Two prominent examples of penalty-based aggregation of circular data*

Raúl Pérez-Fernández[1][0000−0001−6966−1681] and
Bernard De Baets[2][0000−0002−3876−620X]

[1] Departamento de Estadística e I.O. y D.M., Universidad de Oviedo, Spain
perezfernandez@uniovi.es
[2] KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, Belgium bernard.debaets@ugent.be

Abstract. Aggregation processes appear naturally in many fields of application. The formalization of such processes has been a core topic for researchers in the fuzzy set community for decades, mostly focusing on the aggregation of elements of a bounded poset (typically a bounded real interval). Recent work by the present authors has aimed at further generalizing aggregation theory so it can accommodate aggregation processes on more general structures such as multivariate data, ranking data and string data. In this work, the aggregation of circular data is explored and, in particular, two prominent examples of aggregation functions for circular data (namely the circular mean and the circular median) are presented within this revisited aggregation theory framework.

Keywords: Aggregation · Penalty function · Circular data.

1 Introduction

As brought to the attention in [11], the problem of combining several real values into a single one can be traced back to Ancient Greece. Nevertheless, it is not until the nineteenth century that Cauchy [5] introduces the first formal definition of mean as an internal function, i.e., a function that is bounded from below by the minimum and from above by the maximum. Two centuries later countless papers and monographs on the topic have been written studying the more general notion of aggregation function (see, e.g., [1, 10]). Aggregation functions are not necessarily internal but still allow to combine several real values into a single one.

At the same time, different scientific communities have addressed similar aggregation-related tasks, while dealing with structures different from the set of real numbers. For instance, from the very same aggregation theory community,

* This contribution is part of grant PID2022-140585NB-I00 funded by MICIU/AEI/10.13039/501100011033 and "FEDER/UE". BDB received funding from the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" programme.

aggregation on bounded partially ordered sets has been addressed [13]. Also, the aggregation of multivariate data [9, 17] has been quite popular in the field of statistics, whereas the aggregation of ranking data [21] has attracted the attention of different fields such as that of social choice theory. Other examples of structures on which aggregation has been studied are those of imagery data [14], string data [18] and compositional data [20].

A general theory of aggregation on structures was already initiated back in 1993 by Yager [24] based on the notion of penalty function. These penalty functions largely attracted the attention of aggregation theorists in the context of real numbers, see e.g. [2, 4], however the work of Yager was not really picked up on different structures until recent works such as those by Gagolewski [8] and Pérez-Fernández and De Baets [16]. The latter work paved the way towards the development of a theory of aggregation on sets equipped with a betweenness relation [19], besets for short.

The present work aims at positioning two prominent examples of aggregation of circular data (namely the circular mean and the circular median) within the framework of (penalty-based) aggregation on besets. It must be remarked that the aggregation of circular data is an old acquaintance of statisticians, which have routinely addressed aggregation of this type of data in the context of location estimation for circular data [7, 15].

The remainder of the paper is structured as follows. Section 2 presents the general framework for (penalty-based) aggregation on besets. The setting of circular data is introduced in Section 3 by already proposing a natural betweenness relation for this type of data. Section 4 positions the circular mean and the circular median within the framework of penalty-based aggregation of besets, whereas Section 5 shows that both these functions do not fulfill the definition of an aggregation function for circular data. We end with some conclusions in Section 6.

2 Aggregation on besets

A betweenness relation is a ternary relation that formalizes the notion of one element being in between two other elements. Although different axiomatic definitions have been studied (see, e.g., [6, 12, 22]), we consider the one in [16] that has already been considered successfully in the context of penalty-based data aggregation.

Definition 1. *A ternary relation B on a non-empty set X is called a betweenness relation if it satisfies the following three properties:*

(i) Symmetry in the end points: for any $x, y, z \in X$, it holds that

$$(x, y, z) \in B \Leftrightarrow (z, y, x) \in B\,.$$

(ii) Closure: for any $x, y, z \in X$, it holds that

$$\big((x, y, z) \in B \wedge (x, z, y) \in B\big) \Leftrightarrow y = z\,.$$

(iii) End-point transitivity: for any $o, x, y, z \in X$, *it holds that*

$$\big((o, x, y) \in B \wedge (o, y, z) \in B\big) \Rightarrow (o, x, z) \in B\,.$$

A set X *equipped with a betweenness relation* B *is called a beset and denoted by* (X, B).

The notion of set of bounds of a beset [19] generalizes the bounds of an interval to more general structures.

Definition 2. *Given a beset* (X, B)*, a non-empty subset* S *of* X *is called a set of bounds of* (X, B) *if, for any* $y \in S$ *and any* $x, z \in X \backslash S$*, it holds that* $(x, y, z) \notin B$*. We thus refer to* (X, B, S) *as a bounded beset.*

A betweenness relation on a certain space X induces a natural betweenness relation on the product space X^n.

Definition 3. *Given a betweenness relation* B *on a set* X *and* $n \in \mathbb{N}$*, the product betweenness relation on* X^n *induced by* B *is the ternary relation* $B^{(n)}$ *defined as*

$$B^{(n)} = \left\{(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in (X^n)^3 \,|(\forall i \in \{1, \ldots, n\})((x_i, y_i, z_i) \in B)\right\}\,.$$

Aggregation functions [1, 10] can thus be defined on more general structures than the real line by using betweenness relations and, more specifically, bounded besets (see, [19]).

Definition 4. *Consider a bounded beset* (X, B, S) *and* $n \in \mathbb{N}$*. A function* $A : X^n \to X$ *is called an (n-ary) aggregation function on* (X, B, S) *if*

(i) it satisfies the boundary conditions, i.e., $A(o, \ldots, o) = o$*, for any* $o \in S$*;*
(ii) it is monotone, i.e., for any $o \in S$ *and any* $\mathbf{x}, \mathbf{y} \in X^n$*, the fact that* $\big((o, \ldots, o), \mathbf{x}, \mathbf{y}\big) \in B^{(n)}$ *implies that* $\big(A(o, \ldots, o), A(\mathbf{x}), A(\mathbf{y})\big) \in B$*.*

In a similar line of thought, a generalization of the definition of penalty-function [2, 4] was proposed in [16]. Intuitively, a penalty function is a function that assigns a penalty to a given element according to how much it disagrees with a list of elements to be aggregated.

Definition 5. *Consider* $n \in \mathbb{N}$*, a set* X *and a betweenness relation* B *on* X^n*. A function* $P : X \times X^n \to \mathbb{R}^+$ *is called a penalty function (compatible with* B*) if the following four properties hold:*

(P1) $P(y; \mathbf{x}) \geq 0$*, for any* $y \in X$ *and any* $\mathbf{x} \in X^n$*;*
(P2) $P(y; \mathbf{x}) = 0$ *if and only if* $\mathbf{x} = (y, \ldots, y)$*;*
(P3) The set of minimizers of $P(\cdot; \mathbf{x})$ *is non-empty, for any* $\mathbf{x} \in X^n$*.*
(P4) $P(y; \mathbf{x}) \leq P(y; \mathbf{x}')$*, for any* $y \in X$ *and any* $\mathbf{x}, \mathbf{x}' \in X^n$ *such that* $((y, \ldots, y), \mathbf{x}, \mathbf{x}') \in B$*.*

The ultimate use of a penalty function is to define a penalty-based (aggregation) function where the result of the aggregation is defined to be the element that results in the smallest penalty given the list of elements to be aggregated.

Definition 6. *Consider $n \in \mathbb{N}$, a set X, a betweenness relation B on X^n and a penalty function $P : X \times X^n \to \mathbb{R}^+$ compatible with B. The function $f : X^n \to \mathcal{P}(X)$ defined by*

$$f(\mathbf{x}) = \arg\min_{y \in X} P(y; \mathbf{x}) \,,$$

for any $\mathbf{x} \in X^n$, is called the penalty-based function associated with P.

Note that the codomain of the penalty-based function above is the powerset $\mathcal{P}(X)$ of X. If the goal is that there exists a unique aggregated element, the definition above should read "The function $f : X^n \to X$ defined by

$$f(\mathbf{x}) \in \arg\min_{y \in X} P(y; \mathbf{x}) \,,$$

for any $\mathbf{x} \in X^n$".

3 A betweenness relation for circular data

Let $\mathcal{D} = [0, 2\pi[$ denote the set of circular data points, i.e., the set of angles measured in radians. As usual when dealing with circular data, all arithmetic operations are performed modulo 2π.

A natural betweenness relation on $\mathcal{D}$ assures that y is in between x and z if y lies on the shortest arc between x and z.

Proposition 1. *The ternary relation $B_{\mathcal{D}}$ on $\mathcal{D}$, defined as*

$$\begin{aligned} B_{\mathcal{D}} &= \{(x, y, z) \in \mathcal{D}^3 \mid \big(\sin(z - x) \cdot \sin(y - x) \geq 0\big) \wedge \big(\cos(z - x) \leq \cos(y - x)\big)\} \\ &= \{(x, y, z) \in \mathcal{D}^3 \mid \big(0 \leq y - x \leq z - x \leq \pi\big) \vee \big(\pi \leq z - x \leq y - x \leq 2\pi\big)\} \,, \end{aligned}$$

is a betweenness relation on $\mathcal{D}$.

Proof. (i) Symmetry in the end points. Let $(x, y, z) \in B_{\mathcal{D}}$. We distinguish two cases:

- $0 \leq y - x \leq z - x \leq \pi$: It follows that $\pi \leq x - z \leq y - z \leq 2\pi$, which implies that $(z, y, x) \in B_{\mathcal{D}}$.
- $\pi \leq z - x \leq y - x \leq 2\pi$: It follows that $0 \leq y - z \leq x - z \leq \pi$, which implies that $(z, y, x) \in B_{\mathcal{D}}$.

(ii) Closure. $(x, y, z) \in B_{\mathcal{D}}$ and $(x, z, y) \in B_{\mathcal{D}}$ is equivalent to $\sin(z - x) \cdot \sin(y - x) \geq 0$ and $\cos(z - x) = \cos(y - x)$, which at the same time is equivalent to $y = z$.
(iii) End-point transitivity. Let $(o, x, y) \in B_{\mathcal{D}}$ and $(o, y, z) \in B_{\mathcal{D}}$. We distinguish two cases:

- $0 \leq y - o \leq z - o \leq \pi$: Since it also holds that $(o, x, y) \in B_{\mathcal{D}}$, it necessarily holds that $0 \leq x - o \leq y - o \leq z - o \leq \pi$, which ultimately implies that $(o, x, z) \in B_{\mathcal{D}}$.

– $\pi \leq z - o \leq y - o \leq 2\pi$: Since it also holds that $(o, x, y) \in B_{\mathcal{D}}$, it necessarily holds that $\pi \leq z - o \leq y - o \leq x - o \leq 2\pi$, which ultimately implies that $(o, x, z) \in B_{\mathcal{D}}$. □

Proposition 2. *Let $B_{\mathcal{D}}$ be the betweenness relation on $\mathcal{D}$ introduced in Proposition 1. It holds that $(x, y, z) \in B_{\mathcal{D}}$ implies that:*

(i) $(0, y - x, z - x) \in B_{\mathcal{D}}$;
(ii) $(-x, -y, -z) \in B_{\mathcal{D}}$;
(iii) $(x + a, y + a, z + a) \in B_{\mathcal{D}}$.

Proof. (i) follows immediately by definition. (ii) follows from the fact that sin is an odd function and cos is an even function. (iii) follows from the cancellation of the addends a when substracting any two among $x + a$, $y + a$ and $z + a$. □

Figure 1 illustrates the betweenness relation $B_{\mathcal{D}}$. The elements in the red area are in between $o = 0$ and $x = \frac{\pi}{3}$.

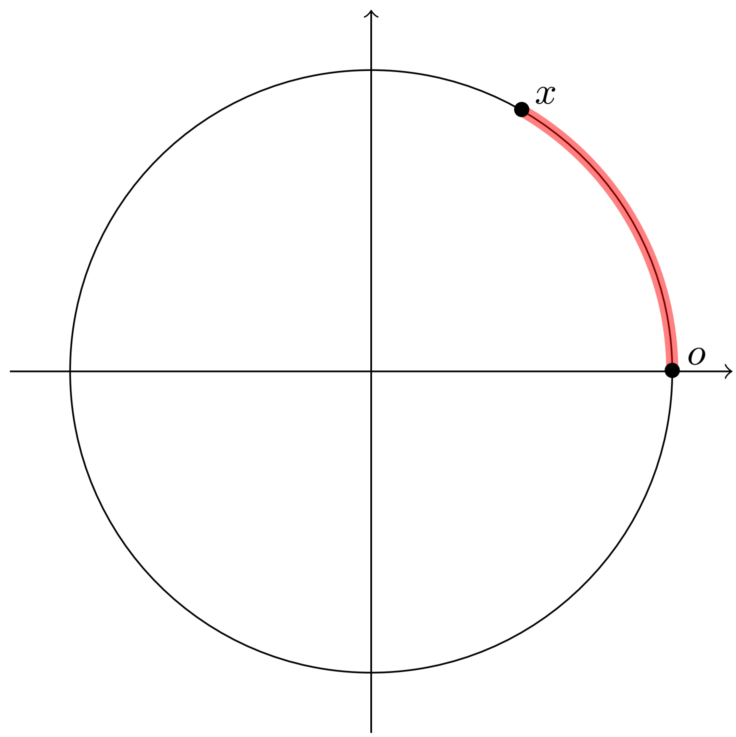

Fig. 1. Graphical representation of $B_{\mathcal{D}}$.

4 Penalty-based aggregation of circular data

In the context of circular data, $\overline{\mathbf{x}}$ is not the arithmetic mean (which is not well defined for circular data). In this setting, the circular mean $\overline{\mathbf{x}}$ of the angles $\mathbf{x} = (x_1, \ldots, x_n)$ is defined as:

$$\overline{\mathbf{x}} = \begin{cases} \arctan(S/C)\,, & \text{if } C > 0\,, \\ \arctan(S/C) + \pi\,, & \text{if } C < 0\,, \\ \frac{\pi}{2}\,, & \text{if } C < 0 \text{ and } S > 0\,, \\ -\frac{\pi}{2}\,, & \text{if } C < 0 \text{ and } S < 0\,, \\ \text{undefined}\,, & \text{if } C = 0 = S\,, \end{cases}$$

where $C = \frac{1}{n}\sum_{i=1}^{n} \cos(x_i)$ and $S = \frac{1}{n}\sum_{i=1}^{n} \sin(x_i)$.

It is known that the circular mean is the unique minimizer of the function (of y):

$$P_1(y; \mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}(1 - \cos(x_i - y))\,.$$

The following result proves that the function P_1 is a penalty function in the sense of Definition 6 and that, therefore, the circular mean is a penalty-based function in the sense of Definition 5.

Proposition 3. *Consider $n \in \mathbb{N}$, $\mathcal{D}$ and the betweenness relation $(B_{\mathcal{D}})^{(n)}$ on $\mathcal{D}^n$. The function $P_1 : \mathcal{D} \times \mathcal{D}^n \to \mathbb{R}^+$ is a penalty function compatible with $(B_{\mathcal{D}})^{(n)}$ and the circular mean $\bar{\cdot} : \mathcal{D}^n \to \mathcal{D}$ is a penalty-based function associated with P_1.*

Proof. (P1) follows from the fact that $\cos(x) \in [-1, 1]$ for any $x \in \mathcal{D}$. (P2) follows from the fact that $P_1(y; \mathbf{x}) = 0$ is equivalent to $\cos(x_i - y) = 1$ for any $i \in \{1, \ldots, n\}$, which also is equivalent to $y = x_i$ for any $i \in \{1, \ldots, n\}$. (P3) follows from the fact that $P_1(\cdot; \mathbf{x})$ is a continuous function on the compact set $\mathcal{D}$. (P4) Let $y \in \mathcal{D}$, $\mathbf{x}, \mathbf{x}' \in \mathcal{D}^n$ be such that $((y, \ldots, y), \mathbf{x}, \mathbf{x}') \in (B_{\mathcal{D}})^{(n)}$. By definition of $(B_{\mathcal{D}})^{(n)}$, it follows that $\cos(x'_i - y) \leq \cos(x_i - y)$ for any $i \in \{1, \ldots, n\}$, which eventually implies that $P_1(y; \mathbf{x}) \leq P_1(y, \mathbf{x}')$. □

The circular median $\tilde{\mathbf{x}}$ of the angles $\mathbf{x} = (x_1, \ldots, x_n)$ is defined as an angle that minimizes (in y)

$$P_2(y; \mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}(\pi - |\pi - |x_i - y||) = \frac{1}{n}\sum_{i=1}^{n} \min(x_i - y, 2\pi - x_i + y)\,.$$

Note that this minimizer may not be unique. If n is odd, then the circular median is an angle in $\mathbf{x}$ such that (i) $\lceil \frac{n}{2} \rceil$ of the angles in $\mathbf{x}$ lie on the arc $[\tilde{\mathbf{x}}, \tilde{\mathbf{x}} + \pi[$; and (ii) the majority of the angles in $\mathbf{x}$ are closer to $\tilde{\mathbf{x}}$ than to $\tilde{\mathbf{x}} + \pi$.

The following result proves that the function P_2 is a penalty function in the sense of Definition 6 and that, therefore, the circular median is a penalty-based function in the sense of Definition 5.

Proposition 4. *Consider $n \in \mathbb{N}$, $\mathcal{D}$ and the betweenness relation $(B_{\mathcal{D}})^{(n)}$ on $\mathcal{D}^n$. The function $P_2 : \mathcal{D} \times \mathcal{D}^n \to \mathbb{R}^+$ is a penalty function compatible with $(B_{\mathcal{D}})^{(n)}$ and the circular median $\tilde{\cdot} : \mathcal{D}^n \to \mathcal{P}(\mathcal{D})$ is a penalty-based function associated with P_2.*

Proof. (P1) follows from the fact that we are considering internal arithmetic operations on $\mathcal{D}$. (P2) follows from the fact that $P_2(y;\mathbf{x}) = 0$ is equivalent to $|\pi - |x_i - y|| = \pi$ for any $i \in \{1,\dots,n\}$, which also is equivalent to $y = x_i$ for any $i \in \{1,\dots,n\}$. (P3) follows from the fact that $P_2(\cdot;\mathbf{x})$ is a continuous function on the compact set $\mathcal{D}$. (P4) Let $y \in \mathcal{D}$, $\mathbf{x},\mathbf{x}' \in \mathcal{D}^n$ be such that $((y,\dots,y), \mathbf{x},\mathbf{x}') \in (B_{\mathcal{D}})^{(n)}$. By definition of $(B_{\mathcal{D}})^{(n)}$, it follows that $0 \leq x_i - y \leq x'_i - y \leq \pi$ or $\pi \leq x'_i - y \leq x_i - y \leq 2\pi$ for any $i \in \{1,\dots,n\}$, which eventually implies that $(\pi - |\pi - |x_i - y||) \leq (\pi - |\pi - |x'_i - y||)$ for any $i \in \{1,\dots,n\}$. Ultimately, the latter implies that $P_2(y;\mathbf{x}) \leq P_2(y,\mathbf{x}')$. □

5 Aggregation functions for circular data

In this section, we firstly prove that the circular mean is not an aggregation function for circular data in the sense of Definition 4.

Proposition 5. *Consider the bounded beset $(\mathcal{D}, B_{\mathcal{D}}, \mathcal{D})$ and $n \in \mathbb{N}$. The circular mean $\bar{\cdot} : \mathcal{D}^n \to \mathcal{D}$ is not an aggregation function on $(\mathcal{D}, B_{\mathcal{D}}, \mathcal{D})$.*

Proof. Even though the circular mean satisfies the boundary conditions (as a result of property (P2) of a penalty function), the monotonicity property does not hold. If $n = 2$, consider $\mathbf{o} = (0,0)$, $\mathbf{x} = (\frac{\pi}{8}, \frac{7\pi}{4})$ and $\mathbf{y} = (\frac{\pi}{4}, \frac{7\pi}{4})$. It follows that $(\mathbf{o},\mathbf{x},\mathbf{y}) \in B^{(n)}$, however $\overline{\mathbf{o}} = \overline{\mathbf{y}} = 0 \neq \overline{\mathbf{x}}$. The same example can be considered if $n > 2$ by considering $\mathbf{o}' = (0,\dots,0)$, $\mathbf{x}' = (\frac{\pi}{8}, \frac{7\pi}{4}, 0, \dots, 0)$ and $\mathbf{x}' = (\frac{\pi}{4}, \frac{7\pi}{4}, 0, \dots, 0)$. □

Secondly, we prove that the circular median is not an aggregation function for circular data in the sense of Definition 4.

Proposition 6. *Consider the bounded beset $(\mathcal{D}, B_{\mathcal{D}}, \mathcal{D})$ and $n \in \mathbb{N}$. The circular median $\tilde{\cdot} : \mathcal{D}^n \to \mathcal{P}(\mathcal{D})$ is not an aggregation function on $(\mathcal{D}, B_{\mathcal{D}}, \mathcal{D})$.*

Proof. Even though the circular median satisfies the boundary conditions (as a result of property (P2) of a penalty function), the monotonicity property does not hold. Consider $\mathbf{o} = (0,0,0,0,0)$, $\mathbf{x} = (0.1, 3.13, 3.14, 6, 6.2)$ and $\mathbf{y} = (3.12, 3.13, 3.14, 3.15, 3.16)$. It follows that $(\mathbf{o},\mathbf{x},\mathbf{y}) \in B^{(n)}$, however $\tilde{\mathbf{o}} = 0$, $\tilde{\mathbf{x}} = 6$ and $\tilde{\mathbf{y}} = 3.14$. □

6 Conclusions

It has been shown that the circular mean and the circular median can be accommodated within the framework of penalty-based data aggregation developed in [16]. Unfortunately, even though these two prominent functions for the aggregation of circular data satisfy the boundary conditions and are actually idempotent (as a result of property (P2) of a penalty function), it is shown that they are not aggregation functions on $(\mathcal{D}, B_{\mathcal{D}}, \mathcal{D})$ in the sense of Definition 4.

It remains as a future study subject to explore weaker properties than monotonicity that could accommodate the mean and circular medians for the aggregation of circular data. This direction will further extend current research in the context of aggregation of real numbers in which weaker types of monotonicity were explored, such as weak monotonicity [23] and directional monotonicity [3].

References

1. Beliakov, G., Pradera, A., Calvo, T.: Aggregation Functions: A Guide for Practitioners, Studies in Fuzziness and Soft Computing, vol. 221. Springer, Berlin, Heidelberg (2007)
2. Bustince, H., Beliakov, G., Dimuro, G.P., Bedregal, B., Mesiar, R.: On the definition of penalty functions in data aggregation. Fuzzy Sets and Systems **323**, 1–18 (2017)
3. Bustince, H., Fernández, J., Kolesárová, A., Mesiar, R.: Directional monotonicity of fusion functions. European Journal of Operational Research **244**, 300–308 (2015)
4. Calvo, T., Beliakov, G.: Aggregation functions based on penalties. Fuzzy Sets and Systems **161**, 1420–1436 (2010)
5. Cauchy, A.L.: Cours d'analyse de l'Ecole Royale Polytechnique, Analyse algébraique, vol. 1. Debure, Paris (1821)
6. Fishburn, P.C.: Betweenness, orders and interval graphs. Journal of Pure and Applied Algebra **1**(2), 159–178 (1971)
7. Fisher, N.I.: Statistical Analysis of Circular Data. Cambridge University Press, Cambridge (1995)
8. Gagolewski, M.: Data Fusion: Theory, Methods, and Applications. Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland (2015)
9. Gagolewski, M.: Penalty-based aggregation of multidimensional data. Fuzzy Sets and Systems **325**, 4–20 (2017)
10. Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: Aggregation Functions. Cambridge University Press, Cambridge (2009)
11. Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: Aggregation functions: Means. Information Sciences **181**, 1–22 (2011)
12. Huntington, E.V., Kline, J.R.: Sets of independent postulates for betweenness. Transactions of the American Mathematical Society **18**(3), 301–325 (1917)
13. Komorníková, M., Mesiar, R.: Aggregation functions on bounded partially ordered sets and their classification. Fuzzy Sets and Systems **175**, 48–56 (2011)
14. Lopez-Molina, C., Bustince, H., Fernández, J., Couto, P., De Baets, B.: A gravitational approach to edge detection based on triangular norms. Pattern Recognition **43**(11), 3730–3741 (2010)
15. Mardia, K.V., Jupp, P.E.: Directional Statistics. Wiley, Chichester (2000)
16. Pérez-Fernández, R., De Baets, B.: On the role of monometrics in penalty-based data aggregation. IEEE Transactions on Fuzzy Systems **27**(7), 1456–1468 (2019)
17. Pérez-Fernández, R., De Baets, B., Gagolewski, M.: A taxonomy of monotonicity properties for the aggregation of multidimensional data. Information Fusion **52**, 322–334 (2019)

18. Pérez-Fernández, R., De Baets, B.: Penalty-based aggregation of strings. In: New Trends in Aggregation Theory. pp. 212–222. No. 10, Springer (2019)
19. Pérez-Fernández, R., De Baets, B.: Aggregation theory revisited. IEEE Transactions on Fuzzy Systems **29**(4), 797–804 (2020)
20. Pérez-Fernández, R., Gagolewski, M., De Baets, B.: On the aggregation of compositional data. Information Fusion **73**, 103–110 (2021)
21. Pérez-Fernández, R., Rademaker, M., De Baets, B.: Monometrics and their role in the rationalisation of ranking rules. Information Fusion **34**, 16–27 (2017)
22. Pitcher, E., Smiley, M.F.: Transitivities of betweenness. Transactions of the American Mathematical Society **52**(1), 95–114 (1942)
23. Wilkin, T., Beliakov, G.: Weakly monotonic averaging functions. International Journal of Intelligent Systems **30**(2), 144–169 (2015)
24. Yager, R.R.: Toward a general theory of information aggregation. Information Sciences **68**, 191–206 (1993)