Short Papers

Sugeno Integrals, H_α, and H^β Indices: How to Compare Scientists From Different Academic Areas

LeSheng Jin, Radko Mesiar, and Andrea Stupňanová

***Abstract*—The usage of the standard h-index to compare scientists from different research fields is often misleading. To avoid this possible failure, based on the Sugeno integrals, three new modified indices H_α, H^β, and H_α^β are proposed, thus, allowing to compensate the lower number of citations or of papers characteristic for a considered field. Our approach has a transparent geometric interpretation, and thus, it offers good candidates to replace the standard h-index H in considered databases, such as Web of Science (WOS), Scopus, or Google Scholar.**

***Index Terms*—Bibliometric databases, compensative h-index, h-index, Sugeno integral.**

I. Introduction

Citation analysis is a tool used worldwide for evaluating scientific performance and impact of various subjects, including scholars (scientists, researchers), journals, research organizations, or affiliations. Surely the most considered index in this area is the Hirsch index H. Recall that though the h-index was introduced to evaluate and compare excellent physicists [15], rather soon it was applied in any research field covered by some database, such as Web of Science (WOS), Scopus, and Google Scholar, to name a few. Almost immediately after the introduction of h-index H, several criticisms appeared; see [1] and [2] and compare also [3] and [4]. One of the major objections is related to the fact that this scientometric index does not distinguish the performance (number of papers and number of citations) typical for a particular field of science.

Recall that several scholars have discussed the problem of comparing scientists from different academic areas by means of the same fixed approach [5]–[10]. Among recent comments on field-dependent normalization of h-index, recall, e.g., [11], [12]. In particular, Podlubný [9], [13] has stressed and enumerated different citation attitudes in different branches. So, for example, one citation in mathematics can be roughly related to 15 citations in chemistry (see [13]). Clearly, then the h-index gives a weak information when comparing scientists in maths and chemistry, for example.

The aim of this contribution is a proposal of some modified versions of the h-index. Our approach is based on the fact that the h-index can be expressed in the form of the Sugeno integral with respect to the counting measure [14]. We also introduce another new Sugeno integral-based approach to the h-index and generalize it. In particular, we propose and study three simple modifications of these representations of the h-index. Our modifications are parametrized by constants $\alpha, \beta \in [0, \infty]$, leading to indices H_α, H^β, and H_α^β. The considered constants α and β reflect the different attitudes related to the (average) number of papers and citations in different scientific branches so that $H = H_1 = H^1$ and $H_\alpha = H_\alpha^1, H^\beta = H_1^\beta$ whenever $\alpha, \beta \in [0, 1]$.

Concerning the bounds, we propose $H_0(\mathbf{x}) = n$ (the number of papers of the considered scientist), $H^0(\mathbf{x}) = x_1$ (the number of citations of the highest cited paper of the considered scientist), $H_0^0(\mathbf{x}) = \max\{n, x_1\}$, and $H_\infty(\mathbf{x}) = H^\infty(\mathbf{x}) = H_\infty^\infty(\mathbf{x}) = 0$. We have in mind that the overall spread of the h-index is related to its simple verbal description and simple geometric interpretation (see, e.g., the H-figures in Scopus). Therefore, our proposals aim to have similar features, and then the only debatable problem remains an appropriate choice of constants α and β.

This paper is organized as follows. In the next section, we recall h-index and its Sugeno integral representations. In Section III, we introduce and discuss the H_α index. Section IV is devoted to the proposal and study of the H^β index. In Section V, H_α^β index is considered and some hints how to choose constants α and β are given in Section VI concludes this paper.

II. h- Index and the Sugeno Integral

We recall first the original Hirsch's introduction of his h-index H.

Definition 2.1 (Hirsch [15]): A scientist has index $H = h$ if h of his or her total of n papers have at least h citations each but there are no $h + 1$ papers having each at least $h + 1$ citations.

For the mathematical evaluation of the h-index, we will consider framework used in [16]. For a considered scholar X having n papers in a considered fixed database (WOS, Scopus, and Google Scholar), the numerical performance $\mathbf{x}$ of X is an n-tuple, $\mathbf{x} = (x_1, \ldots, x_n)$, where the value x_i denotes the number of citations received by the ith most cited paper of the scholar X. Clearly, then $\mathbf{x} \in \mathcal{S}$, where $\mathcal{S}$ is the set of decreasingly ordered non-negative integer sequences of any finite length, i.e.

$$\mathcal{S} = \{(x_1, \ldots, x_n) | n \in \mathbb{N}, x_1, \ldots, x_n \in \mathbb{N}_0, x_1 \geq \cdots \geq x_n\}. \quad (1)$$

Here, $\mathbb{N}$ stands for the set of all positive integers, while $\mathbb{N}_0$ is created by all nonnegative integers, i.e., $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Observe that a scholar X with no relevant publication is not considered in any of the mentioned

Manuscript received September 27, 2018; revised January 29, 2019; accepted April 24, 2019. Date of publication May 2, 2019; date of current version April 1, 2020. This work was supported by the Slovak Research and Development Agency under Contract no. APVV-17-0066 and Grant VEGA 1/0682/16. *(Corresponding author: Andrea Stupňanová.)*

L. Jin is with the Business School, Nanjing Normal University, Nanjing 210017, China (e-mail: jls1980@163.com).

R. Mesiar is with the Department of Mathematics and Descriptive Geometry, FCE Slovak University of Technology in Bratislava, 801 05 Bratislava, Slovakia, and also with the Institute of Information Theory and Automation, Czech Academy of Sciences, 182 08 Praque, Czech Republic (e-mail: radko.mesiar@stuba.sk).

A. Stupňanová is with the Department of Mathematics and Descriptive Geometry, FCE Slovak University of Technology in Bratislava, 810 05 Bratislava, Slovakia (e-mail: andrea.stupnanova@stuba.sk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TFUZZ.2019.2914625

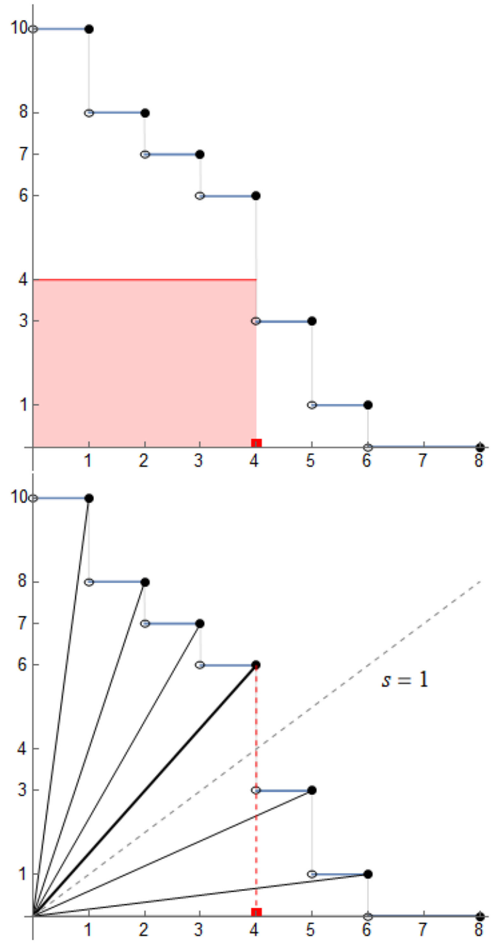

Fig. 1. Illustration of (up) formula (2) and of (down) formula (3) for $H(\mathbf{x}) = 4$.

databases. Note that then the h-index H is a function, $H : \mathcal{S} \to \mathbb{N}_0$, defined by

$$H(\mathbf{x}) = \bigvee_{i=1}^{n} (x_i \wedge i) = \max\{\min\{x_1, 1\}, \dots, \min\{x_n, n\}\}. \quad (2)$$

Here, and also later, we use the alternative notations $\vee = \max$ and $\wedge = \min$ in a way to ensure the transparency of the related formulas. As an alternative approach, one can consider the slopes $s_i = \frac{x_i}{i}$ of the straight lines connecting points (0, 0) and (i, x_i), and then

$$H(\mathbf{x}) = \max\{i \in \{1, \dots, n\} | s_i \geq 1\} \quad (3)$$

with the convention $\max \emptyset = 0$. Two formulas (2) and (3) offer two different geometric interpretations of the index H, see Fig. 1.

In the first case, when formula (2) is considered, the h-index $H(\mathbf{x}) = k$ corresponds to the greatest square $]0, k]^2$, $k \in \mathbb{N}_0$, contained in the endograph of the function $f_{\mathbf{x}} :]0, n] \to \{x_1, \dots, x_n\}$, $f_{\mathbf{x}}(t) = x_i$ whenever $t \in]i-1, i]$, $i \in \{1, \dots, n\}$.

In the second case, considering formula (3), $H(\mathbf{x}) = k$ means that the slope $s_k = \frac{x_k}{k} \geq 1$ (i.e., the segment connecting point (0, 0) and (k, x_k) is in the positive cone determined by the y-axis and the axis of the first quadrant), but $s_{k+1} < 1$ (or $k = n$). We visualize both approaches on a record $\mathbf{x} = (10, 8, 7, 6, 3, 1, 0, 0)$, i.e., $H(\mathbf{x}) = 4$.

The h-index H can be seen as a special Sugeno integral, too. Recall that considering a measurable space $(\mathbb{N}, 2^{\mathbb{N}})$ and a monotone measure $\mu : 2^{\mathbb{N}} \to [0, \infty]$, $\mu(\emptyset) = 0$, and $\mu(A) \leq \mu(B)$ whenever $A \subseteq B \subseteq \mathbb{N}$, the related Sugeno integral $\mathbf{Su}_\mu : [0, \infty]^{\mathbb{N}} \to [0, \infty]$ is given by

$$\mathbf{Su}_\mu(f) = \bigvee_{a=0}^{\infty} \Big(a \wedge \mu(\{n \in \mathbb{N} | f(n) \geq a\}) \Big)$$

$$= \bigvee_{A \subseteq \mathbb{N}} \Big(\mu(A) \wedge \big(\bigwedge_{n \in A} f(n) \big) \Big). \quad (4)$$

For more details concerning the Sugeno integral see [17]–[20].

Note that the set $\mathcal{S}$ given by (1) can be embedded into $[0, \infty]^{\mathbb{N}}$, assigning to $\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{S}$ an infinitary vector $\bar{\mathbf{x}} \in [0, \infty]^{\mathbb{N}}$ given by $\bar{\mathbf{x}} = (x_1, \dots, x_n, 0, \dots)$. Coming back to the h-index, Torra and Narukawa [14] have shown that, for any $\mathbf{x} \in \mathcal{S}$

$$H(\mathbf{x}) = \mathbf{Su}_m(\bar{\mathbf{x}}) \quad (5)$$

where $m : 2^{\mathbb{N}} \to [0, \infty]$ is the counting measure given by $m(A) = \mathrm{card}(A)$. However, there is also a new representation of the h-index based on the Sugeno integral.

Theorem 2.1: Let $\mathbf{x} \in \mathcal{S}$. Then

$$H(\mathbf{x}) = \mathbf{Su}_{\mu_{\mathbf{x}}}(\mathrm{id}) \quad (6)$$

where $\mathrm{id} : \mathbb{N} \to [0, \infty], \mathrm{id}(k) = k$, is the identity function, and $\mu_{\mathbf{x}} : 2^{\mathbb{N}} \to [0, \infty]$, is an additive measure such that, for any $i \in \mathbb{N}, \mu_{\mathbf{x}}(\{i\}) = x_i - x_{i+1}$, using the convention $x_{n+1} = x_{n+2} = \dots = 0$.

Proof: Based on the formula (4)

$$\mathbf{Su}_{\mu_{\mathbf{x}}}(id) = \bigvee_{i=1}^{\infty} \Big(i \wedge \mu_{\mathbf{x}}(\{k \in \mathbb{N} | \mathrm{id}(k) \geq i\}) \Big)$$

$$= \bigvee_{i=1}^{n} \Big(i \wedge \sum_{k=i}^{n} \mu_{\mathbf{x}}(\{k\}) \Big) = \bigvee_{i=1}^{n} (i \wedge x_i) = H(\mathbf{x})$$

where the last equality follows from (2). ■

In the next sections, we will modify formulas (5) and (6) to get a more fair evaluation of scholars from different areas.

III. H_α-Index

To find a more fair evaluation of scholars working in the scientific domains with smaller number of citations, we propose the next index H_α based on a parameter $\alpha \in [0, \infty]$ and formula (6).

For any increasing (not necessarily strictly increasing) mapping $\varphi : [0, \infty] \to [0, \infty], \varphi(0) = 0$, and a monotone measure $\mu : 2^{\mathbb{N}} \to [0, \infty]$, obviously also $\varphi \circ \mu : 2^{\mathbb{N}} \to [0, \infty], \varphi \circ \mu(A) = \varphi(\mu(A))$, is a monotone measure. Therefore, for any $\mathbf{x} \in \mathcal{S}$, the Sugeno integral $\mathbf{Su}_{\varphi \circ \mu}(\mathrm{id})$ is well defined. For $\alpha \in]0, \infty[$, consider $\varphi_\alpha : [0, \infty] \to [0, \infty]$

$$\varphi_\alpha(t) = \frac{t}{\alpha}.$$

Then, modifying formula (6)

$$\mathbf{Su}_{\varphi_\alpha \circ \mu_{\mathbf{x}}}(\mathrm{id}) = \bigvee_{i=1}^{n} \Big(i \wedge \frac{x_i}{\alpha} \Big).$$

Note that the aforementioned Sugeno integral need not have integer value. To ensure this, we propose the next H_α index, $H_\alpha : \mathcal{S} \to \mathbb{N}_0$

$$H_\alpha(\mathbf{x}) = \bigvee_{i=1}^{n} \Big\lfloor \min\Big\{ \frac{x_i}{\alpha}, i \Big\} \Big\rfloor \quad (7)$$

(here, $\lfloor \cdot \rfloor$ is the floor of a real number).

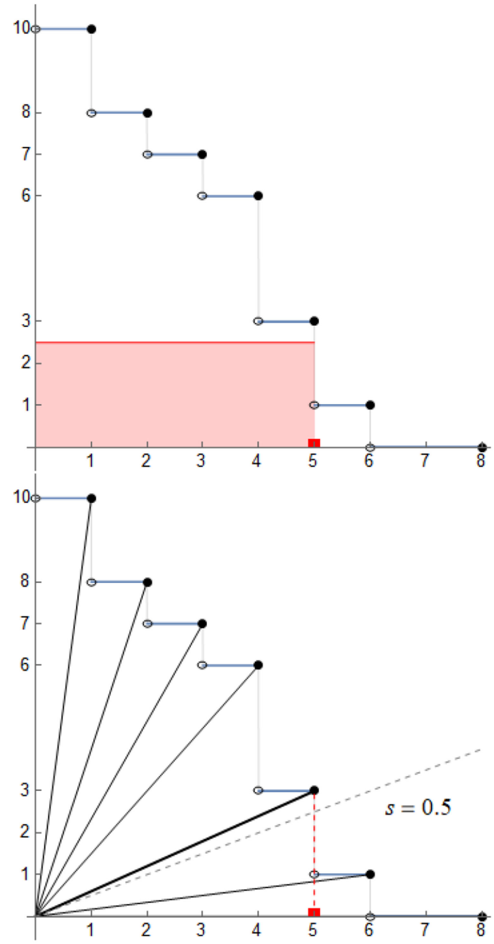

Fig. 2. Illustration of (top) formula (7) and of (bottom) formula (8) for $H_{0.5}(\mathbf{x}) = 5$.

It is not difficult to check that $\lim_{\alpha \to 0^+} H_\alpha(\mathbf{x}) = j$ whenever $\mathbf{x} = (x_1, \ldots, x_n)$ is such that $x_j > 0$ and either $j = n$ or $x_{j+1} = 0$. However, related to the next verbal characterization of H_α, we prefer to adopt a convention $H_0(\mathbf{x}) = n$ for any $\mathbf{x} \in \mathcal{S}, \mathbf{x} = (x_1, \ldots, x_n)$. On the other hand, obviously $\lim_{\alpha \to +\infty} H_\alpha(\mathbf{x}) = 0$ and $H_1(\mathbf{x}) = H(\mathbf{x})$ (i.e., H_1 is the standard h-index) for any $\mathbf{x} \in \mathcal{S}$.

Definition 3.1: A scientist has index $H_\alpha = h$, if he or she has h papers, each having at least $\alpha \cdot h$ citations, but there are no $h+1$ papers having each at least $\alpha \cdot (h+1)$ citations.

Dealing with slopes s_i, it is not difficult to see an other formula for H_α equivalent to (7)

$$H_\alpha(\mathbf{x}) = \max\{i \in \{1 \ldots, n\} | s_i \geq \alpha\}. \tag{8}$$

We have again two possible geometric interpretations. Considering formula (7), we see that $H_\alpha(\mathbf{x}) = k$ whenever the rectangle $]0, k] \times [0, \alpha \cdot k]$ is the greatest rectangle of type $]0, i] \times [0, \alpha \cdot i]$, $i \in \{0, 1, \ldots, n\}$, contained in the endograph of the function $f_\mathbf{x}$ [see Fig. 2(top)]. For this interpretation of H_α, one can compare the approach discussed in [21]. On the other hand, considering the formula (8), $h_\alpha(\mathbf{x}) = k$ means that the slope $s_k = \frac{x_k}{k} \geq \alpha$ but $s_{k+1} < \alpha$ (or $k = n$), see Fig. 2(bottom). The next figures are again related to the record $\mathbf{x} = (10, 8, 7, 6, 3, 1, 0, 0)$, and $\alpha = 0.5$ is considered. Hence, $H_{0.5}(\mathbf{x}) = 5$.

Two next facts easily follow from Definition 2.1, or from formulas (7) or (8).

Proposition 3.1: For any $\mathbf{x}, \mathbf{y} \in \mathcal{S}$, $\mathbf{x} \leq \mathbf{y}$ (i.e., $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_m)$, $n \leq m$ and $x_1 \leq y_1, \ldots, x_n \leq y_n$), and any $\alpha \in [0, 1]$, it holds

$$H_\alpha(\mathbf{x}) \leq H_\alpha(\mathbf{y}).$$

Proposition 3.2: For any $\alpha_1, \alpha_2 \in [0, 1]$, $\alpha_1 < \alpha_2$, and $\mathbf{x} \in \mathcal{S}$, it holds

$$H_{\alpha_1}(\mathbf{x}) \geq H_{\alpha_2}(\mathbf{x}).$$

Proposition 3.1 shows the monotonicity of each index H_α, i.e., for any new citation or paper linked to the scientist X, the updated index H_α cannot decrease, i.e., H_α is a valid scientometric index. On the other hand, Proposition 3.2 shows that smaller parameter α can compensate smaller number of citations.

As an easy corollary of mentioned facts we see that, for any $\alpha \in [0, 1]$, it holds $H_\alpha(\mathbf{x}) \in \{k, \ldots, n\}$, where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{S}$ and $k = H(\mathbf{x})$. Then, for any $i \in \{k, \ldots, n\}$, one can completely describe the set of all parameters α such that $H_\alpha(\mathbf{x}) = i$.

Example 3.1: Consider, as previously, $\mathbf{x} = (10, 8, 7, 6, 3, 1, 0, 0)$. Then $s_1 = 10$, $s_2 = 4, s_3 = \frac{7}{3}$, $s_4 = \frac{3}{2}$, $s_5 = \frac{3}{5}$, $s_6 = \frac{1}{6}$, $s_7 = s_8 = 0$, and thus

$$H_\alpha(\mathbf{x}) = \begin{cases} 0, & \text{if } \alpha > 10 \\ 1, & \text{if } \alpha \in]4, 10] \\ 2, & \text{if } \alpha \in \left]\frac{7}{3}, 4\right] \\ 3, & \text{if } \alpha \in \left]\frac{6}{5}, \frac{7}{3}\right] \\ 4, & \text{if } \alpha \in \left]\frac{3}{5}, \frac{6}{5}\right] \\ 5, & \text{if } \alpha \in \left]\frac{1}{6}, \frac{3}{5}\right] \\ 6, & \text{if } \alpha \in \left]0, \frac{1}{6}\right] \\ 8, & \text{if } \alpha = 0. \end{cases}$$

Remark 3.1: Observe that if $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{x}_1 = (x_1 \wedge n, \ldots, x_n \wedge n)$, then $H(\mathbf{x}) = H(\mathbf{x}_1)$, i.e., for the standard h-index H and for any considered paper, citations over the total number of papers do not count. Based on this observation, one can compute the index $H_\alpha(\mathbf{x})$ as a standard h-index of a modified record $\mathbf{x}_\alpha$, where $\mathbf{x}_\alpha = (/x_1/_{\alpha,n}, \ldots, /x_n/_{\alpha,n})$ with

$$/x_i/_{\alpha,n} = \begin{cases} \lfloor \frac{x_i}{\alpha} \rfloor \wedge n, & \text{if } \alpha \neq 0 \\ x_i \wedge n, & \text{if } \alpha = 0. \end{cases}$$

So, for $\mathbf{x} = (10, 8, 7, 6, 3, 1, 0, 0)$, we have $\mathbf{x}_{0.5} = (8, 8, 8, 8, 6, 2, 0, 0)$, and thus, $H_{0.5}(\mathbf{x}) = H(\mathbf{x}_{0.5}) = 5$.

Remark 3.2: Note that the aforementioned formula $\mathbf{Su}_{\varphi_\alpha \circ \mu_\mathbf{x}}(\mathrm{id})$ is related to h_α-index (not integer valued, in general) proposed and discussed in [22]. The related graphical representation [see also Fig. 2 (down)] can be found also in [23, Fig. 2(left)].

IV. H^β Index

The index H_α introduced in the previous section is able to compensate a lower number of citations. Somewhat similarly, one can attempt to compensate a lower number of papers.

Definition 4.1: A scientist has index $H^\beta = h$, if he or she has at least $\lfloor \beta \cdot h \rfloor \vee 1$ papers, each having at least h citations, but there are no $\lfloor \beta \cdot (h+1) \rfloor \vee 1$ papers having each at least $h+1$ citations.

In this case, the next equivalent formulas can be introduced, supposing $\beta \in]0, \infty[$ (note that $H^0(\mathbf{x}) = x_1$)

$$H^\beta(\mathbf{x}) = \max\left\{j \in \left\{1, \ldots, \left\lfloor \frac{n}{\beta} \right\rfloor\right\} \Big| x^\beta_j \geq j\right\} = H(\mathbf{x}^\beta) \tag{9}$$

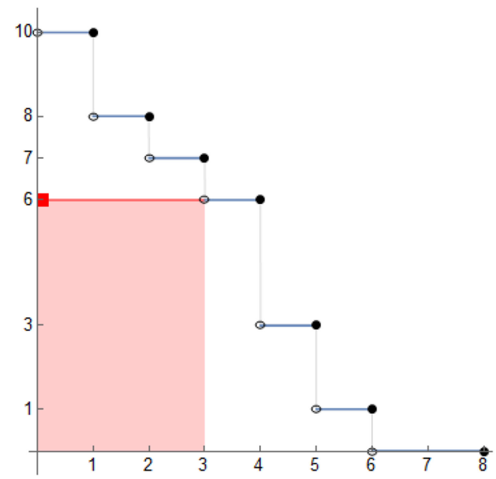

Fig. 3. Illustration of Example 4.1 with $H^{0.5}(\mathbf{x}) = 6$.

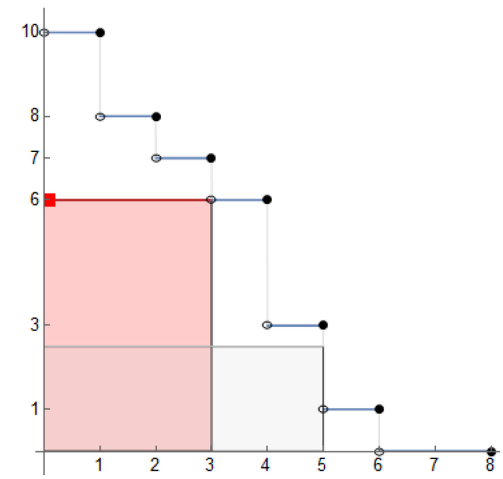

Fig. 4. Illustration of $H^{0.5}_{0.5}(\mathbf{x}) = 6$ from Example 5.1.

where $x^{\beta}_j = f_\beta(j),\ f_\beta : \left]0, \frac{n}{\beta}\right] \to \{x_1, \ldots, x_n\}$

$$f_\beta(t) = x_i \text{ whenever } \frac{i-1}{\beta} < t \le \frac{i}{\beta},\ i \in \{1, \ldots, n\}$$

and

$$H^\beta(\mathbf{x}) = \max\left\{ j \in \left\{1, \ldots, \left\lfloor \frac{n}{\beta} \right\rfloor \right\} \,\middle|\, x_{\lceil \beta \cdot j \rceil} \ge j \right\} \tag{10}$$

where $\lceil . \rceil$ is the ceiling of a real number. Note that if $\beta = \frac{1}{m}$ for some integer $m \in \mathbb{N}$, then, to compute H^β, we have simply to repeat the record $\mathbf{x}$ m-times and reorder the related $n \cdot m$-tuple in decreasing sense, and then, apply the standard h-index H to this extended record. It is not difficult to check that $H^0(\mathbf{x}) = x_1$, while $H^\infty(\mathbf{x}) = \lim_{\beta \to +\infty} H^\beta(\mathbf{x}) = 0$.

Similarly as in the case of the H_α-index, also H^β-index is related to a modification of the Sugeno integral. Considering the representation (5) of the h-index H, we modify first the counting measure m by means of transformation φ_β, i.e., we consider a new monotone measure $\varphi_\beta \circ m$ given by

$$\varphi_\beta \circ m(A) = \frac{\text{card}(A)}{\beta} \quad (\text{for } \beta \in]\,0, \infty[\,).$$

As the Sugeno integral $\mathbf{Su}_{\varphi_\beta \circ m}(\mathbf{x})$ need not have an integer value, we consider now the ceiling of $\mathbf{Su}_{\varphi_\beta \circ m}(\mathbf{x})$, which gives then exactly the H^β-index introduced in Definition 4.1.

Example 4.1: Continuing in Example 3.1, with $\beta = 0.5$, we have

$$\mathbf{x}^{0.5} = (10, 10, 8, 8, 7, 7, 6, 6, 3, 3, 1, 1, 0, 0, 0, 0)$$

and then, $H^{0.5}(\mathbf{x}) = H(\mathbf{x}^{0.5}) = 6$.

Concerning the geometric interpretation, observe that $H^\beta(\mathbf{x}) = k$ means that the rectangle $]0, \beta \cdot k] \times [0, k]$ is the greatest rectangle of type $]0, \beta \cdot i] \times [0, i]$, $i \in \{0, 1, \ldots, x_1\}$ contained in the endograph of the function $f_{\mathbf{x}}$ introduced in Section II; see also Fig. 3 and compare [21].

Also, in this case H^β is a valid scientometric index.

Proposition 4.1: For any $\mathbf{x}, \mathbf{y} \in \mathcal{S}$, $\mathbf{x} \le \mathbf{y}$ and any $\beta \in]0, \infty[$

$$H^\beta(\mathbf{x}) \le H^\beta(\mathbf{y}).$$

Proposition 4.2: For any $\beta_1, \beta_2 \in]0, \infty[,\ \beta_1 < \beta_2$, and $\mathbf{x} \in \mathcal{S}$, it holds

$$H^{\beta_1}(\mathbf{x}) \ge H^{\beta_2}(\mathbf{x}).$$

V. Index H^β_α and the Choice of Parameters α and β

Based on Propositions 3.1 and 4.1, one can aggregate the indices H_α and H^β to obtain a valid scientometric index for authors with compensation of low number of citations and/or low number of papers, related to the standards in the considered scientific domain. We propose the next index H^β_α, though several other possibilities could be considered alternatively.

Definition 5.1: Let $\alpha, \beta \in]0, \infty[$ and $\mathbf{x} \in \mathcal{S}$. Then, the index $H^\beta_\alpha : \mathcal{S} \to \mathbb{N}_0$ is given by

$$H^\beta_\alpha(\mathbf{x}) = \max\{H_\alpha(\mathbf{x}), H^\beta(\mathbf{x})\}.$$

The following facts are easy to be verified:

1) H^β_α is increasing in argument $\mathbf{x}$ and decreasing in both parameters α, β;
2) $H^1_1 = H$, $H^1_\alpha = H_\alpha$, and $H^\beta_1 = H^\beta$ for any $\alpha, \beta \in [0, 1]$;
3) $H^0_0(\mathbf{x}) = \max\{n, x_1\}$;
4) $H^\infty_\infty(\mathbf{x}) = 0$.

Example V.1: Continuing in Examples 3.1 and 4.1, we see that $H^{0.5}_{0.5}(\mathbf{x}) = 6$.

We also have a geometric interpretation of the index H^β_α. Namely, $H^\beta_\alpha(\mathbf{x}) = k$ whenever either $]0, k] \times [0, \alpha \cdot k]$ or $]0, \beta \cdot k] \times [0, k]$ is contained in the endograph of the function $f_{\mathbf{x}}$ introduced in Section II, but this is not more true for the rectangle $]0, k+1] \times [0, \alpha \cdot (k+1)]$ nor for $]0, \beta \cdot (k+1)] \times [0, k+1]$; see also Fig. 4.

Though all introduced indices H_α, H^β, and H^β_α have a compensation ability for a more fair comparison of researchers coming from different scientific domains, their real performance heavily depends on the appropriate choice of constants α and β. A deeper discussion of this topic is a problem for the near future. Here, we present only some direct usages and applications of the methods. One possible choice is to consider an average h-index of scientists in considered research areas then to put, for a chosen area*

$$\alpha = \beta = \frac{h^*_{\text{area}}}{\max\{h_{\text{area}} \mid \text{all considered areas}\}}.$$

Another possibility is to derive α and β based on average (median) impact factors of journals in the considered area. We focus now in more details to a work of Podlubný [9] (see also [13]) where the ratios of average number of citations in a considered scientific area in comparison with Mathematics was considered. Podlubný's results are summarized in Table I.

TABLE I
COMPARISON OF THE NUMBERS OF CITATIONS IN DIFFERENT FIELDS OF SCIENCE

Field	Average ratio of citation number to the number of citations in Mathematics
Clinical medicine	78
Biomedical research	78
Biology	8
Chemistry	15
Physics	19
Earth/space sciences	9
Engineering/technology	5
Mathematics	1
Social/behavioral sciences	13

Based on the data from Science and Engineering Indicators 2004. National Science Foundation, May 04, 2004, for the period 1992–2001, see [9] and [13].

When comparing a scientist X having his/her record $\mathbf{x}$ and working in field i with average ratio r_i coming from Table I, and a scientist Y characterized by a record $\mathbf{y}$ and working in a field j characterized by r_j, suppose $r_i \geq r_j$, and then, we propose to consider $\alpha = \beta = \sqrt{\frac{r_j}{r_i}}$. Subsequently, to compare the scientists X and Y, we have to compare indices $H(\mathbf{x})$ and $H_\alpha^\beta(\mathbf{y})$.

Example 5.2: Consider a mathematician Y characterized by a record $\mathbf{y} = (6, 4, 2, 1, 0)$ and an earth/space sciences researcher X with record $\mathbf{x}$ considered in the previous examples, i.e., $\mathbf{x} = (10, 8, 7, 6, 3, 1, 0, 0)$. Then, $r_j = 1$ and $r_i = 9$ (see Table I), and then, $\alpha = \beta = \sqrt{\frac{r_j}{r_i}} = \frac{1}{3}$. Recall that $H(\mathbf{x}) = 4$. Next, $H_{\frac{1}{3}}(\mathbf{y}) = 3$ and $H^{\frac{1}{3}}(\mathbf{y}) = H(6, 6, 6, 4, 4, 4, 2, 2, 2, 1, 1, 1, 0, 0, 0) = 4$, i.e., $H_{\frac{1}{3}}^{\frac{1}{3}}(\mathbf{y}) = 4$. We can conclude that both scientists have the same scientific performance, though $H(\mathbf{x}) = 4$ and $H(\mathbf{y}) = 2$.

Another possible approach how to estimate reasonable parameters α and β can be based on In Cites Journal Citation Reports, see Table II.

We see an evident difference is citation attitudes in different categories. For comparing scientists working in different categories, we propose to consider $\alpha = \beta = \text{AIF}$ (related to the last considered year), where the aggregate impact factor (AIF) for a subject category is calculated the same way as the impact factor for a journal, but it takes into account the number of citations to all journals in the category and the number of articles from all journals in the category. An aggregate impact factor of 1.0 means that, on average, the articles in the subject category published one or two years ago have been cited one time. Observe that (in this case, each scientist working in subject category i and having n publications, each with $n \times \text{AIF}_i$ number of citations (rounded to an integer) will satisfy $H_{\text{AIF}_i}(\mathbf{x}) = H^{\text{AIF}_i}(\mathbf{x}) = H_{\text{AIF}_i}^{\text{AIF}_i}(\mathbf{x}) = n$, considering n sufficiently large.

Example 5.3: Continuing in Example 5.2, observe that for a mathematician Y the related $\text{AIF}_1 = 0.855$, while for the researcher X (we have considered as the related category Meteorology and Atmospheric sciences), we have $\text{AIF}_2 = 3.143$. Then, $H_{0.855}(\mathbf{y}) = H^{0.855}(\mathbf{y}) = H_{0.855}^{0.855}(\mathbf{y}) = 2$ as well as $H_{3.143}(\mathbf{x}) = H^{3.143}(\mathbf{x}) = H_{3.143}^{3.143}(\mathbf{x}) = 2$, i.e., again we can conclude the same scientific performance of both scientists.

Note that some larger field, cover several categories. In such a case, we can compute the related AIF by means of total cites CT_i and AIF_i related to single categories

$$\text{AIF} = \frac{\sum \text{CT}_i}{\sum \frac{\text{CT}_i}{\text{AIF}_i}}$$

TABLE II
CATEGORIES DATA FILTERED BY: SELECTED JCR YEAR: 2017 SELECTED EDITIONS: SCIE, SSCI

Rank	Category	Edition	♯ Journals	Total Cites	Median IF	Aggregate IF
1	Economics	SSCI	353	905 730	1.112	1.766
2	Mathematics	SCIE	310	494 556	0.704	0.855
3	Biochemistry & Molecular biology	SCIE	293	3 625 819	2.906	4.281
4	Materials science, multidisciplinary	SCIE	285	3 451 318	2.008	4.646
5	Neurosciences	SCIE	261	2 346 383	3.047	4.015
6	Pharmacology & Pharmacy	SCIE	261	1 571 415	2.481	3.148
7	Engineering, Electrical & Electronic	SCIE	260	1 636 339	1.820	2.723
8	Mathematics, applied	SCIE	252	538 241	0.972	1.299
9	Enviromental sciences	SCIE	242	1 893 304	2.071	3.488
10	Education & Education research	SSCI	239	346 922	1.333	1.542
77	Meteorology & Atmospheric sciences	SCIE	86	622 212	1.984	3.143

i.e., the global AIF of this considered field is the weighted harmonic mean of the local AIF_i's.

VI. CONCLUDING REMARKS

Based on the Sugeno integral approach, we have introduced scientometric indices H_α, H^β, and H_α^β allowing to compare the performance of scientists from different research fields. Namely, H_α is capable to compensate missing citations in fields where the standard average number of citations per paper is low. Similarly, H^β allows to compensate a lower number of cited papers in fields with typically lower paper production. Finally, H_α^β considers both forms of compensation, $H_\alpha^\beta = \max\{H_\alpha, H^\beta\}$. Though we have indicated some possible approaches for an appropriate choice of compensative constants α and β, we are aware of the fact that such a choice should be better justified. This open problem is a challenge for the near future and can have different solutions for different databases. We expect that, based on a deeper analysis similar to that of Podlubny [9], [13] when the number of citations in different fields were studied and their constant ratio was verified, for different fields different constants α and β will be approved, considering $\alpha = 1$ ($\beta = 1$) for fields with highest average citation number per paper (clinical medicine or biomedical research?), and assigning the lowest constants α (β) to fields with the lowest average citation number per paper (Mathematics?). Alternatively, we can consider $\alpha = 1$ ($\beta = 1$) for a "median" field (Social/behavioral sciences with $r = 13$) and assign the relevant constants α and β based on Table I, putting $\alpha = \beta = \sqrt{\frac{r_i}{13}}$. As another promising approach, one can consider α or β derived from the AIF, which can be found, for single considered categories, in Journal Citation Reports.

REFERENCES

[1] J. Iglesias and C. Pecharroman, "Scaling the h-index for different ISI fields," *Scientometrics*, vol. 73, no. 3, pp. 303–320, 2007. [Online]. Available: http://dx.doi.org/10.1007/s11192-007-1805-x
[2] L. Bornmann and H. D. Daniel, "What do citation counts measure? A review of studies on citing behavior," *J. Documentation*, vol. 64, no. 1, pp. 45–80, 2008, doi: 10.1108/00220410810844150.

[3] M. V. Anauati, S. Galiani, and R. H. Gálvez, "Quantifying the life cycle of scholarly articles across fields of economic research," *Econ. Inquiry*, vol. 54, no. 2, pp. 1339–1355, 2016, doi: 10.1111/ecin.12292.

[4] A. Yong, "Critique of Hirschs h index: A combinatorial fermi problem," *Notices Amer. Math. Soc.*, vol. 61, no. 9, pp. 1040–1050, 2014. [Online]. Available: http://dx.doi.org/10.1090/noti1164

[5] S. Alonso, F. J. Cabrerizo, E. Herrera-Viedma, and F. Herrera, "h-index: A review focused in its variants, computation and standardization for different scientific fields," *J. Informetrics*, vol. 3, no. 4, pp. 273–289, 2009.

[6] P. Ball, "Index aims for fair ranking of scientists," *Nature*, vol. 436, p. 900, 2005.

[7] W. Glänzel and A. Schubert, "A new classification scheme of science fields and subfields designed for scientometric evaluation purposes," *Scientometrics*, vol. 56, no. 3, pp. 357–367, 2003.

[8] S. Lehmann, A. D. Jackson, and B. E. Lautrup, "Measures for measures," *Nature*, vol. 444, no. 7122, pp. 1003–1004, 2006.

[9] I. Podlubny, "Comparison of scientific impact expressed by the number of citations in different field of science," *Scientometrics*, vol. 64, no. 1, pp. 95–99, 2005.

[10] D. F. Taber, "Quantifying publication impact," *Science*, vol. 309, no. 5744, p. 2166, 2005.

[11] L. Bornmann and L. Leydesdorff, "Count highly-cited papers instead of papers with h citations: Use normalized citation counts and compare "like with like"!," *Scientometrics*, vol. 115, pp. 1119–1123, 2018.

[12] J. A. Teixeira da Silva and J. Dobránszki, "Multiple versions of the h-index: Cautionary use for formal academic purposes," *Scientometrics*, vol. 115, no. 2, pp. 1107–1113, 2018.

[13] I. Podlubny and K. Kassayova, "Towards a better list of citation superstars: Compiling a multidisciplinary list of highly cited researchers," *Res. Eval.*, vol. 15, no. 3, pp. 154–162, 2006.

[14] V. Torra and Y. Narukawa, "The h-index and the number of citations: Two fuzzy integrals," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 3, pp. 795–797, Jun. 2008.

[15] J. E. Hirsch, "An index to quantify an individual's scientific research output," in *Proc. Nat. Acad. Sci.*, 2005, vol. 102, no. 4, pp. 16569–16572. [Online]. Available: www.pnas.org/cgi/doi/10.1073/pnas.0507655102

[16] R. Mesiar and M. Gagolewski, "H-index and other Sugeno integrals: Some defects and their compensation," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 6, pp. 1668–1672, Dec. 2016.

[17] M. Grabisch, J. L. Marichal, R. Mesiar, and E. Pap, *Aggregation Functions.* New York, NY, USA: Cambridge Univ. Press, 2009.

[18] E. Pap, *Null-Additive Set Functions*, *Mathematics and its Applications*, vol. 337. Dordrecht, Netherlands: Kluwer, 1995.

[19] M. Sugeno, "Theory of fuzzy integrals and its applications," Ph.D. dissertation, Tokyo Inst. Technol., Tokyo, Japan, 1974.

[20] Z. Wang and G. J. Klir, *Generalized Measure Theory* (International Series on Systems Science and Engineering), vol. 25. New York, NY, USA: Springer, 2009.

[21] M. Gagolewski and P. Grzegorzewski, "A geometric approach to the construction of scientific impact indices," *Scientometrics*, vol. 81, no. 3, pp. 617–634, 2009.

[22] N. J. van Eck and L. Waltman, "Generalizing the h- and g-indices," *J. Informetrics*, vol. 2, pp. 263–271, 2008.

[23] M. Gagolewski, "A remark on limit properties of generalized h- and g-indices," *J. Informetrics,* vol. 3, pp. 367–368, 2009.