

Learning concurrently data and rule bases of Mamdani fuzzy rule-based systems by exploiting a novel interpretability index

Michela Antonelli · Pietro Ducange ·
Beatrice Lazzerini · Francesco Marcelloni

Published online: 15 June 2010
© Springer-Verlag 2010

Abstract Interpretability of Mamdani fuzzy rule-based systems (MFRBSs) has been widely discussed in the last years, especially in the framework of multi-objective evolutionary fuzzy systems (MOEFSs). Here, multi-objective evolutionary algorithms (MOEAs) are applied to generate a set of MFRBSs with different trade-offs between interpretability and accuracy. In MOEFSs interpretability has often been measured in terms of complexity of the rule base and only recently partition integrity has also been considered. In this paper, we introduce a novel index for evaluating the interpretability of MFRBSs, which takes both the rule base complexity and the data base integrity into account. We discuss the use of this index in MOEFSs, which generate MFRBSs by concurrently learning the rule base, the linguistic partition granularities and the membership function parameters during the evolutionary process. The proposed approach has been experimented on six real world regression problems and the results have been compared with those obtained by applying the same MOEA, with only accuracy and complexity of the rule base as objectives. We show that our approach achieves the best trade-offs between interpretability and accuracy.

Keywords Accuracy-interpretability trade-off · Granularity learning · Interpretability index · Multi-objective evolutionary fuzzy systems · Piecewise linear transformation

1 Introduction

In the last years, multi-objective evolutionary fuzzy systems (MOEFSs) have captured a growing interest in the fuzzy community (Herrera 2008; Ishibuchi 2007). MOEFSs exploit multi-objective evolutionary algorithms (MOEAs) (Coello Coello and Lamont 2004; Deb 2001) to generate fuzzy rule-based systems (FRBSs) with good trade-offs between interpretability and accuracy. MOEAs are particularly suitable in this context since interpretability and accuracy are conflicting objectives: an increase in accuracy usually leads to a decrease in interpretability.

Among the different types of FRBSs, Mamdani fuzzy rule-based systems (MFRBSs) (Mamdani and Assilian 1975) have had a predominant role in MOEFSs, thanks to their feature of being completely defined in linguistic form and therefore particularly comprehensible to the users. MFRBSs consist of a completely linguistic rule base (RB), a data base (DB) containing the fuzzy sets associated with the linguistic terms used in the RB and a fuzzy logic inference engine.

As discussed in Alonso et al. (2008, 2009), Botta et al. (2009), Mencar et al. (2007) and Mencar and Fanelli (2008), since interpretability is a subjective concept, there is no general agreement on its formal definition and therefore there exists a real difficulty in formulating a measure of interpretability shared within the fuzzy community. Thus, researchers have focused their attention on discussing some factors which characterize interpretability

M. Antonelli (✉) · P. Ducange · B. Lazzerini · F. Marcelloni
Dipartimento di Ingegneria dell'Informazione:
Elettronica, Informatica, Telecomunicazioni,
University of Pisa, Via Diotisalvi 2, 56122 Pisa, Italy
e-mail: m1.antonelli@iet.unipi.it

P. Ducange
e-mail: p.ducange@iet.unipi.it

B. Lazzerini
e-mail: b.lazzerini@iet.unipi.it

F. Marcelloni
e-mail: f.marcelloni@iet.unipi.it

and on proposing some constraints which have to be satisfied for these factors (de Oliveira 1999; Guillaume 2001). An interesting survey on interpretability constraints has been recently published in Mencar and Fanelli (2008) with the objective of giving a homogeneous description of semantic and syntactic interpretability issues regarding both the RB and the DB. In Zhou and Gan (2008), a taxonomy of fuzzy model interpretability has been proposed by considering both low- and high-levels interpretability. Low-level interpretability is related to the semantic constraints that ensure fuzzy partition interpretability while high-level interpretability is associated with a set of criteria defined on the RB. Finally, in Alonso et al. (2009), the authors describe a conceptual framework for characterizing interpretability of fuzzy systems: the framework includes a global description of the FRBS structure, on the basis of the taxonomy and constraints discussed in Zhou and Gan (2008) and Mencar and Fanelli (2008), respectively, and a local explanation for understanding the FRBS behaviour. The local explanation considers a number of factors such as inference mechanisms, aggregation, conjunction and disjunction operators, defuzzification and rule type, which affect the FRBS behaviour.

Although a large amount of factors and constraints should be considered to assess the FRBS interpretability, a common approach has been to distinguish between interpretability of the RB, also known as complexity, and interpretability of fuzzy partitions, also known as integrity of the DB (de Oliveira 1999).

Complexity is usually defined in terms of simple measures, such as number of rules in the RB (RB simplicity) (Gacto et al. 2009, 2010) and number of linguistic terms in the antecedent of rules (simplicity of fuzzy rules) (Alcalá et al. 2009; Cococcioni et al. 2007; Ishibuchi 2007). Integrity depends on some properties, such as coverage, distinguishability and normality, which are fully satisfied by strong partitions and in particular by uniform partitions.

In this paper, we introduce a novel and simple interpretability index, which takes both the partition integrity and the RB complexity into consideration. First of all, we introduce a partition dissimilarity measure which computes how much the partitions generated in the evolutionary process are different from the uniform partition. Since uniform partitions are universally considered partitions with a high level of integrity, the more the measure is low, the more the partition is interpretable. Then, for each rule, we sum the dissimilarity measures of all the variables involved in the antecedent of the rule and of the output variable in the consequent. Finally, we define the index as the complement to 1 of the normalized average of these sums computed on all the rules in the RB. We use this index as one of the two objectives of an MOEA which generates a set of MFRBSs by concurrently learning the

rule base, the linguistic partition granularities and the membership function parameters during the evolutionary process.

In the first papers about MOEFSs, the optimization of the RB (rule learning or selection) has been performed considering a prefixed DB (Cococcioni et al. 2007; Ishibuchi and Nojima 2007; Ishibuchi and Yamamoto 2004) and, vice versa, the tuning or learning of the DB have been carried out using a prefixed RB (Botta et al. 2009). On the other hand, the ideal approach to MFRBS generation would be to learn concurrently DB and RB because they are strictly correlated. Some examples in the single-objective framework can be found in Cordon et al. (2001a, b) and Teng and Wang (2004). Some recent works discussed in Gacto et al. (2009, 2010), and in Alcalá et al. (2009) have proposed to exploit MOEAs to perform the learning of MF parameters concurrently with rule selection and rule learning, respectively. Further, in Antonelli et al. (2009), we have carried out the learning of the RB together with the learning of the partition granularities. Finally, in Antonelli et al. (2009) and Pulkkinen and Koivisto (2010) authors discuss two different multi-objective evolutionary approaches to learn concurrently the granularities of the fuzzy partitions, the MF parameters and a compact RB.

In this paper, we extend the approach discussed in Antonelli et al. (2009) by using the purposely defined interpretability index, thus generating a set of MFRBSs with different tradeoffs between accuracy and interpretability index. In particular, RB learning is achieved by exploiting the chromosome coding and mating operators introduced in Cococcioni et al. (2007). MF parameter learning is performed by using the piecewise linear transformation discussed in Klawonn (2006) and Pedrycz and Gomide (2007), which has already allowed to us to obtain a high modelling capability with a limited number of parameters in the MOEFS framework (Antonelli et al. 2009). Granularity learning is obtained by exploiting the concept of virtual RB and the appropriate mapping strategy discussed in Antonelli et al. (2009).

The proposed approach has been experimented on six real world regression problems and the results have been compared with those obtained by applying the same MOEA, with accuracy and complexity of the rule base as objectives. Pareto fronts obtained with the two approaches are almost similar in terms of accuracy, but solutions generated by using the interpretability index in place of the complexity measure are characterized on average by a higher partition integrity and generally a lower complexity.

The paper is organized as follows: in Sect. 2, we briefly describe the granularity and MF parameters learning. Section 3 discusses some interpretability issues and

introduces the interpretability index. In Sect. 4, we describe the two-objective evolutionary approach, including the chromosome coding, the fitness function and the genetic operators, used to generate the MFRBSs. Finally, Sect. 5 shows the experimental results and Sect. 6 draws some final conclusions.

2 Learning Mamdani fuzzy rule-based systems

2.1 Mamdani fuzzy rule-based systems

Let $\mathbf{X} = \{X_1, \dots, X_f, \dots, X_F\}$ be the set of input variables and X_{F+1} be the output variable. Let U_f , with $f = 1, \dots, F + 1$, be the universe of the f th variable. Let $P_f = \{A_{f,1}, \dots, A_{f,T_f}\}$ be a strong fuzzy partition (Ruspini 1969) of T_f fuzzy sets on variable X_f . The DB and the RB of an MRFBS are composed, respectively, of $F + 1$ partitions P_f and of M rules expressed as:

$$R_m : \text{IF } X_1 \text{ is } A_{1,j_{m,1}} \text{ and } \dots \text{ and } X_F \text{ is } A_{F,j_{m,F}} \\ \text{THEN } X_{F+1} \text{ is } A_{F+1,j_{m,F+1}} \quad (m = 1, \dots, M) \quad (1)$$

where $j_{m,f} \in [1, T_f]$ identifies the index of the fuzzy set (among the T_f fuzzy sets of partition P_f), which has been selected for X_f in rule R_m .

We adopt triangular fuzzy sets $A_{f,j}$ defined by the tuple $(a_{f,j}, b_{f,j}, c_{f,j})$, where $a_{f,j}$ and $c_{f,j}$ correspond to the left and right extremes of the support of $A_{f,j}$, and $b_{f,j}$ to the core. Further, since we use fuzzy strong partitions (Ruspini 1969), for $j = 2, \dots, T_f - 1$, $b_{f,j} = c_{f,j-1}$ and $b_{f,j} = a_{f,j+1}$. Finally, we assume that $a_{f,1} = b_{f,1}$ and $b_{f,T_f} = c_{f,T_f}$.

To take the “don’t care” condition into account, a new fuzzy set $A_{f,0}$ ($f = 1, \dots, F$) is added to all the F input partitions P_f . This fuzzy set is characterized by a membership function equal to 1 on the overall universe (Ishibuchi et al. 1997).

The terms $A_{f,0}$ allow generating rules which contain only a subset of the input variables. It follows that $j_{m,f} \in [0, T_f]$, $f = 1, \dots, F$, and $j_{m,F+1} \in [1, T_{F+1}]$. Thus, an MFRBS can be completely described by a matrix $J \in \mathbb{N}^{M \times (F+1)}$ (Cococcioni et al. 2007), where the generic element (m, f) indicates that fuzzy set $A_{f,j_{m,f}}$ has been selected for variable X_f in rule R_m . We adopt the product and the center of gravity method as AND logical operator and defuzzification method, respectively. Since we search for compact rule bases with a reduced number of rules and of conditions in the antecedents, it is possible that the number of distinct labels used for one variable in the rule base is lower than its granularity. Thus, it might occur that some input activates no rule and therefore results to be “covered” by no rule. In these cases, we adopt the inference strategy proposed in Alcalá et al. (2007), which determines

an output for a non-covered input based on the two closest rules to the input. The distance between the point and the rules is calculated considering the cores of the labels used in the rules.

Given a set of N input observations $\mathbf{x}_n = [x_{n,1}, \dots, x_{n,F}]$, with $x_{n,f} \in \mathfrak{R}$, and the set of the corresponding outputs $x_{n,F+1} \in \mathfrak{R}$, $n = 1, \dots, N$, we apply a specific MOEA, namely (2 + 2)M-PAES (Cococcioni et al. 2007), which produces a set of MFRBSs with different trade-offs between accuracy and interpretability by learning simultaneously the RB and the partition granularities and the MF parameters, which define the DB. To this aim, we employ the notion of *virtual partitions* we introduced in Antonelli et al. (2009). This notion derives from the following consideration: according to psychologists, to preserve interpretability, the number of linguistic terms per variable should be small (7 ± 2) due to a limit of human information processing capability (Alonso et al. 2008, 2009). Thus, we fix an upper bound T_{MAX} for the number of fuzzy sets. The virtual partitions are generated by uniformly partitioning each variable with T_{MAX} fuzzy sets.

During the evolutionary process, rule generation and MF parameter tuning are performed on these virtual partitions. The actual granularity is used only in the computation of the fitness. In practice, we generate RBs (*virtual RBs*) and tune MF parameters by using virtual partitions, but assess their quality using each time different “lens” depending on the actual number of fuzzy sets used to partition the single variables. Thus, we do not worry about the actual granularity in applying crossover and mutation operators. Obviously, to compute the fitness we have to transform the virtual MFRBS into the actual MFRBS and this process requires to define appropriate mapping strategies, both for the RB and for the MF parameters. In the following, we will describe these two strategies in detail.

2.2 Granularity learning

To map the virtual RB defined on virtual partitions into a concrete RB defined on variables partitioned with T_f fuzzy sets, we adopt the following mapping strategy. Let $\tilde{P}_f = \{\tilde{A}_{f,1}, \dots, \tilde{A}_{f,T_{\text{MAX}}}\}$ be a virtual partition for a generic variable X_f and “ X_f is $\tilde{A}_{f,h}$ ”, $h \in [0, T_{\text{MAX}}]$, be a generic fuzzy proposition defined in a rule of the virtual RB. Then, the proposition will be mapped to “ X_f is $\hat{A}_{f,s}$ ”, with $s \in [0, T_f]$, where $\hat{A}_{f,s}$ is the fuzzy set more similar to $\tilde{A}_{f,h}$ among the fuzzy sets in the uniform partition $\hat{P}_f = \{\hat{A}_{f,1}, \dots, \hat{A}_{f,T_f}\}$ defined on X_f with the actual granularity T_f . For the sake of simplicity, we have trivially considered as similarity measure the distance between the centroids of the two fuzzy sets. If there are two fuzzy sets in $\hat{P}_f = \{\hat{A}_{f,1}, \dots, \hat{A}_{f,T_f}\}$ with centroids at the same distance from the centroid of $\tilde{A}_{f,h}$, we choose randomly one of the two fuzzy sets.

Note that different rules of the virtual RB can be mapped to equal rules in the concrete RB. This occurs because distinct fuzzy sets defined on the partitions used in the virtual RB can be mapped to the same fuzzy set defined on the partitions used in the concrete RB. In the case of equal rules, only one of these rules is considered in the concrete RB. The original different rules are, however, maintained in the virtual RB. In the following, we denote with M^v and M^c the number of rules in the virtual and concrete RBs, respectively.

2.3 MF parameters learning

We approach the problem of learning the MF parameters by using a piecewise linear transformation (Klawonn 2006; Pedrycz and Gomide 2007). We define the transformation on the virtual partitions. Then, we exploit this transformation to tune the MFs defined on the actual granularity. The transformation is described in Fig. 1 for a generic variable X_f . Here, $\tilde{P}_f = \{\tilde{A}_{f,1}, \dots, \tilde{A}_{f,T_{MAX}}\}$ and $\check{P}_f = \{\check{A}_{f,1}, \dots, \check{A}_{f,T_{MAX}}\}$ denote the initial and the transformed virtual strong partitions, respectively. In the following, we assume that the interval ranges of the two partitions are identical. Further, we consider each variable normalized in $[0, 1]$. Finally, we adopt triangular MFs where $\tilde{b}_{f,1}, \dots, \tilde{b}_{f,T_{MAX}}$ and $\check{b}_{f,1}, \dots, \check{b}_{f,T_{MAX}}$ are the cores of $\tilde{A}_{f,1}, \dots, \tilde{A}_{f,T_{MAX}}$ and $\check{A}_{f,1}, \dots, \check{A}_{f,T_{MAX}}$, respectively. Piecewise linear transformation $t(x_f)$ is defined for $j = 2, \dots, T_f$ as:

$$t(x_f) = \frac{\tilde{b}_{f,j} - \tilde{b}_{f,j-1}}{\check{b}_{f,j} - \check{b}_{f,j-1}}(x_f - \check{b}_{f,j-1}) + \tilde{b}_{f,j-1}, \tag{2}$$

with $\check{b}_{f,j-1} \leq x_f < \check{b}_{f,j}, j = 2, \dots, T_{MAX}$.

Once fixed $T_{MAX}, \tilde{b}_{f,1}, \dots, \tilde{b}_{f,T_{MAX}}$ are fixed and therefore known. Further, $\check{b}_{f,1}$ and $\check{b}_{f,T_{MAX}}$ coincide with the extremes of the universe U_f of X_f . Thus, $t(x_f)$ depends on $T_{MAX} - 2$ parameters, that is, $t(x_f; \check{b}_{f,2}, \dots, \check{b}_{f,T_{MAX}-1})$. Once fixed $\check{b}_{f,2}, \dots, \check{b}_{f,T_{MAX}-1}$, the partition $\check{P}_f = \{\check{A}_{f,1}, \dots, \check{A}_{f,T_{MAX}}\}$ can be obtained simply by transforming the three points $(\tilde{a}_{f,j}, \tilde{b}_{f,j}, \tilde{c}_{f,j})$, which describe the generic fuzzy set $\tilde{A}_{f,h}$, into $(\check{a}_{f,j}, \check{b}_{f,j}, \check{c}_{f,j})$ applying $t^{-1}(x_f)$ (Fig. 1). We observe that the piecewise linear transformation ensures that also \check{P}_f is a strong partition.

Once a granularity, say T_f , is computed by the evolutionary process, we generate the uniform partition $\hat{P}_f = \{\hat{A}_{f,1}, \dots, \hat{A}_{f,T_f}\}$ on X_f by using T_f fuzzy sets. Then, we transform this partition by exploiting the piecewise linear transformation defined on the virtual partitions. In

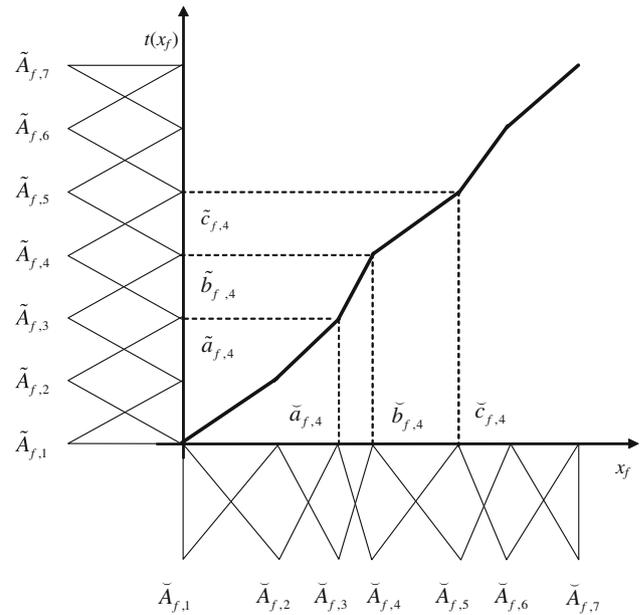


Fig. 1 An example of piecewise linear transformation with $T_{MAX} = 7$

practice, in order to maintain the original shape of the MFs, for $j = 2, \dots, T_f - 1$, we apply t^{-1} to $x_f = \hat{a}_{f,j}, x_f = \hat{b}_{f,j}$ and $x_f = \hat{c}_{f,j}$, where the three points $(\hat{a}_{f,j}, \hat{b}_{f,j}, \hat{c}_{f,j})$ describe the generic fuzzy set $\hat{A}_{f,j}$ in the uniform partition of T_f fuzzy sets. We recall that also the actual transformed partition $P_f = \{A_{f,1}, \dots, A_{f,T_f}\}$ is a strong partition.

Figure 2 shows an example of this transformation for granularity $T_f = 5$ by using the piecewise linear transformation in Fig. 1.

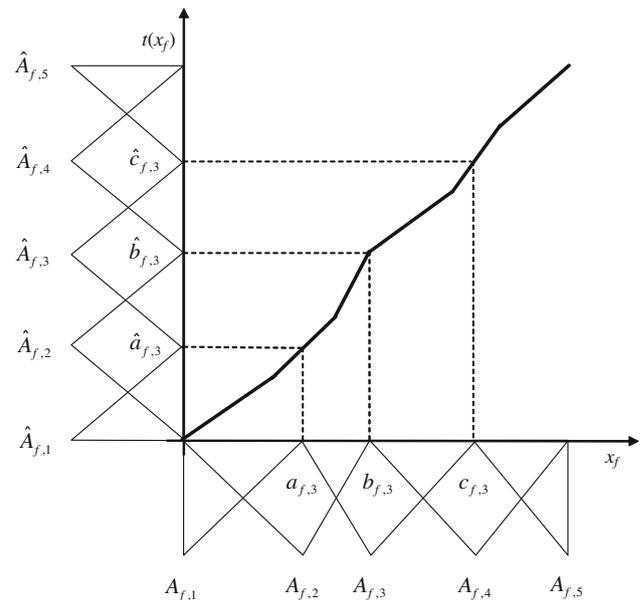


Fig. 2 An example of piecewise linear transformation with $T_f = 5$

3 The problem of the interpretability

3.1 Interpretability: accuracy trade-off

Several methods have been proposed in the literature to generate KBs of MFRBSs from available information (typically, input–output samples) (Casillas et al. 2002; González and Pérez 1999; Wang and Mendel 1992). Generally, these methods aim to maximize the accuracy. Thus, the resulting KBs are usually characterized by a high number of rules and by linguistic fuzzy partitions with a low level of comprehensibility, thus losing that feature which has made MFRBSs preferable to other approaches in real applications, namely interpretability. Only in the last decade, researchers have begun to propose methods to generate MFRBSs taking not only accuracy, but also interpretability into consideration (Casillas et al. 2003).

The interpretability of an MFRBS relies mainly on the simplicity of the fuzzy RB and on the integrity of the fuzzy partitions (Ishibuchi and Yamamoto 2004). To ensure the RB simplicity, both the number of fuzzy rules and the number of antecedent conditions should be maintained low. To this aim, in several works authors have introduced a complexity measure which is optimized together with the accuracy during the evolutionary process. The most popular measures adopted to this aim have been the total number of rules (Alcalá et al. 2007; Gacto et al. 2009, 2010) and the total sum of the conditions in the antecedents of the rules (Alcalá et al. 2009; Antonelli et al. 2009a, b; Cococcioni et al. 2007; Ducange et al. 2009; Ishibuchi and Nojima 2007; Ishibuchi and Yamamoto 2004; Pulkkinen and Koivisto 2010). These measures have been also adopted as two different objectives of the evolutionary process in Ishibuchi and Nojima (2007) and Ishibuchi and Yamamoto (2004).

Integrity of the partitions can be defined in several ways. Here, on the basis of the considerations and discussions made in Alonso et al. (2009), de Oliveira (1999), Guillaume (2001), Mencar and Fanelli (2008) and Zhou and Gan (2008), we state that a fuzzy partition is characterized by a high integrity if it satisfies the following properties:

1. The partition should have a reasonable number of fuzzy sets;
2. The fuzzy sets in the partition should all be normal, i.e., for each fuzzy set there exists at least one point with membership degree equal to 1;
3. Each couple of fuzzy sets should be distinguishable enough, so that there are no two fuzzy sets that represent pretty much the same concept;
4. The overall universe of discourse should be strictly covered, i.e., each point of the universe should belong to at least a fuzzy set with a membership degree over a given reasonable threshold.

As regards property 1, as already discussed in Sect. 2.1, we limit the maximum number of linguistic terms per variable to 7 following psychologists' suggestions derived from considerations on limits of human information processing capability (Alonso et al. 2008). Further, during the evolutionary process the granularity may decrease thanks to the granularity learning.

As regards properties 2, 3 and 4, these properties are fully satisfied by strong partitions and in particular by uniform partitions. On the other hand, uniform partitions are considered to be the most intuitive interpretable partitions. Indeed, each book on fuzzy logic introduces fuzzy partitions by adopting uniform partitions. Thus, we decided to adopt uniform partitions as initial partitions. Obviously, we generate these uniform partitions by using the number of fuzzy sets determined by the granularity.

Often, the MF adaptation process generates partitions which are quite far from the uniform partitions and consequently less interpretable: the more different from being uniform the partition is, the less interpretable it is. Typically, partition integrity has been ensured either by using uniform partitions (Antonelli et al. 2009; Cococcioni et al. 2007; Ducange et al. 2009; Ishibuchi and Nojima 2007; Ishibuchi and Yamamoto 2004) or constraining the variation range of the MF parameters during the evolutionary process (Alcalá et al. 2007a, b, 2009; Antonelli et al. 2009; Gacto et al. 2009, 2010). In the last years, some interesting indices have been proposed in order to control the integrity of the partition during the multi-objective evolutionary process. As an example, in Botta et al. (2009) authors perform a multi-objective evolutionary context adaptation of a predefined RB by concurrently optimizing the accuracy and an integrity index suitable for the specific context operators. Recently, Gacto et al. (2010) have proposed a three-objective evolutionary approach aimed at concurrently selecting rules from an initial rule base and tuning the MF parameters. In order to control the partition integrity, authors introduce an index which considers the MF centroids displacement, the MF lateral amplitude rate and the MF area similarity.

3.2 The interpretability index

In this work, to increase integrity, we force partitions to tend towards the uniform partitions. If we analyze how the MF adaptation process is performed by the piecewise linear transformation, we can realize that the partitions generated during the evolutionary process are similar to the uniform partitions when the piecewise linear transformation tends to be a linear transformation. Actually, the farther from being a line the piecewise linear transformation is, the less similar to the initial partition it is. The piecewise linear

transformation tends to be a linear transformation when $\tilde{b}_{f,j}$ tends to be equal to $\tilde{b}_{f,j}$, $j = 2, \dots, T_{\text{MAX}} - 1$. Thus, to control the linearity of the piecewise linear transformation in the evolutionary learning of the MF parameters, we introduce, for each variable X_f , the following dissimilarity measure:

$$d_f = \frac{2}{T_{\text{MAX}} - 2} \sum_{j=2}^{T_{\text{MAX}}-1} \left| \tilde{b}_{f,j} - \tilde{b}_{f,j} \right|. \tag{3}$$

The highest level of partition integrity occurs when, $\forall j \in \{2, \dots, T_{\text{MAX}} - 1\}$, $\tilde{b}_{f,j} = \tilde{b}_{f,j}$. In this case, $d_f = 0$ and no transformation is performed (actually the piecewise linear transformation is a line). The lowest level of partition integrity occurs when all $\tilde{b}_{f,j}$, $j = 2, \dots, T_{\text{MAX}} - 1$, coincide with one of the extremes of the universe. In this case, $\sum_{j=2}^{T_{\text{MAX}}-1} \left| \tilde{b}_{f,j} - \tilde{b}_{f,j} \right| = \frac{T_{\text{MAX}}-2}{2}$ and thus $d_f = 1$. It follows that $0 \leq d_f \leq 1$, with $d_f = 0$ and $d_f = 1$ corresponding to the highest and lowest partition integrity values, respectively.

Since the piecewise linear transformation only moves the cores and the extremes of the fuzzy sets without deforming their shapes, d_f can be considered a suitable measure for evaluating how much a partition generated by the MF parameter learning is different from the initial partition.

In order to take both the DB integrity and the RB complexity into account, we introduce a purposely defined interpretability index.

First of all, we compute:

$$\text{DC} = \sum_m^{M^v} \sum_f^{F+1} (1 + d_f) \cdot u(j_{m,f}) \tag{4}$$

where $u(j_{m,f}) = \begin{cases} 1 & \text{if } j_{m,f} > 0 \\ 0 & \text{if } j_{m,f} = 0 \end{cases}$. In other words, $u(j_{m,f}) = 1$ only if the index $j_{m,f}$ identifies a fuzzy set different from the *don't care* fuzzy set. We recall that $j_{m,f}$ is the index of the fuzzy set defined on virtual partition \tilde{P}_f which has been selected for X_f in the virtual rule R_m . Thus, DC takes concurrently the number of virtual rules, the number of antecedent conditions and the dissimilarity d_f into account. Obviously, a decrease in the number of rules and antecedent conditions in the virtual RB implies a decrease in the number of rules and antecedent conditions in the concrete RB. The value of DC increases with the increasing of the number of rules and the number of antecedent conditions in the rules, and with the increasing of the values of dissimilarity d_f between the actual and the initial partitions for each linguistic variable X_f . Thus, the higher the value of DC, the lower the MFRBS interpretability.

We note that, since the RB cannot be composed of rules with no condition in the antecedents, DC can never be

equal to zero. From simple mathematical considerations, we derive that $2M_{\text{min}}^v \leq \text{DC} \leq 2M_{\text{max}}^v(F + 1)$, where M_{min}^v and M_{max}^v are the possible minimum and maximum numbers of rules in the virtual RB. Based on DC, we introduce the following interpretability index I to globally evaluate the interpretability of a knowledge base of an MFRBS:

$$I = 1 - \frac{\text{DC} - 2M_{\text{min}}^v}{2M_{\text{max}}^v(F + 1) - 2M_{\text{min}}^v}. \tag{5}$$

Index I varies from 0 (minimum level of interpretability) to 1 (maximum level of interpretability). The maximum value corresponds to an RB composed by the minimum number of rules with only one condition in the antecedent and to a DB with uniform partitions for each linguistic variable.

To increase the value of index I and to improve accuracy are often conflicting objectives. Thus, we approach the generation of MFRBSs by using a two-objective evolutionary algorithm, where the two objectives are the MSE computed as in Antonelli et al. (2009) and the interpretability index I defined in (5), respectively. In particular, the MSE is calculated as:

$$\text{MSE} = \frac{1}{2|E|} \sum_{l=1}^{|E|} (F(x^l) - y^l) \tag{6}$$

where $|E|$ is the size of the dataset, $F(x^l)$ is the output obtained from the MFRBS when the l th input pattern is considered, and y^l is the desired output.

4 The two-objective evolutionary approach

We adopt the (2 + 2)M-PAES proposed in Cococcioni et al. (2007) as MOEA for generating a set of MFRBSs with different trade-offs between MSE and I . In the following, we briefly describe the chromosome coding, the genetic operators and the evolutionary strategy used in the MOEA.

4.1 Chromosome coding

Each solution is codified by a chromosome C composed of three parts (C_1, C_2, C_3), which define the virtual RB, and the granularities and the piecewise linear transformations of all the variables, respectively. In particular, C_1 encodes the virtual RB by considering that each variable X_f is uniformly partitioned by using T_{MAX} fuzzy sets.

As described in Antonelli et al. (2009), C_1 is composed of $M^v(F + 1)$ natural numbers where M^v is the number of rules currently present in the virtual RB. The RB (defined as *concrete RB*) used to compute the MSE is obtained by means of the RB mapping strategy using the actual

granularities fixed by C_2 . C_2 is a vector containing $F + 1$ natural numbers: the f th element of the vector contains the number $T_f \in [2, T_{\text{MAX}}]$ of fuzzy sets which partition the linguistic variable X_f . C_3 is a vector containing $F + 1$ vectors of $T_{\text{MAX}} - 2$ real numbers: the f th vector contains the $[\tilde{b}_{f,2}, \dots, \tilde{b}_{f,T_{\text{MAX}}-1}]$ real values which define the piecewise linear transformation for the f th linguistic variable.

4.2 Genetic operators

In order to generate the offspring populations, we exploit both crossover and mutation. We apply separately the one-point crossover to C_1 and C_2 and the BLX- α crossover, with $\alpha = 0.5$, to C_3 . To constrain the search space, we fix the possible minimum and maximum numbers of rules to M_{min}^V and M_{max}^V , respectively.

Let s_1 and s_2 be two selected parent chromosomes. The common gene for C_1 is selected by extracting randomly a number in $[1, \rho_{\text{min}} - 1]$, where ρ_{min} is the minimum number of rules in s_1 and s_2 . The crossover point is always chosen between two rules and not within a rule. When we apply the one-point crossover to the RB part, we can generate an MFRBS with one or more pairs of equal rules. In this case, we simply eliminate one of the rules from each pair. This allows us to reduce the total number of rules. The common gene for C_2 is extracted randomly in $[1, F]$.

As regards mutation, we apply two mutation operators for C_1 . The first operator adds γ rules to the virtual RB, where γ is randomly chosen in $[1, \gamma_{\text{max}}]$. The upper bound γ_{max} is fixed by the user. The second mutation operator randomly changes δ elements of the matrix J associated with the virtual RB. The number δ is randomly generated in $[1, \delta_{\text{max}}]$. The upper bound δ_{max} is fixed by the user. For each element to be modified, a number is randomly generated in $[0, T_{\text{MAX}}]$.

The mutation applied to C_2 randomly chooses a gene $f \in [1, F + 1]$ and changes the value of this gene by randomly adding or subtracting 1. If the new value is lower than 2 or larger than T_{MAX} , then the mutation is not applied.

The mutation applied to C_3 first chooses randomly a variable $f \in [1, F + 1]$, then extracts a random value $j \in [2, T_{\text{MAX}} - 1]$ and changes the value of $\tilde{b}_{f,j}$ to a random value in $[\tilde{b}_{f,j-1}, \tilde{b}_{f,j+1}]$.

We experimentally verified that these mating operators, together with the appropriate probabilities, ensure a good balancing between exploration and exploitation, thus allowing the MOEA described in the next subsection to create good approximations of the Pareto fronts.

We would like to highlight that the number of rules can change in the virtual RB. Indeed, the crossover operator can decrease the number of rules in the offspring when the

offspring contains two equal rules inherited from the two parents, respectively. In this case, one of the rules is removed from the virtual RB. Further, the first mutation operator adds rules to the virtual RB. On the other hand, the second mutation operator can decrease the number of rules since it can make two rules equal by randomly modifying the selected genes. We would like to remark that rule reduction performed by the crossover operator and the second mutation operator occurs also when the number of input variables is high. Indeed, we have to consider that, during the evolutionary process, some rules are identified as good rules and therefore tend to be included in several solutions. Thus, also in the case of high number of input variables, when we apply the genetic operators we can generate MFRBSs with equal rules and therefore obtain rule reduction.

4.3 The two-objective evolutionary algorithm

We adopted the $(2 + 2)$ M-PAES proposed in Cococcioni et al. (2007). Unlike classical $(2 + 2)$ PAES (Knowles and Corne 2002), which uses only mutation to generate new candidate solutions, $(2 + 2)$ M-PAES exploits both crossover and mutation. Further, in $(2 + 2)$ M-PAES, current solutions are randomly extracted at each iteration rather than maintained until they are not replaced by solutions with particular characteristics.

At the beginning, we generate two solutions s_1 and s_2 and the genes of C_1 , C_2 and C_3 are randomly generated. At each iteration, the application of crossover and mutation operators produces two new candidate solutions from the current solutions s_1 and s_2 . First, we separately apply the three crossover operators with probabilities equal to P_{c1} , P_{c2} and P_{c3} , respectively. Then, we apply the mutation operators to each part of the chromosome. As regards C_1 , if the crossover is not applied, the mutation is always applied; otherwise the mutation is applied with probability P_{m1} . When the mutation is applied, the probabilities of applying the two mutation operators are P_{add} and $1 - P_{\text{add}}$, respectively. The probabilities of applying the mutation to C_2 and C_3 are P_{m2} and P_{m3} , respectively. When the mutation is applied to C_2 the granularity is increased with a probability P_{inc} , otherwise the granularity is decreased.

The candidate solutions are added to the archive only if they are dominated by no solution contained in the archive; possible solutions in the archive dominated by the candidate solutions are removed. Typically, the size of the archive is fixed at the beginning of the execution of the $(2 + 2)$ M-PAES. In this case, when the archive is full and a new solution z has to be added to the archive, if z dominates no solution in the archive, then we insert z into the archive and remove the solution (possibly z itself) that belongs to the region with the highest crowding degree

(Knowles and Corne 2002). If the region contains more than one solution, then the solution to be removed is randomly chosen.

5 Experimental results

5.1 Experimental setup

We tested our method on six regression problem datasets (available at <http://sci2s.ugr.es/keel/datasets.php>). Table 1 summarizes the main characteristics of these datasets. We performed a fivefold cross-validation, using each fold six times with different seeds for the random function generator (30 trials in total).

To assess the advantages of exploiting our interpretability index, we compared the results achieved by our approach with the results obtained by applying the (2 + 2)M-PAES to minimize only the complexity of the concrete RB, together with the MSE, without considering the partition integrity.

The complexity of the concrete RB is computed as:

$$\text{COMP} = \sum_m^{M^c} \sum_f^{F+1} u(j_{m,f}) \quad (7)$$

where $u(j_{m,f}) = \begin{cases} 1 & \text{if } j_{m,f} > 0 \\ 0 & \text{if } j_{m,f} = 0 \end{cases}$. COMP represents the number of propositions used in the antecedents of the rules contained in the concrete RB.

We denote these two approaches as (2 + 2)M-PAES(I) and (2 + 2)M-PAES(C), respectively. In the experiments, we executed the two MOEAs using the parameters shown in Table 2.

In Sect. 5.2, we discuss the results of the MFRBS learning in the MSE-Interpretability plane. With the aim of performing the comparison statistically and not on a single trial, we resort to the concept of average Pareto fronts used in our previous works (Antonelli et al. 2009a, b). First, for each of the 30 trials, we compute the Pareto front approximations for the two MOEAs and order the solutions in these approximations for increasing MSE values. Since

Table 1 Datasets used in the experiments

Datasets	Number of patterns	Number of input variables
Electrical Maintenance (ELE)	1,056	4
Weather Ankara (WA)	1,609	9
Weather Izmir (WI)	1,461	9
Auto-MPG (MPG6)	398	5
Treasury (TR)	1,049	15
Stock (STP)	950	9

Table 2 Values of the parameters used in the experiments

Archive size	64
Total number of evaluations	300,000
Minimum number of virtual rules M_{\min}^V	5
Maximum number of virtual rules M_{\max}^V	50
Crossover probability P_{c1}	0.3
Crossover probability P_{c2}	0.5
Crossover probability P_{c3}	0.5
Mutation probability P_{m1}	0.1
Probability P_{add} of mutation for adding rule	0.75
Mutation probability P_{m2}	0.5
Probability P_{inc} of mutation for increasing granularity	0.85
Mutation probability P_{m3}	0.3
γ_{\max} and δ_{\max}	5

the number of solutions varies from one Pareto front approximation to another, we identify the lowest number of solutions contained in a Pareto front approximation. Then, we retain only the solutions (at most, twenty) with the lowest MSEs for each Pareto front approximation. Finally, we compute the average values, on the 30 Pareto front approximations, of the MSE and of the interpretability index for these solutions. The choice of considering only at most the twenty solutions with the lowest MSEs has been motivated by the observation that the other solutions are in general characterized by quite high MSEs which make these solutions impractical. The number of solutions contained in the average Pareto front is a good measure of the easiness or difficulty met by (2 + 2)M-PAES(I) and (2 + 2)M-PAES(C) in exploring the search space and therefore in generating MFRBSs with different trade-offs.

We also perform a statistical analysis by using the two-sample Kolmogorov–Smirnov test (Massey 1951). This test allows verifying whether there exist statistical differences, in terms of accuracy, between the solutions generated by the two versions of the (2 + 2)M-PAES. The two-sample Kolmogorov–Smirnov test is a non-parametric test which assumes no particular data probability distributions. The test compares the distributions of the values of the MSEs generated by both the versions of (2 + 2)M-PAES. The null hypothesis is accepted if the two distributions are from the same continuous distribution. The alternative hypothesis is that they are from different continuous distributions. We applied the test to three interesting points in the average Pareto fronts: the first (the most accurate), the median and the last (the least accurate) points. We will refer to these average values as FIRST, MEDIAN and LAST, respectively.

The interpretability index introduced in Sect. 3 takes both the RB complexity and the DB integrity into account, thus allowing us to concurrently optimize both aspects of

the interpretability of the global KB. Actually, by only analyzing the interpretability index in the experimental results, it is not easy to directly appreciate its effects in the optimization of the RB complexity and DB integrity. Thus, to make a reliable comparison between (2 + 2)M-PAES(I) and (2 + 2)M-PAES(C) and therefore to appreciate the effects of the use of the interpretability index, in Sect. 5.3 we show and discuss the results in terms of complexity of the concrete RB and in terms of two measures introduced to evaluate the integrities of the concrete and virtual partitions, respectively.

As regards the integrity measure of the concrete partition, we first introduce the dissimilarity d_f^c computed on the concrete partitions as follows:

$$d_f^c = \begin{cases} 0 & \text{if } T_f = 2 \\ \frac{2}{T_f - 2} \sum_{j=2}^{T_f - 1} |b_{f,j} - \hat{b}_{f,j}| & \text{if } T_f > 2 \end{cases} \quad (8)$$

Then, we compute the following average concrete dissimilarity D^c defined as $D^c = \frac{1}{F+1} \sum_{f=1}^{F+1} d_f^c$. D^c expresses how much on average the transformed concrete partitions differ from the uniform concrete partitions, thus providing a measure of the integrity of the concrete partitions: the higher the value of D^c , the lower the partition integrity. As regards the integrity measure of virtual partitions, we calculate the average virtual dissimilarity D^v as $D^v = \frac{1}{F+1} \sum_{f=1}^{F+1} d_f^v$. We recall that

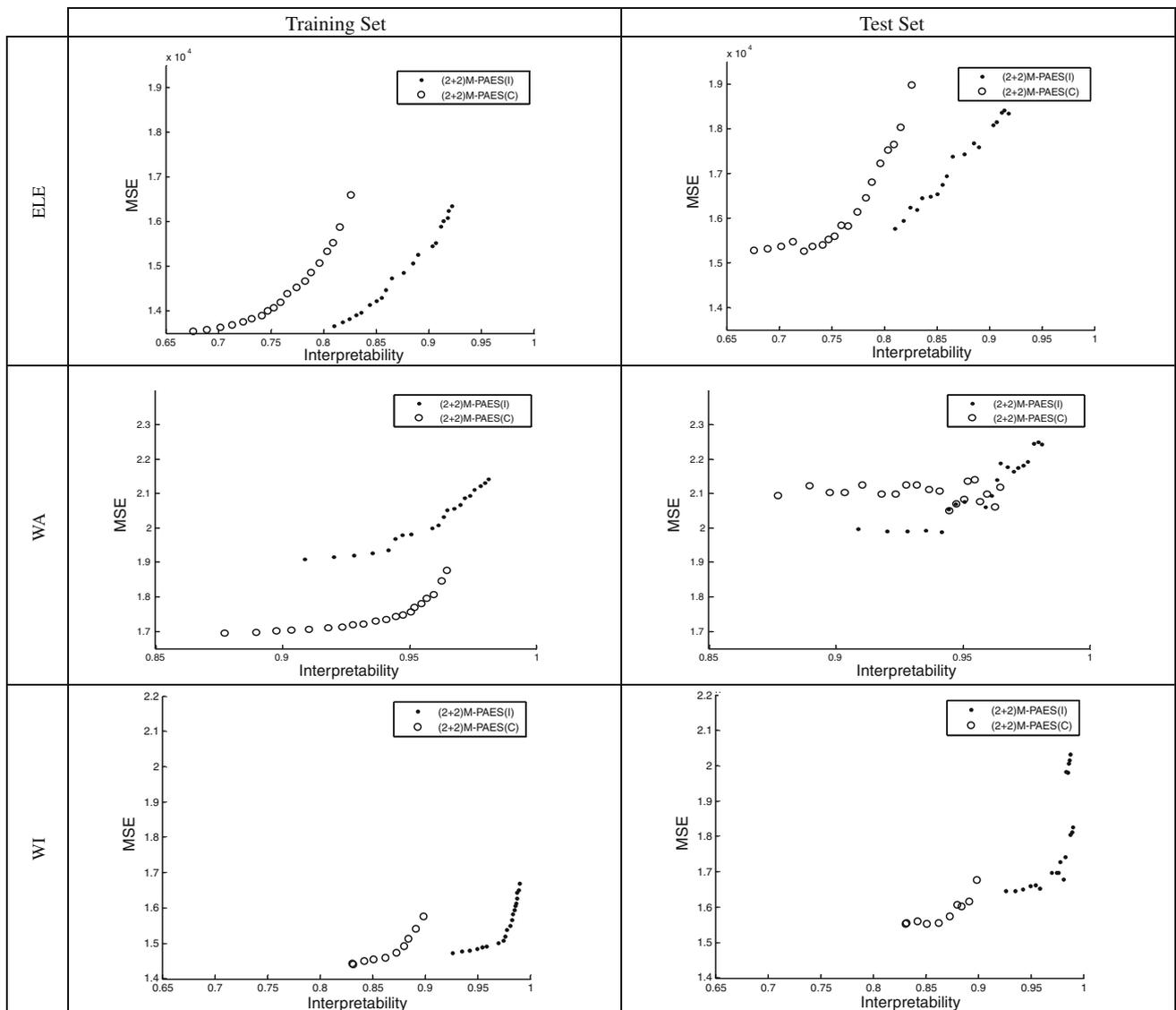


Fig. 3 Average Pareto fronts plotted on the Interpretability-MSE plane for the training and test sets

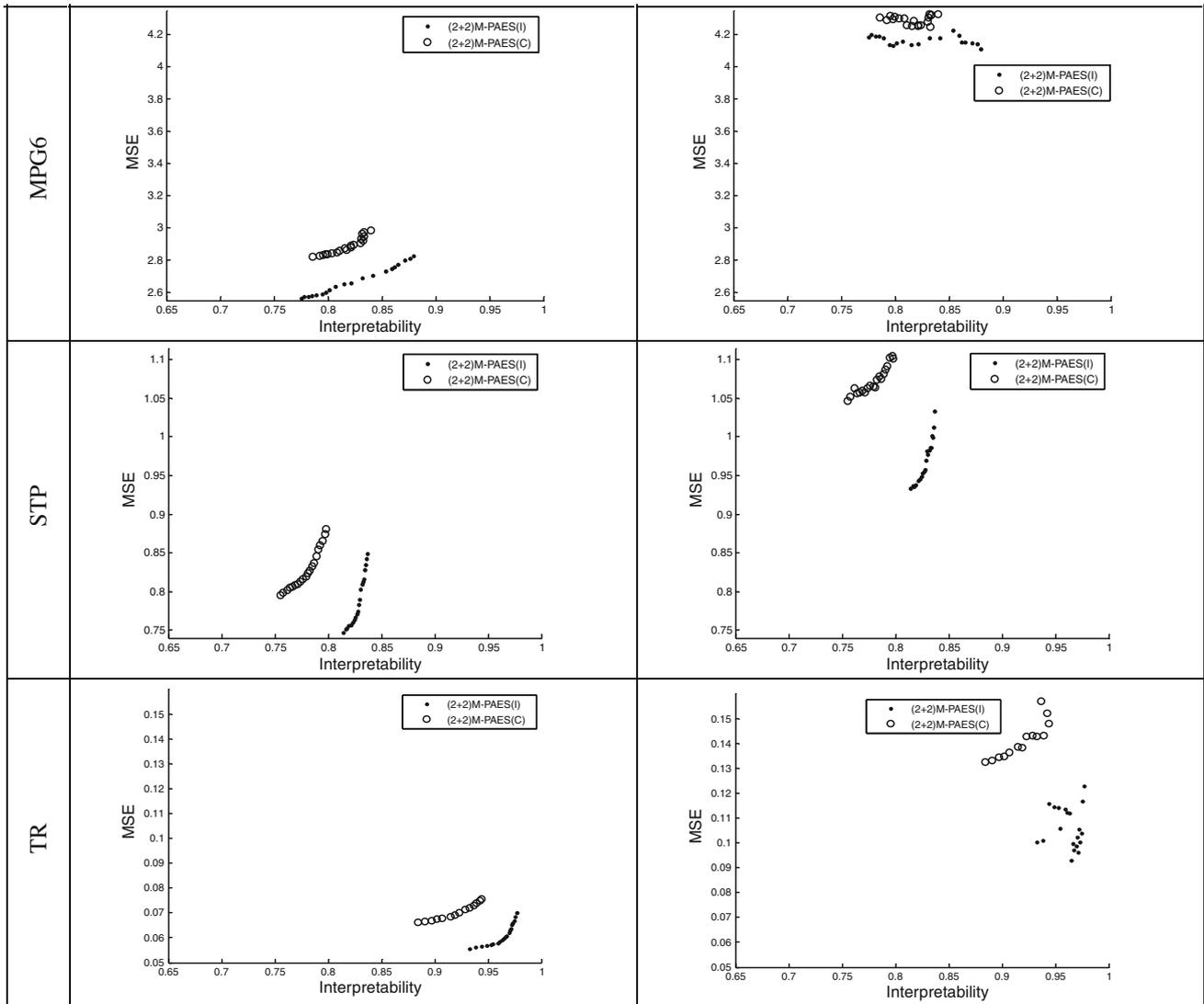


Fig. 3 continued

Table 3 Average MSEs on training and test sets and interpretability index for the FIRST solution

		$\overline{\text{MSE}}_{\text{TR}}(\sigma_{\text{TR}})$	$k-s_{\text{TR}}$	$\overline{\text{MSE}}_{\text{TS}}(\sigma_{\text{TS}})$	$k-s_{\text{TS}}$	$\bar{I}(\sigma_I)$
ELE	(2 + 2)M-PAES(I)	13,660.2 (1,851.5)	=	15,768.6 (3,239.9)	=	0.810 (0.131)
	(2 + 2)M-PAES(C)	13,539.8 (3,764.7)	*	15,278.8 (4,129.4)	*	0.676 (0.090)
WA	(2 + 2)M-PAES(I)	1.911 (0.381)	+	1.997 (0.298)	*	0.909 (0.059)
	(2 + 2)M-PAES(C)	1.694 (0.489)	*	2.094 (0.973)	=	0.877 (0.032)
WI	(2 + 2)M-PAES(I)	1.474 (0.343)	=	1.647 (0.343)	=	0.926 (0.107)
	(2 + 2)M-PAES(C)	1.441 (0.276)	*	1.556 (0.243)	*	0.832 (0.087)
MPG6	(2 + 2)M-PAES(I)	2.565 (0.341)	*	4.185 (1.352)	*	0.776 (0.027)
	(2 + 2)M-PAES(C)	2.820 (0.428)	=	4.304 (0.1365)	=	0.786 (0.045)
STP	(2 + 2)M-PAES(I)	0.748 (0.098)	*	0.934 (0.175)	*	0.814 (0.019)
	(2 + 2)M-PAES(C)	0.795 (0.225)	=	1.046 (0.309)	=	0.755 (0.019)
TR	(2 + 2)M-PAES(I)	0.056 (0.020)	*	0.100 (0.097)	*	0.933 (0.039)
	(2 + 2)M-PAES(C)	0.066 (0.025)	=	0.132 (0.142)	=	0.884 (0.052)

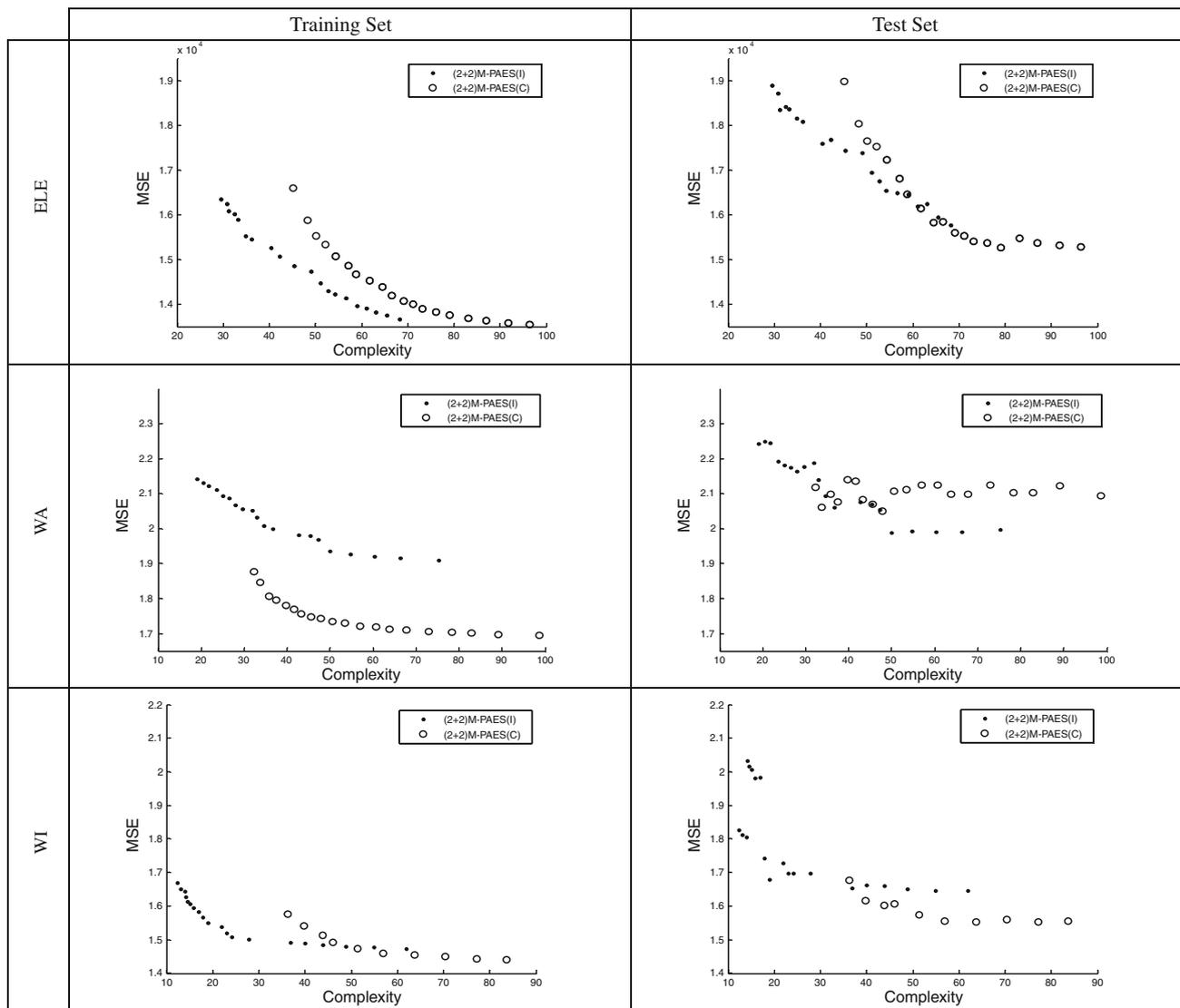


Fig. 4 Average Pareto fronts plotted on the Complexity-MSE plane for the training and test sets

$d_f = \frac{2}{T_{MAX}-2} \sum_{j=2}^{T_{MAX}-1} |\tilde{b}_{f,j} - \tilde{b}_{f,j}|$. The average virtual dissimilarity D^v has the same meaning of D^c but is associated with the transformed virtual partition.

Once extracted the complexity and the average concrete dissimilarity, in Sect. 5.3 we also plot the average Pareto fronts, achieved by the two algorithms on the training and test sets, on the complexity-MSE and D^c -MSE planes. Complexity and number of rules M^c are computed on the concrete RB. In the following, we will discuss the results obtained on the six datasets.

5.2 Analysis of the results on the Interpretability-MSE plane

In this section we report the results obtained by (2 + 2)M-PAES(I) and (2 + 2)M-PAES(C) on the Interpretability-

MSE planes. Figure 3 shows the average Pareto fronts of both the algorithms on the training and test sets. As expected, for all the datasets but WA, the average Pareto fronts generated by (2 + 2)M-PAES(I) dominate the ones generated by (2 + 2)M-PAES(C). For all datasets, the value of the interpretability index decreases (indeed, the complexity of the rule bases and the dissimilarities d_f increase) with the increase of the accuracy.

To statistically compare the results of the two algorithms, in Table 3 we report for the FIRST solution the averages and the standard deviations of the MSEs on training and test sets ($\overline{MSE}_{TR}(\sigma_{TR})$ and $\overline{MSE}_{TS}(\sigma_{TS})$, respectively), and the averages and the standard deviations of the interpretability index $\bar{I}(\sigma_I)$. In “Appendix”, we report the same results for both the MEDIAN and LAST solutions (Tables 6, 7). In order to assess whether the differences between the solutions are statistically

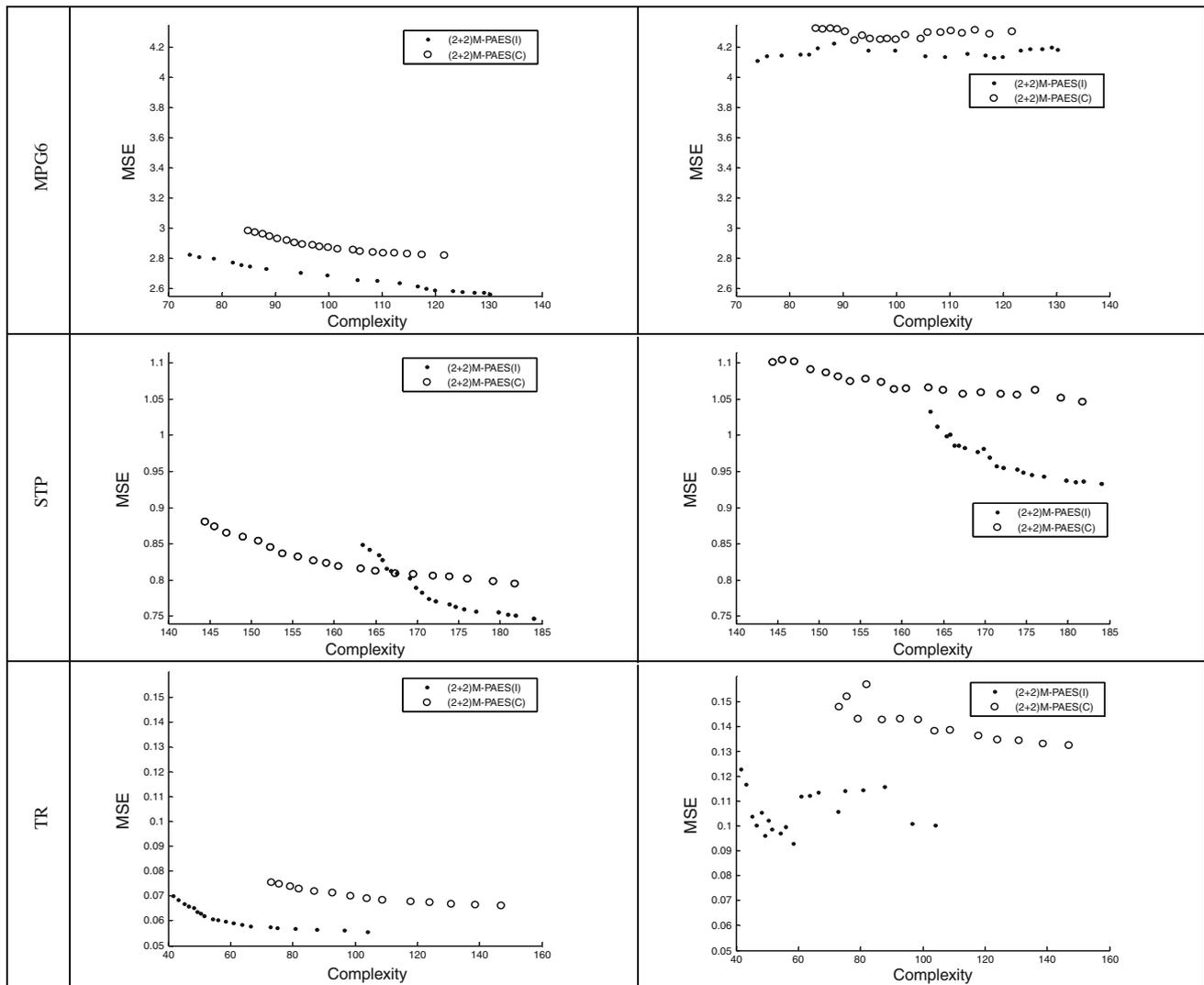


Fig. 4 continued

significant, we also show the results of the Kolmogorov–Smirnov test (column $k-s_{TR}$ and $k-s_{TS}$ for the training and test sets, respectively).

The interpretation of the $k-s$ columns is the following:

- * represents the best result (in bold in the MSE columns);
- + means that the best result has better performance than that of the corresponding row;
- = means that the best result has performance comparable to that of the corresponding row

By analyzing the results of the Kolmogorov–Smirnov test performed on the three representative points of the average Pareto fronts, we observe that the MFRBSs generated by the two approaches are statistically equivalent in terms of both \overline{MSE}_{TR} and \overline{MSE}_{TS} for all datasets except for the \overline{MSE}_{TR} on WA dataset, even though the average Pareto

fronts provided by (2 + 2)M-PAES(I) are characterized by a higher value of \bar{l} . Thus, we can conclude that to take both complexity and integrity into account during the evolutionary process leads to increase the interpretability of the generated MFRBSs without affecting their accuracy.

5.3 Analysis of the results on the complexity-MSE and D^c -MSE planes

Figures 4 and 5 show the average Pareto fronts achieved by (2 + 2)M-PAES(I) and (2 + 2)M-PAES(C) on the training and test sets, plotted on the complexity-MSE and D^c -MSE planes, respectively.

By analyzing Figs. 4 and 5, we can observe that (2 + 2)M-PAES(I) on average generates MFRBSs with lower complexity values than (2 + 2)M-PAES(C). Further, the projections of the average Pareto fronts generated by

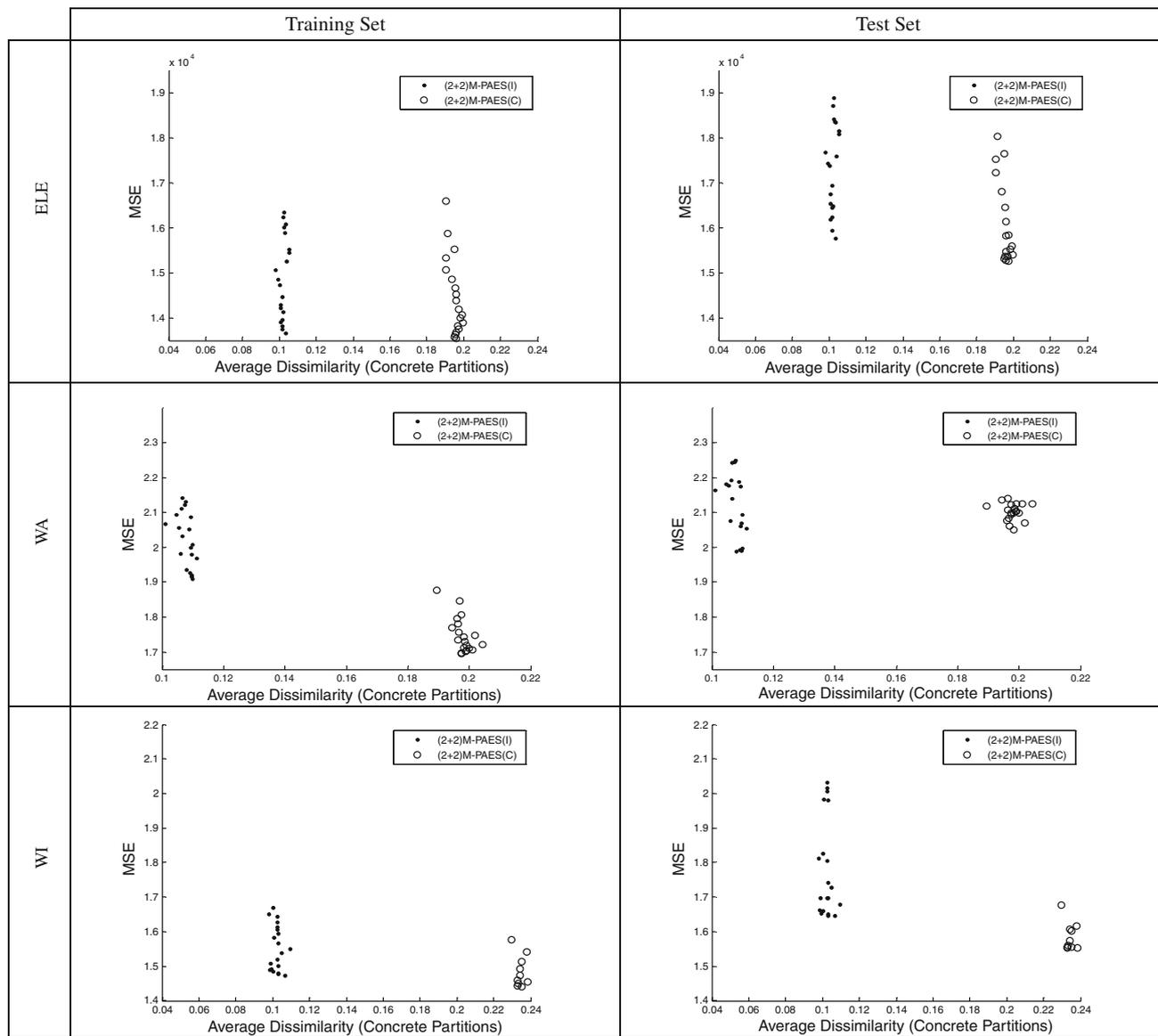


Fig. 5 Average Pareto fronts plotted on the D^c -MSE plane for the training and test sets

$(2 + 2)$ M-PAES(I) and $(2 + 2)$ M-PAES(C) onto the D^c -MSE plane (Fig. 5) are concentrated around 0.08 and 0.24, respectively. As expected, the interpretability index used as objective during the evolutionary process allows increasing the partition integrity. Indeed all the solutions generated by our algorithm are characterized by lower values of D^c .

In Table 4, we show the averages and the standard deviations of the complexity ($\overline{COMP}(\sigma_{COMP})$), of the number of concrete rules ($\overline{M^c}(\sigma_{M^c})$), and of the concrete ($\overline{D^c}(\sigma_{D^c})$) and virtual ($\overline{D^v}(\sigma_{D^v})$) dissimilarities for the FIRST solution. For the sake of commodity, we also report the average and standard deviation of interpretability index $\overline{I}(\overline{I}(\sigma_I))$. The same results for the MEDIAN and LAST solutions are shown in Tables 8 and 9 of “Appendix”,

respectively. The values in Table 4, confirm the trends highlighted by analyzing Figs. 5 and 6: $(2 + 2)$ M-PAES(I) generates MFRBSs which have always lower values of concrete and virtual dissimilarities, thus preserving the partition integrity. Further, these MFRBSs are typically characterized by lower values of complexity and number of rules than the ones generated by $(2 + 2)$ M-PAES(C), except for the MPG6 and STP datasets. Thus, we can conclude that the interpretability index allows optimizing both complexity and integrity during the evolutionary process.

To give a glimpse of the different levels of integrity of the partitions generated by $(2 + 2)$ M-PAES(I) and $(2 + 2)$ M-PAES(C) we plot in Fig. 6a and b two examples of fuzzy partitions for the ELE dataset, characterized by $D^c = 0.099$ and $D^c = 0.19$, respectively. We can observe

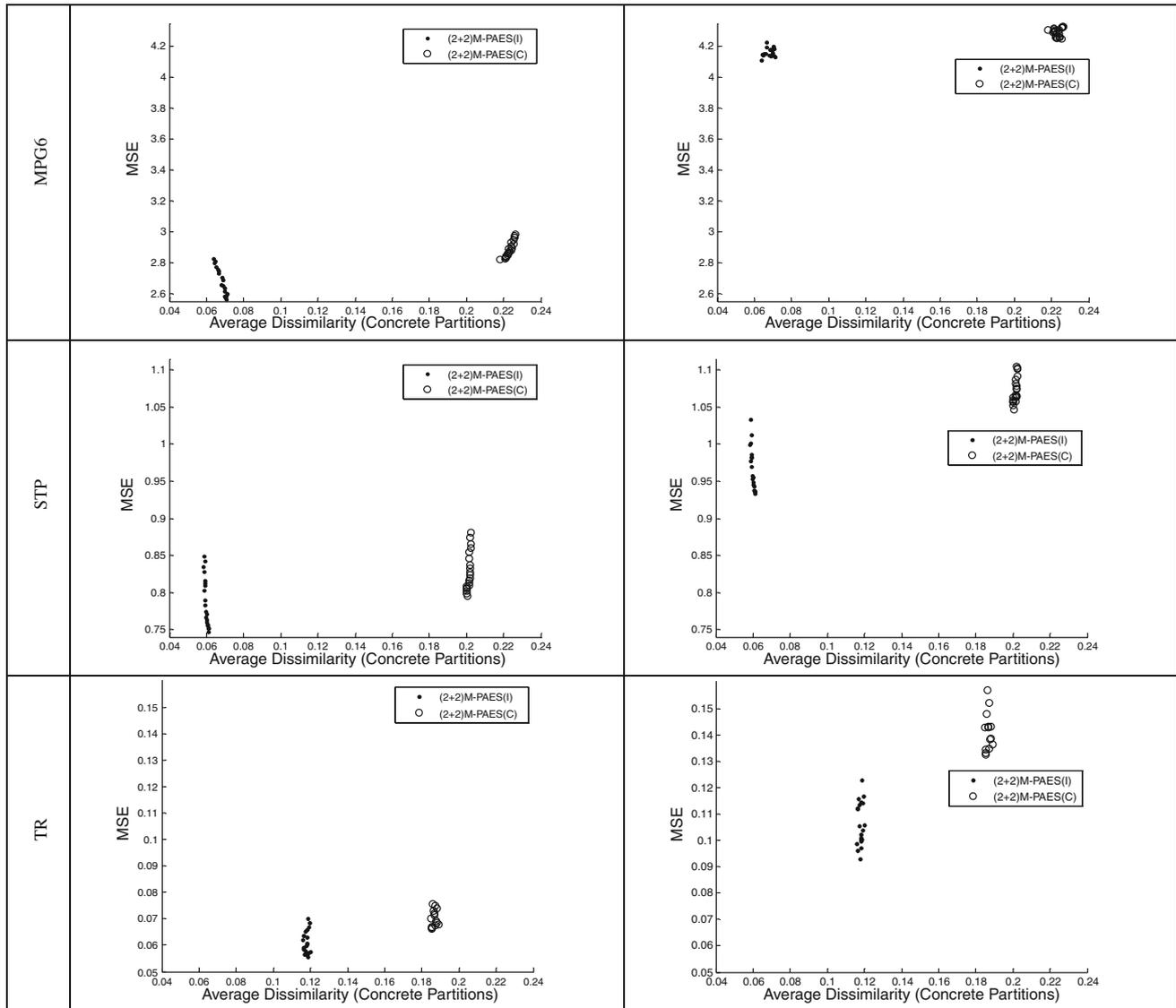


Fig. 5 continued

Table 4 Average interpretability index I , complexity COMP, number M^c of rules and average dissimilarities D^c and D^v for the FIRST solution

		$\bar{I}(\sigma_I)$	$\overline{\text{COMP}}(\sigma_{\text{COMP}})$	$\overline{M^c}(\sigma_{M^c})$	$\overline{D^c}(\sigma_{D^c})$	$\overline{D^v}(\sigma_{D^v})$
ELE	(2 + 2)M-PAES(I)	0.810 (0.131)	68.21 (42.65)	24.24 (12.31)	0.103 (0.048)	0.101 (0.045)
	(2 + 2)M-PAES(C)	0.676 (0.090)	96.48 (27.73)	34.48 (8.97)	0.196 (0.066)	0.241 (0.062)
WA	(2 + 2)M-PAES(I)	0.909 (0.059)	75.16 (46.86)	15.27 (6.43)	0.110 (0.037)	0.115 (0.017)
	(2 + 2)M-PAES(C)	0.877 (0.032)	98.65 (23.11)	20.20 (2.76)	0.197 (0.045)	0.262 (0.037)
WI	(2 + 2)M-PAES(I)	0.926 (0.046)	61.81 (35.95)	13.12 (5.32)	0.107 (0.029)	0.109 (0.025)
	(2 + 2)M-PAES(C)	0.832 (0.087)	83.55 (55.07)	17.83 (8.01)	0.235 (0.054)	0.267 (0.038)
MPG6	(2 + 2)M-PAES(I)	0.776 (0.027)	130.28 (14.67)	48.03 (3.26)	0.071 (0.025)	0.064 (0.013)
	(2 + 2)M-PAES(C)	0.786 (0.045)	121.66 (18.04)	40.36 (5.49)	0.218 (0.107)	0.263 (0.072)
STP	(2 + 2)M-PAES(I)	0.814 (0.019)	184.00 (18.46)	49.42 (1.97)	0.061 (0.017)	0.040 (0.010)
	(2 + 2)M-PAES(C)	0.755 (0.019)	181.73 (13.37)	48.53 (1.25)	0.201 (0.059)	0.268 (0.039)
TR	(2 + 2)M-PAES(I)	0.933 (0.039)	103.92 (52.83)	19.10 (7.31)	0.119 (0.026)	0.129 (0.024)
	(2 + 2)M-PAES(C)	0.884 (0.052)	147.00 (61.97)	25.10 (8.17)	0.185 (0.045)	0.246 (0.033)

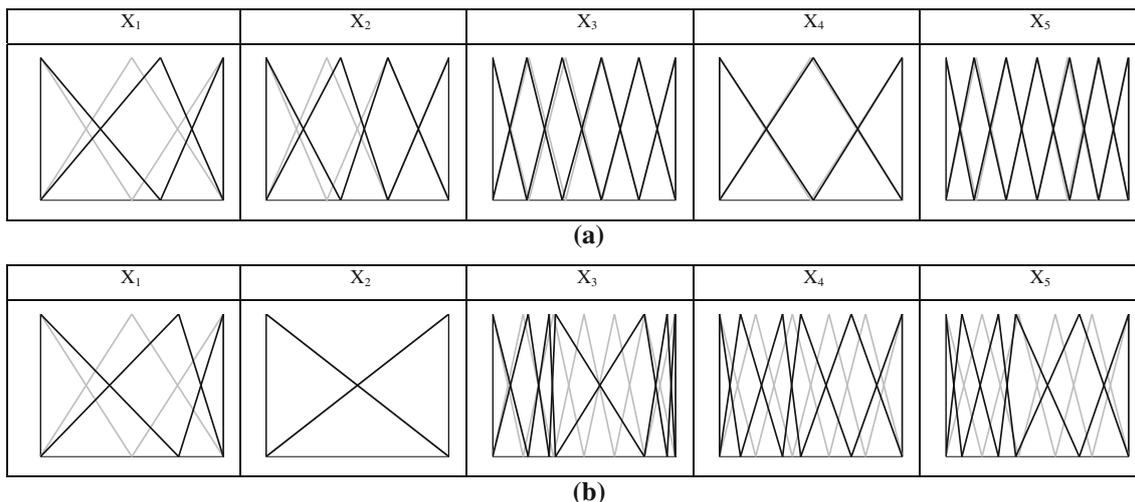


Fig. 6 Two examples of fuzzy partitions for ELE dataset characterized by $D^f = 0.099$ (a) and $D^f = 0.19$ (b), respectively

Table 5 Average values of granularity for all datasets

		$\overline{Gr}(\sigma_{Gr})$
ELE	(2 + 2)M-PAES(I)	4.83 (1.62)
	(2 + 2)M-PAES(C)	4.69 (1.63)
WA	(2 + 2)M-PAES(I)	4.73 (1.77)
	(2 + 2)M-PAES(C)	4.2 (1.54)
WI	(2 + 2)M-PAES(I)	4.35 (1.70)
	(2 + 2)M-PAES(C)	4.68 (1.76)
MPG6	(2 + 2)M-PAES(I)	4.33 (1.78)
	(2 + 2)M-PAES(C)	3.77 (1.63)
STP	(2 + 2)M-PAES(I)	3.85 (1.29)
	(2 + 2)M-PAES(C)	4.02 (1.60)
TR	(2 + 2)M-PAES(I)	4.48 (1.80)
	(2 + 2)M-PAES(C)	4.15 (1.58)

form Fig. 6a that (2 + 2)M-PAES(I) generates partitions practically equal to the initial partitions on three variables (X_3, X_4 and X_5) and very close for the remaining two. On the contrary, in Fig. 6b, we can appreciate that partitions generated by (2 + 2)M-PAES(C) are far from being close to the initial partitions for all the variables but one, X_2 , which has granularity equal to two (and then its partition cannot be moved).

Finally, in Table 5, for each dataset, we report the average granularity ($\overline{Gr}(\sigma_{Gr})$) computed on all the linguistic variables and the 30 trials for the FIRST solution. Although reducing the granularity is not explicitly an objective of the two (2 + 2)M-PAES, we can appreciate how the average granularity is lower than five and so lower than the maximum granularity, thus proving the effectiveness of the granularity learning process. As we

have highlighted in Sect. 3, granularity affects the integrity of a fuzzy partition: the lower the number of fuzzy sets in a partition, the higher the integrity of the partition. Since on average the granularity achieved at the end of the evolutionary process is lower than the maximum value fixed at the beginning, we can conclude that the granularity learning process allows us to increase the level of integrity and consequently to improve interpretability.

6 Conclusions

In this paper we have proposed a novel index for assessing MFRBS interpretability, which takes both the rule base complexity and the partition integrity into account. This index and accuracy have been used as objectives in a two-objective evolutionary algorithm which generates MFRBSs by concurrently learning the rule base, the linguistic partition granularities and the membership function parameters during the evolutionary process. To this aim, we have adopted a modified version of the well-known (2 + 2)PAES and a chromosome consisting of three parts which codify, respectively, the rule base, and, for each linguistic variable, the granularity and the parameters of a piecewise linear transformation of the membership functions.

The proposed approach has been experimented on six real world regression problems and the results have been compared with those obtained by applying the same two-objective evolutionary algorithm, but with accuracy and complexity of the rule base as objectives. We have shown that our approach achieves the best trade-offs between interpretability and accuracy, preserving the partition integrity.

Appendix

In Tables 6, 7, we show the averages and the standard deviations of the MSEs on training and test sets ($\overline{\text{MSE}}_{\text{TR}}(\sigma_{\text{TR}})$ and $\overline{\text{MSE}}_{\text{TS}}(\sigma_{\text{TS}})$, respectively), the results of the Kolmogorov–Smirnov test (column $k\text{-}s_{\text{TR}}$ and $k\text{-}s_{\text{TS}}$ for the training and test sets, respectively), and the averages and the standard deviations of the interpretability index

$\bar{I}(\sigma_I)$ for the MEDIAN and LAST solutions, respectively.

In Tables 8 and 9, we show the averages and the standard deviations of the complexity ($\overline{\text{COMP}}(\sigma_{\text{COMP}})$), of the number of concrete rules ($\overline{M^c}(\sigma_{M^c})$), and of the concrete ($\overline{D^c}(\sigma_{D^c})$) and virtual ($\overline{D^v}(\sigma_{D^v})$) dissimilarities for the MEDIAN and LAST solutions, respectively.

Table 6 Average MSEs on training and test sets and interpretability index for the MEDIAN solution

		$\overline{\text{MSE}}_{\text{TR}}(\sigma_{\text{TR}})$	$k\text{-}s_{\text{TR}}$	$\overline{\text{MSE}}_{\text{TS}}(\sigma_{\text{TS}})$	$k\text{-}s_{\text{TS}}$	$\bar{I}(\sigma_I)$
ELE	(2 + 2)M-PAES(I)	14,729.3 (2,149.0)	=	17,390.1 (3,796.3)	=	0.865 (0.113)
	(2 + 2)M-PAES(C)	14,197.5 (3,850.8)	*	15,849.4 (4,228.4)	*	0.747 (0.074)
WA	(2 + 2)M-PAES(I)	2.009 (0.422)	+	2.0951 (0.3973)	+	0.961 (0.043)
	(2 + 2)M-PAES(C)	1.729 (0.516)	*	2.111 (0.969)	*	0.937 (0.026)
WI	(2 + 2)M-PAES(I)	1.540 (0.385)	=	1.729 (0.438)	=	0.977 (0.023)
	(2 + 2)M-PAES(C)	1.460 (0.282)	*	1.544 (0.242)	*	0.862 (0.070)
MPG6	(2 + 2)M-PAES(I)	2.651 (0.348)	*	4.138 (1.313)	*	0.815 (0.034)
	(2 + 2)M-PAES(C)	2.874 (0.422)	=	4.253 (1.367)	=	0.816 (0.056)
STP	(2 + 2)M-PAES(I)	0.775 (0.101)	*	0.958 (0.183)	*	0.828 (0.024)
	(2 + 2)M-PAES(C)	0.820 (0.223)	=	1.065 (0.313)	=	0.798 (0.024)
TR	(2 + 2)M-PAES(I)	0.060 (0.021)	*	0.093 (0.075)	*	0.965 (0.030)
	(2 + 2)M-PAES(C)	0.069 (0.025)	=	0.138 (0.141)	=	0.919 (0.040)

Table 7 Average MSEs on training and test sets and interpretability index for the LAST solution

		$\overline{\text{MSE}}_{\text{TR}}(\sigma_{\text{TR}})$	$k\text{-}s_{\text{TR}}$	$\overline{\text{MSE}}_{\text{TS}}(\sigma_{\text{TS}})$	$k\text{-}s_{\text{TS}}$	$\bar{I}(\sigma_I)$
ELE	(2 + 2)M-PAES(I)	16,358.5 (2,713.6)	*	18,896.0 (3,672.5)	*	0.922 (0.062)
	(2 + 2)M-PAES(C)	16,595.8 (5,556.4)	=	18,977.3 (5,816.4)	=	0.797 (0.072)
WA	(2 + 2)M-PAES(I)	2.142 (0.449)	+	2.244 (0.529)	=	0.981 (0.031)
	(2 + 2)M-PAES(C)	1.877 (0.733)	*	2.119 (0.937)	*	0.965 (0.021)
WI	(2 + 2)M-PAES(I)	1.670 (0.539)	=	1.827 (0.566)	=	0.99 (0.012)
	(2 + 2)M-PAES(C)	1.577 (0.377)	*	1.678 (0.325)	*	0.899 (0.056)
MPG6	(2 + 2)M-PAES(I)	2.829 (0.350)	*	4.109 (1.321)	*	0.880 (0.026)
	(2 + 2)M-PAES(C)	2.985 (0.457)	=	4.327 (1.410)	=	0.840 (0.060)
STP	(2 + 2)M-PAES(I)	0.849 (0.164)	*	0.958 (0.183)	*	0.828 (0.024)
	(2 + 2)M-PAES(C)	0.881 (0.225)	=	1.102 (0.323)	=	0.798 (0.024)
TR	(2 + 2)M-PAES(I)	0.070 (0.025)	*	0.123 (0.125)	*	0.977 (0.023)
	(2 + 2)M-PAES(C)	0.076 (0.027)	=	0.148 (0.135)	=	0.944 (0.032)

Table 8 Average interpretability index I , complexity COMP, number M^c of rules and average dissimilarities D^c and D^v for the MEDIAN solution

		$\bar{I}(\sigma_I)$	$\overline{\text{COMP}}(\sigma_{\text{COMP}})$	$\overline{M^c}(\sigma_{M^c})$	$\overline{D^c}(\sigma_{D^c})$	$\overline{D^v}(\sigma_{D^v})$
ELE	(2 + 2)M-PAES(I)	0.865 (0.113)	49.08 (36.95)	19.12 (11.17)	0.100 (0.048)	0.096 (0.044)
	(2 + 2)M-PAES(C)	0.761 (0.075)	66.6 (22.61)	26.4 (7.3)	0.197 (0.067)	0.240 (0.056)
WA	(2 + 2)M-PAES(I)	0.961 (0.043)	34.61 (34.89)	9.66 (4.78)	0.110 (0.035)	0.110 (0.017)
	(2 + 2)M-PAES(C)	0.937 (0.026)	53.40 (19.10)	14.10 (2.91)	0.199 (0.044)	0.262 (0.037)
WI	(2 + 2)M-PAES(I)	0.977 (0.023)	21.87 (18.23)	7.75 (3.09)	0.105 (0.037)	0.104 (0.030)
	(2 + 2)M-PAES(C)	0.862 (0.070)	56.94 (43.46)	14.16 (6.39)	0.233 (0.051)	0.268 (0.034)
MPG6	(2 + 2)M-PAES(I)	0.815 (0.034)	109.03 (18.30)	42.25 (5.85)	0.069 (0.023)	0.064 (0.014)
	(2 + 2)M-PAES(C)	0.816 (0.056)	99.90 (17.14)	34.96 (5.30)	0.223 (0.056)	0.264 (0.073)
STP	(2 + 2)M-PAES(I)	0.828 (0.024)	171.39 (23.45)	47.96 (3.89)	0.059 (0.017)	0.037 (0.011)
	(2 + 2)M-PAES(C)	0.780 (0.024)	160.43 (15.87)	45.86 (3.06)	0.202 (0.058)	0.266 (0.040)
TR	(2 + 2)M-PAES(I)	0.965 (0.030)	58.32 (41.40)	13.10 (5.90)	0.118 (0.024)	0.126 (0.023)
	(2 + 2)M-PAES(C)	0.919 (0.040)	103.82 (48.47)	19.65 (6.41)	0.188 (0.043)	0.245 (0.033)

Table 9 Average interpretability index I , complexity COMP, number M^c of rules and average dissimilarities D^c and D^v for the LAST solution

		$\bar{I}(\sigma_I)$	$\overline{\text{COMP}}(\sigma_{\text{COMP}})$	$\overline{M^c}(\sigma_{M^c})$	$\overline{D^c}(\sigma_{D^c})$	$\overline{D^v}(\sigma_{D^v})$
ELE	(2 + 2)M-PAES(I)	0.922 (0.062)	29.56 (19.60)	13.56 (6.34)	0.102 (0.045)	0.090 (0.043)
	(2 + 2)M-PAES(C)	0.826 (0.066)	45.10 (18.25)	20.00 (6.62)	0.191 (0.069)	0.237 (0.057)
WA	(2 + 2)M-PAES(I)	0.981 (0.031)	19.11 (25.08)	7.38 (3.46)	0.107 (0.035)	0.107 (0.020)
	(2 + 2)M-PAES(C)	0.965 (0.021)	32.25 (15.51)	10.70 (2.62)	0.189 (0.045)	0.255 (0.041)
WI	(2 + 2)M-PAES(I)	0.990 (0.012)	12.37 (8.97)	6.37 (1.83)	0.100 (0.033)	0.096 (0.022)
	(2 + 2)M-PAES(C)	0.899 (0.056)	36.27 (32.48)	10.88 (5.20)	0.230 (0.052)	0.264 (0.033)
MPG6	(2 + 2)M-PAES(I)	0.880 (0.026)	73.96 (13.99)	32.60 (4.72)	0.064 (0.025)	0.064 (0.012)
	(2 + 2)M-PAES(C)	0.814 (0.060)	84.90 (15.36)	31.16 (5.02)	0.226 (0.106)	0.258 (0.074)
STP	(2 + 2)M-PAES(I)	0.837 (0.026)	163.42 (26.13)	47.53 (4.52)	0.058 (0.026)	0.036 (0.011)
	(2 + 2)M-PAES(C)	0.798 (0.024)	144.40 (15.14)	44.46 (3.13)	0.203 (0.024)	0.268 (0.035)
TR	(2 + 2)M-PAES(I)	0.977 (0.023)	41.50 (31.54)	10.85 (4.91)	0.118 (0.023)	0.124 (0.023)
	(2 + 2)M-PAES(C)	0.944 (0.032)	73.03 (37.28)	15.65 (5.36)	0.186 (0.043)	0.245 (0.034)

References

Alcalá R, Alcalá-Fdez J, Herrera F, Otero J (2007a) Genetic learning of accurate and compact fuzzy rule based systems based on the 2-Tuples linguistic representation. *Int J Approx Reason* 44:45–64

Alcalá R, Gacto MJ, Herrera F, Alcalá-Fdez J (2007b) A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems. *Int J Uncertain Fuzz Knowl Based Syst* 15(5):521–537

Alcalá R, Ducange P, Herrera F, Lazzarini B, Marcelloni F (2009) A Multi-objective evolutionary approach to concurrently learn rule and data bases of linguistic fuzzy rule-based systems. *IEEE Trans Fuzzy Syst* 17(5):1106–1122

Alonso JM, Magdalena L, Guillaume S (2008) HILK: a new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism. *Int J Intell Syst* 23:761–794

Alonso JM, Magdalena L, González-Rodríguez G (2009) Looking for a good fuzzy system interpretability index: an experimental approach. *Int J Approx Reason* 51(1):115–134

Antonelli M, Ducange P, Lazzarini B, Marcelloni F (2009a) Learning concurrently partition granularities and rule bases of Mamdani fuzzy systems in a multi-objective evolutionary framework. *Int J Approx Reason* 50(7):1066–1080

Antonelli M, Ducange P, Lazzarini B, Marcelloni F (2009b) Multi-objective evolutionary learning of granularity, membership function parameters and rules of Mamdani fuzzy systems. *Evol Intel* 2(1–2):21–37

Botta A, Lazzarini B, Marcelloni F, Stefanescu D (2009) Context adaptation of fuzzy systems through a multi-objective evolutionary approach based on a novel interpretability index. *Soft Comput* 13(5):437–449

Casillas J, Cordon O, Herrera F (2002) COR: a methodology to improve ad hoc data-driven linguistic rule learning methods by inducing cooperation among rules. *IEEE Trans Syst Man Cybern* 32(4):526–537

Casillas J, Cordon O, Herrera F, Magdalena L (eds) (2003) *Interpretability issues in fuzzy modeling*. Springer, Heidelberg

Cococcioni M, Ducange P, Lazzarini B, Marcelloni F (2007) A Pareto-based multi-objective evolutionary approach to the identification of Mamdani fuzzy systems. *Soft Comput* 11(11):1013–1031

- Coello Coello CA, Lamont GB (2004) Applications of multi-objective evolutionary algorithms. World Scientific, Singapore
- Cordon O, Herrera F, Villar P (2001a) Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base. *IEEE Trans Fuzzy Syst* 9(4):667–674
- Cordon O, Herrera F, Magdalena L, Villar P (2001b) A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base. *Inf Sci* 136:85–107
- de Oliveira JV (1999) Semantic constraints for membership function optimization. *IEEE Trans Syst Man Cybern Part A* 29(1):128–138
- Deb K (2001) Multi-objective optimization using evolutionary algorithms. Wiley, Chichester
- Ducange P, Lazzerini B, Marcelloni F (2009) Multi-objective genetic fuzzy classifiers for imbalanced and cost-sensitive datasets. *Soft Comput* 14(7):713–728
- Gacto MJ, Alcalá R, Herrera F (2009) Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems. *Soft Comput* 13(5):419–436
- Gacto MJ, Alcalá R, Herrera F (2010) Integration of an index to preserve the semantic interpretability in the multi-objective evolutionary rule selection and tuning of linguistic fuzzy systems. *IEEE Trans Fuzzy Syst*. doi:10.1109/TFUZZ.2010.2041008
- González A, Pérez R (1999) SLAVE: a genetic learning system based on the iterative approach. *IEEE Trans Fuzzy Syst* 7:176–191
- Guillaume S (2001) Designing fuzzy inference systems from data: An interpretability-oriented review. *IEEE Trans Fuzzy Syst* 9(3):426–443
- Herrera F (2008) Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evol Intel* 1:27–46
- Ishibuchi H (2007) Multiobjective genetic fuzzy systems: review and future research direction. In: Proceedings of FUZZ-IEEE 2007 international conference on fuzzy systems, London, 23–26 July
- Ishibuchi H, Nojima Y (2007) Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning. *Int J Approx Reason* 44(1):4–31
- Ishibuchi H, Yamamoto T (2004) Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining. *Fuzzy Sets Syst* 141(1):59–88
- Ishibuchi H, Murata T, Turksen IB (1997) Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems. *Fuzzy Sets Syst* 89(2):135–150
- Klawonn F (2006) Reducing the number of parameters of a fuzzy system using scaling functions. *Soft Comput* 10(9):749–756
- Knowles JD, Corne DW (2002) Approximating the non dominated front using the Pareto archived evolution strategy. *Evol Comput* 8(2):149–172
- Mamdani EH, Assilian S (1975) An experiment in linguistic synthesis with a fuzzy logic controller. *Int J Man Mach Stud* 7(1):1–13
- Massey FJ (1951) The Kolmogorov-Smirnov test for goodness of fit. *J Am Stat Assoc* 46(253):68–78
- Mencar C, Fanelli AM (2008) Interpretability constraints for fuzzy information granulation. *Inf Sci* 178:4585–4618
- Mencar C, Castellano G, Fanelli AM (2007) Distinguishability quantification of fuzzy sets. *Inf Sci* 177:130–149
- Pedrycz W, Gomide F (2007) Fuzzy systems engineering: toward human-centric computing. Wiley-IEEE Press, NJ
- Pulkkinen P, Koivisto H (2010) A dynamically constrained multiobjective genetic fuzzy system for regression problems. *IEEE Trans Fuzzy Syst* 18(1):161–177
- Ruspini EH (1969) A new approach to clustering. *Inform Control* 15(1):22–32
- Teng Y, Wang W (2004) Constructing a user-friendly ga-based fuzzy system directly from numerical data. *IEEE Trans Syst Man Cybern B* 34(5):2060–2070
- Wang LX, Mendel JM (1992) Generating fuzzy rules by learning from examples. *IEEE Trans Syst Man Cybern* 22(6):1414–1427
- Zhou SM, Gan JQ (2008) Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling. *Fuzzy Sets Syst* 159:3091–3131