

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joiAn informetric model for the *success-index*

Fiorenzo Franceschini*, Maurizio Galetto, Domenico Maisano, Luca Mastrogiacomo

Politecnico di Torino, DIGEP (Department of Management and Production Engineering), Corso Duca degli Abruzzi 24, 10129 Torino, Italy

ARTICLE INFO

Article history:

Received 13 April 2012

Received in revised form 29 August 2012

Accepted 20 September 2012

Keywords:

Lotka's law

h-Index*success-index*

Information production processes

ABSTRACT

Based on an idea by Kosmulski, Franceschini et al. (2012, *Scientometrics* 92(3), 621–641) propose to classify a publication as “successful” when it receives more citations than a specific comparison term (*CT*). In the intention of the authors *CT* should be a suitable estimate of the number of citations that a publication – in a certain scientific context and period of time – should potentially achieve. According to this definition, the *success-index* is defined as the number of successful papers, among a group of publications examined, such as those associated to a scientist or a journal. In the first part of the paper, the *success-index* is recalled, discussing its properties and limitations. Next, relying on the theory of Information Production Processes (IPPs), an informetric model of the index is formulated, for a better comprehension of the index and its properties. Particular emphasis is given to a theoretical sensitivity analysis of the index.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Kosmulski (2011) recently proposed to classify a paper as “successful” when receiving more citations than those made. Developing this idea, he introduced a new indicator called Number of Successful Papers (*NSP*), which makes it possible to “isolate” a subset of publications, defined as *successful papers*, among a group of publications examined—e.g. those associated to a scientist or a journal. Apart from being relatively simple and intuitive, *NSP* enables cross-field normalization at the level of individual publication. In fact, the number of citations (or references) made is used as an indicator of citation propensity – acting as a comparison/normalization term – for normalizing the citations received by the paper itself. One of the greatest advantages of *NSP* is that it can be applied to groups of heterogeneous papers, in terms of scientific (sub-)field, e.g. the overall scientific output of researchers affiliated to different departments or different research institution, etc.

Considering *NSP* from a broader perspective, it can be seen that—given a generic set of publications—this indicator allows to select an “elite” subset. This selection can also be made by other indicators in the literature: e.g. let us consider the *h*-core approach (Hirsch, 2005), the selection by the π -indicator (Vinkler, 2011), the *characteristic scores and scales* (CSS) method (Glänzel, 2011), the ESI's Highly Cited Papers method (ISI Web of Knowledge, 2012), or many others (Bornmann et al., 2012; Plomp, 1990; Tijssen, Visser, & Van Leeuwen, 2002). We remark that, differently from *NSP*, the aforementioned methods require that the sets of publications examined are preliminarily divided into scientific disciplines. Also, some of these methods rely on superimposed classifications such as ISI subject categories.

Unfortunately, the formulation of *NSP* has two main drawbacks: (1) for a certain publication, the estimation of the propensity to cite of analogous papers within the same scientific field is ineffective in terms of statistical significance, and (2) *NSP* is prone to manipulation since an author, for having the largest possible *NSP*, may be tempted to “economize” on

* Corresponding author. Tel.: +39 011 090 7225; fax: +39 011 090 7299.
E-mail address: fiorenzo.franceschini@polito.it (F. Franceschini).

the list of references (Franceschini et al., 2007). For more information on these issues, we refer the reader to Franceschini, Galetto, Maisano, and Mastrogioacomo (2012).

Franceschini et al. (2012) suggested to modify *NSP* into the *success-index* (hereafter abbreviated as *s-index*) by introducing a comparison term (*CT*) which is an indicator of propensity to cite, determined on the basis of a reference sample of publications. For a generic group of papers, the *s-index* is defined as the number of papers with a number of citations greater than or equal to CT_i , i.e. a generic comparison term associated to the *i*-th publication. We underline that CT_i is an estimate of the number of citations that a publication—in a certain scientific context and period of time—should potentially achieve.

With the aim of formalizing this definition, a score is associated to each of the (*P*) publications of interest:

$$\begin{cases} score_i = 1 \text{ when } c_i \geq CT_i \\ score_i = 0 \text{ when } c_i < CT_i \end{cases} \quad (1)$$

where c_i are the citations received by the *i*-th publication. This achievement determines the condition of success. In this sense, CT_i embodies the concept of *potential citation impact* by Moed (2010b). The *s-index* is given by:

$$s\text{-index} = \sum_{i=1}^P score_i \quad (2)$$

Of course, determining the “(non-)success status” of an individual paper should not be intended as a comprehensive assessment of quality. Nevertheless, this does not mean that counting the number of papers from a set above/below some appropriate citation thresholds could not provide useful information. Also, this is the basic idea of the *highly cited publications* indicator, theorized by Waltman and van Eck (2012).

From the above definition, it is also evident how critical the construction of CT_i is. Three are the major “margins of actions” in this operation:

1. Defining the procedure for selecting the reference sample of publications. Possible approaches are: (i) the selection of papers published by the same journal, (ii) the use of superimposed classifications such as ISI subject categories (Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011; Waltman, Yan, & van Eck, 2011), (iii) or the implementation of “adaptive” techniques in which the sample is determined considering the “neighbourhood” of the publication(s) of interest—typically consisting of the set of publications citing or being cited by them (Jackson & Rogers, 2007; Leydesdorff & Shin, 2011; Moed, 2010b, 2010a)—etc.
2. Deciding whether to consider (i) the distribution of the number of references made or (ii) the distribution of the citations received by the publications of the reference sample.
3. Identifying a suitable (central tendency) indicator for obtaining CT_i from the distribution of interest, e.g. mean, median, harmonic mean, percentiles, etc.

It is worth remarking that these three issues are valid for the construction of a generic field-normalized indicator, not necessarily the *s-index*. We also remark that the first issue is particularly critical and currently much debated among bibliometricians.

There are several possible options for defining CT_i , for instance by:

- the average (or median) number of references made by the articles published in the same journal and year of the publication concerned. The assumption behind this construction is that articles published in the same journal are reasonably similar in terms of citation propensity.
- the average (or median) number of references made/received by a sample of publications representing the “neighbourhood” of the publication concerned. In this case, the assumption is that “neighbouring” articles are reasonably similar in terms of citation propensity.

Retracing the steps taken with other bibliometric indices (in particular the *h-index*), this paper proposes an informetric model for the *s-index* relying on the theory of Information Production Processes (Egghe & Rousseau, 2006).

The model presented is based on Lotka’s law and may be useful for a deeper understanding of the indicator features and potential.

The remaining of the paper is organized into four sections. Section 2 recalls Lotka’s law and other basic notions used for developing the model. Section 3 presents a simplified informetric model of the *s-index*, under the assumption that the publications of interest are all within the same discipline, with the same citation propensity. Particular attention is given to (i) the differences between the *h* and the *s*-indices and (ii) a method for estimating the sensitivity of the *s-index*. In Section 4, the model is refined and extended to a more realistic case of a set of heterogeneous papers in terms of scientific fields. Finally, the conclusions are given, summarizing the original contributions of the paper.

2. Law of Lotka

Information Production Processes (IPPs) are all those general frameworks in which there are sources “producing” items (Egghe, 1990; Egghe & Rousseau, 1990). Authors and their publications, journals and their articles, article and their citations, books and borrowings, websites and hyperlinks are just few examples of IIPs. In general, each IIP can be described by means of a size-frequency (or a rank-frequency) density function $f(j)$ expressing the number of sources “producing” j items (Egghe & Rousseau, 1990; Egghe, 2005a). One of the most popular frequency function is the law of Lotka:

$$f(j) = \frac{K}{j^a} \tag{3}$$

$K > 0, a > 1$ (Lotka, 1926). In IIP theory, the Lotka function expresses the number of sources (articles) with exactly j items (citations). We remark that, according to the definition in Eq. (3), uncited publications are not modelled, i.e. $f(j=0)$ is not defined.

Although many other informetric laws which apparently can guarantee better fits in particular cases have been proposed, the Lotka function (Eq. (3)) is the simplest and probably the most used one. Being able to explain many regularities, it is considered a sort of “axiom” within IIP theory (Egghe, 2009).

From Eq. (3) it can be deduced that K represents the number of sources (articles) with exactly 1 item (citations). In other terms

$$f(1) = \frac{K}{1^a} = K \tag{4}$$

In bibliometrics, according to its parameters, Lotka’s law allows to describe the distribution of citations of different groups of papers.

3. The s-index in Lotkaian systems

This section introduces a simplified modelling of the s -index, assuming to analyze a homogeneous group of papers, i.e. a group of papers belonging to the same discipline of interest. Under this hypothesis the CT is the same for all the papers. This assumption will be relaxed later in Section 4.

According to its definition, in a Lotkaian system the s -index can be seen as

$$s = \sum_{j=CT}^{\infty} f(j) = \sum_{j=CT}^{\infty} \frac{K}{j^a} \tag{5}$$

In the theory of Lotkaian informetrics a common choice is to replace the discrete Lotka function with a continuous variant of Eq. (3), generally denoted with φ . The continuous Lotka function φ is given by a power function of a continuous variable $j \geq 1$:

$$\varphi(j) = \frac{k}{j^\alpha} \tag{6}$$

where $k > 0, \alpha > 1$. The use of the continuous instead of the discrete function comes to the fact that it is easier to evaluate derivatives and integrals than discrete differences or sums (Egghe, 2005b). While recognizing the benefits of this approximation, Egghe (2005b) warns about its implications.

Conforming to this convention, the s -index can thus be expressed in an integral form as

$$s = \int_{CT}^{\infty} \varphi(j) dj = \int_{CT}^{\infty} \frac{k}{j^\alpha} dj \tag{7}$$

Eq. (7) can be easily solved as

$$s = \int_{CT}^{\infty} \frac{k}{j^\alpha} dj = \frac{k}{1-\alpha} j^{(1-\alpha)} \Big|_{CT}^{\infty} = \frac{k}{\alpha-1} CT^{(1-\alpha)} \tag{8}$$

Even if k is very similar to K (see Eq. (4)), compared to this latter it loses any “physical” meaning. In order to include in the formula parameters and variables with a practical significance, it can be useful to operate a change of variable. Let P be the total number of productive sources (articles), i.e. the number of sources (articles) with at least one item (citation). Then P can be evaluated as

$$P = \int_1^{\infty} \varphi(j) dj = \int_1^{\infty} \frac{k}{j^\alpha} dj = \frac{k}{\alpha-1} \tag{9}$$

Eq. (8) can be combined with Eq. (9) providing that

$$s = \frac{P(\alpha-1)}{\alpha-1} CT^{(1-\alpha)} = \frac{P}{CT^{(\alpha-1)}} \tag{10}$$

As might be expected, this equation shows how the s -index is directly proportional to P , i.e. the number of productive sources. In contrast, it can be noticed that when CT grows, the s -index obviously tends to decrease ($\alpha > 1$).

As an alternative, the model of Eq. (10) can be revised considering the total number of items (citations) instead of the number (P) of productive sources (articles). The sources with exactly j items are $f(j)$, thus the total number (C) of items (citations) can be obtained as the summation from one to infinity of the product $j \cdot f(j)$. If $\alpha > 2$, in the continuous case C can be estimated as (Egghe, 2005b, 2007):

$$C = \int_1^{\infty} j\varphi(j) \, dj = \int_1^{\infty} \frac{k}{j^{\alpha-1}} \, dj = \frac{k}{\alpha-2} \tag{11}$$

Eq. (11) can be combined with Eq. (9), modifying Eq. (10) into

$$s = \frac{C(\alpha-2)}{\alpha-1} CT^{(1-\alpha)} = \frac{C(\alpha-2)}{(\alpha-1)CT^{\alpha-1}} \tag{12}$$

It is worth remarking that Eq. (12) is valid if and only if $\alpha > 2$. However, Egghe relaxes this constraint proposing a more general relation between C and P . For further details refer to Theorem II.3.1 in Egghe (2007).

Another possibility to relate C and P is to use Heaps' law or Herdan's law (Ye, 2011). However, even if the latter approach brought fair results, it has been proved to be valid only for large values of sources, i.e. when $P \gg 0$ (Egghe, 2007).

3.1. The relation between the h -index and the s -index

One of the most debated and widely used bibliometric indexes is the h -index (Hirsch, 2005; Egghe et al., 2006; Franceschini & Maisano, 2010a, 2010b). It is defined as "the number h such that, for a general group of papers, h papers received at least h citations while the other papers received no more than h citations" (Hirsch, 2005). Egghe and Rousseau (2006) extended the definition of the index to the more general framework of IPPs, providing a model of the index based on the Lotka's law:

$$h = P^{1/\alpha} \tag{13}$$

This model allows to easily find the relation between h and s . Combining Eqs. (10) and (13), the s -index can be stated as:

$$s = \frac{h^\alpha}{CT^{\alpha-1}} \tag{14}$$

This model is not surprising: s increases with the h -index through an exponential relation:

$$ds = \frac{\alpha h^{\alpha-1}}{CT^{\alpha-1}} dh \tag{15}$$

Under the assumption of continuity of the two indexes, Eq. (15) says that the substitution rate between the two indexes is not constant over the h domain, since it depends on the h -value.

Even if the modelling relies on the assumption of continuity, here follows a discrete example aimed at further clarifying the meaning of Eq. (15). Consider the case in which $CT=20$ and $\alpha=2$. If $h=15$, Eq. (15) says that a unitary increase in the h -index corresponds to a 1.5 increase in the s -index. On the other hand, when $h=10$, a unitary increase in the h -index results into an equal increase of the s -index.

This can be also deduced by Corollary C presented by Egghe et al. (2006), where the h -index is related to the number of citations (C) as:

$$h = \left(\frac{\alpha-2}{\alpha-1} C \right)^{1/\alpha} \tag{16}$$

While the s -index linearly depends on the number of citations (see Eq. (12)), h is proportional to its α -th root (Eq. (16)). Considering CT and α as constant and differentiating Eqs. (12) and (16) one obtains:

$$\begin{aligned} dC &= CT^{\alpha-1} \frac{\alpha-1}{\alpha-2} ds = b \cdot ds \\ dC &= \alpha h^{\alpha-1} \frac{\alpha-1}{\alpha-2} dh = a \cdot h^{\alpha-1} \cdot dh \end{aligned} \tag{17}$$

with $a = \alpha(\alpha-1)/(\alpha-2)$ and $b = CT^{\alpha-1}(\alpha-1)/(\alpha-2)$. Eq. (17) says that the higher h , the greater is the effort (in terms of citations) to increase it, whereas this does not hold for the s -index. Fig. 1 qualitatively synthesizes this concept, in the discrete domain.

3.2. Sensitivity analysis of the model

This section proposes a sensitivity analysis of the model. A question arising from the definition of the s -index is: "What are the effects of variations in CT ?"

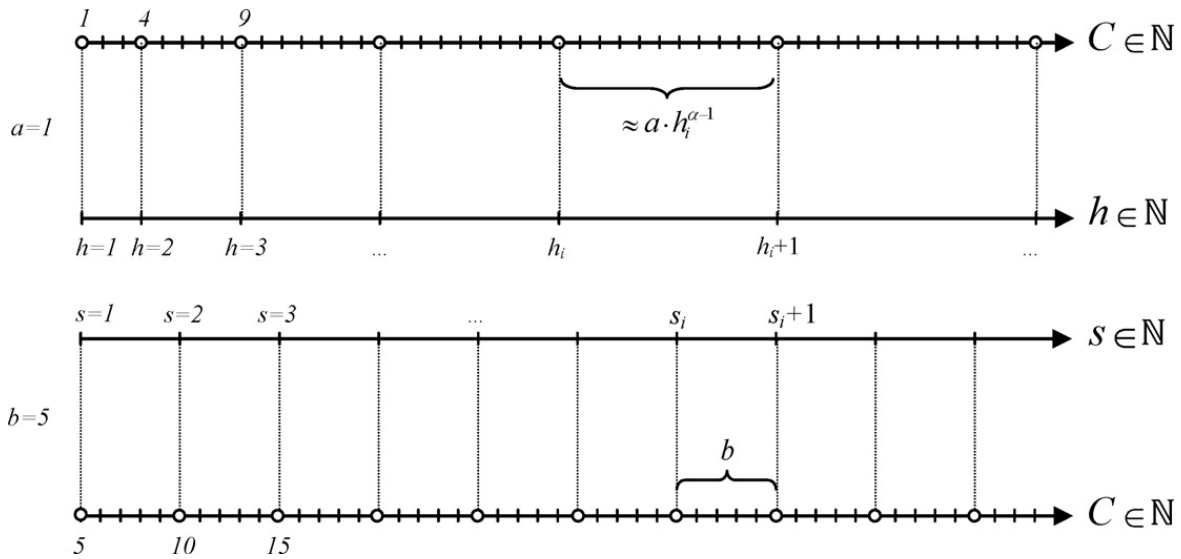


Fig. 1. h and s values represented on the C axis. The figure presents an example with $a=1$ and $b=5$. While the s -index linearly varies with the effort (in terms of C), the higher the h -index the stronger the effort required to increase it. In other terms whereas the distance (in terms of citations) between two consecutive s -classes is constant, the same does not hold for h .

In fact, the estimation of CT has been shown to be the most delicate aspect in the definition of the s -index. Treating CT as a random variable with a certain standard deviation, it is emblematic to see how the uncertainty related to its estimation propagates on the s -index.

The analysis herein presented relies on a well-known statistical method used in many field of engineering: the so-called *delta method* (also referred as law of propagation of uncertainty or error-transmission formula) (Casella & Berger, 2001; JCGM100:2008, 2008; Montgomery, 2009). Through a linearization of the problem, this approach allows the calculation of the expected value and variance of a random variable, which is a nonlinear function of other independent random variables.

In the case of analysis, the s -index can be seen as a non-linear function of three random independent variables: CT , α and P . The model presented in Eq. (10) is continuous, differentiable and relatively smooth with respect to the independent variables, therefore the equation can be expanded into a first order Taylor series about the expected values of three variables, i.e. $(\mu_{CT}, \mu_\alpha, \mu_P)$ as:

$$s = s(CT, \alpha, P) = s(\mu_{CT}, \mu_\alpha, \mu_P) + \frac{\partial s}{\partial CT}(CT - \mu_{CT}) + \frac{\partial s}{\partial \alpha}(\alpha - \mu_\alpha) + \frac{\partial s}{\partial P}(P - \mu_P) + R \quad (18)$$

where R represents the higher order terms.

Neglecting the terms of higher order, it is possible to apply the expected value and variance operators to Eq. (18) to obtain:

$$\mu_s \approx \frac{\mu_P}{\mu_{CT}^{\mu_\alpha - 1}} \quad (19)$$

and

$$\sigma_s^2 \approx \left(\frac{\mu_P(1 - \mu_\alpha)}{\mu_{CT}^{\mu_\alpha}} \right)^2 \sigma_{CT}^2 + \left(\frac{-\mu_P \ln(\mu_{CT})}{\mu_{CT}^{\mu_\alpha - 1}} \right)^2 \sigma_\alpha^2 + (\mu_{CT}^{1 - \mu_\alpha})^2 \sigma_P^2 \quad (20)$$

being σ_{CT} , σ_α , σ_P the standard deviations related to CT , α and P . It can be noticed that, in Eq. (19), the random variables of Eq. (10) are replaced by their expected values.

Focusing on the uncertainty contributions due to CT , and assuming an average value for α around 2, Eq. (20) can be simplified into:

$$\sigma_s^2 \approx \left(\frac{\mu_P}{\mu_{CT}^2} \right)^2 \sigma_{CT}^2 \quad (21)$$

The variance of s depends on that of CT through a propagation coefficient $(\mu_P/\mu_{CT}^2)^2$. The analysis of this coefficient suggests some interesting considerations:

- the higher the expected value of CT , the lower the uncertainty of s -index;
- conversely, the greater the number of productive papers (P), the greater the uncertainty of the s -index.

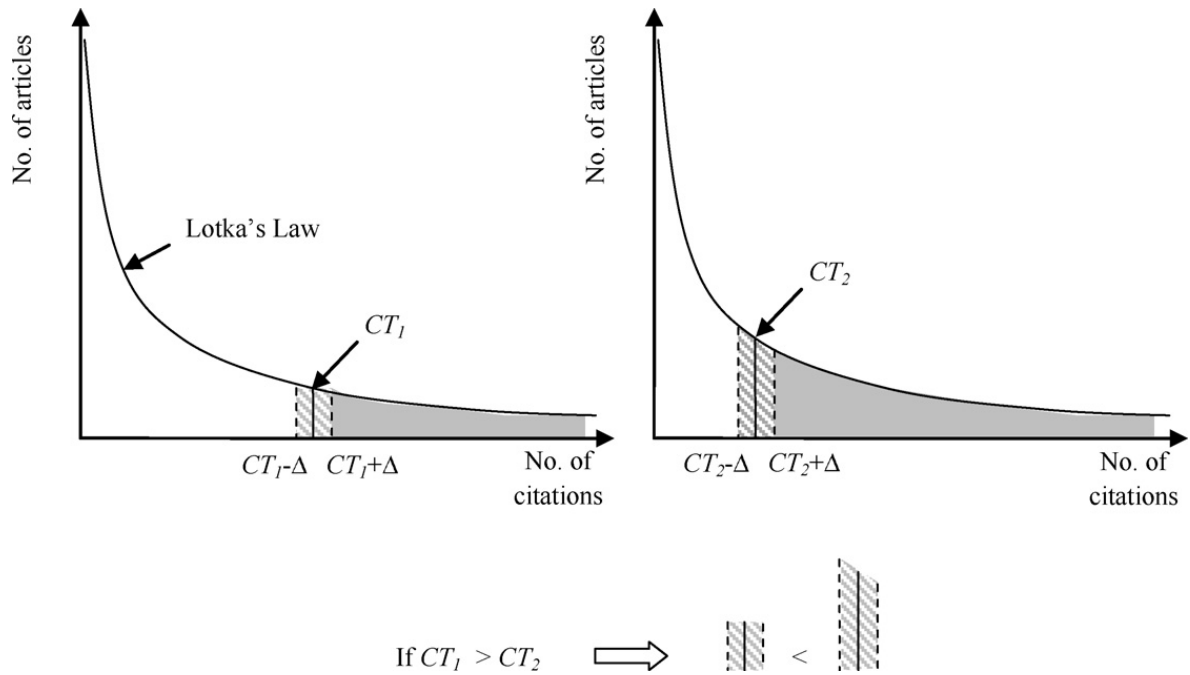


Fig. 2. Qualitative reading of the *s*-index: the *s*-index can be seen as the area subtended by the Lotka's Law curve from *CT* to infinity. Given a fixed variability range of *CT*, the higher is *CT* the lower is its propagation.

Table 1

Parameters used for simulating a Lotkaian distribution of publications/citations. These are the parameters representing author's institution. The values are obtained from Scopus limiting the results to the articles published in 2008.

Parameter	Value
<i>P</i>	718
<i>C</i>	6641
α	2.12

The first consideration is probably the most interesting since it establishes a relationship between the variability of the *s*-index and that related to *CT*. This consideration may also be qualitatively explained by observing Fig. 2: the *s*-index can be interpreted as the area of the “right tail” of the Lotka's law curve. Given a certain variability range ($\pm\Delta$) in the definition of *CT*, the corresponding range in the definition of the *s*-index may be read as the area subtended by the Lotka's law curve from $CT - \Delta$ to $CT + \Delta$. From Fig. 2 it is evident that the higher the value of *CT*, the lower this area.

To confirm the results obtained, a simple Monte Carlo simulation is proposed. A Lotkaian distribution of publications/citations is numerically generated, according to the parameters reported in Table 1.

Next, referring to this distribution, the *s*-index is calculated, varying the *CT*. Since *CT* is generally defined as the average number of citations received or made by a reference sample of papers, we assume that it is normally distributed (cf. central limit theorem), with mean equal to $\mu_{CT} \in \{10, 11, \dots, 30\}$ and standard deviation equal to one. For each value of μ_{CT} , five thousand *CT* values are simulated and, for each one of them, the *s*-index is calculated. This way, it is possible to estimate the variance of *s*. The comparison between the variance obtained through the Monte Carlo simulation and that obtained by Eq. (20) is shown in Fig. 3.

Fig. 3 shows a relatively good agreement between the σ_s^2 values obtained by the numerical simulation and those obtained by Eq. (20).

Similar results can be obtained by numerically generating a Paretian (instead of Lotkaian) distribution of publications/citations.

It has to be remarked that these results are obtained keeping constant all the parameters of Lotka's Law (i.e. *P*, *C* and α), except *CT*.

Both the Monte Carlo simulation and the analytical approach give evidence that the higher *CT*, the lower the variance of the *s*-index. It is worth remarking that a too large *CT* value would tend to reduce the *s*-index's discrimination power. For example, in a paradoxical scenario in which *CT* is infinitive, the *s*-index will virtually be zero with no uncertainty (see Eqs. (10) and (21)). This aspect is also clear observing the partial derivative of *s* with respect to *CT*, obtained from Eq. (10):

$$\frac{\partial s}{\partial CT} = \frac{(1 - \alpha)P}{CT^\alpha} \tag{22}$$

Eq. (22) suggests that when increasing *CT*, *s* tends to asymptotically stabilize around the value of zero.

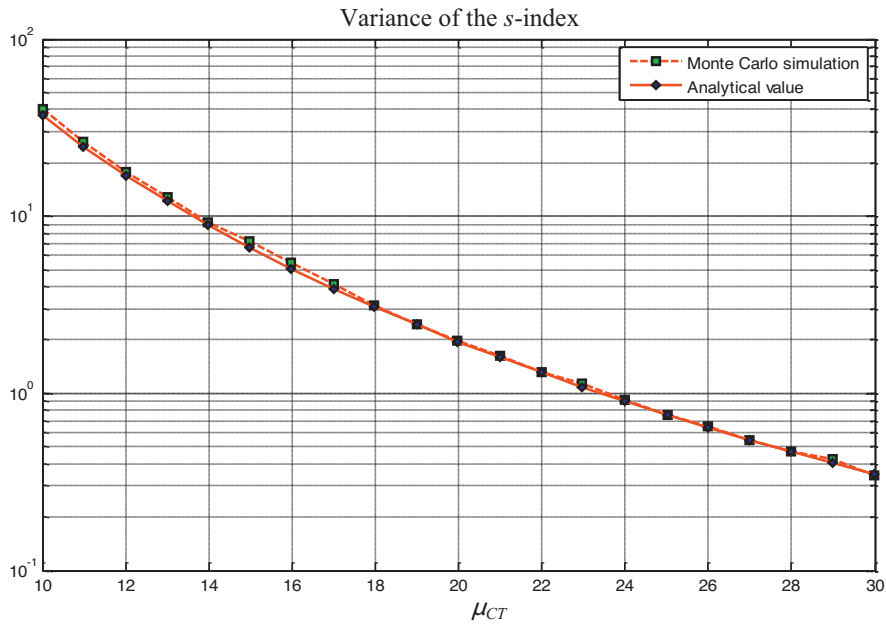


Fig. 3. Comparison between σ_s values obtained by (1) a numerical Monte Carlo simulation and (2) the analytical model in Eq. (20).

These remarks highlight how delicate is the choice of a proper CT when constructing the s -index: CT has to be sufficiently large in order to reduce the uncertainty of s , but not too much, otherwise this would make the s -index useless.

4. Multi-field model

Unlike the h -index, the s -index can be applied to multidisciplinary groups of papers, not necessarily requiring that they are preliminarily divided into scientific fields (Franceschini et al., 2012). There is a large evidence, in the literature, of the fact that different research disciplines show different propensity to the practice of citation (Garfield, 1979). To this purpose, the s -index adapts the concept of successful article to the definition of a specific comparison term CT . Articles from different fields are therefore likely to have different CT s, which – by definition – are intended to be “estimates of the citations that a new publication would be likely to receive in the field of interest” (Franceschini et al., 2012).

How can we compare scientists belonging to different scientific fields? The s -index allows this comparison being CT variable from field to field. Even a single researcher, in his professional life, can do research in different fields. In this case, the s -index can be seen as the sum of the successful papers belonging to each field of interest. Thus, the model proposed in Eq. (10), can be extended as:

$$s = \sum_{i=1}^N s_i = \sum_{i=1}^N \frac{P_i}{CT_i^{\alpha_i - 1}} \tag{23}$$

where s_i is the number of successful paper in the i -th field of interest, P_i is the total number of productive sources (i.e. articles with at least one citation) in the i -th field of interest, CT_i is the comparison term related to the i -th field of interest, α_i is the constant of Lotka’s law applied to each field of interest and N is the number of different fields of interest.

In a more realistic case in which scientific fields are not necessarily delineated and each of the (P) articles of interest has its specific CT_i , Eq. (23) turns out to be:

$$s = \sum_{i=1}^P \frac{1}{CT_i^{\alpha_i - 1}} \tag{24}$$

Similar considerations to those expressed in Section 3 hold for Eqs. (23) and (24).

Being the “multidisciplinary” s -index a summation of P different terms, the variance operator can be applied to Eq. (24), resulting in:

$$\sigma_s^2 \approx \sum_{i=1}^P \left(\frac{1}{\mu_{CT}^2} \right)^2 \sigma_{CT_i}^2 \tag{25}$$

In other terms, the assessment of the s -index is naturally affected by the combined uncertainty on the estimate of all CT_i s.

5. Conclusions

This paper considers the s -index in the framework of Information Production Processes (IIPs) with particular reference to the context of articles (sources)-citations (items). In Lotkaian systems the index can be modelled as $s = \sum_{i=1}^N P_i / CT_i^{\alpha_i - 1}$, where P_i is the total number of productive sources (articles), CT_i is the comparison term related to the i -th field of interest, α_i is the constant of Lotka's law applied to each field of interest and N is the number of different fields of interest.

The proposed model allows an operational comparison between the s -index and the h -index highlighting a particularly interesting aspect: while the h -index tends to vary exponentially with the number of citations (C) and publications (P), the s -index depends linearly on the two variables.

The sensitivity analysis of the s -index is one of the interesting aspects of the research herein presented. It shows that the uncertainty of CT can affect the value of s -index significantly. In particular, the lower the value of CT , the more serious may be the effects on the evaluation of s .

Future developments go towards constructing an hybrid model of the indicator, relying on other laws such as Herdan/Heaps' or Zipf's law. At the price of complicating the analysis, such approaches have proven to be effective for modelling other bibliometric indicators (Schubert & Glänzel, 2007; Ye, 2011).

References

- Bornmann, L., de Moya-Anegón, F., & Leydesdorff, L. (2012). The new excellence indicator in the world report of the SCImago Institutions Rankings 2011. *Journal of Informetrics*, 6(2).
- Casella, G., & Berger, R. L. (2001). *Statistical inference* (2nd edition). Duxbury Press., pp. 240–245, June 18, 2001.
- Egghe, L. (1990). The duality of informetric systems with applications to the empirical laws. *Journal of Information Science*, 16(1), 17–27.
- Egghe, L. (2005a). *Power laws in the information production process: Lotkaian informetrics*. Academic Press.
- Egghe, L. (2005b). Relations between the continuous and the discrete Lotka power function. *Journal of the American Society for Information Science and Technology*, 56(7), 664–668.
- Egghe, L. (2007). Untangling Herdan's law and Heaps' law: Mathematical and informetric arguments. *Journal of the American Society for Information Science and Technology*, 58(5), 702–709.
- Egghe, L. (2009). Lotkaian informetrics and applications to social networks. *Bulletin of the Belgian Mathematical Society-Simon Stevin*, 16(4), 689–703.
- Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics: Quantitative methods in library, documentation and information science*. Elsevier Science Publishers.
- Egghe, L., & Rousseau, R. (2006). An informetric model for the Hirsch-index. *Scientometrics*, 69(1), 121–129.
- Franceschini, F., Galetto, M., & Maisano, D. (2007). *Management by measurement: Designing key indicators and performance measurement systems*. Berlin: Springer.
- Franceschini, F., Galetto, M., Maisano, D., & Mastrogiacomo, L. (2012). The success-index: an alternative approach to the h -index for evaluating an individual's research output. *Scientometrics*, 92(3), 621–641.
- Franceschini, F., & Maisano, D. (2010a). Analysis of the Hirsch index's operational properties. *European Journal of Operational Research*, 203(2), 494–504.
- Franceschini, F., & Maisano, D. (2010b). The citation triad: An overview of a scientist's publication output based on Ferrers diagrams. *Journal of Informetrics*, 4(4), 503–511.
- Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics*, 1(4), 359–375.
- Glänzel, W. (2011). The application of characteristic scores and scales to the evaluation and ranking of scientific journals. *Journal of Information Science*, 37(1), 40–48.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569–16572.
- Jackson, M. O., & Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *The American Economic Review*, 97(3), 890–915.
- JCGM100:2008. (2008). *Evaluation of measurement data—Guide to the expression of uncertainty in measurement*.
- Kosmulski, M. (2011). Successful papers: A new idea in evaluation of scientific output. *Journal of Informetrics*, 5, 481–485.
- ISI Web of Knowledge. (2012). *Essential science indicators*. <<http://thomsonreuters.com>>.
- Leydesdorff, L., & Shin, J. C. (2011). How to evaluate universities in terms of their relative citation impacts: Fractional counting of citations and the normalization of differences among disciplines. *Journal of the American Society for Information Science and Technology*, 62(6), 1146–1155.
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of Washington Academy Sciences*, 16, 317–323.
- Moed, H. F. (2010a). CWTS crown indicator measures citation impact of a research group's publication oeuvre. *Journal of Informetrics*, 3(3), 436–438.
- Moed, H. F. (2010b). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265–277.
- Montgomery, D. C. (2009). *Statistical quality control: A modern introduction*. Wiley&Sons.
- Plomp, R. (1990). The significance of the number of highly cited papers as an indicator of scientific prolificacy. *Scientometrics*, 19(3), 185–197.
- Schubert, A., & Glänzel, W. (2007). A systematic analysis of Hirsch-type indices for journals. *Journal of Informetrics*, 1(3), 179–184.
- Tijssen, R. J. W., Visser, M. S., & Van Leeuwen, T. N. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics*, 54(3), 381–397.
- Vinkler, P. (2011). Application of the distribution of citations among publications in scientometric evaluations. *Journal of the American Society for Information Science and Technology*, 62(10), 1963–1978.
- Waltman, L., & van Eck, N. J. (2012). The inconsistency of the h -index. *Journal of the American Society for Information Science and Technology*, 63(2), 406–415.
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37–47.
- Waltman, L., Yan, E., & van Eck, N. J. (2011). A recursive field-normalized bibliometric performance indicator: An application to the field of library and information science. *Scientometrics*, 89(1), 301–314.
- Ye, F. Y. (2011). A unification of three models for the h -index. *Journal of the American Society for Information Science and Technology*, 62(1), 205–207.