Information Sciences 263 (2014) 166–174

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Monotone measures and universal integrals in a uniform framework for the scientific impact assessment problem

Marek Gagolewski [a,b,*], Radko Mesiar [c,d]

[a] Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6, 01-447 Warsaw, Poland
[b] Faculty of Mathematics and Information Science, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warsaw, Poland
[c] Faculty of Civil Engineering, Department of Mathematics, Slovak University of Technology, 81 368 Bratislava, Slovakia
[d] National Supercomputing Center IT4Innovations, Division University of Ostrava IRAFM, 30. Dubna 22, 701 03 Ostrava 1, Czech Republic

ARTICLE INFO

Article history:
Received 12 June 2013
Received in revised form 21 October 2013
Accepted 7 December 2013
Available online 12 December 2013

Keywords:
Universal integral
Monotone measures
Aggregation
Indices of scientific impact
Bibliometrics
h-index

ABSTRACT

The Choquet, Sugeno, and Shilkret integrals with respect to monotone measures, as well as their generalization – the universal integral, stand for a useful tool in decision support systems. In this paper we propose a general construction method for aggregation operators that may be used in assessing output of scientists. We show that the most often currently used indices of bibliometric impact, like Hirsch's h, Woeginger's w, Egghe's g, Kosmulski's MAXPROD, and similar constructions, may be obtained by means of our framework. Moreover, the model easily leads to some new, very interesting functions.

1. Introduction

Many practical situations, especially in decision making, face us with the problem of aggregating sequences having not necessarily equal lengths. In such cases not only the aggregated elements have impact on the overall evaluation, but also the length of the sequence. This is the case of the bibliometric impact assessment problem, which concerns aggregating the number of citations received by articles published by authors having different productiveness. Of course, "raw" citations are not the only way to measure the quality of a paper: we can use other indicators, like impact factors of their journals, or citations which are normalized by the scientific domain and the number of authors, cf. [3,11,20]. Other instances of this issue include, e.g. manufacturing, quality engineering, webometrics, evaluation of open source software packages, see e.g. [7,10].

Let us assume that the whole information on the "producer's" (e.g. the author's) performance is represented by a vector $\mathbf{x} \in \mathbb{I}^{1,2,\ldots} = \bigcup_{n=1}^{\infty} \mathbb{I}^n$, where $x_i \in \mathbb{I} = [0, \infty]$ denotes the quality of his/her ith "product" (e.g. paper; of course, how to measure its quality is a problem on its own). Our interest here lies in finding a *method* that may be used to synthesize $\mathbf{x}$ so that his/her performance may be *described* with a single numeric value. Bibliometricians generally agree, see [8,10,17–19,24–26], that such an aggregation function should be (a) nondecreasing with respect to the quality of individual papers, e.g. after increasing the number of citations of a single article one should not get lower overall evaluation; (b) nondecreasing

* Corresponding author at: Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6, 01-447 Warsaw, Poland. Tel.: +48 22 38 10 393; fax: +48 22 38 10 105.
E-mail addresses: gagolews@ibspan.waw.pl (M. Gagolewski), mesiar@math.sk (R. Mesiar).

0020-0255/$ - see front matter
http://dx.doi.org/10.1016/j.ins.2013.12.004

with respect to the number of papers, e.g. an author cannot be penalized if he/she publishes yet another paper, even if it has 0 citations at the beginning; and (c) symmetric, i.e. not depending on the order in which the aggregated papers are being presented.

One example of such an aggregation method is the famous *h*-index [12]. Torra and Narukawa showed that the *h*-index is equivalent to the discrete Sugeno integral with respect to the counting measure, see [23]. In fact, we will see that many of the indices of scientific impact may be expressed as monotone (fuzzy, not necessarily additive) integrals, like Choquet [5], Sugeno [22], or Shilkret [21] ones which have been known for over forty years. However, to the best of our knowledge, except for [23], there are only two other papers which describe the connection between the monotone integrals and scientometrics. Beliakov and James [2] consider applications of the Choquet integral-based classifiers to the problem of ranking of scientific journals. Additionally, Bras-Amorós et al. [4] mention that their bibliometric index based on the collaboration distance between cited and citing authors corresponds to a Sugeno integral w.r.t. some fuzzy measure.

The aim of this paper is to present a uniform model for the scientific impact assessment problem (as well as other similar domains) via monotone measures and integrals. Such a framework not only is very flexible, provides intuitive graphical interpretations for the aggregation process, and allows for constructing many new and interesting classes that may be used to describe the scientific record of a scientist. It also stands for another successful application of the fuzzy measure theory.

The paper is structured as follows. In the next section we present an axiomatization of the aggregation operators that are most often used in the post-Hirsch bibliometric impact assessment of individuals. After reviewing the most prominent impact functions, we recall the notion of a monotone measure and universal integral. In Section 3 we propose the uniform model that is based on universal integrals, and in Section 4 we show how to obtain the indices currently applied in practice, and also how to generate and compute very interesting new ones. Section 5 concludes the paper, indicating some important issues concerning the impact assessment task.

2. Preliminaries

One of the problems with the aggregation of vectors in $\mathbb{I}^{1,2,\ldots}$ is that in fact we are required to introduce a *family* of functions, each operating on fixed-length vectors. This is because if $\mathsf{F}: \mathbb{I}^{1,2,\ldots} \to \mathbb{I}$, then F may be written as $\mathsf{F} = (\mathsf{F}^{(1)}, \mathsf{F}^{(2)}, \ldots)$, where $\mathsf{F}^{(n)}: \mathbb{I}^n \to \mathbb{I}$. Thus, to achieve the main aim of the paper, we would have to consider *families* of monotone measures and *families* of integrals. The resulting model, although being interesting from the theoretical viewpoint, would be far too complex for practitioners.

We will therefore focus on the so-called zero-insensitive aggregation operators, cf. e.g. [10]. In such case, each uncited paper is treated as non-existing. Even though this setting may seem quite limiting, in fact most of the currently used bibliometric impact indices do obey this property, see [1] and also Section 2.1.

Let us consider the space $\mathcal{S}$ of infinite nonincreasing sequences with elements in $\mathbb{I}$. Let $\tilde{\cdot}: \mathbb{I}^{1,2,\ldots} \to \mathcal{S}$ be an operator such that $\tilde{\mathbf{x}} = (x_{\{1\}}, x_{\{2\}}, \ldots, x_{\{n\}}, 0, 0, \ldots)$, where $x_{\{i\}}$ denotes the ith greatest value in $\mathbf{x}$.

Proposition 1. *Let $\mathsf{F}: \mathbb{I}^{1,2,\ldots} \to \mathbb{I}$ be an aggregation function. Then F is a zero-insensitive impact function, i.e. it fulfills the following properties:*

1. $\mathsf{F}(0) = 0$ *(lower bound);*
2. $(\forall n)\ (\forall \mathbf{x}, \mathbf{y} \in \mathbb{I}^n)\ (\forall i)\ x_i \leqslant y_i \Rightarrow \mathsf{F}(\mathbf{x}) \leqslant \mathsf{F}(\mathbf{y})$ *(nondecreasingness);*
3. $(\forall \mathbf{x} \in \mathbb{I}^{1,2,\ldots})\ (\forall y \in \mathbb{I})\ \mathsf{F}(\mathbf{x}) \leqslant \mathsf{F}(\mathbf{x}, y)$ *(arity-monotonicity);*
4. $(\forall n)\ (\forall \mathbf{x} \in \mathbb{I}^n)\ (\forall \sigma \in \mathfrak{S}_n)\ \mathsf{F}(\mathbf{x}) = \mathsf{F}(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$*, where $\mathfrak{S}_n$ denotes the set of all permutations of $\{1, \ldots, n\}$ (symmetry);*
5. $(\forall \mathbf{x} \in \mathbb{I}^{1,2,\ldots})\ \mathsf{F}(\mathbf{x}, 0) = \mathsf{F}(\mathbf{x})$ *(zero-insensitivity);*

if and only if there exists a nondecreasing function $\mathsf{E}: \mathcal{S} \to \mathbb{I}$, $\mathsf{E}(0, 0, \ldots) = 0$, such that for all $\mathbf{x} \in \mathbb{I}^{1,2,\ldots}$ we have $\mathsf{F}(\mathbf{x}) = \mathsf{E}(\tilde{\mathbf{x}})$.

The proof is straightforward and therefore omitted.

2.1. A review of impact functions

From now on we consider only vectors in $\mathcal{S}$. Some of the notable examples of zero-insensitive impact functions are listed below.

- Total number of citations:

$$\mathsf{S}(\mathbf{x}) = \sum_{i=1}^{\infty} x_i, \tag{1}$$

or, more generally, a weighted sum of elements of $\mathbf{x} \in \mathcal{S}$. This includes, e.g. "the total number of citations of 5 most cited papers".

- The h-index [12]:

$$\mathsf{H}(\mathbf{x}) = \max\{h \in \mathbb{N} : x_h \geqslant h\}. \tag{2}$$

with convention $x_0 = x_1$.
- The MaxProd-index [15]:

$$\mathsf{MP}(\mathbf{x}) = \max\{i \cdot x_i : i \in \mathbb{N}\}. \tag{3}$$

This index is a particular case of the (projected) l_p-indices, $p \geqslant 1$, see [9].
- The g-index [6]:

$$\mathsf{G}(\mathbf{x}) = \max\left\{g \in \mathbb{N} : \sum_{i=1}^{g} x_i \geqslant g^2\right\}, \tag{4}$$

with convention $\sum_{i=1}^{0} \cdots = 0$.
- The w-index [26]:

$$\mathsf{W}(\mathbf{x}) = \max\{w \in \mathbb{N} : x_i \geqslant w - i + 1 \text{ for all } i \leqslant w\}. \tag{5}$$

The h- and w-index is generalized by e.g. the class of r_p-indices, $p \geqslant 1$, see [9].
- The $h(2)$-index [14]:

$$\mathsf{H2}(\mathbf{x}) = \max\{h \in \mathbb{N} : x_h \geqslant h^2\}. \tag{6}$$

Note that the $h(2)$-index is one of the many examples of very simple, direct modifications of the h-index. Many authors considered other settings than "h^2" on the right side of (6), e.g. "$2h$", "αh" for some $\alpha > 0$, or "h^β", $\beta \geqslant 1$, cf. [1].

and so on.

Note that, originally, many proposals for the bibliometric indices assumed that we aggregate the number of papers' citations, i.e. sequences with elements in $\mathbb{N}$. Generally, however, the paper quality measures may be arbitrary real numbers, for example when citations are normalized according to the number of coauthors, paper's time of publication, quality of a journal, etc., see e.g. [11].

2.2. Monotone measures and integrals

Let $(Z, \mathcal{A})$ be a measurable space, i.e. a nonempty set Z equipped with a σ-algebra. We call $\mu : \mathcal{A} \to \mathbb{I}$ a *monotone measure* (a capacity), if (a) $\mu(\emptyset) = 0$, (b) $\mu(Z) > 0$, and (c) $\mu(U) \leqslant \mu(V)$ for $U \subseteq V$. Note that a monotone measure is not necessarily (σ-) additive. Moreover, let $\mathcal{M}^{(Z,\mathcal{A})}$ denote the set of all monotone measures.

A function $\mathsf{f} : Z \to \mathbb{I}$ is called $\mathcal{A}$-*measurable* if for each T in the σ-algebra of Borel subsets of $\mathbb{I}$, the inverse image $\mathsf{f}^{-1}(T) \in \mathcal{A}$. Let $\mathcal{F}^{(Z,\mathcal{A})}$ denote the set of all $\mathcal{A}$-measurable functions $\mathsf{f} : Z \to \mathbb{I}$.

Please note that for both $\mathcal{M}^{(Z,\mathcal{A})}$ and $\mathcal{F}^{(Z,\mathcal{A})}$ natural partial orders $\preceq_{\mathcal{M}}$ and $\preceq_{\mathcal{F}}$ may be constructed: we have, e.g. $\mathsf{f} \preceq_{\mathcal{F}} \mathsf{f}'$ if and only if for all $z \in Z$ it holds $\mathsf{f}(z) \leqslant \mathsf{f}'(z)$. Moreover, the spaces $(\mathcal{M}^{(Z,\mathcal{A})}, \preceq_{\mathcal{M}})$ and $(\mathcal{F}^{(Z,\mathcal{A})}, \preceq_{\mathcal{F}})$ are lattices.

For further discussion we will also need the notion of pseudomultiplication.

Definition 2. A function $\otimes : \mathbb{I}^2 \to \mathbb{I}$ is called a *pseudomultiplication*, if:

1. it is nondecreasing in each variable, i.e. for $0 \leqslant a_1 \leqslant a_2$ and $0 \leqslant b_1 \leqslant b_2$, we have $a_1 \otimes b_1 \leqslant a_2 \otimes b_2$,
2. it has 0 as the annihilator element, i.e. for all $a \in \mathbb{I}, a \otimes 0 = 0 \otimes a = 0$,
3. it has a neutral element $e > 0$, i.e. for all $a \in \mathbb{I}, a \otimes e = e \otimes a = a$.

Note that $\otimes$ is not necessarily associative or commutative. Standard multiplication $\cdot$ ($e = 1$) and minimum $\wedge$ ($e = \infty$) are particular examples of pseudomultiplication. On the other hand, e.g. maximum $\vee$ does not annihilate at 0, thus does not fall into this class.

Which function shall be called an integral of $\mathsf{f} \in \mathcal{F}^{(Z,\mathcal{A})}$ is still a disputable issue. Generally, it is agreed that an integral should map the space $\mathcal{M}^{(Z,\mathcal{A})} \times \mathcal{F}^{(Z,\mathcal{A})}$ into $\mathbb{I}$, should be at least nondecreasing with respect to each coordinate, and for $\mathsf{f} \equiv 0$ it should "output" the value 0.

Here, we will use the notion of a universal integral, thoroughly discussed in [13]. Let $\{u : \mathsf{f}(u) \geqslant t\} \in \mathcal{A}$ denote the so-called t-level set of $\mathsf{f}, t \in \mathbb{I}$. It is easily seen that $\{u : \mathsf{f}(u) \geqslant t\}_{t \in \mathbb{I}}$ forms a left-continuous, nonincreasing chain (w.r.t. t). Thus, $\mathsf{h}^{(\mu,\mathsf{f})}(t) := \mu(\{u : \mathsf{f}(u) \geqslant t\})$ is a nonincreasing function of t.

The following characterization of a universal integral was given in [13, Proposition 2.7].

Definition 3. A *universal integral* corresponding to the pseudomultiplication $\otimes$ is a function $\mathcal{I} : \mathcal{M}^{(Z,\mathcal{A})} \times \mathcal{F}^{(Z,\mathcal{A})} \to \mathbb{I}$ given by:

$$\mathcal{I}(\mu, \mathsf{f}) = \mathcal{J}\left(\mathsf{h}^{(\mu,\mathsf{f})}\right),$$

where $\mathcal{J}: \mathcal{F}^{(\mathbb{I},\mathcal{B}(\mathbb{I}))} \to \mathbb{I}$ is nondecreasing, and such that for each $c, d \in \mathbb{I}$ we have $\mathcal{J}(d \cdot \mathbf{I}_{(0,c]}) = c \otimes d$.

Given a measurable space $(Z, \mathcal{A})$, here are some well-known examples of universal integrals of $\mathsf{f} \in \mathcal{F}^{(Z,\mathcal{A})}$ w.r.t. a monotone measure $\mu \in \mathcal{M}^{(Z,\mathcal{A})}$:

- The Choquet integral [5]:

$$\mathrm{Ch}(\mu, \mathsf{f}) = \int_{\mathbb{I}} \mathsf{h}^{(\mu,\mathsf{f})}(t)dt. \tag{7}$$

Note that this is formulated the same as the Lebesgue integral, but with respect to an arbitrary monotone measure. We have $\otimes = \cdot$.

- The Sugeno integral [22]:

$$\mathrm{Su}(\mu, \mathsf{f}) = \sup_{t \in \mathbb{I}}\{t \wedge \mathsf{h}^{(\mu,\mathsf{f})}(t)\} \tag{8}$$

where $\wedge$ denotes the minimum operator. We have $\otimes = \wedge$.

- The Shilkret integral [21]:

$$\mathrm{Sh}(\mu, \mathsf{f}) = \sup_{t \in \mathbb{I}}\{t \cdot \mathsf{h}^{(\mu,\mathsf{f})}(t)\}, \tag{9}$$

with convention $0 \cdot \infty = 0$. We have $\otimes = \cdot$.

3. The uniform model for bibliometric impact assessment

In order to introduce our uniform framework for the bibliometric impact assessment problem, we will need a transformation from the vector space $\mathcal{S}$ into the space $\mathcal{F}^{(Z,\mathcal{A})}$ for some $(Z, \mathcal{A})$. Although the most straightforward choice is of course the measurable space $(\mathbb{N}, 2^{\mathbb{N}})$, it is not necessarily the most convenient one. Thus, we fix the space to $(\mathbb{I}, \mathcal{B}(\mathbb{I}))$.

Given $\mathbf{x} \in \mathcal{S}$, let $\langle \mathbf{x} \rangle \in \mathcal{F}^{(\mathbb{I},\mathcal{B}(\mathbb{I}))}$ such that

$$\langle \mathbf{x} \rangle(t) = x_{\lfloor t+1 \rfloor}, \quad t \in \mathbb{I}.$$

It is easily seen that $\langle \mathbf{x} \rangle$ is a nonincreasing step function with steps possible only in points from $\mathbb{N}$. Thanks to this setting, each vector gains a nice graphical interpretation, cf. Fig. 1. In fact, $\langle \mathbf{x} \rangle$ is often called by bibliometricians the *citation function* for the vector $\mathbf{x}$.

Let us consider the family Φ of aggregation operators $\mathsf{F}: \mathcal{S} \to \mathbb{I}$ given by the equation:

$$\mathsf{F}(\mathbf{x}) = \eta(\mathcal{I}(\mu, \langle \varphi(\mathbf{x}) \rangle)) \tag{10}$$

where

- $\varphi: \mathcal{S} \to \mathcal{S}$ – a function nondecreasing in each variable, $\varphi(0, 0, \ldots) = (0, 0, \ldots)$,
- $\mu: \mathcal{B}(\mathbb{I}) \to [0, \infty]$ – a monotone measure,
- $\mathcal{I}$ – a universal integral on $\mathcal{M}^{(\mathbb{I},\mathcal{B}(\mathbb{I}))} \times \mathcal{F}^{(\mathbb{I},\mathcal{B}(\mathbb{I}))}$,
- $\eta: \mathbb{I} \to \mathbb{I}$ – an increasing function, $\eta(0) = 0$.

We have what follows.

Theorem 4. *Each aggregation operator* F *given by (10) is a zero-insensitive impact function.*

Sketch of the proof. We apply Proposition 1 and the fact that each integral is, among others, a nondecreasing function of $\mathsf{f} \in \mathcal{F}^{(X,\mathcal{A})}$ for a fixed nondecreasing measure μ.

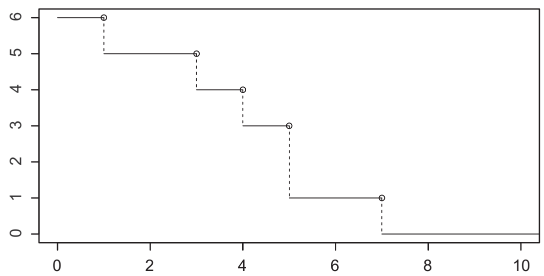

Fig. 1. Citation function for $\mathbf{x} = (6, 5, 5, 4, 3, 1, 1, 0, 0, \ldots) \in \mathcal{S}$.

On the other hand, each zero-insensitive impact function E belongs to Φ. It is because we may set $\varphi(\mathbf{x}) = (\mathsf{E}(\mathbf{x}), 0, 0, \ldots)$. Then, evidently, $\mathsf{E}(\mathbf{x}) = \mathrm{Sh}(\lambda, \langle\varphi(\mathbf{x})\rangle) = \mathrm{Ch}(\lambda, \langle\varphi(\mathbf{x})\rangle)$, where λ denotes the Lebesgue measure, $\lambda((a,b)) = \lambda([a,b]) = (b-a)$.

Please, note that the φ function may be used, e.g. to normalize citation records, and often will be set by extending a function of one variable φ' to $\mathcal{S}$, that is $\varphi(\mathbf{x}) = (\varphi'(x_1), \varphi'(x_2), \ldots)$. Many classical (citation-based) bibliometric indices assume that $\varphi'(x) = \lfloor x \rfloor$ or $\varphi'(x) = x$.

The η function may be used to "calibrate" the output values, especially if we would like to compare the values of different impact functions. It may be neglected if we consider ranking (instead of assessment) problems.

The measure μ will in turn be often set to be the Lebesgue measure λ or some monotonic transformation of λ.

Example 1. It is easily seen that:

- $\mathsf{S}(\mathbf{x}) = \mathrm{Ch}(\lambda, \langle\mathbf{x}\rangle)$,
- $\mathsf{H}(\mathbf{x}) = \mathrm{Su}(\lambda, \langle\lfloor\mathbf{x}\rfloor\rangle) = \lfloor \mathrm{Su}(\lambda, \langle\mathbf{x}\rangle)\rfloor$, cf. also [23],
- $\mathsf{MP}(\mathbf{x}) = \mathrm{Sh}(\lambda, \langle\mathbf{x}\rangle)$.

4. Computational issues and examples

It is important to discuss the implications of choosing different $\varphi, \mu, \mathcal{I}$, and η to the aggregation process and explain how to compute (10) on given input data.

4.1. Algorithms

First of all, we know that universal integrals are calculated only using $\mathsf{h}^{(\mu,\mathsf{f})}$, that is according to measures of level sets of the integrated function. Note that in our case, as we integrate nondecreasing step functions, then $\mathsf{h}^{(\lambda,\langle\varphi(\mathbf{x})\rangle)}(t)$ is in fact the pseudoinverse of $\langle\varphi(\mathbf{x})\rangle$ at t (also a nonincreasing step function). If we would set $(Z, \mathcal{A}) = (\mathbb{N}, 2^{\mathbb{N}})$, this transformation would lead us to functions measurable w.r.t. to a different space. In our case, however, we still are in $\mathcal{F}^{(\mathbb{I}, \mathcal{B}(\mathbb{I}))}$.

Note that the level sets of $\langle\varphi(\mathbf{x})\rangle$ are always of the form $[0,i)$ or $[0,i], i \in \mathbb{N}$. Thus, it does not make much sense in putting $\mu(Z)$ other than $\psi(\lambda(Z))$, where ψ is some nondecreasing function, $\psi(0) = 0$. However, as we shall see in Example 2, by choosing different ψ we put different weights for the "productivity" aspect of an assessed entity, see Fig. 2.

In fact, it may easily be shown that if ψ is invertible, then in such case the pseudoinverse of $\mathsf{h}^{(\psi\circ\lambda,\langle\varphi(\mathbf{x})\rangle)}$ at the point $u \in \mathbb{I}$ is given by $\mathsf{p}_{\mathbf{x}}(u) = \varphi(\mathbf{x})_{\lfloor\psi^{-1}(u)+1\rfloor}$. Using this property, we may propose the algorithm to compute (10), which is given in Fig. 3.

Moreover, note that $\mathcal{J}(\mathsf{h})$ may for some universal integrals be more easily and efficiently calculated directly on $\mathbf{x}'', \mathbf{y}$ (note that $\mathbf{x}''$ is ordered nondecreasingly, and $\mathbf{y}$ – nonincreasingly). For example, in case of the Choquet integral we have:

$$\mathcal{J}(\mathsf{h}) = \sum_{i=1}^{n} (y_i - y_{i+1}) \cdot x''_i,$$

for the Shilkret integral it holds:

$$\mathcal{J}(\mathsf{h}) = \max\{x''_i \cdot y_i : \ i = 1, \ldots, n\}.$$

Both equations may of course be computed in $O(n)$ time. On the other hand, for the Sugeno integral in most cases we seek for $i = \min\{i : \ y_i - x''_i < 0\}$ (which may be performed by a modified binary search algorithm, $O(\log n)$ time complexity), and then return:

$$\mathcal{J}(\mathsf{h}) = (x''_{i-1} \wedge y_{i-1}) \vee (x''_i \wedge y_i).$$

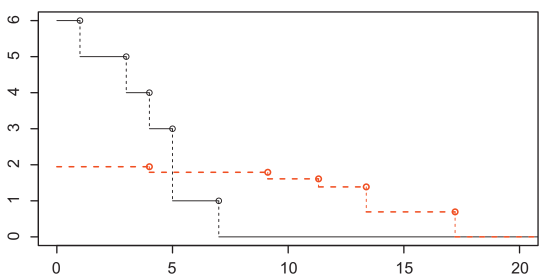

Fig. 2. The solid lines correspond to the citation function for $\mathbf{x} = (6,5,5,4,3,1,1,0,0,\ldots) \in \mathcal{S}$. The dashed lines constitute for the plot of $\mathsf{p}_{\mathbf{x}}(u) = \log(1+\mathbf{x})_{\lfloor (u/4)^{4/3}+1\rfloor}$, which is the pseudoinverse of $\mathsf{h}^{(\psi\circ\lambda,\langle\varphi((x))\rangle)}(t) = 4\lambda(\{u : \langle\log(1+\mathbf{x})\rangle(u) \geqslant t\})^{0.75}$, with $\psi(t) = 4t^{0.75}$ and $\varphi(\mathbf{x}) = (\log(1+x_1), \log(1+x_2), \ldots)$.

1. $\mathbf{x}'$ <- $\varphi(\mathbf{x}) \in \mathcal{S}$;
2. $\mathbf{x}''$ <- $(x'_n, \ldots, x'_1) \in \mathbb{I}^n$, where $n = \max\{i : x_i > 0\}$;
3. $\mathbf{y}$ <- $(\psi(n), \psi(n-1), \ldots, \psi(1), 0) \in \mathbb{I}^{n+1}$;
4. `h <- stepfun(x'', y, right=TRUE)`, i.e. a mapping such that

$$\mathsf{h}(u) = \begin{cases} y_i & \text{if } u \in [x''_{i-1}, x''_i) \text{ for some } i \in \{1, \ldots, n\}, \\ y_{n+1} & \text{otherwise,} \end{cases}$$

   with convention $x''_0 = 0$;
5. Let $\mathcal{J}$ be a function such that $\mathcal{I}(\mu, \mathsf{f}) = \mathcal{J}\left(\mathsf{h}^{(\mu,\mathsf{f})}\right)$, cf. Definition 3;
6. Return $\eta(\mathcal{J}(\mathsf{h}))$ as result;

Fig. 3. An R language-based pseudocode of the algorithm to compute $\mathsf{F}(\mathbf{x}) = \eta(\mathcal{I}(\psi \circ \lambda, \langle\varphi(\mathbf{x})\rangle))$ for a given $\mathbf{x} \in \mathcal{S}$, universal integral $\mathcal{I}$, and functions ψ, φ, and η.

4.2. Choosing $\mathcal{I}, \mu, \varphi$, and η

First let us study the effects of choosing different monotone measures.

Example 2. Let $\varphi = \mathrm{id}$, $\mathcal{I} = \mathrm{Ch}$, and $\eta = \mathrm{id}$.

1. If $\mu = \lambda$, then we get of course $\mathcal{I}(\lambda, \langle\mathbf{x}\rangle) = \sum_i x_i$.
2. For $\mu(A) = \lambda(A)^2$ (a convex transformation), we obtain $\mathcal{I}(\lambda^2, \langle\mathbf{x}\rangle) = \sum_i (i^2 - (i-1)^2) \cdot x_i = 1x_1 + 3x_2 + 5x_3 + 7x_5 + 9x_6 + \ldots$. Thus, we put bigger weight for productivity here.
3. If $\mu(A) = \sqrt{\lambda(A)}$ (a concave transformation), then $\mathcal{I}(\sqrt{\lambda}, \langle\mathbf{x}\rangle) = \sum_i (\sqrt{i} - \sqrt{i-1}) \cdot x_i \simeq 1.00x_1 + 0.41x_2 + 0.32x_3 + 0.27x_4 + 0.24x_5 + 0.21x_6 + \ldots$. In consequence, the top-cited papers are of greater significance.

For instance, consider two vectors $\mathbf{y} = (60, 30, 10, 4, 0, 0, \ldots)$ (bigger quality) and $\mathbf{z} = (15, 13, 11, 11, 9, 8, 7, 7, 6, 5, 3, 3, 2, 1, 1, 1, 1, 0, 0, \ldots)$ (bigger productivity). We have $\mathcal{I}(\lambda, \langle\mathbf{y}\rangle) = \mathcal{I}(\lambda, \langle\mathbf{z}\rangle) = 104$, $\mathcal{I}(\lambda^2, \langle\mathbf{y}\rangle) \simeq 228 < \mathcal{I}(\lambda^2, \langle\mathbf{z}\rangle) \simeq 1050$, and $\mathcal{I}(\sqrt{\lambda}, \langle\mathbf{y}\rangle) \simeq 76.7 > \mathcal{I}(\sqrt{\lambda}, \langle\mathbf{z}\rangle) \simeq 36.9$.

Now let us discuss the effect of selecting different φ that are simple extensions of functions of one variable to $\mathcal{S}$.

Example 3. Let $\mathcal{I} = \mathrm{Su}, \mu = \lambda, \eta = \mathrm{id}$. We know that by choosing $\varphi(\mathbf{x}) = \lfloor\mathbf{x}\rfloor$ we obtain the *h*-index, H. It is easily seen that, e.g. $\mathrm{Su}(\lambda, \lfloor\sqrt{\mathbf{x}}\rfloor) = \mathsf{H2}(\mathbf{x})$. As we indicated in Section 2.1, many other Hirsch-based indices in fact use simple transformations of the input vector, like the one above. Moreover, by dropping the floor function we obtain the generalization of the *h*-index that is real-valued.

The φ function may be used, e.g. to change the impact of extremely high-cited publications, like when we choose $\varphi(\mathbf{x}) = \log(\mathbf{x} + 1)$.

Consideration of more complex $\varphi : \mathcal{S} \to \mathcal{S}$ functions may lead us to other notable aggregation operators: for example, the *g*- and *w*-index. Let $\mathtt{cummin}, \mathtt{cumsum} : \mathbb{R}^\infty \to \mathbb{R}^\infty$ denote the cumulative minimum and sum, respectively, i.e.:

$$\mathtt{cummin}(\mathbf{x}) = (x_1, x_1 \wedge x_2, x_1 \wedge x_2 \wedge x_3, \ldots),$$
$$\mathtt{cumsum}(\mathbf{x}) = (x_1, x_1 + x_2, x_1 + x_2 + x_3, \ldots).$$

Additionally, assume that operations $+, -$, and $\vee$ applied on vectors are performed element-wise.

Proposition 5. *Given $\mathbf{x} \in \mathcal{S}$ it holds:*

$$\mathsf{G}(\mathbf{x}) = \mathrm{Su}\Big(\lambda, \langle\lfloor 0 \vee \mathtt{cummin}(\mathtt{cumsum}(\mathbf{x}) - (1^2, 2^2, \ldots) + (1, 2, \ldots))\rfloor\rangle\Big),$$

Proof. Fix $\mathbf{x} \in \mathcal{S}$. Let $\mathbf{y} = \mathtt{cumsum}(\mathbf{x})$. By definition of the *g*-index (4):

$$\begin{aligned} \mathsf{G}(\mathbf{x}) = &\ \max\{g \in \mathbb{N}_0 :\ y_g - g^2 \geqslant 0\} \\ = &\ \max\left\{g \in \mathbb{N}_0 :\ \min_{i \leqslant g}\{y_i - i^2\} \geqslant 0\right\} \\ = &\ \max\left\{g \in \mathbb{N}_0 :\ 0 \vee \min_{i \leqslant g}\{y_i - i^2 + i\} \geqslant g\right\}. \end{aligned}$$

Thus, if we set $\mathbf{z} = 0 \vee \mathtt{cummin}((y_i - i^2 + i)_{i \in \mathbb{N}})$, then $\mathbf{z} \in \mathcal{S}$ and we see that $\mathsf{G}(\mathbf{x}) = \mathsf{H}(\mathbf{z}) = \mathrm{Su}(\lambda, \langle\lfloor\mathbf{z}\rfloor\rangle)$. $\square$

Proposition 6. *Given* $\mathbf{x} \in \mathcal{S}$ *it holds:*

$$\mathsf{W}(\mathbf{x}) = \mathsf{Su}(\lambda, \langle \lfloor \mathtt{cummin}(\mathbf{x} + (1, 2, \ldots) - 1) \rfloor \rangle),$$

Proof. Fix $\mathbf{x} \in \mathcal{S}$. By the definition of the w-index (5):

$$\begin{aligned} \mathsf{W}(\mathbf{x}) &= \max\{w \in \mathbb{N}_0 : \ x_i + i - 1 \geqslant w \text{ for all } i \leqslant w\} \\ &= \max\left\{w \in \mathbb{N}_0 : \ \min_{i \leqslant w}\{x_i + i - 1\} \geqslant w\right\}. \end{aligned}$$

By setting $\mathbf{z} = \mathtt{cummin}((x_i + i - 1)_{i \in \mathbb{N}})$ it holds $\mathbf{z} \in \mathcal{S}$ and we see that $\mathsf{W}(\mathbf{x}) = \mathsf{H}(\mathbf{z}) = \mathsf{Su}(\lambda, \langle \lfloor \mathbf{z} \rfloor \rangle)$. □

In some applications, the η function may introduce new "added value" to the aggregation process.

Example 4. Let $\mathcal{I} = \mathsf{Sh}, \mu = \lambda, \varphi = \mathrm{id}$. By setting $\eta = \mathrm{id}$ we of course get the MAXPROD-index, $\mathsf{M}P$. We may note, however, that the valuations generated by this index cannot be easily compared to that of the h-index. For example, we get $\mathsf{H}(n * n, 0, 0, \ldots) = n$ and $\mathsf{M}P(n * n, 0, 0, \ldots) = n^2$, where $(n * n) = (n, n, \ldots, n) \in \mathbb{I}^n$. Thus, by setting $\eta(x) = \sqrt{x}$ we may obtain the "calibrated" version of the MAXPROD index.

Of course, integrals other than the classical Choquet, Sugeno, or Shilkret, may also lead to interesting indices. Here are special types of decomposition integrals presented in [16, Def. 4.4]:

$$\mathcal{I}_{(k)}(\mu, \mathsf{f}) = \sup_{a_1, \ldots, a_k \geqslant 0} \left\{ \sum_{i=1}^{k} a_i \mu\left(\left\{u : \ \mathsf{f}(u) \geqslant \sum_{j=1}^{i} a_j\right\}\right)\right\}, \tag{11}$$

where $k \in \mathbb{N}$. It has been shown that these are also universal integrals in the sense of [13].

Example 5. We have $\mathcal{I}_{(1)} = \mathsf{Sh}$, which gives for $\mu = \lambda$ the maximal area of a rectangle with sides parallel to axes, that belongs to the graph $\mathrm{Gr}(\lambda, \mathsf{f}) = \mathrm{cl}(\{(x, y) \in \mathbb{I}^2 : \ y \leqslant \mathsf{h}^{(\lambda, \mathsf{f})}(x)\})$. Additionally, $\lim_{k \to \infty} \mathcal{I}_{(k)} = \mathsf{Ch}$. However, e.g. in the $\mathcal{I}_{(2)}(\lambda, \cdot)$ case, we get the maximal sum of areas for two non-overlapping rectangles.

Table 1
Ranks (the lower the better) for 11 bibliometricians assessed with $\mathcal{I}(\psi \circ \lambda, \langle \varphi(\mathbf{x}) \rangle)$.

#	$\mathcal{I}$	ψ	φ	A	B	C	D	E	F	G	H	I	J	K
1	Ch	id	id	7	2	3	8	1	9	6	5	4	10	11
2	Ch	id	$\sqrt{\cdot}$	5	4	2	8	1	10	7	3	6	9	11
3	Ch	id	$\cdot^2$	7	1	4	3	2	9	6	8	5	10	11
4	Ch	$\sqrt{\cdot}$	id	6	1	3	4	2	9	7	8	5	10	11
5	Ch	$\sqrt{\cdot}$	$\sqrt{\cdot}$	6	2	3	8	1	9	7	4	5	10	11
6	Ch	$\sqrt{\cdot}$	$\cdot^2$	5	1	7	3	2	4	8	9	6	10	11
7	Ch	$\cdot^2$	id	4	6	2	8	1	10	7	3	5	9	11
8	Ch	$\cdot^2$	$\sqrt{\cdot}$	4	5	2	8	1	10	7	3	6	9	11
9	Ch	$\cdot^2$	$\cdot^2$	8	3	2	7	1	9	5	6	4	10	11
10	Su	id	id	7	6	2	8.5	1	10.5	4	5	3	8.5	10.5
11	Su	id	$\sqrt{\cdot}$	8	3	1	4	2	11	6	7	5	9	10
12	Su	id	$\cdot^2$	6	7	2	8	1	10	5	4	3	9	11
13	Su	$\sqrt{\cdot}$	id	6	7	2	8	1	10	5	4	3	9	11
14	Su	$\sqrt{\cdot}$	$\sqrt{\cdot}$	7	6	2	8.5	1	10.5	4	5	3	8.5	10.5
15	Su	$\sqrt{\cdot}$	$\cdot^2$	4	5	2	8	1	10	7	3	6	9	11
16	Su	$\cdot^2$	id	8	3	1	4	2	11	6	7	5	9	10
17	Su	$\cdot^2$	$\sqrt{\cdot}$	8	1	2	7	6	9	5	4	3	10	11
18	Su	$\cdot^2$	$\cdot^2$	7	6	2	8.5	1	10.5	4	5	3	8.5	10.5
19	Sh	id	id	8	1	2	7	3	9	5	6	4	10	11
20	Sh	id	$\sqrt{\cdot}$	6	7	2	8	1	10	4	3	5	9	11
21	Sh	id	$\cdot^2$	5	1	6	3	2	4	9	8	7	10	11
22	Sh	$\sqrt{\cdot}$	id	5	1	6	3	2	4	9	8	7	10	11
23	Sh	$\sqrt{\cdot}$	$\sqrt{\cdot}$	8	1	2	7	3	9	5	6	4	10	11
24	Sh	$\sqrt{\cdot}$	$\cdot^2$	4	1	7	5	2	3	8	9	6	10	11
25	Sh	$\cdot^2$	id	6	7	2	8	1	10	4	3	5	9	11
26	Sh	$\cdot^2$	$\sqrt{\cdot}$	4	5	2	8	1	10	7	3	6	9	11
27	Sh	$\cdot^2$	$\cdot^2$	8	1	2	7	3	9	5	6	4	10	11

Table 2
Basic summary statistics of ranks in Table 1.

	A	B	C	D	E	F	G	H	I	J	K
Min.	4	1	1	3	1	3	4	3	3	8.50	10
1st Qu.	5	1	2	4.50	1	9	5	3.50	4	9	11
Median	6	3	2	8	1	10	6	5	5	9	11
Mean	6.18	3.48	2.78	6.57	1.70	8.87	6	5.37	4.74	9.43	10.87
3rd Qu.	7.50	6	3	8	2	10	7	7	6	10	11
Max.	8	7	7	8.50	6	11	9	9	7	10	11

4.3. A case study

For the sake of illustration, let us apply a few cases of the derived model on data presented in [11]. This data set consist of citation sequences of 11 prominent bibliometricians (A = Egghe L., B = Garfield E., C = Glänzel W., D = Ingwersen P., E = Leydesdorff L., F = McCain K.W., G = Moed H.F., H = Rousseau R., I = Van Raan A.F.J., J = Vinkler P., K = Zitt M.) and is available at M. Gagolewski's webpage.

Let us focus only on authors' rankings that are generated through impact functions defined by Eq. (10). Thus, the choice of η is meaningless here. Moreover, we shall consider all assessment tools that result in taking all the combinations of integrals from the set $\{\mathrm{Ch}, \mathrm{Su}, \mathrm{Sh}\}$, vector transformations $\varphi \in \{x \mapsto x, x \mapsto \sqrt{x}, x \mapsto x^2\}$, as well as monotone measures $\mu = \psi \circ \lambda$, with $\psi \in \{x \mapsto x, x \mapsto \sqrt{x}, x \mapsto x^2\}$.

Table 1 lists rankings obtained for all the 27 cases.

All Kendall's correlation coefficients between pairs of rankings are positive, but some are insignificant at significance level $\alpha = 0.05$ (i.e. uncorrelated, for cases 20–6, 25–6, 24–20, and 25–24). On the other hand, $\tau = 1$, i.e. we have exactly the same rankings, for case pairs 15–8, 26–8, 14–10, 18–10, 16–11, 13–12, 18–14, 26–15, 23–19, 27–19, 25–20, 27–23.

In overall, there is a considerable variability of the results, see Table 2 for a summary. Only the authors J and K consistently gain low valuations, with the maximum rank difference of 1.5 and 1, respectively. We see that without expert knowledge indicating which $\mathcal{I}, \psi, \varphi$ should be selected, one obtains too many possibilities. Moreover, please keep in mind that such a choice should be performed a priori (before calculating the rankings, to prevent any manipulations) and with great care.

5. Concluding remarks

We have introduced a model for the construction of zero-insensitive impact functions based on well-known tools from fuzzy/monotone measure theory, which is their another successful practical application. In our approach, in order to generate an aggregation operator, a decision maker has to select four objects (functions), each having a different, but clear and well-defined role. Of course, when it comes to assessment of a set of vectors, each should be treated with the same settings.

Recall that we have decided to focus on zero-insensitive impact functions. Future work should definitely concern models taking into account families of monotone measures and families of integrals, and establish as well as investigate interesting relationships between them. Moreover, we shall seek for a method that supports automated or semi-automated selection of $\eta, \mu, \mathcal{I}$, and φ.

Another extension, somewhere between zero-insensitiveness and the above-mentioned general case may take into account some kind of compensation, especially valid for assessing young researchers. As we know, at the start of an academic career, it is hard to expect that they gain immediately a big number of papers having a large number of citations. Thus, one may consider a different setting for the $\tilde{\cdot}$ operator. Let $\tilde{\mathbf{x}} = (x_{\{1\}} + c, x_{\{2\}} + c, \ldots, x_{\{n\}} + c, 0, 0, \ldots), c \geqslant 0$. For $c = 0$ we have our original setting. Now, for example, for $c = 1, 2, \ldots$ researchers with $i \leqslant c$ non-cited papers will have, e.g. the h-index equal to i, and the increment in the number of citations will affect the overall valuation only in the case of authors with productivity exceeding c papers.

Acknowledgments

The authors would like to thank anonymous reviewers for valuable suggestions on the first version of the manuscript.

The contribution of M. Gagolewski was partially supported by the European Union from resources of the European Social Fund, Project PO KL "Information technologies: Research and their interdisciplinary applications", Agreement UDA-POKL.04.01.01-00-051/10-00, and by the FNP START Scholarship from the Foundation for Polish Science.

The contribution of R. Mesiar has been done in connection with project APVV-0073-10 and also with IT4Innovations Center of Excellence, Reg. No. CZ.1.05/1.1.00/02.0070 supported by Research and Development for Innovations Operational Program financed by Structural Founds of Europe Union and from the means of state budget of the Czech Republic.

References

[1] S. Alonso, F.J. Cabrerizo, E. Herrera-Viedma, F. Herrera, h-index: A review focused on its variants, computation and standardization for different scientific fields, J. Informetrics 3 (2009) 273–289.

[2] G. Beliakov, S. James, Citation-based journal ranks: the use of fuzzy measures, Fuzzy Sets Syst. 167 (2011) 101–119.
[3] L. Bornmann, H.D. Daniel, What do citation counts measure? A review of studies on citing behavior, J. Doc. 64 (2008) 45–80.
[4] M. Bras-Amorós, J. Domingo-Ferrer, V. Torra, A bibliometric index based on the collaboration distance between cited and citing authors, J. Informetrics 5 (2011) 248–264.
[5] G. Choquet, Theory of capacities, Annales de l'institut Fourier 5 (1954) 131–295.
[6] L. Egghe, An improvement of the H-index: the G-index, ISSI Newsl. 2 (2006) 8–9.
[7] F. Franceschini, D.A. Maisano, The Hirsch index in manufacturing and quality engineering, Qual. Reliab. Eng. Int. 25 (2009) 987–995.
[8] F. Franceschini, D.A. Maisano, Structured evaluation of the scientific output of academic research groups by recent h-based indicators, J. Informetrics 5 (2011) 64–74.
[9] M. Gagolewski, P. Grzegorzewski, A geometric approach to the construction of scientific impact indices, Scientometrics 81 (2009) 617–634.
[10] M. Gagolewski, P. Grzegorzewski, Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals, Int. J. Approx. Reason. 52 (2011) 1312–1324.
[11] M. Gagolewski, R. Mesiar, Aggregating different paper quality measures with a generalized h-index, J. Informetrics 6 (2012) 566–579.
[12] J.E. Hirsch, An index to quantify individual's scientific research output, Proc. Natl. Acad. Sci. 102 (2005) 16569–16572.
[13] E. Klement, R. Mesiar, E. Pap, A universal integral as common frame for Choquet and Sugeno integral, IEEE Trans. Fuzzy Syst. 18 (2010) 178–187.
[14] M. Kosmulski, A new Hirsch-type index saves time and works equally well as the original h-index, ISSI Newsl. 2 (2006) 4–6.
[15] M. Kosmulski, MAXPROD — a new index for assessment of the scientific output of an individual, and a comparison with the h-index, Cybermetrics (2007) 11.
[16] R. Mesiar, A. Stupnňanová, Decomposition integrals, Int. J. Approx. Reason. 54 (2013) 1252–1259.
[17] A. Quesada, Monotonicity and the Hirsch index, J. Informetrics 3 (2009) 158–160.
[18] A. Quesada, More axiomatics for the Hirsch index, Scientometrics 82 (2010) 413–418.
[19] R. Rousseau, Woeginger's axiomatisation of the h-index and its relation to the g -index, the h(2)-index and the R^2-index, J. Informetrics 2 (2008) 335–340.
[20] M. Schreiber, A case study of the modified Hirsch index h_m accounting for multiple coauthors, J. Am. Soc. Inf. Sci. Technol. 60 (2009) 1274–1282.
[21] N. Shilkret, Maxitive measure and integration, Indag. Math. 33 (1971) 109–116.
[22] M. Sugeno, Theory of Fuzzy Integrals and Its Applications, Ph.D. thesis, Tokyo Institute of Technology, 1974.
[23] V. Torra, Y. Narukawa, The h-index and the number of citations: two fuzzy integrals, IEEE Trans. Fuzzy Syst. 16 (2008) 795–797.
[24] G.J. Woeginger, A symmetry axiom for scientific impact indices, J. Informetrics 2 (2008) 298–303.
[25] G.J. Woeginger, An axiomatic analysis of Egghe's g-index, J. Informetrics 2 (2008) 364–368.
[26] G.J. Woeginger, An axiomatic characterization of the Hirsch-index, Math. Soc. Sci. 56 (2008) 224–232.