Preaggregation Functions: Construction and an Application

Giancarlo Lucca, José Antonio Sanz, Graçaliz Pereira Dimuro, Benjamín Bedregal, Radko Mesiar, Anna Kolesárová, and Humberto Bustince, *Senior Member, IEEE*

***Abstract*—In this paper, we introduce the notion of preaggregation function. Such a function satisfies the same boundary conditions as an aggregation function, but, instead of requiring monotonicity, only monotonicity along some fixed direction (directional monotonicity) is required. We present some examples of such functions. We propose three different methods to build preaggregation functions. We experimentally show that in fuzzy rule-based classification systems, when we use one of these methods, namely, the one based on the use of the Choquet integral replacing the product by other aggregation functions, if we consider the minimum or the Hamacher product t-norms for such construction, we improve the results obtained when applying the fuzzy reasoning methods obtained using two classical averaging operators such as the maximum and the Choquet integral.**

***Index Terms*—Aggregation functions, Choquet integral, directional monotonicity, fuzzy measures, fuzzy reasoning method (FRM), fuzzy rule-based classification systems (FRBCSs).**

I. Introduction

Aggregation functions [1], [2] are crucial tools nowadays to deal with many computation problems [3]–[7]. The key property for defining them, apart from the boundary conditions, is monotonicity and, more specifically, monotone increasingness. However, some other statistical tools, such as the mode, are not included in this family, although they are useful, since, even if they are properly defined as functions, monotonicity is violated.

The problem of relaxing the definition of monotonicity has recently attracted a lot of interest. In [8], Wilkin and Beliakov proposed the notion of weak monotonicity in order to extend the usual monotonicity property. In this case, monotonicity is required only along the direction of the first quadrant diagonal. This concept of weak monotonicity has been further extended by Bustince *et al.* [9] by introducing the notion of directional monotonicity, which allows monotonicity along (some) fixed ray. In particular, directionally monotone functions encompass weak monotone functions, as well as the mode and any aggregation function.

In particular, in this paper, we consider the following objectives.

1) To introduce the concept of preaggregation functions.
2) To study the first properties of these new functions.
3) To introduce three different methods for building preaggregation functions.
4) To show an application where the introduction of the new concept of preaggregation function is justified.

To achieve these goals, we use the notion of directional monotonicity. Moreover, for one of the construction methods that we propose, in the definition of the Choquet integral, we replace the product by the minimum or the Hamacher product t-norm, and this way, we obtain preaggregation functions that are not aggregation functions. We show that using these new functions in a fuzzy rule-based classification system (FRBCS) and, in particular, in the fuzzy reasoning method (FRM) of fuzzy association rule-based classification model for high dimensional problems (FARC-HD) [10], which is currently one of the most accurate FRBCSs, the obtained results are better than both applying the classical Choquet integral and the well-known FRM of the winning rule (WR).

This paper is organized as follows. In Section II, we present some related preliminary concepts that are necessary to understand the paper. In Section III, we introduce the notion of preaggregation function and discuss some properties. Three methods of construction of preaggregation functions are described in Section IV. The generalization of the FRM using preaggregation functions is described in detail in Section V. The experimental framework and the analysis of the obtained results when considering some of our preaggregation functions are reported in Section VI. In Section VII, we draw the main conclusions, and the detailed results of the experiments are available in the Appendix.

II. Preliminaries

A. Aggregation Functions

An important class of operators that are used in this paper are the *aggregation functions* [1], [11].

Manuscript received March 2, 2015; revised May 14, 2015; accepted June 2, 2015. Date of publication June 17, 2015; date of current version March 29, 2016. This work was supported in part by the Spanish Ministry of Science and Technology under Projects TIN2008-06681-C06-01, TIN2010-15055, TIN2013-40765-P, and TIN2011-29520, by the Brazilian funding agency CNPQ under Processes 481283/2013-7, 306970/2013-9, 232827/2014-1, and 307681/2012-2, and by Grants VEGA 1/0419/13 and VEGA 1/0420/15.

G. Lucca, J. A. Sanz, and H. Bustince are with the Departamento of Automática y Computación and the Institute of Smart Cities, Universidad Publica de Navarra, Navarra 31006, Spain (e-mail: giancarlo.lucca88@gmail.com; joseantonio.sanz@unavarra.es; bustince@unavarra.es).

G. Pereira Dimuro is with the PPGCOMP, Centro de Ciencias Computacionais, Universidade Federal do Rio Grande, Rio Grande 96200-350, Brazil (e-mail: gracalizdimuro@furg.br).

B. Bedregal is with the Departamento de Informática e Matemática Aplicada, Universidade Federal do Rio Grande do Norte, Natal 59072-970, Brazil (e-mail: bedregal@dimap.ufrn.br).

R. Mesiar is with the Slovak University of Technology, 81005, Bratislava, Slovakia, and also with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, 18208 Prague, Czech Republic (e-mail: mesiar@math.sk).

A. Kolesárová is with the Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology, 81237 Bratislava, Slovakia (e-mail: anna.kolesarova@stuba.sk).

Digital Object Identifier 10.1109/TFUZZ.2015.2453020

TABLE I
T-NORMS USED IN THIS PAPER

Name	Definition
Minimum	$T_M(x,y) = \min\{x,y\}$
Algebraic Product	$T_P(x,y) = xy$
Łukasiewicz	$T_{Ł}(x,y) = \max\{0, x+y-1\}$
Drastic Product	$T_{\mathrm{DP}}(x,y) = \begin{cases} x, & \text{if } y=1 \\ y, & \text{if } x=1 \\ 0, & \text{otherwise} \end{cases}$
Nilpotent Minimum	$T_{\mathrm{NM}}(x,y) = \begin{cases} \min\{x,y\}, & \text{if } x+y>1 \\ 0, & \text{otherwise} \end{cases}$
Hamacher Product	$T_{\mathrm{HP}}(x,y) = \begin{cases} 0, & \text{if } x=y=0 \\ \frac{xy}{x+y-xy}, & \text{otherwise} \end{cases}$

Definition 2.1: A function $A : [0,1]^n \to [0,1]$ is said to be an n-ary aggregation function if the following conditions hold.

A1) A is increasing[1] in each argument: For each $i \in \{1,\ldots,n\}$, if $x_i \leq y$, then $A(x_1,\ldots,x_n) \leq A(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)$.

A2) A satisfies the boundary conditions: $A(0,\ldots,0) = 0$ and $A(1,\ldots,1) = 1$.

Definition 2.2: A bivariate aggregation function $T : [0,1]^2 \to [0,1]$ is a t-norm if, for all $x,y,z \in [0,1]$, it satisfies the following properties:

T1) Commutativity: $T(x,y) = T(y,x)$.

T2) Associativity: $T(x,T(y,z)) = T(T(x,y),z)$.

T3) Boundary condition: $T(x,1) = x$.

If T satisfies (T3) (and also $T(1,x) = x$) only, then it is called a semicopula.

Since t-norms are associative, it is possible to extend each t-norm T in a unique way to an n-ary operation in the usual way by induction [12]. The bivariate t-norms that are used in this paper are presented in Table I. Observe that a convex combination of t-norms is a (commutative) semicopula, but not a t-norm, in general, since associativity may be violated.

B. Fuzzy Measures

In this section, we recall the notion of fuzzy measure, which is going to be a key tool for constructing some of our examples of preaggregation functions.

In the following, consider the set $N = \{1,\ldots,n\}$ for an arbitrary positive integer n.

Definition 2.3: A function $\mathfrak{m} : 2^N \to [0,1]$ is a fuzzy measure if, for all $X,Y \subseteq N$, it satisfies the following properties.

($\mathfrak{m}$1) *Increasingness:* If $X \subseteq Y$, then $\mathfrak{m}(X) \leq \mathfrak{m}(Y)$.

($\mathfrak{m}$2) *Boundary conditions:* $\mathfrak{m}(\emptyset) = 0$ and $\mathfrak{m}(N) = 1$.

In the context of aggregation functions, fuzzy measures are used to evaluate the relationship among the elements to be aggregated, which represents the importance of a coalition. The fuzzy measures considered in this paper, defined for $A \subseteq N$, are the following:

[1] In this paper, an increasing (decreasing) function does not need to be strictly increasing (decreasing).

Uniform measure:

$$\mathfrak{m}_U(A) = \frac{|A|}{n}. \tag{1}$$

Dirac's measure: For a previously fixed $i \in N$,

$$\mathfrak{m}_D^i(A) = \begin{cases} 1, & \text{if } i \in A \\ 0, & \text{if } i \notin A. \end{cases} \tag{2}$$

Additive measure (Wmean): Take $W = (w_1,\ldots,w_n) \in [0,1]^n$ such that $\sum_{i=1}^n w_i = 1$. Consider

$$\mathfrak{m}_W(\{i\}) = w_i$$

Then, for $|A| > 1$, define

$$\mathfrak{m}_W(A) = \sum_{i \in A} w_i. \tag{3}$$

Symmetric measure (OWA): Take $W = (w_1,\ldots,w_n) \in [0,1]^n$ such that $\sum_{i=1}^n w_i = 1$. Then, for any nonempty subset A, define

$$\mathfrak{m}_{sW}(A) = \sum_{i=1}^{|A|} w_i. \tag{4}$$

Note that this expression is different from (3), since in this case, only the cardinal of each subset A is taken into account.

Power measure:

$$\mathfrak{m}_{PM}(A) = \left(\frac{|A|}{n}\right)^q, \text{ with } q > 0. \tag{5}$$

Observe also that from the considered fuzzy measures, $\mathfrak{m}_U$, $\mathfrak{m}_D^i$, and $\mathfrak{m}_W$ are additive, and $\mathfrak{m}_U$, $\mathfrak{m}_{sW}$, and $\mathfrak{m}_{PM}$ are symmetric, that is, the measure of any subset A only depends on the cardinality of A.

The Choquet integral generalizes the Lebesgue integral, which is defined with respect to additive measures. However, the Choquet integral is defined with respect to fuzzy measures. In this paper, we consider only the discrete Choquet integral, related to fuzzy measures, which are defined on finite spaces.

Definition 2.4: [1, Def. 1.74] Let $\mathfrak{m} : 2^N \to [0,1]$ be a fuzzy measure. The discrete Choquet integral of $\mathbf{x} = (x_1,\ldots,x_n) \in [0,1]^n$ with respect to $\mathfrak{m}$ is defined as a function $C_{\mathfrak{m}} : [0,1]^n \to [0,1]$, given by

$$C_{\mathfrak{m}}(\mathbf{x}) = \sum_{i=1}^n \left(x_{(i)} - x_{(i-1)}\right) \cdot \mathfrak{m}\left(A_{(i)}\right) \tag{6}$$

where $\left(x_{(1)},\ldots,x_{(n)}\right)$ is an increasing permutation on the input $\mathbf{x}$, that is, $0 \leq x_{(1)} \leq \cdots \leq x_{(n)}$, with the convention that $x_{(0)} = 0$, and $A_{(i)} = \{(i),\ldots,(n)\}$ is the subset of indices of $n-i+1$ largest components of $\mathbf{x}$.

The Choquet integral combines the inputs in such a way that the importance of the different groups of inputs (coalitions) may be taken into account. Allowing to assign importance to all possible groups of criteria, the Choquet integral offers greater flexibility in the aggregation modeling. Since the weighted arithmetic mean and OWA operators are special cases of the Choquet integral, with respect to additive and symmetric fuzzy measures, respectively, Choquet integral-based aggregation functions represent a larger class of aggregation functions [1], [13], [14].

Note that the Choquet integral with respect to $\mathfrak{m}_W$ is a weighted arithmetic mean and, with respect to $\mathfrak{m}_{sW}$, is an OWA operator.[2] These facts explain the acronyms we have chosen in this study for these measures.

C. Directional Monotonicity

This section is devoted to recalling the basic concept for our definition of preaggregation function, that of directional monotonicity [9].

Definition 2.5: Let $\vec{r} = (r_1, \ldots, r_n)$ be a real n-dimensional vector, $\vec{r} \neq \vec{0}$. A function $F : [0,1]^n \to [0,1]$ is $\vec{r}$-increasing if for all points $(x_1, \ldots, x_n) \in [0,1]^n$ and for all $c > 0$ such that $(x_1 + cr_1, \ldots, x_n + cr_n) \in [0,1]^n$, it holds

$$F(x_1 + cr_1, \ldots, x_n + cr_n) \geq F(x_1, \ldots, x_n) .$$

That is, an $\vec{r}$-increasing function is a function that is increasing along the ray (direction) determined by the vector $\vec{r}$. For this reason, we say that F is directionally monotone, or, more specifically, directionally increasing. Note that every increasing function (in the usual sense) is, in particular, $\vec{r}$-increasing, for every nonnegative real vector $\vec{r}$. However, the class of directionally increasing functions is much wider than that of aggregation functions. For instance:

1) Fuzzy implication functions (see [21]) are $(-1, 1)$-increasing functions. This implies that many other functions, which are widely used in applications and which can be obtained from implication functions, are also directionally increasing. This is the case, for instance, of some subsethood measures (see [22]).
2) Many functions used for comparison of data are also directionally increasing. In particular, this is the case of those based on componentwise comparison by means of the Euclidean distance $|x - y|$, as for restricted equivalence functions [23];
3) Weakly increasing functions (see[8]) are a particular case of directionally increasing functions, with $\vec{r} = (1, \ldots, 1)$.

III. Preaggregation Functions

In this section, we introduce the notion of preaggregation function and discuss some properties and construction methods.

Definition 3.1: A function $F : [0,1]^n \to [0,1]$ is said to be an n-ary preaggregation function if the following conditions hold.

PA1) There exists a real vector $\vec{r} \in [0,1]^n$ $(\vec{r} \neq \vec{0})$ such that F is $\vec{r}$-increasing.

PA2) F satisfies the boundary conditions: $F(0, \ldots, 0) = 0$ and $F(1, \ldots, 1) = 1$.

Example 3.1: Some examples of preaggregation functions are the following.

i) Consider the mode, $\mathrm{Mod}(x_1, \ldots, x_n)$, defined as the function that gives back the value that appears most times in the considered n-tuple or the smallest of the values that appears most times, in case there is more than one. Then, the mode is $(1, \ldots, 1)$-increasing, and it is a particular case of preaggregation function, which is not an aggregation function.

ii) $F(x, y) = x - (\max\{0, x - y\})^2$ is, for instance, $(0, 1)$-increasing, and it is an example of a preaggregation function, which is not an aggregation function.

iii) Weakly increasing functions satisfying the boundary conditions (PA2) are also preaggregation functions, which need not be aggregation functions.

iv) Take $\lambda \in]0, 1[$. The weighted Lehmer mean $L_\lambda : [0,1]^2 \to [0,1]$, given by

$$L_\lambda(x, y) = \frac{\lambda x^2 + (1 - \lambda) y^2}{\lambda x + (1 - \lambda) y}$$

(with convention $0/0 = 0$) is $(1 - \lambda, \lambda)$-increasing; therefore, it is a preaggregation function.

v) Define $A, B : [0,1]^2 \to [0,1]$ by

$$A(x, y) = \begin{cases} x(1 - x), & \text{if } y \leq 3/4 \\ 1, & \text{otherwise} \end{cases}$$

and

$$B(x, y) = \begin{cases} y(1 - y), & \text{if } x \leq 3/4 \\ 1, & \text{otherwise.} \end{cases}$$

Then, both A and B are preaggregation functions that are not aggregation functions. In fact, A is $(0, a)$-increasing for any $a > 0$, but for no other direction $\vec{r} = (a, b)$, $b > 0$, while B is $(b, 0)$-increasing for any $b > 0$, but for no other direction $\vec{r} = (a, b)$, $a > 0$. However, $C = (A + B)/2$ is not a preaggregation function, just illustrating the fact that the class of all preaggregation functions with a fixed dimension n is not a convex class.

If F is a preaggregation function with respect to a vector $\vec{r}$, we just say that F is an $\vec{r}$-preaggregation function.

Remark 3.1: Note that if $A : [0,1]^n \to [0,1]$ is an aggregation function, then A is also a preaggregation function. In fact, if, for a nonzero vector $\vec{r} \in [0,1]^n$, we denote by $PA_{\vec{r}}$ the class of all $\vec{r}$-increasing preaggregation functions, then the class of all preaggregation functions PA is the union of all these classes $PA_{\vec{r}}$, while the class of all aggregation functions is the intersection of all the classes $PA_{\vec{r}}$. The latter intersection is the same as the intersection over $PA_{\vec{e_i}}$, where $\vec{e_i} = (0, ...1, ...0)$, $i \in \{1, \ldots, n\}$, is the vector having 1 as ith value, and all other coordinates are equal to zero.

Note that the reverse of the first claim of Remark 3.1 does not hold, as the cases considered in Example 3.1(i) and (ii) show. Preaggregation functions that are not aggregation functions will be called proper preaggregation functions. However, we can use aggregation functions to obtain directionally increasing functions as follows.

The next results were proved for directionally monotone functions in our recent paper [9].

Proposition 3.1: Let $A : [0,1]^m \to [0,1]$ be an aggregation function. Let $F_i : [0,1]^n \to [0,1]$ $(i \in \{1, \ldots, m\})$ be a family of m $\vec{r}$-preaggregation functions for the same vector $\vec{r} \in [0,1]^n$.

[2]The OWA operators were first introduced by Yager [15], and several forms and usage of OWA operators have been discussed in the literature [16]–[20].

Then, the function $A(F_1,\ldots,F_m):[0,1]^n \to [0,1]$, defined as

$$A(F_1,\ldots,F_m)(x_1,\ldots,x_n) = A(F_1(x_1,\ldots,x_n),\ldots,F_m(x_1,\ldots,x_n))$$

is also an $\vec{r}$-preaggregation function.

Proof: Due to [9, Proposition 3], only the boundary conditions for the functions $(F_1,\ldots,F_m)$ should be guaranteed. However, their validity is obvious. ■

The following corollary is straightforward.

Corollary 3.1: Let $F_1, F_2 : [0,1]^n \to [0,1]$ be two $\vec{r}$-preaggregation functions for the same vector $\vec{r} \in [0,1]^n$. Then, we have the following.

1) $\frac{F_1+F_2}{2}$ is also an $\vec{r}$-preaggregation function.
2) F_1F_2 is also an $\vec{r}$-preaggregation function.

Regarding duality, we can state the following.

Proposition 3.2: Let $F:[0,1]^n \to [0,1]$ be an $\vec{r}$-preaggregation function for $\vec{r} \in [0,1]^n$. Then, the function

$$F^d(x_1,\ldots,x_n) = 1 - F(1-x_1,\ldots,1-x_n)$$

is also an $\vec{r}$-preaggregation function.

Proof: Obviously, $F^d(0,\ldots,0) = 0$ and $F^d(1,\ldots,1) = 1$. Now, the result follows from [9, Proposition 3]. ■

The following corollary is now straight.

Corollary 3.2: Let F be an $\vec{r}$-preaggregation function. Then, the function $\frac{F+F^d}{2}$ is a self-dual $\vec{r}$-preaggregation function.

IV. Three Methods of Constructing Preaggregation Functions

In this section, we introduce and illustrate three methods of constructing preaggregation functions. The first method is based on the composition of appropriate functions, the second one is inspired by the construction of the discrete Choquet integral, and the third of the proposed methods is inspired by the construction of the discrete Sugeno integral.

A. Construction of Preaggregation Functions by Composition

Fix $n \in \mathbb{N}$. Let I be a proper subset of $N = \{1,\ldots,n\}$ and consider that $I = \{i_1,\ldots,i_k\}$ with $i_1 < \ldots < i_k$. For an n-tuple $\mathbf{x} = (x_1,\ldots,x_n) \in [0,1]^n$, its I-projection is a k-tuple $\mathbf{x}_I = (x_{i_1},\ldots,x_{i_k})$, where k is the cardinality of I. We will use I-projections $\mathbf{x}_I$ of points $\mathbf{x} \in [0,1]^n$ and I-projections $\vec{r}_I$ of (geometrical) vectors $\vec{r} \in [0,1]^n$ as well. Finally, for a function $F:[0,1]^n \to [0,1]$, let $\mathcal{D}^{\uparrow}(F) = \{\vec{r} \in [0,1]^n \mid F \text{ is } \vec{r}\text{-increasing}\}$. Note that the zero vector is not excluded now.

Proposition 4.1: Let $\{I_1,\ldots,I_k\}$ be a partition of N, $k > 1$. For $j \in \{1,\ldots,k\}$, let $n_j = |I_j|$ and consider functions $F_j : [0,1]^{n_j} \to [0,1]$ such that $F_j(1,\ldots,1) = 1$. Then, for any aggregation function $G:[0,1]^k \to [0,1]$, the composite function $H:[0,1]^n \to [0,1]$ defined by

$$H(\mathbf{x}) = G\left(F_1\left(\mathbf{x}_{I_1}\right),\ldots,F_k\left(\mathbf{x}_{I_k}\right)\right)$$

is $\vec{r}$-increasing for any vector $\vec{r} \in [0,1]^n$ such that $\vec{r}_{I_j} \in \mathcal{D}^{\uparrow}(F_j)$, $j = 1,\ldots,k$, and $H(\mathbf{1}) = 1$. Moreover, if there is a $j_0 \in \{1,\ldots,k\}$ such that F_{j_0} is a preaggregation function, and 0 is an annihilator of G, then the function H is a preaggregation function.

Proof: Clearly, $H(\mathbf{1}) = G\left(F_1\left(\mathbf{1}_{I_1}\right),\ldots,F_k\left(\mathbf{1}_{I_k}\right)\right) = G(1,\ldots,1) = 1$. Moreover, if $F_{j_0}(0,\ldots,0) = 0$ for some $j_0 \in \{1,\ldots,k\}$ and 0 is an annihilator of G, then

$$\begin{aligned} H(\mathbf{0}) &= G\left(F_1\left(\mathbf{0}_{I_1}\right),\ldots,F_{j_0}\left(\mathbf{0}_{I_{j_0}}\right),\ldots,F_k\left(\mathbf{0}_{I_k}\right)\right) \\ &= G\left(F_1\left(\mathbf{0}_{I_1}\right),\ldots,0,\ldots,F_k\left(\mathbf{0}_{I_k}\right)\right) = 0. \end{aligned}$$

Next, consider a vector $\vec{r} \in [0,1]^n$ such that $\vec{r}_{I_j} \in \mathcal{D}^{\uparrow}(F_j)$ for each $j = 1,\ldots,k$. Then, for any $c > 0$ and $\mathbf{x} \in [0,1]^n$ such that also $\mathbf{x} + c\vec{r} \in [0,1]^n$, it holds that

$$\begin{aligned} H(\mathbf{x} + c\vec{r}) &= G\left(F_1\left(\mathbf{x}_{I_1} + c\vec{r}_{I_1}\right),\ldots,F_k\left(\mathbf{x}_{I_k} + c\vec{r}_{I_k}\right)\right) \\ &\geq G\left(F_1\left(\mathbf{x}_{I_1}\right),\ldots,F_k\left(\mathbf{x}_{I_k}\right)\right) = H(\mathbf{x}) \end{aligned}$$

where the inequality follows from the increasing monotonicity of the aggregation function G, and the fact that $F_j\left(\mathbf{x}_{I_j} + c\vec{r}_{I_j}\right) \geq F_j\left(\mathbf{x}_{I_j}\right)$, $j = 1,\ldots,k$.

Now, suppose that F_{j_0} is a preaggregation function, i.e., $F_{j_0}(0,\ldots,0) = 0$ and F_{j_0} is $\vec{v}$-increasing for some nonzero vector $\vec{v} \in [0,1]^{n_{j_0}}$. Due to the aforementioned facts, H satisfies the boundary conditions and is directionally increasing in the direction of a nonzero vector $\vec{r} \in [0,1]^n$ such that $\vec{r}_{I_{j_0}} = \vec{v}$ and $\vec{r}_{N \setminus I_{j_0}} = (0,\ldots,0)$, which proves that H is a preaggregation function. ■

Example 4.1: Let $n = 2$ and $\vec{v} = (v_1, v_2) \in]0,1]^2$. For obtaining a proper preaggregation function which is $\vec{v}$-increasing, it is enough to consider the weighted Lehmer mean $L_\lambda : [0,1]^2 \to [0,1]$ with $\lambda = \frac{v_2}{v_1+v_2}$ [see Example 3.1(iv)] given by

$$L_\lambda(x,y) = \frac{v_2x^2 + v_1y^2}{v_2x + v_1y}.$$

This fact and Proposition 4.1 allow us to construct a preaggregation function H, which is directionally increasing in the direction of any *a priori* given vector $\vec{0} \neq \vec{r} \in [0,1]^n$.

Consider, for example, $n = 4$ and $\vec{r} = (0.5, 0.4, 0.3, 0.7)$. Let $G = T_M$, $I_1 = \{1,3\}$, $I_2 = \{2,4\}$, $F_1 = L_{3/8}$, $F_2 = L_{7/11}$. Then, $H:[0,1]^4 \to [0,1]$ given by

$$H(x_1,x_2,x_3,x_4) = \min\left\{\frac{3x_1^2 + 5x_3^2}{3x_1 + 5x_3}, \frac{7x_2^2 + 4x_4^2}{7x_2 + 4x_4}\right\}$$

is an $\vec{r}$-increasing proper preaggregation function.

B. Choquet-Like Construction Method of Preaggregation Functions

This method is inspired in the way the Choquet integral is built, replacing the product operation in (6) by other aggregation functions.

Let $\mathfrak{m} : 2^N \to [0,1]$ be a fuzzy measure and $M:[0,1]^2 \to [0,1]$ be a function such that $M(0,x) = 0$ for every $x \in [0,1]$. Taking as basis the Choquet integral, we define the function $C_{\mathfrak{m}}^M : [0,1]^n \to [0,n]$ by

$$C_{\mathfrak{m}}^M(\mathbf{x}) = \sum_{i=1}^{n} M\left(x_{(i)} - x_{(i-1)}, \mathfrak{m}\left(A_{(i)}\right)\right) \qquad (7)$$

TABLE II
SOME PREAGGREGATION FUNCTIONS OBTAINED USING THE T-NORMS

T-Norm	Resulting pre-aggregation function
Minimum	$C_{\mathfrak{m}}^{T_M}(\mathbf{x}) = \sum_{i=1}^{n} \min\left\{x_{(i)} - x_{(i-1)}, \mathfrak{m}\left(A_{(i)}\right)\right\}$
Łukasiewicz	$C_{\mathfrak{m}}^{T_Ł}(\mathbf{x}) = \sum_{i=1}^{n} \max\left\{0, x_{(i)} - x_{(i-1)} + \mathfrak{m}\left(A_{(i)}\right) - 1\right\}$
Drastic Product	$C_{\mathfrak{m}}^{DP}(\mathbf{x}) = \sum_{i=1}^{n} \begin{cases} x_{(1)}, & \text{if } i = 1 \\ \mathfrak{m}\left(A_{(i)}\right), & \text{if } x_{(i)} - x_{(i-1)} = 1 \\ 0, & \text{otherwise} \end{cases}$
Nilpotent Minimum	$C_{\mathfrak{m}}^{NM}(\mathbf{x}) = \sum_{i=1}^{n} \begin{cases} \min\left\{x_{(i)} - x_{(i-1)}, \mathfrak{m}\left(A_{(i)}\right)\right\}, & \text{if } x_{(i)} - x_{(i-1)} + \mathfrak{m}\left(A_{(i)}\right) > 1 \\ 0, & \text{otherwise} \end{cases}$
Hamacher Product	$C_{\mathfrak{m}}^{HP}(\mathbf{x}) = \sum_{i=1}^{n} \begin{cases} 0, & \text{if } x_{(i)} = x_{(i-1)} \text{ and } \mathfrak{m}\left(A_{(i)}\right) = 0 \\ \dfrac{\left(x_{(i)} - x_{(i-1)}\right) \cdot \mathfrak{m}\left(A_{(i)}\right)}{x_{(i)} - x_{(i-1)} + \mathfrak{m}\left(A_{(i)}\right) - \left(x_{(i)} - x_{(i-1)}\right) \cdot \mathfrak{m}\left(A_{(i)}\right)}, & \text{otherwise} \end{cases}$

where $N = \{1, \ldots, n\}$, $(x_{(1)}, \ldots, x_{(n)})$ is an increasing permutation on the input $\mathbf{x}$, that is, $0 \le x_{(1)} \le \cdots \le x_{(n)}$, with the convention that $x_{(0)} = 0$, and $A_{(i)} = \{(i), \ldots, (n)\}$ is the subset of indices of $n - i + 1$ largest components of $\mathbf{x}$. Note that $C_{\mathfrak{m}}^{M}$ is well defined by (7) even if the permutation is not unique.

Now, we have the following result.

Theorem 4.1: Let $M : [0,1]^2 \to [0,1]$ be a function such that for all $x, y \in [0,1]$ it satisfies $M(x,y) \le x$, $M(x,1) = x$, $M(0,y) = 0$ and M is (1,0)-increasing. Then, for any fuzzy measure $\mathfrak{m}$, $C_{\mathfrak{m}}^{M}$ is a preaggregation function that is idempotent and averaging, i.e.,

$$\min(x_1, \ldots, x_n) \le C_{\mathfrak{m}}^{M}(x_1, \ldots, x_n) \le \max(x_1, \ldots, x_n).$$

Proof: Note that

$$\begin{aligned} C_{\mathfrak{m}}^{M}(x_1, \ldots, x_n) &= \sum_{i=1}^{n} M\left(x_{(i)} - x_{(i-1)}, \mathfrak{m}\left(A_{(i)}\right)\right) \\ &\le \sum_{i=1}^{n} (x_{(i)} - x_{(i-1)}) \\ &= x_{(n)} = \max(x_1, \ldots, x_n). \end{aligned}$$

From these two inequalities, idempotency follows. Besides

$$\begin{aligned} \min(x_1, \ldots, x_n) = x_{(1)} &= M\left(x_{(1)} - x_{(0)}, \mathfrak{m}\left(A_{(1)}\right)\right) \\ &\le C_{\mathfrak{m}}^{M}(x_1, \ldots, x_n). \end{aligned}$$

Finally, take $\vec{r} = \vec{1} = (1, \ldots, 1)$. Note that in (7), for $i \ge 2$, it follows that, for any $c > 0$,

$$\begin{aligned} &M\left(x_{(i)} + c - (x_{(i-1)} + c), \mathfrak{m}\left(A_{(i)}\right)\right) \\ &= M\left(x_{(i)} - x_{(i-1)}, \mathfrak{m}\left(A_{(i)}\right)\right) \end{aligned}$$

whereas, for $i = 1$

$$\begin{aligned} M\left(x_{(1)} + c - x_{(0)}, \mathfrak{m}\left(A_{(1)}\right)\right) &= M\left(x_{(1)} + c, \mathfrak{m}\left(A_{(1)}\right)\right) \\ &\ge M\left(x_{(1)}, \mathfrak{m}\left(A_{(1)}\right)\right). \end{aligned}$$

Therefore, $C_{\mathfrak{m}}^{M}$ is $\vec{1}$-increasing. ∎

Remark 4.1: Under the constraints of Theorem 4.1, we cannot ensure the monotonicity of $C_{\mathfrak{m}}^{M}$, i.e., $C_{\mathfrak{m}}^{M}$ is, in general, a proper preaggregation function. To see it, observe the following.

1) Take $M(x,y) = T_M(x,y)$. Consider $N = \{1,2,3,4\}$ and the uniform measure $\mathfrak{m} = \mathfrak{m}_U$ given in (1). Then, we have that

$$\begin{aligned} C_{\mathfrak{m}}^{T_M}(0.05, 0.1, 0.7, 0.9) &= 0.8, \text{ whereas} \\ C_{\mathfrak{m}}^{T_M}(0.05, 0.1, 0.8, 0.9) &= 0.7. \end{aligned}$$

Therefore, $C_{\mathfrak{m}}^{T_M}$ is not an increasing function, and hence, it is not an aggregation function.

2) Consider the Łukasiewicz t-norm $T_Ł(x,y) = \max\{0, x + y - 1\}$. Again, for $N = \{1,2,3,4\}$ and the uniform measure $\mathfrak{m} = \mathfrak{m}_U$, we have that

$$\begin{aligned} C_{\mathfrak{m}}^{T_Ł}(0.05, 0.1, 0.7, 0.9) &= 0.15\ , \quad \text{whereas} \\ C_{\mathfrak{m}}^{T_Ł}(0.05, 0.2, 0.7, 0.9) &= 0.05\ . \end{aligned}$$

Therefore, $C_{\mathfrak{m}}^{T_Ł}$ is not an increasing function, and hence, it is not an aggregation function. Analogous counterexamples can be found for the cases of the drastic product, the Hamacher product, or the nilpotent minimum t-norms.

Consider $N = \{1, \ldots, n\}$ and a fuzzy measure $\mathfrak{m} : 2^N \to [0,1]$. In Table II, we present the value of $C_{\mathfrak{m}}^{T}$, which are preaggregation functions but not aggregation functions, for the different t-norms given in Table I.

C. Sugeno-Like Construction Method of Preaggregation Functions

In this section, we follow the notation of Definition 2.4. Recall that the formula for the discrete Sugeno integral $S_{\mathfrak{m}} : [0,1]^n \to [0,1]$ can be written as

$$S_{\mathfrak{m}}(\mathbf{x}) = \bigvee_{i=1}^{n} \min\left\{x_{(i)}, \mathfrak{m}\left(A_{(i)}\right)\right\}.$$

Inspired by this formula, for any function $M : [0,1]^2 \to [0,1]$, we define the function $S_{\mathfrak{m}}^M : [0,1]^n \to [0,1]$ by the formula

$$S_{\mathfrak{m}}^M(\mathbf{x}) = \bigvee_{i=1}^{n} M\left(x_{(i)}, \mathfrak{m}\left(A_{(i)}\right)\right). \tag{8}$$

We prove a sufficient condition for M ensuring that $S_{\mathfrak{m}}^M$ is a preaggregation function for any fuzzy measure $\mathfrak{m}$.

Proposition 4.2: Let $M : [0,1]^2 \to [0,1]$ be a function increasing in the first variable, and let for each $y \in [0,1]$, $M(0,y) = 0$ and $M(1,1) = 1$. Then, $S_{\mathfrak{m}}^M$ defined in (8) is a preaggregation function for any fuzzy measure $\mathfrak{m}$.

Proof: It is easy to check that, for any $\mathfrak{m}$,

$$S_{\mathfrak{m}}^M(\mathbf{0}) = \bigvee_{i=1}^{n} M\left(0, \mathfrak{m}\left(A_{(i)}\right)\right) = 0$$

and

$$S_{\mathfrak{m}}^M(\mathbf{1}) = \bigvee_{i=1}^{n} M\left(1, \mathfrak{m}\left(A_{(i)}\right)\right)$$
$$= M(1, \mathfrak{m}(A_{(1)})) = M(1,1) = 1.$$

Moreover, for vector $\vec{1} = (1,\ldots,1)$, we get

$$S_{\mathfrak{m}}^M(\mathbf{x} + c\vec{1}) = \bigvee_{i=1}^{n} M\left(x_{(i)} + c, \mathfrak{m}\left(A_{(i)}\right)\right)$$
$$\geq \bigvee_{i=1}^{n} M\left(x_{(i)}, \mathfrak{m}\left(A_{(i)}\right)\right) = S_{\mathfrak{m}}^M(\mathbf{x})$$

i.e., $S_{\mathfrak{m}}^M$ is $\vec{1}$-increasing, which completes the proof that $S_{\mathfrak{m}}^M$ is a preaggregation function. ■

Note that any function M satisfying the constraints of Proposition 4.2 is, in fact, a binary $(1,0)$-increasing preaggregation function that satisfies $M(0,y) = 0$ for each $y \in [0,1]$.

Example 4.2: (i) Let $M : [0,1]^2 \to [0,1]$ be any aggregation function. Then, $S_{\mathfrak{m}}^M : [0,1]^n \to [0,1]$ is also an aggregation function, independently of $\mathfrak{m}$.

(ii) Consider the function F, $F(x,y) = x|2y-1|$. Note that F is a proper preaggregation function which satisfies the constraints of Proposition 4.2, and thus, for any $\mathfrak{m}$, the function $S_{\mathfrak{m}}^F : [0,1]^n \to [0,1]$, $S_{\mathfrak{m}}^F(\mathbf{x}) = \bigvee_{i=1}^{n} F\left(x_{(i)}, \mathfrak{m}\left(A_{(i)}\right)\right)$ is a preaggregation function (even an aggregation function thought F is not).

For example, for $n = 2$, $\mathfrak{m}(\{1\}) = 1/3$, $\mathfrak{m}(\{2\}) = 3/4$, we get

$$S_{\mathfrak{m}}^F(x,y) = \begin{cases} x \vee \dfrac{y}{2}, & \text{if } x \leq y \\ y \vee \dfrac{x}{3}, & \text{if } x > y. \end{cases}$$

V. Fuzzy Reasoning Method Using Preaggregation Functions

In this section, we present a generalization of the FRM proposed by Barrenechea *et al.* [24], using the proposed preaggregation functions, which are the result of combining different t-norms and fuzzy measures. To do so, we first explain the components of standard FRBCSs, and then, the new FRM is introduced.

A classification problem consists of m training examples $\mathbf{x}_p = (x_{p1},\ldots,x_{pn}, y_p)$, with $p = 1,\ldots,m$, where x_{pi}, with $i = 1,\ldots,n$, is the value of the ith attribute variable and $y_p \in \mathbb{C} = \{C_1, C_2, \ldots, C_M\}$ is the label of the class of the pth training example.

Among all the techniques used to face classification problems, one of the most used are the FRBCSs [25], since they allow the inclusion of all the available information in the system modeling, generating an interpretable model, and providing accurate results. The two main components of FRBCSs are the following.

1) The Knowledge Base containing the Rule Base and the Data Base, where the fuzzy inference rules and the membership functions are stored, respectively.
2) The Fuzzy Reasoning Mechanism, which is used to classify examples using the information available in the Knowledge Base.

The choice of the aggregation function plays a crucial role in FRBCSs [26], [27], since it determines the behavior of the FRM [28]. This is due to the fact that in the FRM, the local information given by each fuzzy rule is aggregated to provide global information, which is associated with each class of the problem [27]–[31]. Finally, the example is assigned to the class having the *maximum* global information.

The usage of the *maximum* as the aggregation function in the FRM to obtain the global information is very common in the literature, which is known as the FRM of the WR [27], [28], [32], [33]. However, whenever one considers, for each class, just the information given by a single fuzzy rule having the highest compatibility with the example, the available information provided by the remaining fuzzy rules of the system is ignored.

Denote by $x_p = (x_{p1},\ldots,x_{pn})$ the n-dimensional vector of attribute values corresponding to an example $\mathbf{x}_p$. The fuzzy rules that are used in this study are of the following form:

Rule R_j :

$$\text{If } x_{p1} \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_{pn} \text{ is } A_{jn} \text{ then } x_p \text{ in } C_j^k \text{ with } RW_j \tag{9}$$

where R_j is the label of the jth rule, A_{ji} is an antecedent fuzzy set modeling a linguistic term, C_j^k is the label of the consequent fuzzy set C^k modeling the class associated with the rule R_j, with $k \in \{1,\ldots,M\}$, and $RW_j \in [0,1]$ is the rule weight [34].

Let $x_p = (x_{p1},\ldots,x_{pn})$ be a new example to be classified, L the number of rules in the rule base, and M the number of classes of the problem. The new FRM using preaggregation functions presents the following steps.

Matching degree: It is the strength of the activation of the if-part of the rules for the example x_p, which is computed using a t-norm $T' : [0,1]^n \to [0,1]$:

$$\mu_{A_j}(x_p) = T'(\mu_{A_{j1}}(x_{p1}),\ldots,\mu_{A_{jn}}(x_{pn})), \text{ with } j = 1,\ldots,L. \tag{10}$$

Association degree: It is the association degree of the example x_p with the class of each rule in the rule base, given by

$$b_j^k(x_p) = \mu_{A_j}(x_p) \cdot RW_j^k, \text{ with } k = \text{Class}(R_j), \quad j = 1, \ldots, L. \tag{11}$$

Example classification soundness degree for all classes: In this step, we apply preaggregation functions [see (7)] to combine the association degrees calculated in the previous step, obtaining the classification soundness degrees, defined by

$$Y_k(x_p) = C_{\mathfrak{m}}^T \left(b_1^k(x_p), \ldots, b_L^k(x_p)\right), \text{ with } k = 1, \ldots, M \tag{12}$$

where $C_{\mathfrak{m}}^T$ is the obtained preaggregation, which is the result of combining a bivariate t-norm $T : [0,1]^2 \to [0,1]$ and a fuzzy measure $\mathfrak{m} : 2^N \to [0,1]$.

Since, whenever $b_i^k(x_p) = 0$, it holds that

$$C_{\mathfrak{m}}^T(b_1^k(x_p), \ldots, b_L^k(x_p)) = C_{\mathfrak{m}}^T(b_1^k(x_p), \ldots, b_{j-1}^k(x_p), b_{j+1}^k(x_p), \ldots, b_L^k(x_p))$$

then, for practical reasons, only $b_j^k > 0$ are considered in (12).

Classification: A decision function $F : [0,1]^M \to \{1, \ldots, M\}$ defined over the example classification soundness degrees of all classes and determining the class corresponding to the maximum soundness degree is given by

$$F(Y_1, \ldots, Y_M) = \min_{k=1\ldots M} k \text{ such that } Y_k = \max_{w=1,\ldots,M}(Y_w). \tag{13}$$

In practical applications, it is sufficient to consider

$$F(Y_1, \ldots, Y_M) = \arg\max_{k=1,\ldots,M}(Y_k). \tag{14}$$

Barrenechea *et al.* proposed to use the classical Choquet integral (product t-norm) instead of preaggregation in (12). They also considered tuning the exponent of the power measure using an evolutionary algorithm [24]. Specifically, they used the CHC evolutionary model [35], which was used to define the most suitable exponent to be used for each class.[3] We denote this proposal as power measure genetically adjusted (Power_GA).

VI. Analysis of the Application of Preaggregation Functions in Classification Problems

This section is aimed at providing an application of preaggregation functions in real-world problems. Specifically, as introduced in Section V, we consider to introduce this new theory to extend the FRM proposed by Barrenechea *et al.* [24], which was applied to tackle classification problems.

The aim of the experimental study is to see whether the usage of preaggregation functions in this FRM allows the results of the classical Choquet integral (product t-norm) to be enhanced. To do so, we test the performance of the FRM using 30 different preaggregation functions, which are all the possible combinations among the six t-norms shown in Table I and the five fuzzy measures (see Section II) considered in this paper. Finally, as it was done in [24], we also analyze if the best FRM (the best preaggregation) is able to enhance the results of the well-known FRM of the WR, that is, the usage of the maximum to aggregate the information in the third step of the FRM described in Section V. Consequently, we want to show that the usage of preaggregation functions allows the results obtained with two classical averaging operators to be enhanced.

[3] See [24] for a detailed explanation of the evolutionary algorithm.

TABLE III
Datasets Used in This Study

Id.	Dataset	#Inst	#Att	#Class
App	Appendiciticis	106	7	2
Bal	Balance	625	4	3
Ban	Banana	5300	2	2
Bnd	Bands	365	19	2
Bup	Bupa	345	6	2
Cle	Cleveland	297	13	5
Eco	Ecoli	336	7	8
Gla	Glass	214	9	6
Hab	Haberman	306	3	2
Hay	Hayes-Roth	160	4	3
Iri	Iris	150	4	3
Led	Led7digit	500	7	10
Mag	Magic	1902	10	2
New	Newthyroid	215	5	3
Pag	Pageblocks	5472	10	5
Pho	Phoneme	5404	5	2
Pim	Pima	768	8	2
Rin	Ring	740	20	2
Sah	Saheart	462	9	2
Sat	Satimage	6435	36	7
Seg	Segment	2310	19	7
Tit	Titanic	2201	3	2
Two	Twonorm	740	20	2
Veh	Vehicle	846	18	4
Win	Wine	178	13	3
Wis	Wisconsin	683	11	2
Yea	Yeast	1484	8	10

In the remainder of this section, we first explain the adopted experimental framework (see Section VI-A), and then, we present the results as well as their corresponding analysis (see Section VI-B).

A. Experimental Framework

We use 27 real-world datasets selected from the KEEL dataset repository [36]. Table III summarizes the properties of these datasets, showing, for each dataset, the identifier (Id.) as well as the name (Dataset), the number of instances (#Inst), the number of attributes (#Att), and the number of classes (#Class). The *magic*, *page-blocks*, *penbased*, *ring*, *satimage*, and *twonorm* datasets have been stratified sampled at 10% in order to reduce their size for training. Examples with missing values have been removed, e.g., in the *wisconsin* dataset.

We adopt the model proposed in [24], [37], and [38], that is, a fivefold cross-validation model, where a dataset is split in five partitions randomly, each partition with 20% of the examples, and a combination of four of them is then used for training and the other is used for testing. This process is repeated five times by using a different partition to test the system each time. For each partition, the output is computed as the mean of the numbers of correctly classified examples divided by the total number of examples for each partition, that is, the accuracy rate. Then, we consider the average result of the five partitions as the final classification rate of the algorithm. This procedure is a standard for testing the performance of classifiers [39], [40].

We use FARC-HD [10] to accomplish the fuzzy rule learning process. We have considered the following configuration: the product t-norm as the conjunction operator T', the Certainty Factor as the rule weight RW_j, five linguistic labels per variable, 0.05 for the minimum support, 0.8 as the threshold for the confidence, the depth of the search trees is limited to 3, and the parameter determining the number of fuzzy rules that cover each example, k_t, is set to 2. For the genetic process, we have used populations composed of 50 individuals, 30 bits per gen for the Gray codification (for incest prevention), and 20,000 as the maximum number of iterations. Finally, for the Dirac fuzzy measure, the value of the variable *i* used to decide if $i \in A$, for $A \subseteq N = \{0, \ldots, n\}$, we adopt the median value, is given by

$$i = \begin{cases} \dfrac{n+1}{2}, \text{ if } n \text{ is odd} \\ \dfrac{n}{2} + 1, \text{ if } n \text{ is even.} \end{cases}$$

In order to give statistical support to the analysis of the results, we consider the usage of hypothesis validation techniques [41], [42]. Specifically, we use nonparametric tests, since the initial conditions that guarantee the reliability of the parametric tests cannot be performed [43].

In fact, we use the aligned Friedman test [44] to detect statistical differences among a group of results and to show how good a method is with respect to the others. In this method, the algorithm achieving the lowest average ranking is the best one. Furthermore, we apply the post-hoc Holm's test [45] to study whether the best method rejects the equality hypothesis with respect to its partners. The post-hoc procedure allows us to know whether a hypothesis of comparison could be rejected at a specified level of significance α. Specifically, we compute the adjusted *p*-value (APV) to take into account that multiple tests are conducted. As a result, we can directly compare the APV with the level of significance α so as to be able to reject the null hypothesis.

Finally, we also consider the usage of the Wilcoxon test [46] in order to perform pairwise comparisons.

B. Experimental Results

The summary of the results provided by all the different configurations of the FRM, i.e., all the preaggregation functions, is introduced in Table IV. Each column of this table shows the results obtained using the fuzzy measure reported in its top cell using the six t-norms, which are shown by rows. The number in each cell is the average of the accuracy rate obtained in the 27 datasets by the corresponding preaggregation function. The best result for each fuzzy measure is highlighted in **boldface**. The number in brackets is the number of datasets in which each t-norm has obtained the best performance for each fuzzy measure (ties are excluded). The detailed results obtained in each dataset are available in the Appendix.

TABLE IV
AVERAGED RESULTS OBTAINED BY THE DIFFERENT PREAGGREGATION FUNCTIONS CONSIDERED IN THE STUDY

	Uniform	Dirac	Wmean	OWA	Power_GA
Product	78.68 (7)	**78.01 (3)**	78.12 (4)	77.33 (4)	78.55 (5)
Minimum	78.85 (7)	77.81 (0)	78.75 (7)	78.33 (10)	79.00 (7)
Łukasiewicz	76.61 (1)	77.81 (1)	76.92 (0)	76.44 (1)	78.14 (0)
Drastic	76.66 (0)	77.81 (0)	76.66 (1)	76.66 (2)	76.66 (1)
Nilpotent	76.88 (1)	77.81 (0)	76.76 (3)	76.60 (1)	78.78 (5)
Hamacher	**79.16 (8)**	77.81 (1)	**79.19 (10)**	**78.61 (7)**	**79.42 (7)**

From these results, we can observe two situations.

1) The performance of the product, minimum, and Hamacher is, in general, clearly better than that of Łukasiewicz, Drastic product, and Nilpotent minimum.
2) The performance of all the t-norms using the Dirac's measure is almost the same.

The reason implying the low performance of Łukasiewicz, Drastic product, and Nilpotent product is that after aggregating a set of values, the obtained one is similar to that obtained if we aggregated them using the minimum function (not the preaggregation associated with the minimum), which usually obtains poor results. The explanation is as follows: Let x and y be the result of the fuzzy measure and the subtraction of the elements to be aggregated using the Choquet integral, respectively.

1) *Łukasiewicz:* $x + y - 1$ is lower than 0 on half of its domain. Therefore, most of the time we do not add anything, which implies obtaining the minimum or a value close to it.
2) *Drastic product:* The value of the fuzzy measure is never 1 (except when we have all the elements), and it is very difficult to have a difference between two values to be aggregated equal to 1. Therefore, most of the time we add 0.
3) *Nilpotent minimum:* In the same way as Łukasiewicz, on half of the domain, $x + y$ is greater than 1. Consequently, we also add 0 most of the times.

Regarding the behavior of the Dirac's measure, the similar behavior among all the t-norms is due to the fact that this measure returns always either 1 or 0. Furthermore, it is known that $T(x_{(i)} - x_{(i-1)}, 0) = 0$ and $T(x_{(i)} - x_{(i-1)}, 1) = x_{(i)} - x_{(i-1)}$, for any t-norm T. Consequently, the selected t-norm T does not have a great influence on the results of the preaggregation functions.

Due to the aforementioned poor results obtained when applying Łukasiewicz, Drastic product, and Nilpotent minimum, we focus the remainder of the analysis on the product, minimum, and Hamacher t-norms.

From the results in Table IV, we can observe that, with the exception of the Dirac's fuzzy measure, the results of the Hamacher t-norm are better than those of the minimum t-norm, which in turn are better than the ones of the product. This trend is also present, in general, on the number of datasets in which each of these t-norms obtain the best result.

TABLE V
ALIGNED FRIEDMAN AND HOLM TESTS TO COMPARE THE DIFFERENT PREAGGREGATION FUNCTIONS

	Uniform	Dirac	WMean	OWA	Power_GA
Product	42.94 (0.21)	**38.13**	51.09 (<u>0.002</u>)	53.91 (<u>0.003</u>)	50.78 (<u>0.004</u>)
Minimum	45.13 (0.21)	43.38 (0.771)	42.13 (<u>0.054</u>)	35.24 (0.828)	41.20 (0.112)
Hamacher	**50.22**	41.18 (0.771)	**29.78**	**33.85**	**31.02**

TABLE VI
WILCOXON TEST TO COMPARE THE PRODUCT (R^+) VERSUS THE MINIMUM (R^-)

Comparison	R^+	R^-	p-value
Uniform+Prod versus Uniform+Min	195.5	182.5	0.925
Dirac+Prod versus Dirac+Min	214	164	0.625
WMean+Prod versus WMean+Min	135.5	242.5	0.200
OWA+Prod versus OWA+Min	107.5	270.5	0.004
Power_GA+Prod versus Power_GA+Min	132	249	0.148

In order to support the previous findings, we carry out a statistical test to compare, for each fuzzy measure, the product, minimum, and Hamacher t-norms. To do so, we have used the Aligned Friedman test as well as the Holm's post-hoc test. The results of these statistical techniques are reported in Table V, where in each column, we find the different fuzzy measures, whereas the three t-norms are shown in rows. The number in each cell is the average rank computed with the aligned Friedman test, and the number in brackets is the APV computed with the Holm's test. The best t-norm for each fuzzy measure is the one with the less rank, which stressed in **boldface**, whereas the APV is <u>underlined</u> in case of statistical differences in favor to the best t-norm.

From the results in Table V, we can observe that the usage of the Hamacher t-norm provides the best behavior for all the fuzzy measures, with the exception of the one defined by Dirac due to the previous mentioned behavior. In fact, we find statistical differences with respect to the product when using the additive (WMean), symmetric (OWA), and Power_GA fuzzy measures and a low APV when using the uniform measure. Therefore, we can conclude that the usage of the Hamacher t-norm allows us to enhance the results of the product.

Furthermore, we also want to analyze if the minimum is also appropriate when compared with the usage of the product. To do so, we compare, for each fuzzy measure, the results provided by the product versus the ones of the minimum. To perform these comparisons, we have applied the Wilcoxon's test to conduct such pairwise comparisons. The obtained results are introduced in Table VI, where we can observe that when using the additive (WMean), symmetric (OWA), and Power_GA fuzzy measures there is a trend in favor to the minimum, whereas in the two remainder fuzzy measures, the behavior of these two t-norms is similar.

Finally, we want to study whether the results obtained by the best preaggregation function are able to improve those provided by the well-known FRM of the WR, that is, the usage of the maximum to aggregate the information. According to Table IV, we select the preaggregation function resulting of the combination among the Power_GA fuzzy measure and the Hamacher t-norm (Power_GA+Ham), since it provides the best average result. The results provided by this preaggregation function as well as those obtained with the WR are reported in Table VII, where the best result for each dataset is highlighted in **boldface**. From these results, it can be observed that the global behavior of Power_GA+Ham is better than that of the WR. This is due to the fact that Power_GA+Ham provides the best result in 17 out of the 27 datasets considered in the study. We also apply

TABLE VII
RESULTS IN TESTING PROVIDED BY CARD_GA+HAM AND WR

Dataset	WR	Power_GA+Ham
App	**84.89**	82.99
Bal	82.08	**82.72**
Ban	84.30	**85.96**
Bnd	68.56	**72.13**
Bup	61.16	**65.80**
Cle	55.23	**55.58**
Eco	75.61	**80.07**
Gal	**63.11**	63.10
Hab	71.22	**72.21**
Hay	79.46	**79.49**
Iri	**94.67**	93.33
Led	**69.80**	68.60
Mag	79.60	**79.76**
New	94.42	**95.35**
Pag	**94.52**	94.34
Pho	82.01	**83.83**
Pim	**75.38**	73.44
Rin	**90.00**	88.79
Sah	67.31	**70.77**
Sat	**80.40**	**80.40**
Seg	92.99	**93.33**
Tit	**78.87**	**78.87**
Two	84.32	**85.27**
Veh	67.62	**68.20**
Win	94.36	**96.63**
Wis	96.49	**96.78**
Yea	**56.54**	56.53
Mean	78.70	**79.42**

TABLE VIII
WILCOXON TEST TO COMPARE THE POWER MEASURE GENETICALLY ADJUSTED METHOD WITH THE HAMACHER T-NORM (R^+) VERSUS THE CLASSICAL FRM OF THE WR (R^-)

Comparison	R^+	R^-	p-value
Power_GA+Ham vs. WR	267.5	110.5	0.06

the Wilcoxon's test to support these findings, whose obtained results are shown in Table VIII. According to the statistical results, we can confirm with a high level of confidence that the usage of Power_GA+Ham is better than that of the WR.

VII. Conclusion

In this paper, based on the notion of an aggregation function, we have introduced the concept of a preaggregation function. We have described three construction methods for such functions. In particular, one of them derives from the Choquet integral by using other t-norms in the place of the product t-norm considered in the standard definition of the Choquet integral. Furthermore, we have proposed to apply this specific instance of preaggregation in the FRM of FRBCSs to aggregate the local information given by each fuzzy rule of the system.

In the experimental study, we have shown that the usage of the Hamacher or even the minimum t-norms allows one to improve the results obtained when applying the classical Choquet integral, that is, when using the product t-norm. Moreover, we have checked that the preaggregation providing the best results, which is obtained combining the Hamacher t-norm and the power measure genetically learnt, enhances the results achieved by the well-known FRM of the WR, that is, applying the maximum as the aggregation function. Therefore, the preaggregation functions introduced in this paper can offer greater flexibility for FRBCSs, enlarging the scope of the application of the approach proposed by Barrenechea *et al.* [24].

Future work is concerned with the study of the properties satisfied by the preaggregation functions, and the usage of overlap functions [6], [7], [47]–[49] for the generalization of the Choquet integral, also using a fuzzy interval approach [50]–[54], as, e.g., in [31], [33], and [55].

Appendix

The tables in this appendix present the obtained results in each dataset considering the different t-norms, for each fuzzy measure. Each table contains the results obtained with a different fuzzy measure:

1) Table IX: results of the uniform measure for the six t-norms.
2) Table X: results of the Dirac's fuzzy measure for the six t-norms.
3) Table XI: results of an additive fuzzy measure for the six t-norms.
4) Table XII: results of the ordered weighted averaging fuzzy measure for the six t-norms.
5) Table XIII: results of the genetic uniform fuzzy measure for the six t-norms.

The structure of these five tables is as follows: In each row, we find a dataset, and in each column, we introduce a different t-norm. The best result for each dataset is stressed in **boldface**.

TABLE IX
Detailed Results in Testing Using the Uniform Measure

Dataset	Product	Minimum	Lukasiewicz	Drastic	Nilpotent	Hamacher
App	86.80	84.89	**87.75**	83.03	82.12	85.89
Bal	78.24	**82.24**	75.04	76.80	77.12	80.96
Ban	**84.45**	83.38	82.70	82.72	81.91	84.19
Bnd	64.00	**70.24**	64.07	65.56	63.81	69.96
Bup	64.35	63.19	63.77	63.48	65.22	**65.80**
Cle	**57.57**	55.55	55.24	56.89	52.51	56.90
Eco	78.28	76.20	72.91	75.61	75.61	**79.17**
Gal	**65.90**	63.58	62.62	62.17	62.16	64.47
Hab	**74.50**	72.53	73.51	73.20	73.84	72.87
Hay	**81.00**	78.69	78.77	78.77	79.52	79.49
Iri	94.00	94.00	94.00	92.67	**94.67**	93.33
Led	68.20	68.80	67.40	67.00	68.40	**69.00**
Mag	79.02	79.49	76.50	77.28	76.97	**80.65**
New	94.42	**95.35**	93.02	92.56	92.56	94.88
Pag	94.16	93.80	93.61	**94.34**	94.16	94.34
Pho	83.14	81.92	80.18	79.70	79.81	**83.33**
Pim	72.26	**74.74**	71.62	72.65	72.40	74.48
Rin	85.81	**88.24**	78.38	78.11	79.59	87.43
Sah	**70.97**	70.55	68.83	68.61	69.70	69.68
Sat	**84.50**	81.80	77.76	78.38	76.36	79.47
Seg	92.60	93.07	90.00	90.69	89.74	**93.25**
Tit	**78.87**	**78.87**	**78.87**	**78.87**	**78.87**	**78.87**
Two	80.54	**83.24**	77.84	77.70	76.22	82.70
Veh	63.53	68.56	66.78	64.89	65.72	**69.03**
Win	94.37	93.81	88.71	88.73	95.49	**95.51**
Wis	95.90	**96.05**	95.02	95.32	94.44	95.76
Yea	56.94	56.26	53.44	53.97	**56.94**	56.00
Mean	78.68	78.85	76.61	76.66	76.88	**79.16**

TABLE X
Detailed Results in Testing Using the Dirac's Measure

Dataset	Product	Minimum	Lukasiewicz	Drastic	Nilpotent	Hamacher
Dataset	Product	Minimum	Lukasewitz	Drastic	Nilpo	Hamacher
App	**80.17**	**80.17**	**80.17**	**80.17**	**80.17**	**80.17**
Bal	**78.24**	**78.24**	77.60	**78.24**	78.08	**78.24**
Ban	**84.09**	**84.09**	**84.09**	**84.09**	**84.09**	**84.09**
Bnd	**70.67**	65.97	65.97	65.97	65.97	65.97
Bup	**64.06**	**64.06**	**64.06**	**64.06**	**64.06**	**64.06**
Cle	**55.56**	**55.56**	**55.56**	**55.56**	**55.56**	**55.56**
Eco	**77.70**	**77.70**	**77.70**	**77.70**	**77.70**	**77.70**
Gal	**64.98**	**64.98**	**64.98**	**64.98**	**64.98**	**64.98**
Hab	**71.23**	**71.23**	**71.23**	**71.23**	**71.23**	**71.23**
Hay	78.69	78.69	**79.46**	78.69	78.69	78.69
Iri	**93.33**	**93.33**	**93.33**	**93.33**	**93.33**	**93.33**
Led	68.00	68.00	68.00	68.00	68.00	**68.20**
Mag	**77.86**	**77.86**	**77.86**	**77.86**	**77.86**	**77.86**
New	**93.02**	**93.02**	**93.02**	**93.02**	**93.02**	**93.02**
Pag	**94.52**	**94.52**	**94.52**	**94.52**	**94.52**	**94.52**
Pho	**82.33**	**82.33**	**82.33**	**82.33**	**82.33**	**82.33**
Pim	**72.52**	**72.52**	**72.52**	**72.52**	**72.52**	**72.52**
Rin	**84.59**	**84.59**	**84.59**	**84.59**	**84.59**	**84.59**
Sah	**70.97**	68.82	68.82	68.82	68.82	68.82
Sat	**79.84**	78.85	78.85	78.85	78.85	78.85
Seg	**91.04**	**91.04**	**91.04**	**91.04**	**91.04**	**91.04**
Tit	**79.06**	**79.06**	**79.06**	**79.06**	**79.06**	**79.06**
Two	**82.30**	**82.30**	**82.30**	**82.30**	**82.30**	**82.30**
Veh	62.35	**64.66**	**64.66**	**64.66**	**64.66**	**64.66**
Win	**96.06**	**96.06**	**96.06**	**96.06**	**96.06**	**96.06**
Wis	**95.90**	**95.90**	**95.90**	**95.90**	**95.90**	**95.90**
Yea	**57.21**	**57.21**	**57.21**	**57.21**	**57.21**	**57.21**
Mean	**78.01**	77.81	77.81	77.81	77.80	77.81

TABLE XI
DETAILED RESULTS IN TESTING USING AN ADDITIVE MEASURE (WMEAN)

Dataset	Product	Minimum	Lukasiewicz	Drastic	Nilpotent	Hamacher
Dataset	Product	Minimum	Lukasewitz	Drastic	Nilpo	Hamacher
App	82.08	83.94	82.08	83.03	83.98	**85.84**
Bal	78.08	**81.60**	75.52	76.80	74.56	81.12
Ban	83.85	84.02	83.30	82.72	82.11	**84.47**
Bnd	61.33	**71.32**	69.83	65.56	68.20	67.99
Bup	64.35	61.16	65.22	63.48	**65.80**	65.51
Cle	57.56	55.24	54.86	56.89	56.22	**57.92**
Eco	**79.46**	78.86	73.53	75.61	76.19	76.49
Gal	63.54	**64.05**	62.62	62.17	63.57	64.02
Hab	72.54	70.91	73.19	**73.20**	70.24	72.21
Hay	77.98	78.69	78.77	78.77	**79.52**	79.49
Iri	93.33	94.00	93.33	92.67	**94.00**	93.33
Led	68.40	68.20	67.80	67.00	67.80	**69.40**
Mag	80.55	**80.76**	76.08	77.28	76.97	80.02
New	93.95	**94.88**	92.56	92.56	92.56	94.42
Pag	94.34	94.16	93.97	94.34	**94.71**	94.34
Pho	**82.51**	82.11	79.44	79.70	79.90	82.25
Pim	73.56	74.86	72.40	72.65	71.75	**75.78**
Rin	85.68	88.24	76.89	78.11	78.65	**88.78**
Sah	65.59	69.27	70.57	68.61	67.97	**71.21**
Sat	**81.40**	78.23	78.85	78.38	77.91	79.78
Seg	92.12	92.21	90.00	90.69	89.61	**92.86**
Tit	**78.87**	**78.87**	**78.87**	**78.87**	**78.87**	**78.87**
Two	82.03	83.65	77.97	77.70	75.14	**85.41**
Veh	**70.00**	68.67	64.89	64.89	64.42	69.86
Win	94.40	**95.48**	94.37	88.73	92.11	93.81
Wis	95.76	95.75	95.46	95.32	95.02	**97.07**
Yea	56.00	**57.21**	54.58	53.97	54.79	56.00
Mean	78.12	78.75	76.92	76.66	76.76	**79.19**

TABLE XII
DETAILED RESULTS IN TESTING USING A SYMMETRIC MEASURE (OWA)

Dataset	Product	Minimum	Lukasiewicz	Drastic	Nilpotent	Hamacher
Dataset	Product	Minimum	Lukasewitz	Drastic	Nilpo	Hamacher
App	83.03	82.99	82.12	83.03	**88.66**	84.85
Bal	78.88	**82.56**	77.28	76.80	74.72	80.80
Ban	**84.55**	83.23	82.21	82.72	82.79	83.23
Bnd	61.33	68.26	64.95	65.56	66.61	**68.56**
Bup	62.90	61.74	65.51	63.48	63.19	**66.67**
Cle	53.20	55.23	53.54	**56.89**	55.56	56.21
Eco	**76.20**	75.90	73.82	75.61	75.03	74.12
Gal	63.09	62.64	61.23	62.17	64.03	**67.74**
Hab	73.19	71.89	**74.48**	73.20	72.88	71.57
Hay	78.75	**79.49**	78.77	78.77	78.77	79.49
Iri	92.00	**93.33**	92.00	92.67	91.33	92.00
Led	67.60	68.20	67.00	67.00	67.00	**68.40**
Mag	79.49	79.18	77.71	77.28	76.97	**80.13**
New	91.63	90.70	92.09	**92.56**	91.63	91.16
Pag	94.34	**95.25**	94.16	94.34	94.16	94.34
Pho	81.98	81.92	79.03	79.70	79.87	**82.72**
Pim	72.65	**75.00**	73.05	72.65	73.18	73.56
Rin	81.89	**86.76**	75.54	78.11	77.97	86.49
Sah	69.89	**70.99**	68.18	68.61	69.03	69.25
Sat	**79.84**	78.54	77.29	78.38	76.98	79.00
Seg	**92.21**	91.56	90.00	90.69	90.39	92.03
Tit	**78.87**	**78.87**	**78.87**	**78.87**	**78.87**	**78.87**
Two	81.35	86.35	77.03	77.70	74.59	**88.78**
Veh	64.12	**66.08**	64.30	64.89	62.29	66.07
Win	93.25	**94.37**	94.35	88.73	91.59	94.37
Wis	95.17	**96.34**	95.46	95.32	94.88	96.05
Yea	56.47	**57.68**	53.91	53.97	55.12	56.13
Mean	77.33	78.33	76.44	76.66	76.60	**78.61**

TABLE XIII
DETAILED RESULTS IN TESTING USING THE POWER MEASURE GENETICALLY ADJUSTED (POWER_GA)

Dataset	Product	Minimum	Lukasiewicz	Drastic	Nilpotent	Hamacher
App	80.13	81.17	81.17	83.03	**83.98**	82.99
Bal	82.40	**82.72**	80.32	76.80	81.28	82.72
Ban	**86.32**	85.28	84.40	82.72	84.21	85.96
Bnd	64.00	70.25	71.02	65.56	69.37	**72.13**
Bup	**66.96**	61.16	65.22	63.48	64.06	65.80
Cle	55.58	56.26	55.20	**56.89**	54.54	55.58
Eco	76.51	78.57	74.72	75.61	77.41	**80.07**
Gal	64.02	**64.96**	64.49	62.17	64.51	63.10
Hab	72.52	71.87	73.18	73.20	**73.84**	72.21
Hay	**79.49**	77.95	77.98	78.77	78.75	**79.49**
Iri	91.33	92.67	92.00	92.67	**94.00**	93.33
Led	68.20	**68.80**	68.20	67.00	68.40	68.60
Mag	78.86	**80.23**	79.39	77.28	79.55	79.76
New	94.88	93.95	93.49	92.56	93.95	**95.35**
Pag	94.16	94.16	94.70	94.34	**94.89**	94.34
Pho	82.98	82.61	81.25	79.70	81.11	**83.83**
Pim	74.60	**76.04**	74.09	72.65	74.47	73.44
Rin	**90.95**	90.27	88.65	78.11	89.19	88.78
Sah	68.82	**71.65**	70.56	68.61	70.55	70.77
Sat	79.84	79.47	78.07	78.38	79.63	**80.40**
Seg	**93.46**	92.42	90.74	90.69	90.39	93.33
Tit	**78.87**	**78.87**	**78.87**	**78.87**	**78.87**	**78.87**
Two	84.46	84.86	83.78	77.70	85.00	**85.27**
Veh	64.71	**68.44**	62.53	64.89	64.90	68.20
Win	93.79	95.51	93.78	88.73	**97.16**	96.63
Wis	**97.22**	96.63	95.76	95.32	96.63	96.78
Yea	55.73	56.33	56.33	53.97	56.47	**56.53**
Mean	78.55	79.00	78.14	76.66	78.78	**79.42**

REFERENCES

[1] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*. Berlin, Germany: Springer, 2007.

[2] M. Grabisch, J. Marichal, R. Mesiar, and E. Pap, *Aggregation Functions*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[3] H. Bustince, E. Barrenechea, T. Calvo, S. James, and G. Beliakov, "Consensus in multi-expert decision making problems using penalty functions defined over a Cartesian product of lattices," *Inf. Fusion*, vol. 17, pp. 56–64, 2014.

[4] H. Bustince, M. Pagola, R. Mesiar, E. Hüllermeier, and F. Herrera, "Grouping, overlaps, and generalized bientropic functions for fuzzy modeling of pairwise comparisons," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 3, pp. 405–415, Jun. 2012.

[5] H. Bustince, J. Montero, and R. Mesiar, "Migrativity of aggregation operators," *Fuzzy Sets Syst.*, vol. 160, no. 6, pp. 766–777, 2009.

[6] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, and R. Orduna, "Overlap functions," *Nonlinear Anal.*, vol. 72, no. 3-4, pp. 1488–1499, 2010.

[7] A. Jurio, H. Bustince, M. Pagola, A. Pradera, and R. Yager, "Some properties of overlap and grouping functions and their application to image thresholding," *Fuzzy Sets Syst.*, vol. 229, pp. 69–90, 2013.

[8] T. Wilkin and G. Beliakov, "Weakly monotone aggregation functions," *Int. J. Intell. Syst.*, vol. 30, pp. 144–169, 2015.

[9] H. Bustince, A. Kolesárová, J. Fernandez, and R. Mesiar, "Directional monotonicity of fusion functions," *Eur. J. Operational Research*, vol. 244, pp. 300–308, 2015.

[10] J. Alcala-Fdez, R. Alcala, and F. Herrera, "A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 5, pp. 857–872, Oct. 2011.

[11] G. Mayor and E. Trillas, "On the representation of some aggregation functions," in *Proc. IEEE Int. Symp. Multiple-Valued Logic*, 1986, pp. 111–114.

[12] E. P. Klement, R. Mesiar, and E. Pap, *Triangular Norms*. Dordrecht, The Netherlands: Kluwer, 2000.

[13] M. Grabisch and C. Labreuche, "A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid," *Ann.. Oper. Res.*, vol. 175, no. 1, pp. 247–286, 2010.

[14] J.-L. Marichal, “Aggregation of interacting criteria by means of the discrete Choquet integral,” in *Aggregation Operators* (ser. Studies in Fuzziness and Soft Computing), vol. 97, T. Calvo, G. Mayor, and R. Mesiar, Eds. Heidelberg, Germany: Physica-Verlag HD, 2002, pp. 224–244.

[15] R. R. Yager, “On ordered weighted averaging aggregation operators in multicriteria decision making,” *IEEE Trans. Syst. Man Cybern.*, vol. 18, no. 1, pp. 183–190, Jan./Feb. 1988.

[16] B. S. Ahn, “A priori identification of preferred alternatives of OWA operators by relational analysis of arguments,” *Inf. Sci.*, vol. 180, no. 23, pp. 4572–4581, 2010.

[17] Y.-L. He, J. N. Liu, Y.-X. Hu, and X.-Z. Wang, “OWA operator based link prediction ensemble for social network,” *Exp. Syst. Appl.*, vol. 42, no. 1, pp. 21–50, 2015.

[18] B. Llamazares, “Constructing choquet integral-based operators that generalize weighted means and OWA operators,” *Inf. Fusion*, vol. 23, pp. 131–138, 2015.

[19] J. M. Merigó and A. M. Gil-Lafuente, “Decision-making in sport management based on the OWA operator,” *Expert Syst. Appl.*, vol. 38, no. 8, pp. 10408–10413, 2011.

[20] X. Sang and X. Liu, “An analytic approach to obtain the least square deviation OWA operator weights,” *Fuzzy Sets Syst.*, vol. 240, pp. 103–116, 2014.

[21] H. Bustince, P. Burillo, and F. Soria, “Automorphisms, negations and implication operators,” *Fuzzy Sets Syst.*, vol. 134, no. 2, pp. 209–229, 2013.

[22] H. Bustince, M. Pagola, and E. Barrenechea, “Construction of fuzzy indices from fuzzy DI-subsethood measures: Application to the global comparison of images,” *Inf. Sci.*, vol. 177, no. 3, pp. 906–929, 2007.

[23] H. Bustince, E. Barrenechea, and M. Pagola, “Image thresholding using restricted equivalence functions and maximizing the measures of similarity,” *Fuzzy Sets Syst.*, vol. 158, no. 5, pp. 496–516, 2007.

[24] E. Barrenechea, H. Bustince, J. Fernandez, D. Paternain, and J. A. Sanz, “Using the Choquet integral in the fuzzy reasoning method of fuzzy rule-based classification systems,” *Axioms*, vol. 2, no. 2, pp. 208–223, 2013.

[25] H. Ishibuchi, T. Nakashima, and M. Nii, *Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*. Berlin, Germany: Springer-Verlag, 2004.

[26] H. Ishibuchi, T. Nakashima, and M. Nii, *Classification and Modeling with Linguistic Information Granules, Advanced Approaches to Linguistic Data Mining* (ser. Advanced Information Processing). Berlin, Germany: Springer, 2005.

[27] H. Ishibuchi and Y. Nojima, “Pattern classification with linguistic rules,” in *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models* (ser. Studies in Fuzziness and Soft Computing), vol. 220, H. Bustince, F. Herrera, and J. Montero, Eds. Berlin, Germany: Springer, 2008, pp. 377–395.

[28] O. Cordon, M. J. del Jesus, and F. Herrera, “A proposal on reasoning methods in fuzzy rule-based classification systems,” *Int. J. Approx. Reason.*, vol. 20, no. 1, pp. 21–45, 1999.

[29] M. Elkano, M. Galar, J. Sanz, A. Fernandez, E. Barrenechea, F. Herrera, and H. Bustince Sola, “Enhancing multi-class classification in FARC-HD fuzzy classifier: On the synergy between n-dimensional overlap functions and decomposition strategies,” *IEEE Trans. Fuzzy Syst.*, 2014, to be published.

[30] S. R. Samantaray, K. El-Arroudi, G. Joos, and I. Kamwa, “A fuzzy rule-based approach for islanding detection in distributed generation,” *IEEE Trans. Power Del.*, vol. 25, no. 3, pp. 1427–1433, Jul. 2010.

[31] J. A. Sanz, M. Galar, A. Jurio, A. Brugos, M. Pagola, and H. Bustince, “Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system,” *Appl. Soft Comput.*, vol. 20, pp. 103–111, 2014.

[32] J. Sanz, A. Fernández, H. Bustince, and F. Herrera, “A genetic tuning to improve the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets: Degree of ignorance and lateral position,” *Int. J. Approx. Reason.*, vol. 52, no. 6, pp. 751–766, 2011.

[33] J. A. Sanz, A. Fernández, H. Bustince, and F. Herrera, “Improving the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets and genetic amplitude tuning,” *Inf. Sci.*, vol. 180, no. 19, pp. 3674–3685, 2010.

[34] H. Ishibuchi and T. Nakashima, “Effect of rule weights in fuzzy rule-based classification systems,” *IEEE Trans. Fuzzy Syst.*, vol. 9, no. 4, pp. 506–515, Aug. 2001.

[35] L. J. Eshelman, “The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination,” in *Foundations of Genetic Algorithms*, G. J. E. Rawlings, Ed. San Francisco, CA, USA: Morgan Kaufmann, 1991, pp. 265–283.

[36] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera, “KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework,” *J. Multiple-Valued Logic Soft Comput.*, vol. 17, nos. 2/3, pp. 255–287, 2011.

[37] G. Lucca, J. Sanz, R. Vargas, G. P. Dimuro, B. Bedregal, and H. Bustince, “Using t-norms for generalizing Choquet integrals with an application to fuzzy rule-based classification systems,” in *Proc. III Brazilian Cong. Fuzzy Syst. Recent Adv. Fuzzy Syst.*, 2014, pp. 300–310.

[38] J. Sanz, A. Fernández, H. Bustince, and F. Herrera, “IVTURS: A linguistic fuzzy rule-based classification system based on a new interval-valued fuzzy reasoning method with tuning and rule selection,” *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 3, pp. 399–411, Jun. 2013.

[39] M. Galar, A. Fernández, E. Barrenechea, H. Bustince, and F. Herrera, “An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes,” *Pattern Recog.*, vol. 44, no. 8, pp. 1761–1776, 2011.

[40] M. Galar, A. Fernández, E. Barrenechea, H. Bustince, and F. Herrera, “A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches,” *IEEE Trans. Syst. Man, Cybern. C, Appl. Rev.*, vol. 42, no. 4, pp. 463–484, Jul. 2012.

[41] S. García, A. Fernández, J. Luengo, and F. Herrera, “A study of statistical techniques and performance measures for genetics-based machine learning: Accuracy and interpretability,” *Soft Comput.*, vol. 13, no. 10, pp. 959–977, 2009.

[42] D. Sheskin, *Handbook of parametric and nonparametric statistical procedures*, 2nd ed. London, U.K.: Chapman & Hall/CRC, 2006.

[43] J. Demšar, “Statistical comparisons of classifiers over multiple data sets,” *J. Mach. Learning Res.*, vol. 7, pp. 1–30, 2006.

[44] J. L. Hodges and E. L. Lehmann, “Ranks methods for combination of independent experiments in analysis of variance,” *Ann.. Math. Statist.*, vol. 33, pp. 482–497, 1962.

[45] S. Holm, “A simple sequentially rejective multiple test procedure,” *Scandinavian J. Statist.*, vol. 6, pp. 65–70, 1979.

[46] F. Wilcoxon, “Individual comparisons by ranking methods,” *Biometrics*, vol. 1, pp. 80–83, 1945.

[47] B. C. Bedregal, G. P. Dimuro, H. Bustince, and E. Barrenechea, “New results on overlap and grouping functions,” *Inf. Sci.*, vol. 249, pp. 148–170, 2013.

[48] G. P. Dimuro and B. Bedregal, “Archimedean overlap functions: The ordinal sum and the cancellation, idempotency and limiting properties,” *Fuzzy Sets Syst.*, vol. 252, pp. 39–54, 2014.

[49] G. P. Dimuro, B. Bedregal, and R. H. N. Santiago, “On (G, N)-implications derived from grouping functions,” *Inf. Sci.*, vol. 279, pp. 1–17, 2014.

[50] B. C. Bedregal, G. P. Dimuro, and R. H. S. Reiser, “An approach to interval-valued R-implications and automorphisms,” in *Proc. Joint 2009 Int. Fuzzy Syst. Assoc. World Cong. 2009 Eur. Soc. Fuzzy Logic Technol. Conf.*, 2009, pp. 1–6.

[51] B. C. Bedregal, G. P. Dimuro, R. H. N. Santiago, and R. H. S. Reiser, “On interval fuzzy S-implications,” *Inf. Sci.*, vol. 180, no. 8, pp. 1373–1389, 2010.

[52] G. Dimuro, “On interval fuzzy numbers,” in *Proc. Workshop-School Theoretical Comput. Sci.*, 2011, pp. 3–8.

[53] G. P. Dimuro, B. C. Bedregal, R. H. N. Santiago, and R. H. S. Reiser, “Interval additive generators of interval t-norms and interval t-conorms,” *Inf. Sci.*, vol. 181, no. 18, pp. 3898–3916, 2011.

[54] G. P. Dimuro, B. R. C. Bedregal, R. H. S. Reiser, and R. H. N. Santiago, “Interval additive generators of interval t-norms,” in *Proc. 15th Int. Workshop Logic, Lang. Inf. Comput.*, 2008, pp. 123–135.

[55] H. Bustince, M. Galar, B. Bedregal, A. Kolesárová, and R. Mesiar, “A new approach to interval-valued Choquet integrals and the problem of ordering in interval-valued fuzzy set applications,” *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 6, pp. 1150–1162, Dec. 2013.

Giancarlo Lucca received the Undergraduate and Master’s degrees from the Universidade Federal do Rio Grande, Natal, Brazil. He is currently working toward the Ph.D. degree with the Universidad Publica de Navarra, Pamplona, Spain, under the advising of Prof. H. Bustince.

José Antonio Sanz received the M.Sc. degree in computer sciences and the Ph.D. degree both from the Public University of Navarre, Pamplona, Spain, in 2008 and 2011, respectively.

He is currently an Associate Lecturer with the Department of Automatics and Computation, Public University of Navarre. He is the author of 16 published original articles in international journals and is involved in teaching artificial intelligence for students of computer sciences. His research interests include fuzzy techniques for classification problems, interval-valued fuzzy sets, genetic fuzzy systems, data mining and medical applications of soft computing techniques.

Dr. Sanz received the Best Paper Award in the FLINS 2012 international conference and the Pepe Millá Award in 2014. He is a Member of the European Society for Fuzzy Logic and Technology.

Graçaliz Pereira Dimuro received the Undergraduate degree from Universidade Catlica de Pelotas, Pelotas, Brazil, and the M.Sc. and Ph.D. degrees from the Instituto de Informtica of Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.

She is currently an Adjunct Professor with the Centro de Cincias Computacionais, Universidade Federal do Rio Grande, Rio Grande, Brazil, a Researcher of the Brazilian Research Funding Agency CNPQ, and is in a sabbatic research year with the Universidad Publica de Navarra, Pamplona, Spain, with a grant by the Science Without Borders Program (CNPq/Brazil).

Benjamín Bedregal was born in Arica, Chile. He received the Engineering degree in computing from the University of Tarapacá, Arica, Chile, the M.Sc. degree in informatics and the Ph.D. degree in computer sciences from the Federal University of Pernambuco, Recife, Brazil, in 1987 and 1996, respectively.

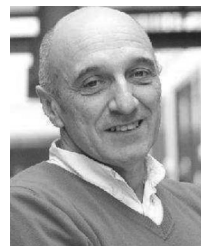
In 1996, he became an Assistant Professor with the Department of Informatics and Applied Mathematics, Federal University of Rio Grande do Norte, Natal, Brazil, where he is currently a Full Professor. His research interests include nonstandard fuzzy sets theory, aggregation functions, fuzzy connectives, clustering, fuzzy lattices, and fuzzy automata.

Radko Mesiar received the Graduate degree from the Faculty of Mathematics and Physics, Comenius University, Bratislava, Slovakia, in 1974, and the Ph.D. degree from the same faculty in 1979 with Ph.D. thesis "Subadditive martingale processes." He also received D.Sc. degree from the Academy of Sciences of the Czech Republic, Prague, Czech Republic, in 1996.

Since 1978, he has been with the Department of Mathematics, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, where he is currently the Head. He became an Associate Professor in 1983 and a Full Professor in 1998. He is a Fellow Member of the Institute of Information and Automation, Czech Academy of Sciences, Prague, in 1995, and of IRAFM, University of Ostrava, Ostrava, Czech Republic, in 2006. He is the coauthor of two scientific monographs: *Triangular Norms* (Norwell, MA, USA: Kluwer, 2000) and *Aggregation Functions* (Cambridge, U.K.: Cambridge Univ. Press, 2009), five edited volumes, and an author of more than 200 papers in WoS journals. He is the Founder and organizer of the International Conference on Fuzzy Set Theory and Applications and the International Summer School on Aggregation Operators.

Anna Kolesárová was born in Slovakia. She received the M.Sc. degree in mathematics and physics from Comenius University, Bratislava, Slovakia, and the Ph.D. degree from the Slovak Academy of Sciences, Bratislava.

In 2008, she became a Full Professor with the Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology, Bratislava. Her current research interests include aggregation functions, with a special stress to copulas, measures and integrals, decision making, and fuzzy arithmetic.



Humberto Bustince (M'08–SM'15) received the Ph.D. degree in mathematics from the Public University of Navarra, Pamplona, Spain, in 1994.

He is currently a Full Professor with the Department of Automatics and Computation, Public University of Navarra. He is the author of more than 65 published original articles and is involved in teaching artificial intelligence for students of computer sciences. His research interests include fuzzy logic theory, extensions of fuzzy sets (type-2 fuzzy sets, interval-valued fuzzy sets, Atanassov's intuitionistic fuzzy sets), fuzzy measures, aggregation functions, and fuzzy techniques for image processing.

Dr. Bustince is a Board Member of the European Society for Fuzzy Logic and Technology. He is the Editor-in-Chief of the *Mathware and Soft Computing Magazine*.