



HAL
open science

Block similarity in fuzzy tuples

Carl Frelicot, Hoel Le Capitaine

► **To cite this version:**

Carl Frelicot, Hoel Le Capitaine. Block similarity in fuzzy tuples. Fuzzy Sets and Systems, 2013, 220, pp.53-68. hal-00732835

HAL Id: hal-00732835

<https://hal.science/hal-00732835>

Submitted on 4 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Block similarity in fuzzy tuples

Carl Frélicot^a, Hoel Le Capitaine^{b,1,*}

^aUniversité La Rochelle, Laboratoire de Mathématiques, Image et Applications, EA 3165, F-17000 La Rochelle, France

^bUniversité Nantes, LINA, UMR CNRS 6241, Rue C. Pauc, F-44306 Nantes, France

Abstract

A common problem in decision-making is to analyze a tuple of numerical values associated with options, such as the degree of satisfaction assigned by experts to alternatives or probability values for hypotheses computed from data. With no loss of generality, it is assumed that the tuple contains values in the unit interval. For post-processing of typical value(s), singular values that may arise from noise in the data, or from unreliable experts, must not be taken into account. We present the concept of block similarity to address the problem of detecting subset(s) of typical values instead of extracting singular ones. The concept relies on suitable aggregation operators that combine the tuple components. Three different block similarity operators are proposed and discussed. These rely on Sugeno integrals of the tuple with respect to three different measures, namely cardinal weighting, symmetric kernel weighting and non-linear weighting. Numerical examples demonstrate their behaviors and their ability to detect blocks of similar values.

Keywords: Fuzzy connectives, aggregation operators, measures of information

1. Introduction

Many data mining processes involve algorithms that rely on the (dis)similarity of objects. Depending on the application domain or particular data, this notion varies [1] and many similarity measures have been defined using various mathematical properties or assumptions. We can first distinguish fuzzy comparison and characterization measures. On the one hand, comparison measures generally take two fuzzy tuples as inputs and then output a value representing the compatibility (e.g. distance, similarity, inclusion) of the inputs. Three main approaches that use fuzzy set theory have been reported so far. The first is based on metrics [2], mainly the Minkowski family of metrics. This approach originates from studies on how to measure the distance between two real functions and does not consider any specific interpretation of the values compared in that a metric comparison may not have any sense in the space considered. In particular, it is not restricted to inputs belonging to some specific metric spaces in which L_p norms are valid. The second way to compare fuzzy values comes from some set theory considerations whereby union, intersection and complementation are defined in a fuzzy set framework. Measures based on cardinality and possibility belong to this category [3, 4, 5]. The third is based on fuzzy implication functions or more

*Corresponding author

Email addresses: `carl.frelicot@univ-lr.fr` (Carl Frélicot), `hoel.lecapitaine@univ-nantes.fr` (Hoel Le Capitaine)

¹This work was mainly carried out while at Laboratoire de Mathématiques, Image et Applications.

generally on logical concepts [6, 7, 8]. It relies on the following assumption: two fuzzy quantities are considered as similar if one implies the other, and vice versa. By contrast, characterization measures take a single fuzzy tuple as the input and return values assessing the entropy, the ambiguity or the uncertainty that exists within the tuple. Measures of fuzziness are well-known examples of such measures that output a real value to reflect the fuzziness of a fuzzy set. The above-mentioned comparison measures can be used to define characterization measures. Given a suitable reference set Ω and a fuzzy set A , a characterization of A in terms of its similarity to Ω can be obtained. Cross and Sudkamp provide a general discussion on both similarity and compatibility measures within fuzzy set theory [9].

In this paper, we address the problem of measuring similarities within fuzzy tuples. Unlike conventional measures of similarity that compare two fuzzy sets, we are interested in comparing values within a single fuzzy set. We attach great importance to detection of specific information detection. Once information is extracted from raw data (e.g. a set of samples in pattern classification, a set of features for image analysis, a set of attributes of interest in multi-criteria decision-making), it is often associated with membership degrees through a convenient mapping. However, the resulting fuzzy tuple describing a given observation can be normal or not,² and may reveal different particular situations, such as some groups of values are similar whatever their magnitude, one value is much greater/less than the others, almost all values are large/small, and so on. Depending on the application, these situations may be considered positive or negative. For instance, in pattern classification, having two large and roughly equal membership degrees means that there is a doubt about the decision to be taken, while in multi-criteria decision making, it means that the two experts agree on the decision. Although the final decision and post-processing are completely different, it is interesting to extract such information in both applications. More generally, there is a need to define operators for detecting all blocks of similar values within a fuzzy tuple rather than only detecting how many and which values are high or low.

In order to illustrate this concept, we present in the following a numerical example that will be used in all the propositions of the paper.

Example 1. *Consider a coalition of six experts $\{A, B, C, D, E, F\}$ who assign a degree of satisfaction of 0.96, 0.89, 0.55, 0.45, 0.40 and 0.09, respectively, to one alternative. Overall similarity among expert judgments, indicated by blocks of similar values, gives higher confidence in the final decision.*

From this example, it is clear that some experts may be over-optimistic (experts A and B) or over-pessimistic (expert F), so it may be more appropriate to evaluate the similarity of judgments regardless of the highest and lowest scores (based on only C, D and E).

The remainder of the paper is organized as follows. Preliminary notions for aggregation functions and operators are first recalled in Section 2, followed by the introduction of bounded ordered weighted operators. Section 3 presents the concept of block similarity, along with three novel operators based on i) cardinal weighting, ii) kernel weighting and iii) non-linear weighting. For each of these operators, numerical examples are provided in Section 4 to demonstrate their advantages and drawbacks. Finally, Section 5 concludes and draws some perspectives.

2. Aggregation operators

Aggregation operators or functions are used to associate several real input values, say an n -tuple $\mathbf{u} = \{u_1, \dots, u_n\}$, with a representative single value. The values to be aggregated are defined on a finite real interval I or on an ordinal scale.

²A fuzzy tuple is said to be normal if its components sum up to 1.

Definition 1. [10] An aggregation function is a function $A : I^n \rightarrow I$, that

- (i) Is nondecreasing in each variable,
- (ii) Fulfills boundary conditions, and
- (iii) Is the identity function for the unary case.

Many families of aggregation functions exist and they take a wide variety of forms [11, 12]. These include triangular norms [13], ordered weighted averaging (OWA) operators [14], compensatory operators [15] and Choquet and Sugeno integrals [16]. Aggregation functions are generally classified either according to some mathematical properties they share (symmetry, associativity, monotonicity [10]) or the way the values are aggregated (conjunctive, disjunctive, compensatory [17]). They are used in many fields, such as decision-making and pattern recognition, for which various properties can be requested according to some specific applications. Although the values to be combined are different in multi-criteria decision-making and pattern classification, many authors consider that fuzzy membership functions play a role similar to utility functions [18]. We now focus on aggregation functions for which $I = [0, 1]$.

2.1. Triangular norms and conorms

Minimum and maximum operators were the earliest operators used to model set-theory operations on fuzzy sets, and they are now usually extended by triangular norms and conorms. These operators are commonly used because of their ability to generalize the logical AND and OR crisp operators to fuzzy sets [19].

Definition 2. [12] A triangular norm (*t-norm* for short) is an aggregation function $\top : [0, 1]^2 \rightarrow [0, 1]$ that is symmetric, associative and a conjunctor.³

According to the associative property, a t-norm is an n -ary aggregation operator. It is easy to show that the minimum operator

$$\top_M(\mathbf{u}) = \min(u_1, \dots, u_n) \quad (1)$$

is the largest t-norm. In the following, we use two other basic t-norms, namely the product \top_P and the Łukasiewicz t-norm \top_L , defined by

$$\top_P(\mathbf{u}) = \prod_{i=1}^n u_i \quad (2)$$

and

$$\top_L(\mathbf{u}) = \max\left(0, \sum_{i=1}^n u_i - (n-1)\right). \quad (3)$$

Some t-norms are Archimedean⁴. A useful representation theorem [20] of continuous Archimedean t-norms is given by

$$\top(\mathbf{u}) = f^{(-1)}\left(\sum_{i=1}^n f(u_i)\right), \quad (4)$$

³A conjunctor has a neutral element $e = 1$.

⁴A continuous t-norm is said to be Archimedean if $\top(u, u) < u$ for all $u \in [0, 1]$.

where $f : [0, 1] \rightarrow [0, \infty]$ is the strictly decreasing additive generator of \top , such that $f(1) = 0$, and $f^{(-1)} : [0, \infty] \rightarrow [0, 1]$ is the pseudo-inverse of f , defined as

$$f^{(-1)}(y) = \begin{cases} \sup\{x \in [0, 1] : f(x) < y\} & \text{for } f \text{ non-decreasing} \\ \sup\{x \in [0, 1] : f(x) > y\} & \text{for } f \text{ non-increasing.} \end{cases}$$

According to any fuzzy complement,⁵ we can define dual operators of t-norms.

Definition 3. [12] A triangular conorm (t-conorm, or s-norm for short) is an aggregation function $\perp : [0, 1]^2 \rightarrow [0, 1]$ that is symmetric, associative and a disjunctive.⁶

According to the associative property, an s-norm is an n -ary aggregation operator. It is easy to show that if we restrict ourselves to dual operators with respect to the strict negation,⁷ then the maximum operator

$$\perp_M(\mathbf{u}) = \max(u_1, \dots, u_n) \quad (5)$$

is the lowest s-norm. The s-norms dual to \top_P and \top_L are defined by

$$\perp_P(\mathbf{u}) = 1 - \prod_{i=1}^n (1 - u_i) \quad (6)$$

and

$$\perp_L(\mathbf{u}) = \min\left(1, \sum_{i=1}^n u_i\right). \quad (7)$$

2.2. Conjunctive and disjunctive combinations

Since triangular norms are not necessarily adequate for modeling generic aggregation problems because of their lack of compensation, numerous propositions involving a combination of t-norms and s-norms have been proposed [21]. These operators are generally called compensatory operators; uninorms [22], nullnorms [23], and exponential and linear convex combinations [15] are the most typical examples of this type of combination. Some special combinations of operators are of particular interest in the context of this paper. The first is a conjunctive combination of the disjunctive combinations applied to specific subsets of a n -tuple.

Definition 4. [24] Let \mathcal{P} be the power set of $N = \{1, 2, \dots, n\}$ and $\mathcal{P}_k = \{A \in \mathcal{P} : \text{card}(A) = k\}$. The k -order fuzzy OR (k fOR for short) is the aggregation function $\perp^k : [0, 1]^n \rightarrow [0, 1]$, defined by

$$\perp^k(\mathbf{u}) = \bigwedge_{A \in \mathcal{P}_{k-1}} \left(\perp_{j \in N \setminus A} u_j \right). \quad (8)$$

This operator evaluates whether k values in \mathbf{u} are high, and has been successfully used in pattern recognition problems for issues such as reject options in pattern classification [24] and cluster validity [25]. Additional properties are as follows [24]:

⁵A fuzzy complement is a continuous, monotonic and involutive function $C : [0, 1] \rightarrow [0, 1], u \mapsto C(u)$ that satisfies the boundary conditions $C(0) = 1$ and $C(1) = 0$.

⁶A disjunctive has a neutral element $e = 0$.

⁷The strict negation is the fuzzy complement defined by $N : [0, 1] \rightarrow [0, 1], u \mapsto N(u) = 1 - u$.

(i) $\stackrel{1}{\perp}(\mathbf{u}) = \perp(\mathbf{u})$ and $\stackrel{n}{\perp}(\mathbf{u}) = \top(\mathbf{u})$.

(ii) $\stackrel{k}{\perp}(\mathbf{u})$ is exactly the k th largest element of \mathbf{u} if the standard norms ($\top_M = \min$, $\perp_M = \max$) are taken.

A more general approach using (4) has been proposed [26]:

$$\stackrel{k}{\perp}(\mathbf{u}) = f^{(-1)}\left(\sum_{A \in \mathcal{P}_{k-1}} f\left(g^{(-1)}\left(\sum_{j \in N \setminus A} g(u_j)\right)\right)\right), \quad (9)$$

where f and g are the additive generators of the t-norm and s-norm, respectively. Note that f and g are linked by the relationship $g(u) = f(1 - u)$ if the t-norm and the s-norm are dual. The authors defined the dual of $kfOR$ (8) to evaluate whether k values in \mathbf{u} are low [26].

Definition 5. [26] Let \mathcal{P} be the power set of $N = \{1, 2, \dots, n\}$ and $\mathcal{P}_k = \{A \in \mathcal{P} : \text{card}(A) = k\}$. The k -order fuzzy AND ($kfAND$ for short) is the aggregation function $\stackrel{k}{\top} : [0, 1]^n \rightarrow [0, 1]$, defined by

$$\stackrel{k}{\top}(\mathbf{u}) = \perp_{A \in \mathcal{P}_{k-1}} \left(\top_{j \in N \setminus A} u_j \right). \quad (10)$$

Eq. (4) allows us to equivalently express $kfAND$ as:

$$\stackrel{k}{\top}(\mathbf{u}) = g^{(-1)}\left(\sum_{A \in \mathcal{P}_{k-1}} g\left(f^{(-1)}\left(\sum_{j \in N \setminus A} f(u_j)\right)\right)\right). \quad (11)$$

Additional properties are as follows [26]:

(i) $\stackrel{1}{\top}(\mathbf{u}) = \top(\mathbf{u})$ and $\stackrel{n}{\top}(\mathbf{u}) = \perp(\mathbf{u})$.

(ii) $\stackrel{k}{\top}(\mathbf{u})$ is exactly the k th lowest element of \mathbf{u} if the standard norms ($\top_M = \min$, $\perp_M = \max$) are taken.

A clear analogy between the $kfAND$ and the k th-order statistic (denoted OS_k) is evident. Moreover, some recent propositions, such as generalized medians [27] and k -tolerant capacities [28], exhibit close links to $kfAND$. Note that in a previous paper, we proposed a weighted variant of these operators in which each subset is associated with specific weights [26]. For clarity and brevity, such models are not discussed here.

One purpose of OWA operators is to relax the constraints of conjunctive and disjunctive operators. Like OWA operators, ordered weighted extrema operators allow association of weights to positions rather than to elements of a n -tuple.

Definition 6. [29] Let $\mathbf{w} \in [0, 1]^n$ be weights satisfying $\min(\mathbf{w}) = 0$. The ordered weighted minimum with respect to \mathbf{w} ($OWmin$ for short) is the aggregation function: $[0, 1]^n \rightarrow [0, 1]$, defined by

$$OWmin(\mathbf{u}) = \min_{k=1, n} (\max(u_{(k)}, w_k)), \quad (12)$$

where $u_{(\cdot)}$ is a permutation such that $u_{(1)} \geq \dots \geq u_{(n)}$.

Definition 7. [29] Let $\mathbf{w} \in [0, 1]^n$ be weights satisfying $\max(\mathbf{w}) = 1$. The ordered weighted maximum with respect to \mathbf{w} ($OWmax$ for short) is the aggregation function: $[0, 1]^n \rightarrow [0, 1]$, defined by

$$OWmax(\mathbf{u}) = \max_{k=1, n} (\min(u_{(k)}, w_k)). \quad (13)$$

Note that in the original definition, $u_{(k)}$ are sorted in increasing order, and special cases of original operators are retrieved up to a convenient permutation on the weights. In particular, it can be shown that if $w_{n-i+1} = 1$ and $w_k = 0$ for $k < n - i + 1$, then $OWmax$ is the i th-order statistic. An analogy can be found between the weighted extrema operators (12)–(13) and the k -order fuzzy connectives (8)–(11) provided the standard norms ($\top_M = \min$, $\perp_M = \max$) are taken. However, there are also some clear differences. Similarities and differences are illustrated by the following example.

Example 2. Consider $N = \{1, 2, 3, 4\}$ and $\mathbf{u} = (u_{(1)}, u_{(2)}, u_{(3)}, u_{(4)})$ such that $u_{(k)} \geq u_{(k+1)}$ for $k = 1, 2, 3$.⁸ $OWmin(\mathbf{u})$ is then obtained by computing

$$\min\left(\max(u_{(1)}, w_1), \max(u_{(2)}, w_2), \max(u_{(3)}, w_3), \max(u_{(4)}, w_4)\right). \quad (14)$$

By choosing $k = 3$ for the disjunctive combination of the $kfOR$, we have $\mathcal{P}_2 = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}$. Applying (8) with min–max norms, we obtain

$$\min\left(\max(u_{(3)}, u_{(4)}), \max(u_{(2)}, u_{(4)}), \max(u_{(2)}, u_{(3)}), \max(u_{(1)}, u_{(4)}), \max(u_{(1)}, u_{(3)}), \max(u_{(1)}, u_{(2)})\right). \quad (15)$$

According to (14) and (15), it is easy to see that:

- The number of disjunctive combinations to be conjunctively combined is differs, $card(\mathcal{P}_{k-1}) = \binom{n}{k-1}$ for $kfOR$ and n for $OWmin$.
- $OWmin$ combines values and weights, while $kfOR$ combines values and values. Setting weights in (12) so that $OWmin$ is exactly $kfOR$ is not trivial in the general case. In this example, in which $kfOR$ gives the third largest value $u_{(3)}$, we may refer to the case in which $OWmin$ reduces to the k th-order statistic.
- Since $kfOR$ does not require specification of weights, it can be freely used without additional considerations. Conversely, specific information obtained from data cannot be taken into account.

Contrary to a standard approach based on distances between values in which dimensions are considered one at a time, the k -order fuzzy connectives (8)–(11) can be used to consider and evaluate the similarity of consecutive (ordered) values. Unfortunately, one of these values is, by construction, the largest (for $kfOR$) or smallest (for $kfAND$) value. In some applications in which the values of interest do not include one of these special cases, it is necessary to relax this boundary constraint.

2.3. Bounded ordered weighted operators

To get rid of the largest (or smallest) values, we first introduce a slight modification of the ordered weighted extrema operators (12)–(13).

Definition 8. Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. Let $\mathbf{w} \in [0, 1]^n$ be weights satisfying $\max_{k \in \{i, \dots, j\}} w_k = 1$. The (i, j) -bounded ordered weighted maximum with respect to \mathbf{w} ($OWmax_{(i,j)}$ for short) is the aggregation function $[0, 1]^n \rightarrow [0, 1]$, defined by

$$OWmax_{(i,j)}(\mathbf{u}) = \max_{k=i,j}(\min(u_{(k)}, w_k)). \quad (16)$$

⁸Note that \mathbf{u} does not need to be sorted; we adopt this notation for clarity.

It is easy to prove that $OWmax_{(i,j)}$ is idempotent and symmetric and has the following compensative property:

$$\min(\mathbf{u}) \leq u_{(j)} = \min(u_{(i)}, \dots, u_{(j)}) \leq OWmax_{(i,j)}(\mathbf{u}) \leq \max(u_{(i)}, \dots, u_{(j)}) = u_{(i)} \leq \max(\mathbf{u}).$$

We now look at some special cases obtained by setting different bounds (i, j) and weights \mathbf{w} :

- If $(i, j) = (1, n)$, then $OWmax_{(i,j)}$ reduces to $OWmax$.
- If $i = j$, then $OWmax_{(i,i)} = \min(u_{(i)}, w_i) = u_{(i)}$ because of the maximum constraint on \mathbf{w} .
- If $w_i = 1$ and $w_k = 0$ for all $k \neq i$, then $OWmax_{(i,j)}(u_1, \dots, u_n) = u_{(i)}$.
- If $w_j = 1$ and $w_k = 0$ for all $k \neq j$, then $OWmax_{(i,j)}(u_1, \dots, u_n) = u_{(j)}$; more generally, if $w_l = 1$ and $w_k = 0$ for all $k \neq l$, then $OWmax_{(i,j)}(u_1, \dots, u_n) = u_{(l)}$.

This last property means that if \mathbf{w} focuses on a particular index k , the overall aggregation will be based only on the values w_k and $u_{(k)}$. Note that if $w_k = 1$, then for any weights w_l such that $j \geq l > k \geq i$, $OWmax_{(i,j)}$ remains unchanged. If $w_k = 0$, then it can be considered that the k th largest value is disregarded. However, it plays a role in the ordering process, in the sense that the associated weights will vary with the ordered values. Similar properties hold for the bounded ordered weighted minimum operator below.

Definition 9. Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. Let $\mathbf{w} \in [0, 1]^n$ be weights satisfying $\min_{k \in \{i, \dots, j\}} w_k = 0$. The (i, j) -bounded ordered weighted minimum with respect to \mathbf{w} ($OWmin_{(i,j)}$ for short) is the aggregation function $[0, 1]^n \rightarrow [0, 1]$, defined by

$$OWmin_{(i,j)}(\mathbf{u}) = \min_{k=i,j} (\max(u_{(k)}, w_k)). \quad (17)$$

Because of the non-compensative behavior of min–max norms, we next extend the bounded ordered weighted extrema operators (17)–(16) by means of t-norms and s-norms, respectively. Use of Archimedean triangular norm couples (\top, \perp) then allows us to obtain an overall value that is more representative of all the values considered.

Definition 10. Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. Let $\mathbf{w} \in [0, 1]^n$ be weights satisfying $\min_{k \in \{i, \dots, j\}} w_k = 0$. The (i, j) -bounded ordered weighted t-norm with respect to \mathbf{w} ($OWtnorm_{(i,j)}$ for short) is the aggregation function $[0, 1]^n \rightarrow [0, 1]$, defined by

$$OWtnorm_{(i,j)}(\mathbf{u}) = \top_{k=i}^j (\perp(u_{(k)}, w_k)). \quad (18)$$

Definition 11. Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. Let $\mathbf{w} \in [0, 1]^n$ be weights satisfying $\max_{k \in \{i, \dots, j\}} w_k = 1$. The (i, j) -bounded ordered weighted s-norm with respect to \mathbf{w} ($OWsnorm_{(i,j)}$ for short) is the aggregation function $[0, 1]^n \rightarrow [0, 1]$, defined by

$$OWsnorm_{(i,j)}(\mathbf{u}) = \perp_{k=i}^j (\top(u_{(k)}, w_k)). \quad (19)$$

These operators are still symmetric, but not idempotent or compensatory.

3. Block similarity operators

3.1. Cardinal weighting

To evaluate the similarity of two consecutive values $u_{(k+1)}$ and $u_{(k)}$, we can use the ratio $u_{(k+1)}/u_{(k)}$. The output is large if the two values are close to each other and continuously decreases while $u_{(k+1)}$ decreases. Since the *kfOR* operator generalizes the k th largest value of \mathbf{u} , the ratio $\perp^{k+1}(\mathbf{u})/\perp^k(\mathbf{u})$ is a good candidate for assessing the similarity of two consecutive values. However, we are interested in the similarity of not just two but several consecutive values, such as the similarity of values between $u_{(i)}$ and $u_{(j>i)}$. Can the ratio $\perp^j(\mathbf{u})/\perp^i(\mathbf{u})$ be used as a measure of similarity for all norm couples (\top, \perp) ?

Definition 12. [30] Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. A block similarity operator ($BS_{(i,j)}$ for short) is an aggregation function $[0, 1]^n \rightarrow [0, 1]$ that

(i) Has lower and upper boundaries given by

$$BS_{(i,j)}(\mathbf{u}) = 0 \text{ whenever } u_{(i)} = 1 \text{ and } u_{(j)} = 0 \quad (20)$$

$$BS_{(i,j)}(\mathbf{u}) = 1 \Leftrightarrow u_{(i)} = u_{(j)}; \text{ and} \quad (21)$$

(ii) Is left-decreasing and right-increasing, i.e. $\forall 0 \leq \varepsilon \leq u_{(i-1)} - u_{(i)}$,

$$BS_{(i,j)}(u_{(1)}, \dots, u_{(i)} + \varepsilon, \dots, u_{(c)}) \leq BS_{(i,j)}(u_{(1)}, \dots, u_{(i)}, \dots, u_{(c)}) \quad (22)$$

and $\forall 0 \leq \varepsilon \leq u_{(j-1)} - u_{(j)}$,

$$BS_{(i,j)}(u_{(1)}, \dots, u_{(j)} + \varepsilon, \dots, u_{(c)}) \geq BS_{(i,j)}(u_{(1)}, \dots, u_{(j)}, \dots, u_{(c)}). \quad (23)$$

It is easy to show that $\perp^j(\mathbf{u})/\perp^i(\mathbf{u})$ is a block similarity operator for (\top_M, \perp_M) , but not for (\top_P, \perp_P) , since (22) is not satisfied. We overcome this drawback using a bounded ordered weighted s-norm, which is somewhat analogous to using a Sugeno integral.

Definition 13. [16] The Sugeno integral $\mathcal{S}_\mu(\mathbf{u})$ of \mathbf{u} with respect to a fuzzy measure μ is defined by

$$\mathcal{S}_\mu(\mathbf{u}) = \perp_{i=1}^n (u_{(i)} \top \mu(A_i)), \quad (24)$$

where $A_i = \{1, \dots, i\}$.

In the context of this paper, the most natural choice for a fuzzy measure is a cardinal one, in particular the following:

$$\mu_k(A_i) = \begin{cases} 0 & \text{if } \text{Card}(A_i) < k \\ 1 & \text{else} \end{cases}, \quad (25)$$

so that rewriting (24) as

$$\mathcal{S}_\mu(\mathbf{u}) = \left(\perp_{i=1}^{k-1} (u_{(i)} \top \mu_k(A_i)) \right) \perp \left(\perp_{i=k}^c (u_{(i)} \top \mu_k(A_i)) \right) \quad (26)$$

allows us to see that $\mathcal{S}_\mu(\mathbf{u})$ is

$$\mathcal{S}_\mu(\mathbf{u}) = \begin{cases} \perp_{i=k}^n u_{(i)} & \text{if } u_{(k-1)} > u_{(k)} \\ \perp_{i=j}^n u_{(i)} & \text{else, where } j \text{ is such that } u_{(j-1)} > u_{(j)} = \dots = u_{(k)} \end{cases}, \quad (27)$$

because of the absorbing (neutrality) property of 0 (1) for any t-norm (s-norm). The fuzzy measure μ_k satisfies the constraints on the weights as defined in the previous section: it lies in the unit interval and reaches a maximum value of 1 when $\text{card}(A_i) \geq k$. Note that the fuzzy measure (25) corresponds to k -tolerant capacities as proposed by Marichal [28]. However, the measure is used with a generalized Sugeno integral and therefore gives different properties.

Proposition 1. *Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. The (i, j) -bounded cardinal weighting measure ($CW_{(i,j)}$ for short) defined by*

$$CW_{(i,j)}(\mathbf{u}) = \mathcal{S}_{\mu}^j(\mathbf{u}) / \mathcal{S}_{\mu}^i(\mathbf{u}) \quad (28)$$

is a block similarity operator (Definition 12) for the standard (\top_M, \perp_M) and the product (\top_P, \perp_P) norm couples.

The proof is given in Appendix A.

Such block similarity operators are not very satisfactory because of the cardinal measure. Indeed, values outside of the block $\{i, \dots, j\}$ contribute asymmetrically to the similarity of values within the block. The contribution of values $\{u_{(j+1)}, \dots, u_{(n)}\}$ is equal to one, while values $\{u_{(1)}, \dots, u_{(i-1)}\}$ are not taken into account. Moreover, the cardinal measure equally weights all values to 1. One important point of the approach is that we want to be able to precisely set the different contributions of internal values. Consequently, we propose another weighting approach that enables us to take into account just the values of the block.

3.2. Symmetric block-similarity operator

To make the contributions decrease as the index moves away from the block bounds, we propose the use of a symmetric kernel function involving a free parameter $\lambda \in \mathbb{R}^+$ that allows us to set the area of influence of the kernel.

Definition 14. *Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. A bounded $OWsnorm$ of \mathbf{u} with respect to a symmetric kernel function $\mathcal{K}_{\lambda}(k, l)$ centered around $l \in \{i, \dots, j\}$ is defined by*

$$\mathcal{S}_{\mathcal{K}_{\lambda}}^k(\mathbf{u}) = \bigoplus_{\ell=\frac{i+j}{2}}^k (u_{(\ell)} \top \mathcal{K}_{\lambda}(\ell, k)) \quad (29)$$

if $(j - i)$ is even, and by

$$\mathcal{S}_{\mathcal{K}_{\lambda}}^{k^{\pm}}(\mathbf{u}) = \bigoplus_{\ell=\frac{i+j^{\pm}}{2}}^k (u_{(\ell)} \top \mathcal{K}_{\lambda}(\ell, k)) \quad (30)$$

if $(j - i)$ is odd, where k^{\pm} denotes the two possible values k^+ and k^- , meaning that the control variable ℓ starts from $\frac{i+j+1}{2}$ or $\frac{i+j-1}{2}$.

Many symmetric functions are available, such as uniform, Gaussian, Epanechnikov and triangular functions [31]. We impose a normalization so that $\mathcal{K}(l, l) = 1$ and the corresponding $u_{(l)}$ are of maximum weight in the $OWsnorm$ operators (29)–(30). This normalization is imposed because the weights for the value of the bounds, i.e. $u_{(i)}$ and $u_{(j)}$, are the most important ones and decrease when moving toward the middle index, e.g. $\frac{i+j}{2}$ if $(j - i)$ is even.

Proposition 2. Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. The (i, j) -bounded symmetric kernel weighting measure ($KW_{(i,j)}^{\mathcal{K}_\lambda}$ for short) defined by

$$KW_{(i,j)}^{\mathcal{K}_\lambda}(\mathbf{u}) = \begin{cases} \frac{j}{i} \frac{\mathcal{S}_{\mathcal{K}_\lambda}^i(\mathbf{u})}{\mathcal{S}_{\mathcal{K}_\lambda}^j(\mathbf{u})} & \text{if } (j-i) \text{ is even} \\ \frac{j^+}{i^-} \frac{\mathcal{S}_{\mathcal{K}_\lambda}^i(\mathbf{u})}{\mathcal{S}_{\mathcal{K}_\lambda}^j(\mathbf{u})} & \text{if } (j-i) \text{ is odd,} \end{cases} \quad (31)$$

with the convention $KW_{(i,j)}^{\mathcal{K}_\lambda}(\mathbf{u}) = 1$ if $u_{(i)} = 0$, is a block similarity operator (Definition 12) for strictly continuous norm couples (\top, \perp) .

The proof is given in Appendix B.

Among the potential kernel functions, we focus, without loss of generality, on the Gaussian function used for numerical examples in next section, defined by

$$\mathcal{N}_\lambda(k, l) = \exp\left(-\frac{(k-l)^2}{\lambda}\right), \quad (32)$$

where λ controls the contribution width. It is easy to show that when the width λ tends to zero, then (32) tends to a Dirac function δ_l centered around l . Conversely, when λ tends to infinity, then (32) becomes the constant value 1, implying an equal contribution by all values within the block.

Proposition 3. As defined by (31), any (i, j) -bounded symmetric kernel weighting measure for any continuous couple (\top, \perp) is such that

$$\lim_{\lambda \rightarrow 0} KW_{(i,j)}^{\mathcal{K}_\lambda}(\mathbf{u}) = \begin{cases} 1 & \text{if } u_{(i)} = 0 \\ \frac{u_{(j)}}{u_{(i)}} & \text{else} \end{cases} \quad (33)$$

$$\lim_{\lambda \rightarrow +\infty} KW_{(i,j)}^{\mathcal{K}_\lambda}(\mathbf{u}) = \begin{cases} \frac{\prod_{k=\frac{i+j}{2}}^j u_{(k)}}{\prod_{k=\frac{i+j}{2}}^i u_{(k)}} & \text{if } (j-i) \text{ is even} \\ \frac{\prod_{k=\frac{i+j+1}{2}}^j u_{(k)}}{\prod_{k=\frac{i+j-1}{2}}^i u_{(k)}} & \text{if } (j-i) \text{ is odd} \\ 1 & \text{if } u_{(i)} = 0. \end{cases} \quad (34)$$

The proof is trivial and is left to the reader.

It is evident that the tuple $\{u_{(i+1)}, \dots, u_{(j-1)}\}$ is not taken into account when the width is close to zero, whereas all values from $u_{(i)}$ to $u_{(j)}$ are considered when λ is large. In other terms the contribution of intermediate values increases with λ . Note that $KW_{(i,i+1)}^{\mathcal{K}_\lambda}(\mathbf{u}) = \frac{u_{(i+1)}}{u_{(i)}}$ does not depend on λ , whatever the kernel function. This means that increasing λ will not make two consecutive (ordered) values in \mathbf{u} more similar but may increase the similarity of blocks of larger size.

3.3. Non-linear weighting through fuzzy measures

The concept of measure is very important in mathematics, in particular for integrals. Measures are usually additive, in the sense that the measure of two independent criteria is the sum of the individual measures. However, it would be very interesting to consider non-additive measures. For instance, consider the work carried out by a group of people. If we represent the working efficiency of each individual by a measure, the overall work efficiency will not be the sum of the measures, but depends on the interaction between the group members. This tool provides additional flexibility compared to operators such as weighted means and OWA operators. In the following we use the term fuzzy measure, or capacity, to denote a measure for which the additivity constraint is relaxed and replaced by monotony.

Definition 15. [16] A discrete fuzzy measure (or capacity) on a finite set $X = \{x_1, \dots, x_n\}$ is a function $\mu : 2^X \rightarrow [0, 1]$ that

- (i) Fulfills boundary conditions $\mu(\emptyset) = 0$ and $\mu(X) = 1$, and
- (ii) Is monotonic, i.e. $\mu(A) \leq \mu(B)$ if $A \subset B$.

The triplet $(X, \mathcal{P}(X), \mu)$ is said to be a fuzzy measure space.

In the following, we consider specific fuzzy measures defined iteratively.

Definition 16. [16] Let $X = \{x_1, \dots, x_n\}$ be a finite set and let $\lambda > -1$. A Sugeno λ -fuzzy measure on X is a discrete fuzzy measure μ satisfying the following:

If $A, B \in 2^X$ with $A \cap B = \emptyset$, then

$$\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\mu(A)\mu(B). \quad (35)$$

Given X with the boundary condition $\mu(X) = 1$, a numerical value for λ can be obtained by solving

$$1 + \lambda = \prod_{i=1}^n (1 + \lambda\mu(\{x_i\})). \quad (36)$$

The block similarity operators proposed in the previous subsections are defined as a ratio of two specific Sugeno integrals, one for each bounding index of the (i, j) block. We begin with the upper index j . Depending on the parity of $(i - j)$, we consider X as the set of indices $\{\frac{i+j+1}{2}, \dots, j\}$ or $\{\frac{i+j}{2}, \dots, j\}$, and define the fuzzy measures of singletons by:

$$\mu_1(\{k\}) = \frac{u(k)}{\sum_{\ell=i}^j u(\ell)}. \quad (37)$$

Then, if $(i - j)$ is even (the odd case is straightforward), the Sugeno integral of \mathbf{u} with respect to μ_1 can be written as:

$$\mathcal{S}_{\mu_1}^j(\mathbf{u}) = \bigwedge_{\ell=\frac{i+j}{2}}^j (u(\ell) \top \mu_1(A_\ell)), \quad (38)$$

where A_ℓ is a subset of X obtained by

$$A_\ell = \{f \in X \mid f \leq \ell\}. \quad (39)$$

For the lower index i , X is set of indices $\{i, \dots, \frac{i+j+1}{2}\}$ or $\{i, \dots, \frac{i+j}{2}\}$ depending on the parity of $(i - j)$. We define similar individual fuzzy measures of singletons by:

$$\mu_2(\{k\}) = \frac{u(k)}{\sum_{\ell=i}^j u(\ell)}. \quad (40)$$

Once λ has been obtained for the measure μ_2 , weights are sorted in decreasing order so that the maximum weight is associated with $u_{(i)}$, and the weights decrease thereafter. Then, if $(i - j)$ is even (the odd case is straightforward), the Sugeno integral of \mathbf{u} with respect to μ_1 can be written as:

$$\mathcal{S}_{\mu_2}^i(\mathbf{u}) = \bigwedge_{\ell=\frac{i+j}{2}}^i (u(\ell) \top \mu_2(A_\ell)), \quad (41)$$

where A_ℓ is defined by (39).

Proposition 4. Let $N = \{1, \dots, n\}$ and $(i, j) \in N \times N$, $i \leq j$. The (i, j) -bounded non linear weighting measure ($NLW_{(i,j)}$ for short) defined by

$$NLW_{(i,j)}(\mathbf{u}) = \begin{cases} \mathcal{S}_{\mu_1}^j(\mathbf{u})/\mathcal{S}_{\mu_2}^i(\mathbf{u}) & \text{if } (j-i) \text{ is even} \\ \mathcal{S}_{\mu_1}^{j^+}(\mathbf{u})/\mathcal{S}_{\mu_2}^{i^-}(\mathbf{u}) & \text{if } (j-i) \text{ is odd,} \end{cases} \quad (42)$$

with the convention $NLW_{(i,j)}(\mathbf{u}) = 1$ if $u_{(i)} = 0$, is a block similarity operator (Definition 12) for strictly continuous norm couples (\top, \perp) .

The proof is straightforward when considering the proofs already given in Appendix A and Appendix B.

With this last proposition, a user does not need to specify a resolution parameter as is the case with kernel-based operators. Instead, the fuzzy measure is built from the data:

- Weights for singletons are set according to their relative importance in the block.
- Weights for a coalition of subsets are set according to the updating formula (36).

The computational complexity of the proposed operators is similar to that of k -order statistics, since only sorting of complexity $n \log(n)$ is required if n values are considered. All subsequent operations are linear in time, regardless of the operator.

4. Numerical study

We illustrate how the three proposed (i, j) -bounded measures, namely cardinal weighting $CW_{(i,j)}$ (28), symmetric kernel weighting $KW_{(i,j)}^{\mathcal{K}_1}$ (31) and non-linear weighting $NLW_{(i,j)}$ (42), behave in the detection of block similarities within the tuple $\mathbf{u} = \{0.96, 0.89, 0.55, 0.45, 0.40, 0.09\}$ of Example 1. This tuple is sufficiently generic to show the behavior of the operators when faced with mixed inputs. In particular, the same results would be obtained using only $\mathbf{u} = \{0.96, 0.55, 0.45, 0.40\}$ or $\mathbf{u} = \{0.45, 0.40, 0.09\}$, and we let the reader imagine other tuples that could be used for further illustration purposes.

When considering the components of the tuple $\mathbf{u} = \{0.96, 0.89, 0.55, 0.45, 0.40, 0.09\}$, we would reasonably expect that applying a threshold to the output of the measure would allow detection of:

- A first block of two similar values $\{0.96, 0.89\}$ indexed by $\{1, 2\}$, and
- A second block of three similar values $\{0.55, 0.45, 0.40\}$ indexed by $\{3, 4, 5\}$.

The outputs of each measure considered are presented in tables comparing all possible bounding index couples ($i = 1, \dots, 6$ and $j = i, \dots, 6$). Results are given for two norm couples, namely the standard (\top_M, \perp_M) norms (left tables) and the product (\top_P, \perp_P) norms (right tables). Two different thresholds, $t_1 = 0.8$ and $t_2 = 0.7$, are used. Output values exceeding t_1 are denoted by a gray cell, while output values exceeding t_2 are in bold font. Note that the thresholds are not compared to the values in \mathbf{u} , but to the output of the operators (28), (31) and (42), so that there is no relationship between \mathbf{u} and the thresholds. However, a large operator output implies large block similarity, easing selection of the block for a given threshold in the expected block.

As shown in Table 1, the cardinal weighting measure CW does not allow proper evaluation of block similarities, whatever the norm couple (\top, \perp) , even if some can be accepted, as expected. Use of standard norms gives three blocks, $\{0.96, 0.89\}$, $\{0.55, 0.45\}$ and $\{0.45, 0.40\}$, for t_1 , while the two last blocks are fused, resulting in the two expected blocks for t_2 . However, for product norms, the unexpected blocks

\mathbf{u}	0.96	0.89	0.55	0.45	0.40	0.09	\mathbf{u}	0.96	0.89	0.55	0.45	0.40	0.09
$\downarrow i, j \rightarrow$	1	2	3	4	5	6	$\downarrow i, j \rightarrow$	1	2	3	4	5	6
1	1.000	0.927	0.573	0.469	0.417	0.094	1	1.000	0.986	0.865	0.700	0.454	0.090
2	–	1.000	0.618	0.506	0.449	0.101	2	–	1.000	0.878	0.710	0.461	0.091
3	–	–	1.000	0.818	0.727	0.164	3	–	–	1.000	0.809	0.525	0.104
4	–	–	–	1.000	0.889	0.200	4	–	–	–	1.000	0.649	0.129
5	–	–	–	–	1.000	0.225	5	–	–	–	–	1.000	0.198
6	–	–	–	–	–	1.000	6	–	–	–	–	–	1.000

Table 1: Values of the cardinal weighting operator $CW_{(i,j)}(\mathbf{u})$ for the standard (τ_M, \perp_M) (left) and product (τ_P, \perp_P) (right) norms

$\{0.96, 0.89, 0.55\}$ and $\{0.96, 0.89, 0.55, 0.45\}$ are detected. The reason for this, as already mentioned, is two-fold: the asymmetric contribution of external values and the equal weights of internal values, independent of their relative position in relation to the block bounds. The two other proposed measures correct these points.

Tables 2 and 3 show the importance of choosing the t-norm couple and the kernel parameter λ for the symmetric kernel weighting measure KW . Only the Gaussian kernel \mathcal{N}_λ is considered for simplicity. According to Table 2, for $\lambda = 0.5$ a relatively small value of the kernel parameter results in detection of the two expected blocks $\{0.96, 0.89\}$ and $\{0.55, 0.45, 0.40\}$ for t_2 and the three blocks $\{0.96, 0.89\}$, $\{0.55, 0.45\}$ and $\{0.45, 0.40\}$ for t_1 , whatever the norm couple.

A left–right comparison of Tables 2 and 3 (obtained for $\lambda = 10$) shows that product norms result in larger output values than standard norms for almost all entries. Since the norm couple affects the measure values, it is therefore possible to obtain blocks of larger size for a given threshold by changing the norm, provided the threshold is carefully chosen.

\mathbf{u}	0.96	0.89	0.55	0.45	0.40	0.09	\mathbf{u}	0.96	0.89	0.55	0.45	0.40	0.09
$\downarrow i, j \rightarrow$	1	2	3	4	5	6	$\downarrow i, j \rightarrow$	1	2	3	4	5	6
1	1.000	0.927	0.573	0.469	0.417	0.141	1	1.000	0.927	0.626	0.509	0.453	0.144
2	–	1.000	0.618	0.506	0.449	0.152	2	–	1.000	0.618	0.547	0.486	0.155
3	–	–	1.000	0.818	0.727	0.246	3	–	–	1.000	0.818	0.756	0.241
4	–	–	–	1.000	0.889	0.301	4	–	–	–	1.000	0.889	0.290
5	–	–	–	–	1.000	0.225	5	–	–	–	–	1.000	0.225
6	–	–	–	–	–	1.000	6	–	–	–	–	–	1.000

Table 2: Values of the kernel weighting operator $KW_{(i,j)}^{\mathcal{N}_{\lambda=0.5}}(\mathbf{u})$ for the standard (τ_M, \perp_M) (left) and product (τ_P, \perp_P) (right) norms.

Given a norm couple, increasing the kernel parameter λ increases the output values, as expected, and thus potentially the size of blocks detected for a fixed threshold, as revealed by comparison of Tables 2 and 3. Note that for Table 3 the kernel parameter is too large ($\lambda = 10$) to demonstrate how its tuning is important. If λ is too large, the KW measure will fail, regardless of the fine tuning of the threshold. For instance, it is evident that the unexpected blocks $\{0.96, 0.89, 0.55\}$ and $\{0.45, 0.40, 0.09\}$ are necessarily detected when using the standard norm couple. Nevertheless, the expected blocks are detected for convenient values of λ (Table 2). The next measure proposed does not require an external parameter, so it is more convenient for end users.

The outputs of the non-additive non-linear weighting measure NLW are given in Table 4. The two

\mathbf{u}	0.96	0.89	0.55	0.45	0.40	0.09	\mathbf{u}	0.96	0.89	0.55	0.45	0.40	0.09
$\downarrow i, j \rightarrow$	1	2	3	4	5	6	$\downarrow i, j \rightarrow$	1	2	3	4	5	6
1	1.000	0.927	0.927	0.573	0.573	0.469	1	1.000	0.927	0.920	0.729	0.779	0.597
2	–	1.000	0.618	0.618	0.506	0.506	2	–	1.000	0.618	0.766	0.682	0.618
3	–	–	1.000	0.818	0.818	0.727	3	–	–	1.000	0.818	0.879	0.572
4	–	–	–	1.000	0.889	0.889	4	–	–	–	1.000	0.889	0.646
5	–	–	–	–	1.000	0.225	5	–	–	–	–	1.000	0.225
6	–	–	–	–	–	1.000	6	–	–	–	–	–	1.000

Table 3: Values of the kernel weighting operator $KW_{(i,j)}^{N_{i=10}}(\mathbf{u})$ for the standard (\top_M, \perp_M) (left) and product (\top_P, \perp_P) (right) norms.

expected blocks $\{0.96, 0.89\}$ and $\{0.55, 0.45, 0.40\}$ are retrieved for t_2 , whatever the norm couple, and the three blocks $\{0.96, 0.89\}$, $\{0.55, 0.45\}$ and $\{0.45, 0.40\}$ are detected for t_1 . Compared to the outputs in Tables 2 and 3, it is evident that this measure provides a trade-off between a small λ -based KW measure and a large λ -based one. However, this trade-off is not uniform due to the non-linear weighting: depending on the block considered, differences can be very large or very small. Furthermore, NLW seems to be less sensitive to the norm couple than CW and KW .

\mathbf{u}	0.96	0.89	0.55	0.45	0.40	0.09	\mathbf{u}	0.96	0.89	0.55	0.45	0.40	0.09
$\downarrow i, j \rightarrow$	1	2	3	4	5	6	$\downarrow i, j \rightarrow$	1	2	3	4	5	6
1	1.000	0.927	0.572	0.469	0.469	0.417	1	1.000	0.927	0.677	0.515	0.576	0.359
2	–	1.000	0.618	0.506	0.449	0.449	2	–	1.000	0.618	0.552	0.482	0.379
3	–	–	1.000	0.818	0.727	0.218	3	–	–	1.000	0.818	0.748	0.229
4	–	–	–	1.000	0.889	0.266	4	–	–	–	1.000	0.889	0.278
5	–	–	–	–	1.000	0.225	5	–	–	–	–	1.000	0.225
6	–	–	–	–	–	1.000	6	–	–	–	–	–	1.000

Table 4: Values of the kernel weighting operator $NLW_{(i,j)}(\mathbf{u})$ for the standard (\top_M, \perp_M) (left) and product (\top_P, \perp_P) (right) norms.

The proposed operators can be used to assess the underlying ambiguity of a membership vector in pattern recognition problems. Consider a cluster validity problem that consists of finding an optimal number of clusters for a data set. Given a data point for which a tuple \mathbf{u} of cluster membership degrees are computed (using a clustering algorithm), standard approaches do not take into account the relative importance of consecutive (ordered) values, such as the entropy of the tuple given by $E(\mathbf{u}) = -\sum_{i=1}^n u_i \log(u_i)$. However, having two (or more) consecutive high values implies that the data point lies in an area in which several clusters overlap. This information, which has a large impact on the cluster validity measure, is not considered in standard approaches, but is in the proposed operators [25, 30]. Another interesting application is a multi-class selection problem in supervised pattern classification, in which a data point can be assigned to several classes according to a tuple of membership degrees obtained from a labeling process (e.g. Bayesian inference, support vector machine). The proposed operators allow selection of subsets of classes, as they present a block of similar membership values. Moreover, their output value can be used as a confidence measure to be associated with each selected subset of classes [32].

5. Conclusion

This paper presents the concept of block similarity in a tuple whose components lie in the unit interval. A block is defined as a subset of similar values. For detection of blocks in a tuple, three suitable aggregation

operators that combine the ordered components of the tuple are proposed and discussed. They all rely on a ratio of Sugeno integrals of the tuple with respect to a particular weighting measure and offer many combination possibilities using triangular norms. The first proposed measure, called cardinal weighting, equally weights the internal values of a possible block and suffers from asymmetric contribution by external values. Using symmetric kernel functions, these drawbacks are overcome in the second proposed measure, called kernel weighting. The kernel parameter that controls the area of influence of contributing values allows detection of blocks at different resolution levels, but its tuning may be delicate in practice. The third proposed measure, called non-linear weighting, uses two proper fuzzy measures for weighting. It may be more convenient for end users because it is parameter-free, but it loses some resolution potential. The behavior of each proposed measure was illustrated for a numerical example. The advantages and drawbacks of each proposed measure and its ability to detect blocks of similar values were demonstrated. The proposed aggregation operators can be used in any decision-making application in which a tuple of numerical values is associated with some options. Two potential uses in pattern recognition are the cluster validity problem, which deals with selection of the best number of groups from a clustering perspective [30], and the problem of rejecting some samples based on a confidence measure assigned in supervised classification [32].

In future research we would like to study the free parameters of the operators. In particular, for all operators the user must choose a suitable t-norm and s-norm. Intuitively, the minimum and maximum, for which the compensative property does not hold, should not be used when the inputs are interdependent. In this paper we restricted the focus to dual t-norms and s-norms, and it would be interesting to theoretically and experimentally analyze the behavior of the operators when the fuzzy connectives are not dual. Therefore, we plan a more extensive study of the behavior of the parameters and proposal of some general user guidelines when choosing 1) the connectives and 2) the kernel (function and resolution parameter), depending on the data to be analyzed. Finally, we would also like to investigate threshold learning so that similar blocks retrieved by the operator match user- or expert-defined blocks of similar values.

Appendix A. Proof of Proposition 1

We need to prove that $CW_{(i,j)}(\mathbf{u})$ satisfies boundary conditions (20) and (21) and monotony conditions (22) and (23).

Proof.

(20): If $\mathbf{u} = (u_{(1)}, \dots, u_{(i-1)}, 1, \dots, u_{(j-1)}, 0, \dots, u_{(n)})$, then $\mathcal{S}_{\mu}^j(\mathbf{u}) = 0$ and $\mathcal{S}_{\mu}^i(\mathbf{u}) = 1$ because 0 and 1 are the neutral and absorbing elements, respectively, of any t-conorm \perp , and therefore $CW_{(i,j)}(\mathbf{u}) = 0$.

(21) (\Leftarrow): If $u_{(i)} = u_{(j)}$, then $\mathcal{S}_{\mu}^j(\mathbf{u}) = \mathcal{S}_{\mu}^i(\mathbf{u})$ according to (27), and therefore $CW_{(i,j)}(\mathbf{u}) = 1$.

(21) (\Rightarrow):

- For $\perp = \perp_M = \max$, it is straightforward that if $CW_{(i,j)}(\mathbf{u}) = 1$, then $u_{(i)} = u_{(j)}$.
- For $\perp = \perp_P$, we use its strict monotony property.

If $CW_{(i,j)}(\mathbf{u}) = 1$, then there exists k in $\{1, \dots, j\}$ such that $\mathcal{S}_{\mu}^j(\mathbf{u}) = u_{(k)} \perp \dots \perp u_{(j)} \perp \dots \perp u_{(n)}$ and there exists $l \leq k$ such that $\mathcal{S}_{\mu}^i(\mathbf{u}) = u_{(l)} \perp \dots \perp u_{(k)} \perp \dots \perp u_{(n)} = u_{(l)} \perp \dots \perp u_{(k-1)} \perp \mathcal{S}_{\mu}^j(\mathbf{u})$.

Thus, $CW_{(i,j)}(\mathbf{u}) = \mathcal{S}_{\mu}^j(\mathbf{u}) / (u_{(l)} \perp \dots \perp u_{(k-1)} \perp \mathcal{S}_{\mu}^j(\mathbf{u})) = 1$.

For any strictly monotonic \perp , this implies $l = k$ and therefore $u_{(k)} = u_{(l)}$, which completes the proof.

(22): Let $u_{(i-1)} - u_{(i)} \geq \varepsilon > 0$. For all \perp , the following inequalities hold:

$$\begin{aligned} \mathcal{S}_\mu^j(u_{(1)}, \dots, u_{(i-1)}, u_{(i)} + \varepsilon, \dots, u_{(j)}, \dots, u_{(n)}) &\leq \mathcal{S}_\mu^j(u_{(1)}, \dots, u_{(i-1)}, u_{(i)}, \dots, u_{(j)}, \dots, u_{(n)}) \\ \mathcal{S}_\mu^i(u_{(1)}, \dots, u_{(i-1)}, u_{(i)} + \varepsilon, \dots, u_{(j)}, \dots, u_{(n)}) &\geq \mathcal{S}_\mu^i(u_{(1)}, \dots, u_{(i-1)}, u_{(i)}, \dots, u_{(j)}, \dots, u_{(n)}). \end{aligned}$$

Therefore,

$$CW_{(i,j)}(u_{(1)}, \dots, u_{(i-1)}, u_{(i)} + \varepsilon, \dots, u_{(j)}, \dots, u_{(n)}) \leq CW_{(i,j)}(u_{(1)}, \dots, u_{(i-1)}, u_{(i)}, \dots, u_{(j)}, \dots, u_{(n)}).$$

(23): Let $u_{(j-1)} - u_{(j)} \geq \varepsilon > 0$. For all \perp , we have to compare

$$\left(u_{(l)} \perp \dots \perp u_{(j)} + \varepsilon \perp \dots \perp u_{(n)} \right) / \left(u_{(k)} \perp \dots \perp u_{(i)} \perp \dots \perp u_{(l)} \perp \dots \perp u_{(j)} + \varepsilon \perp \dots \perp u_{(n)} \right) \text{ and}$$

$$\left(u_{(j)} \perp \dots \perp u_{(n)} \right) / \left(u_{(k)} \perp \dots \perp u_{(i)} \perp \dots \perp u_{(l)} \perp \dots \perp u_{(j)} \perp \dots \perp u_{(n)} \right), \text{ where } l > i.$$

Without loss of generality, we suppose that $u_{(i)} > u_{(j)} + \varepsilon$, because if $u_{(i)} = u_{(j)} + \varepsilon$, then

$$CW_{(i,j)}(u_{(1)}, \dots, u_{(j)} + \varepsilon, \dots, u_{(n)}) = 1, \text{ so that (23) is satisfied.}$$

• For $\perp = \perp_M = \max$, (23) is satisfied because

$$CW_{(i,j)}(u_{(1)}, \dots, u_{(j)} + \varepsilon, \dots, u_{(n)}) = CW_{i,j}(u_{(1)}, \dots, u_{(j)}, \dots, u_{(n)}).$$

• For $\perp = \perp_P$, we first extend the case to $[0, 1] \times \mathbb{R}^+$, so that given $a \in [0, 1], b, c \in \mathbb{R}^+, b \leq c, \gamma \geq 1$, we can write

$$a \perp_P \gamma b = a + \gamma b - \gamma a b \leq \gamma(a + b - ab) = \gamma(a \perp_P b), \quad (\text{A.1})$$

$$a \perp_P b \leq a \perp_P c \text{ by the monotony of } \perp. \quad (\text{A.2})$$

Let $\alpha \geq 1$ such that $(u_{(j)} + \varepsilon) \perp_P \dots \perp_P u_{(n)} = \alpha (u_{(j)} \perp_P \dots \perp_P u_{(n)})$. Consequently,

$$u_{(l)} \perp_P \dots \perp_P (u_{(j)} + \varepsilon) \perp_P \dots \perp_P u_{(n)} \geq \alpha (u_{(j)} \perp_P \dots \perp_P u_{(n)}). \quad (\text{A.3})$$

Let $v = u_{(j)} \perp_P \dots \perp_P u_{(n)}$; then $u_{(j-1)} \perp_P u_{(j)} \perp_P \dots \perp_P u_{(n)} = u_{(j-1)} \perp_P v$. Therefore,

$$u_{(j-1)} \perp_P \alpha v \leq \lambda (u_{(j-1)} \perp_P v) \text{ and } u_{(j-2)} \perp_P (u_{(j-1)} \perp_P \alpha v) \leq u_{(j-2)} \perp_P \alpha (u_{(j-1)} \perp_P v).$$

Let $w = u_{(j-1)} \perp_P v$. We similarly have $u_{(j-2)} \perp_P \alpha w \leq \alpha (u_{(j-2)} \perp_P w)$, and hence

$$u_{(j-2)} \perp_P (u_{(j-1)} \perp_P \alpha v) \leq \alpha (u_{(j-2)} \perp_P u_{(j-1)} \perp_P v).$$

Consequently, the following inequality holds:

$$\begin{aligned} u_{(k)} \perp_P \dots \perp_P u_{(i)} \perp_P \dots \perp_P u_{(l)} \perp_P \dots \perp_P (u_{(j)} + \varepsilon) \perp_P \dots \perp_P u_{(n)} \\ \leq \alpha (u_{(k)} \perp_P \dots \perp_P u_{(i)} \perp_P \dots \perp_P u_{(l)} \perp_P \dots \perp_P \dots \perp_P v). \end{aligned} \quad (\text{A.4})$$

Setting $A = \frac{u_{(l)} \perp_P \dots \perp_P (u_{(j)} + \varepsilon) \perp_P \dots \perp_P u_{(n)}}{u_{(i)} \perp_P \dots \perp_P u_{(j)} \perp_P \dots \perp_P u_{(l)} \perp_P \dots \perp_P (u_{(j)} + \varepsilon) \perp_P \dots \perp_P u_{(n)}}$, we have

$$\begin{aligned} A &\geq \frac{\alpha v}{u_{(k)} \perp_P \dots \perp_P u_{(j-1)} \perp_P \dots \perp_P \alpha v} \text{ by (A.3) and} \\ A &\geq \frac{\alpha v}{\alpha (u_{(i)} \perp_P \dots \perp_P u_{(j)} \perp_P \dots \perp_P u_{(l)} \perp_P \dots \perp_P u_{(j)} \perp_P \dots \perp_P u_{(n)})} \text{ by (A.4)} \\ &= \frac{v}{u_{(i)} \perp_P \dots \perp_P u_{(j)} \perp_P \dots \perp_P u_{(l)} \perp_P \dots \perp_P u_{(j)} \perp_P \dots \perp_P u_{(n)}}. \end{aligned}$$

It follows that $CW_{(i,j)}(u_{(1)}, \dots, (u_{(j)} + \varepsilon), \dots, u_{(n)}) \geq CW_{(i,j)}(u_{(1)}, \dots, u_{(j)}, \dots, u_{(n)})$, which completes the proof. \square

Appendix B. Proof of Proposition 2

We need to prove that $KW_{(i,j)}^{\mathcal{K}_\lambda}(\mathbf{u})$ satisfies the boundary conditions (20) and (21) and monotony conditions (22) and (23).

Proof.

(20) and

(21) (\Leftrightarrow) are straightforward because of the properties of (\top, \perp) and the symmetry of \mathcal{K}_λ .

(21) (\Rightarrow) Assume that $(j - i)$ is odd (the proof is the same for an even $(j - i)$) and let $x = \left(u_{\left(\frac{i+j}{2}\right)} \top \mathcal{K}\left(\frac{i+j}{2}, j\right)\right) \perp \dots \perp \left(u_{(j-1)} \top \mathcal{K}(j-1, j)\right)$ and $y = \left(u_{(i+1)} \top \mathcal{K}(i+1, i)\right) \perp \dots \perp \left(u_{\left(\frac{i+j}{2}\right)} \top \mathcal{K}\left(\frac{i+j}{2}\right)\right)$. Since $u_{(i)} > u_{(j)}$, we have $\frac{x \perp u_{(j)}}{u_{(i)} \perp y} = 1 \Leftrightarrow x \perp u_{(j)} = u_{(i)} \perp y$. The strict monotonicity of \perp implies that $x = y = 1$ because 1 is absorbing for any \perp . Therefore, $x = 1$ ($y = 1$) $\Leftrightarrow \mathcal{K}_\lambda = 1$ and $u_{\left(\frac{i+j}{2}\right)} = \dots = u_{(j-1)} = 1$ ($u_{(i+1)} = \dots = u_{\left(\frac{i+j}{2}\right)} = 1$). Finally, if the couple (\top, \perp) is strictly monotonic or if $\mathcal{K}_\lambda \neq 1$ or $(u_{(i+1)}, \dots, u_{(j-1)}) \neq (1, \dots, 1)$, then $KW_{(i,j)}^{\mathcal{K}_\lambda}(\mathbf{u}) \neq 1$, which completes the proof by contraposition.

Note that for norm couples (\top, \perp) that are not strictly continuous, such as $(\top, \perp)_S$ and $(\top, \perp)_L$, (21) (\Rightarrow) does not hold. However, in many applications such as pattern classification, management of a vector \mathbf{u} such as $(u_{(i+1)}, \dots, u_{(j-1)}) = (1, \dots, 1)$ is seldom required.

(22) and

(23) are straightforward because of the properties of (\top, \perp) and the symmetry of \mathcal{K}_λ . □

- [1] R. M. Gregson, *Psychometrics of Similarity*, Academic Press, New-York, 1975.
- [2] B. M. Ayyub, G. J. Klir, *Uncertainty Modeling and Analysis in Engineering and the Sciences*, Chapman & Hall, 2006.
- [3] B. Bouchon-Meunier, M. Rifqi, S. Bothorel, Towards general measures of comparison of objects, *Fuzzy Sets and Systems* 84 (1996) 143–153.
- [4] Y. Toliaas, S. Panas, L. Tsoukalas, Generalized fuzzy indices for similarity matching, *Fuzzy Sets and Systems* 120 (2001) 255–270.
- [5] B. D. Baets, S. Janssens, H. D. Meyer, On the transitivity of a parametric family of cardinality-based similarity measures, *International Journal of Approximate Reasoning* 50 (2009) 104 – 116.
- [6] W. Bandler, L. Kohout, Fuzzy power sets and fuzzy implication operators, *Fuzzy Sets and Systems* 4 (1980) 13–30.
- [7] H.-Y. Zhang, W.-X. Zhang, Hybrid monotonic inclusion measure and its use in measuring similarity and distance between fuzzy sets, *Fuzzy Sets and Systems* 160 (2009) 107–118.
- [8] H. Le Capitaine, C. Frélicot, Towards a unified logical framework of fuzzy implications to compare fuzzy sets, in: 13th International Fuzzy Systems Association World Congress, IFSA, Lisboa, Portugal, pp. 1200–1205.
- [9] V. Cross, T. Sudkamp, *Similarity and Compatibility in Fuzzy Set Theory*, Physica-Verlag; Heidelberg, 2002.
- [10] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation functions: Means, *Information Sciences* 181 (2011) 1–22.
- [11] C. Calvo, G. Mayor, R. Mesiar, *Aggregation Operators: New Trends and Applications*, Physica-Verlag, 2002.
- [12] M. Grabisch, J. Marichal, R. Mesiar, E. Pap, *Aggregation Functions*, Number 127 in *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, 2009.
- [13] K. Menger, Statistical metrics, *Proc. National Academy of Science USA* 28 (1942) 535–537.
- [14] R. Yager, Ordered weighted averaging aggregation operators in multi-criteria decision making, *IEEE Transactions on Systems, Man and Cybernetics* 18 (1988) 183–190.
- [15] H.-J. Zimmermann, P. Zysno, Latent connectives in human decision making, *Fuzzy Sets and Systems* 4 (1980) 37–51.
- [16] M. Sugeno, *Theory of fuzzy integrals and its applications*, Ph.D. thesis, Tokyo Institute of Technology, 1974.
- [17] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation functions: Construction methods, conjunctive, disjunctive and mixed classes, *Information Sciences* 181 (2011) 23–43.

- [18] R. E. Bellman, L. A. Zadeh, Decision-making in fuzzy environment, *Management Science* 17 (1970) 141–164.
- [19] E. P. Klement, R. Mesiar, Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms, Elsevier, 2005.
- [20] C. M. Ling, Representation of associative functions, *Publicationes Mathematica Debrecen* 12 (1965) 189–212.
- [21] A. Pradera, E. Trillas, T. Calvo, A general class of triangular norm-based aggregation operators: quasi-linear t - s operators, *International Journal of Approximate Reasoning* 30 (2002) 57–72.
- [22] R. Yager, A. Rybalov, Uninorm aggregation operators, *Fuzzy Sets and Systems* 80 (1996) 111–120.
- [23] T. Calvo, B. De Baets, J. Fodor, The functional equations of frank and alsina for uninorms and nullnorms, *Fuzzy Sets and Systems* 137 (2001) 3–10.
- [24] L. Mascarilla, M. Berthier, C. Frélicot, A k -order fuzzy or operator for pattern classification with k -order ambiguity rejection, *Fuzzy Sets and Systems* 159 (2008) 2011–2029.
- [25] H. Le Capitaine, C. Frélicot, A cluster validity index combining an overlap measure and a separation measure based on fuzzy aggregation operators, *IEEE Trans. Fuzzy Systems* 19 (2011) 580–588.
- [26] H. Le Capitaine, C. Frélicot, On (weighted) k -order fuzzy connectives, in: *19th IEEE International Conference on Fuzzy Systems*, 2010, Barcelona, Spain, pp. 1–8.
- [27] C. Calvo, R. Mesiar, Generalized medians, *Fuzzy Sets and Systems* 124 (2001) 59–64.
- [28] J.-L. Marichal, k -intolerant capacities and choquet integrals, *European Journal of Operational Research* 177 (2007) 1453–1468.
- [29] D. Dubois, H. Prade, C. Testemale, Weighted fuzzy pattern matching, *Fuzzy Sets and Systems* 28 (1988) 313–331.
- [30] H. Le Capitaine, T. Batard, C. Frélicot, M. Berthier, Blockwise similarity in $[0,1]$ via triangular norms and sugeno integrals – application to cluster validity, in: *16th IEEE International Conference on Fuzzy Systems*, 2007, London, England, pp. 835–840.
- [31] M. P. Wand, M. C. Jones, Kernel Smoothing, *Monographs on Statistics and Applied Probability*, Chapman & Hall, 1994.
- [32] H. Le Capitaine, C. Frélicot, A family of measures for best top- n class-selective decision rules, *Pattern Recognition* 45 (2012) 552–562.