of representative rule sets for each number of rules in the populations. 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 decision process 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 reduced by the technique of section III. So, we do not define an explicit objective function to minimize L.

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

$$\begin{array}{ll} Minimize & f_1 = MSE \\ Minimize & f_2 = \max(g_s, S) \end{array} \qquad (17)$$

The output of the algorithm is finally calculated according to:

$$\begin{array}{ll} Minimize & f_1 = MSE \\ Minimize & f_2 = S \\ Minimize & f_3 = M \end{array} \qquad (18)$$

VII. EXPERIMENTS AND RESULTS

We consider the modeling of the rule base given in [22]: The corresponding surface is shown in Figure VII. In [20] a model with four rules was identified from sampled data ($N = 546$) by the supervised clustering algorithm, which was initialized with 12 clusters. This model was optimized using a Genetic Algorithm to result in a MSE of 1.6. Table I shows results obtained with the Pareto based multi-objective neuro-EAs.The following values for the parameters were used in the simulations: $PS = 100$, $minNS = 5$, $maxNS = 30$, cross probability 0.9, mutation probability 0.9, $g_S = 0.25$, and $max = 5$.

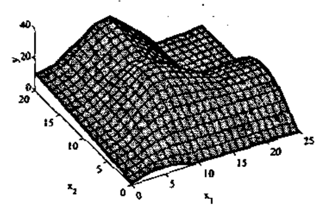


Fig. 1. Real surface for the example in Sugeno and Kang (1988).

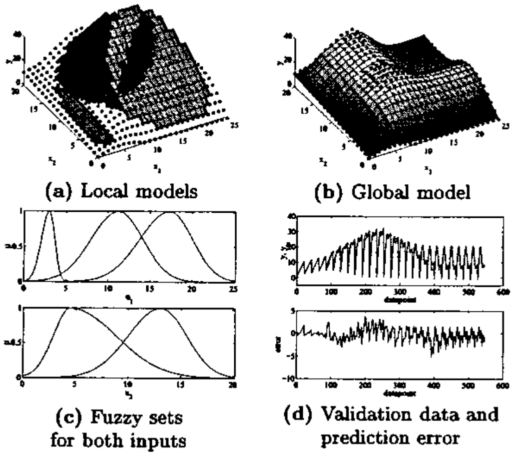
TABLE I
NON-DOMINATED SOLUTIONS ACCORDING TO (18) OBTAINED WITH THE MULTI-OBJECTIVE NEURO-EA.

M	L	MSE	S
1	2	125.391	0.0
2	4	25.606	0.348
3	5	2.187	0.349
4	5	1.017	0.349
5	5	0.910	0.350

We finally choose a compromise solution (4-rules fuzzy model). Figure 2 shows the local model, the surface generated by the model, fuzzy sets for each variable and the prediction error.

VIII. CONCLUSIONS AND FUTURE RESEARCH

In this paper we present a Pareto-based multi-objective neuro-evolutionary algorithm to obtain interpretable fuzzy models. Criteria such as accuracy, transparency and compactness have been propose and are taken into account in the optimization process. Some of these criteria have been partially incorporated into the EA by means of ad hoc techniques. Advantages of gaussian fuzzy sets arise with the possibility of training the RBF neural networks associated with the fuzzy models in order to obtain more accuracy. In addition, several new ideas to reduce computatinal load and improve the global search capabilities, have been incorporated in the evolutionary algorithm. 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. 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. The main difference between the proposed EAs and other approaches for fuzzy modeling is the reduced complexity because



(a) Local models (b) Global model

(c) Fuzzy sets for both inputs (d) Validation data and prediction error

Fig. 2. Accurate, transparent and compact fuzzy model for the example in Sugeno and Kang (1988).

we use a single EA for generating, and tuning the fuzzy model. Moreover, human intervention is only required at the end of the run in choosing one of the multiple non-dominated solutions. Results obtained are good in comparison with other iterative techniques reported in literature, with the advantage that the proposed techniques identifies a set of alternative solutions.

In future works, more complex fuzzy modeling test problems are going to be consider 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 and applications in the real world by means of research projects.

References

[1] C.A. Coello *List of references on multi-objective evolutionary optimization.* http://www.lania.mx/~ccoello/EMOO/EMOObib.html, 2000.

[2] O. Cordón and F. Herrera. *A proposal for improving the accuracy of linguistic modeling.* IEEE Transactions on Fuzzy Systems, 8(3):335-344, 2000.

[3] C.M. Fonseca and P.J. Fleming. *An overview of evolutionary algorithms in multiobjective optimization*, Evolutionary Computation, 3(1):1-16, 1995

[4] D.E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, 1989.

[5] A.F. Gómez-Skarmeta and F. Jiménez. *Fuzzy modeling with hybrid systems.* Fuzzy Sets and Systems, 104:199-208, 1999.

[6] J. Horn and N. Nafpliotis. *Multiobjective optimization using the niched pareto genetic algorithm.* IlliEAL Report No. 93005, July 1993.

[7] H.S. Hwang. *Control strategy for optimal compromise between trip time and energy comsumption in a high-speed railway.* IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 28(6):791-802, 1998.

[8] H. Ishibuchi, T. Murata, and I. Trksen. *Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems.* Fuzzy Sets and Systems, 89:135-150, 1997.

[9] H. Ishibuchi, T. Nakashima, and T. Murata. *Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems.* IEEE Transactions on Systems, Man, and Cubernetics - Part B: Cybernetics, 29/5):601-618, 1999.

[10] 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.

[11] F. Jiménez, J.L. Verdegay, A.F. Gómez-Skarmeta . *Evolutionary techniques for constrained multi-objective optimization problems.* Proceeding of the Genetic and Evolutionary Computation Conference (GECCO-99), Workshop on Multi-Criterion Optimization Using Evolutionary Methods, Orlando, Florida, USA, 115-116, 1999.

[12] F. Jiménez, A.F. Gmez-Skarmeta, H. Roubos, and R. Babuska. *Accurate, Transparent, and Compact Fuzzy Models for Function Approximation and Dynamic Modeling Through Multi-objective Evolutionary Optimization.* Lectures Notes in Computer Science, Evolutionary Multi-Criterion Optimization, 1993:653-667, 2001.

[13] Y. Jin. *Fuzzy Modeling of High-Dimensional Systems.* Transactions on Fuzzy Systems: Complexity Reduction and Interpretability Improvement. 8:212-221, 2000.

[14] T.A. Johansen and 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.

[15] J. Valente de Oliveira. *Semantic constraints for membership function optimization.* IEEE Transactions on Fuzzy Systems, 19(1):128-138, 1999

[16] H. Pomares and I. Rojas and J. Ortega and 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.

[17] J.A. Roubos and M. Setnes. *Compact and transparent fuzzy models and classifiers through iterative complexity reduction.* IEEE Transactions on Fuzzy Systems, 9(4):516-524, 2001

[18] M. Russo. *FuGeNeSys - a fuzzy genetic neural system for fuzzy modeling.* IEEE Transactions on Fuzzy Systems, 6(3):373-388, 1998.

[19] Setnes, M. and Babuška, R. and Verbruggen, H. B. *Rule-Based Modeling: Precision and Transparency.* IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications & Reviews, 28:165-169, 1998

[20] M. Setnes and J.A. Roubos. *Transparent Fuzzy Modeling using Fuzzy Clustering and GA's.* 18th International Conference of the North American Fuzzy Information Processing Society, 198-202, 1999.

[21] N. Srinivas and K. Deb. *Multiobjective optimization using nondominated sorting in genetic algorithms.* Evolutionary Computation, 2(3),221-248, 1995.

[22] M. Sugeno and G.T. Kang. *Structure identification of fuzzy model.* Fuzzy Sets and Systems, 28:15-23, 1998.

[23] 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.

[24] J. Valente de Oliveira. *Semantic constraints for membership function optimization.* IEEE Transactions on Fuzzy Systems, 19(1):128-138, 1999.

[25] 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.