Scientometrics (2014) 101:963–983
DOI 10.1007/s11192-014-1250-6

The *citer-success*-index: a citer-based indicator to select a subset of elite papers

Fiorenzo Franceschini · Domenico Maisano · Luca Mastrogiacomo

Received: 18 October 2013 / Published online: 18 February 2014

Abstract The goal of this paper is introducing the *citer-success*-index (*cs*-index), i.e. an indicator that uses the number of different citers as a proxy for the impact of a generic set of papers. For each of the articles of interest, it is defined a comparison term—which represents the number of citers that, on average, an article published in a certain period and scientific field is expected to "infect"—to be compared with the actual number of citers of the article. Similarly to the recently proposed *success*-index (Franceschini et al. Scientometrics 92(3):621–6415, 2011), the *cs*-index allows to select a subset of "elite" papers. The *cs*-index is analyzed from a conceptual and empirical perspective. Special attention is devoted to the study of the link between the number of citers and cited authors relating to articles from different fields, and the possible correlation between the *cs*- and the *success*-index. Some advantages of the *cs*-index are that (i) it can be applied to multidisciplinary groups of papers, thanks to the field-normalization that it achieves at the level of individual paper and (ii) it is not significantly affected by self citers and recurrent citers. The main drawback is its computational complexity.

Keywords Impact · Citations · References · Citation propensity · Success-index · Citers · Cited authors

F. Franceschini · D. Maisano (✉) · L. Mastrogiacomo
Department of Management and Production Engineering, Politecnico di Torino, DIGEP,
Corso Duca degli Abruzzi 24, 10129 Turin, Italy
e-mail: domenico.maisano@polito.it

F. Franceschini
e-mail: fiorenzo.franceschini@polito.it

L. Mastrogiacomo
e-mail: luca.mastrogiacomo@polito.it

Introduction and literature review

In bibliometrics, one of the main analysis dimensions is the impact of scientific publications, which is commonly estimated by counting the number of citations that they accumulate over time (Egghe and Rousseau 1990). As an alternative to citations, Dieks and Chang (1976) and Braun et al. (1985) suggested to use the total number of different citers (or citing authors), i.e. the members of the scientific community who are “infected” by a certain paper. The number of different citers is a proxy which is harder to compute, but more elegant, as only marginally affected by citations from self citers and recurrent citers.

More than 10 years ago, White (2001) carried out an investigation at the level of citers, in which the habit of citing other authors is seen as a characteristic of the writing style of scientists.

The idea of citers was recently dug up by Ajiferuke and Wolfram (2010), who proposed and implemented an indicator based on citers, without encountering the computational obstacles of the past, thanks to the current evolution of databases and information management tools. The indicator is the *ch*-index, which is a variant of the very well-known *h*-index (Hirsch, 2005). The *ch*-index was empirically analyzed by Franceschini et al. (2010), showing (i) the general correlation between *ch* and *h*, and (ii) the potential of *ch* in complementing the information given by *h*. A theoretical interpretation of the correlation between *ch* and *h* was recently provided by Egghe (2012).

In this article we focus the attention on the *success*-index (*s*-index), i.e. a recent indicator that, for a generic set of articles, allows to select an “elite” subset, according to a logic different from that of *h* (Franceschini et al. 2012). 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 with the i-th publication.* CT_i is a conventional proxy for the number of citations that articles of the same scientific area and period of time of the article of interest (i.e. the *i*-th publication) are likely to obtain.

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

$$\begin{cases} \text{score}_i = 1 & \text{when } c_i \geq CT_i \\ \text{score}_i = 0 & \text{when } c_i < CT_i \end{cases}, \tag{1}$$

where c_i are the citations obtained by the *i*-th publication. The *s*-index is therefore given by:

$$s\text{-index} = \sum_{i=1}^{P} \text{score}_i. \tag{2}$$

Apart from *s*, there are other indicators in the literature that allow to select an elite subset (Vinkler 2010), based on the comparison between the number of citations accumulated by each paper and a threshold. E.g. let us consider the selection by $P_{\text{top 10 \%}}$-indicator (Bornmann 2013), that by π-indicator (Vinkler 2009, 2011), the characteristic scores and scales method (Glänzel 2011) or the ESI’s Highly Cited Papers method (ISI Web of Knowledge 2012). We remark that, differently from *s*, the aforementioned methods require that the set of publications examined are preliminarily categorized into scientific (sub-)disciplines.

As regards the *s*-index, there are several options for constructing the CT_i related to an *i*-th paper of interest. The more accurate methods are also the more complex and computationally burdensome. Therefore, the conventional choice of the option to be adopted depends on the needs of the specific case. In general, three issues are crucial (Franceschini et al. 2013a, b):

1. Defining the procedure for selecting a reference sample of homologous publications. Possible approaches are: (i) the selection of papers of same age, type (e.g. research article, review, letter, etc.) and published by the same journal of the i-th paper of interest, (ii) the use of superimposed classifications such as ISI subject categories, (iii) the implementation of "adaptive" techniques in which the sample is determined considering the "neighbourhood" of the paper of interest—typically consisting of the set of papers citing or being cited by it.
2. Deciding whether to consider (i) the distribution of the number of references given or (ii) the citations obtained by the publications of the 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.

For the purpose of example, a possible option for constructing CT_i is using the mean value of the distribution of the number of references given by the articles that cite a sample of articles, in the same ISI subject category of the article of interest. For more information on the strategies for constructing CT_i, we refer the reader to (Franceschini et al. 2013a, b).

Regarding point (2), Franceschini et al. (2012, 2013a) state that indicators based on the distribution of references given—rather than citations obtained—have several advantages:

- The number of references is fixed over time, while the number of citations obtained tends to increase and requires a certain accumulation period to stabilize.
- This stability is also derived by the fact that the number of references is likely to be less variable than the number of citations obtained.
- Bibliographic references are less influenced by journal particularities, such as the average citation impact of articles.

Conceptually, the link between references given (by the papers of the reference sample) and citations obtained (by the papers of interest) originates from a simple consideration: focussing on the totality of the scientific literature in a certain field and according to a simplified model configuration of *isolated* fields—i.e. excluding transfers of citations between different disciplines (see Fig. 1)—the following relationship applies:

$$\sum_{i=1}^{P} c_i = \sum_{i=1}^{P} r_i, \tag{3}$$

where P is the total number of articles (that can cite each other) in the isolated field; c_i is the number of citations obtained by the i-th paper; r_i is the number of citations given by the i-th paper.

The equality of Eq. 3 can also be expressed in terms of average values:

$$\frac{1}{P}\sum_{i=1}^{P} c_i = \frac{1}{P}\sum_{i=1}^{P} r_i \quad \Rightarrow \quad \bar{c} = \bar{r}. \tag{4}$$

For more detailed and rigorous information on the relation between the $\bar{c}$ and $\bar{r}$ values concerning a set of documents, we refer the reader to (Egghe and Rousseau 1990).

Returning to the s-index, apart from the simplicity of meaning, a great advantage is that it implements a field-normalization at the level of single paper and can therefore be applied

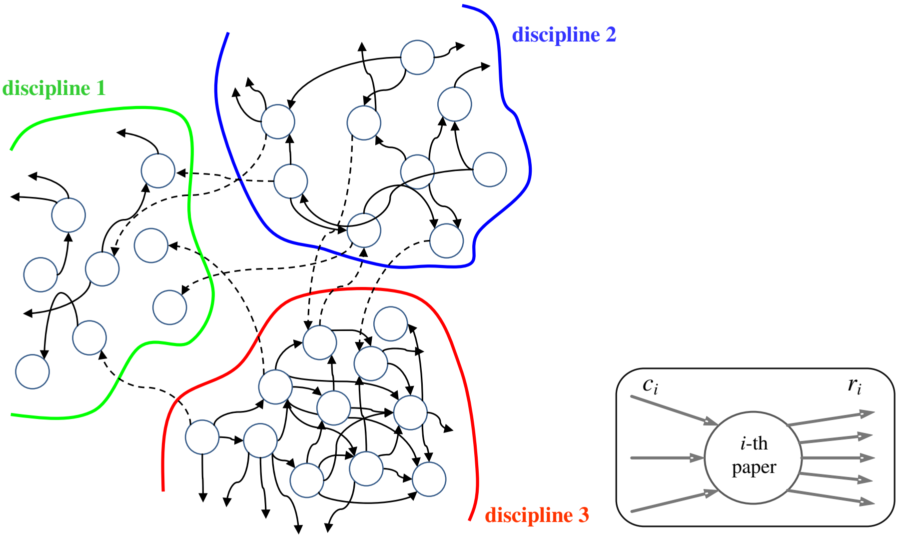

Fig. 1 Schematic representation of scientific disciplines (delimited by *solid lines*) associated with the papers in the scientific literature (represented by *circles*). *Dotted arrows* represent the citations exchanged between papers of different fields. Regarding a generic i-th paper (in the lower-right inset), c_i denotes the total citations obtained (*incoming arrows*), while r_i denotes the total citations given (*outgoing arrows*)

to multidisciplinary groups of articles, for instance the whole production output of a research institution.

Another important quality of the s-index is that it is defined on a *ratio* scale. This feature has several practical implications that make this indicator more versatile than others—such as the h-index, which is defined on an *ordinal* scale (Franceschini et al. 2012):

- The s-index reflects compositions of the input publication sets (with the corresponding citations). In other terms, the union of two groups of publications with s-index of 2 and 5 (with no common publications) will always originate a third group of publications with s-index of $2 + 5 = 7$. This simple property is very useful for extending the use of the s-index to multi-disciplinary institutions, e.g. joining groups of publications from different scientific fields.
- The s-index eases normalizations aimed at obtaining the so-called size-independency (Franceschini et al. 2013a). Given a general group of papers and the same capacity of producing successful papers, it is reasonable to assume that thr s-index should increase proportionally with the different types of "resources" deployed. In fact, several normalized indicators can be obtained dividing the s-index by the resource unit of interest; e.g. the staff number of a research institution, the age of a researcher, the number of articles of a journal, the amount of funding received in a certain period, etc.

The purpose of the paper is introducing the *citer-success*-index (or *cs*-index), i.e. a variant of the s-index, which is based on citers instead of citations, according to a logic similar to that of ch. Before getting into the problem, Fig. 2 introduces the reader to the indicators and notation that will be used in the remaining of the paper.

Given a set of articles, the cs-index identifies a subset for which the number of different citers of an i-th article exceeds a specified comparison term cCT_i. Formalizing, a score is associated with each i-th of the (P) publications of interest:

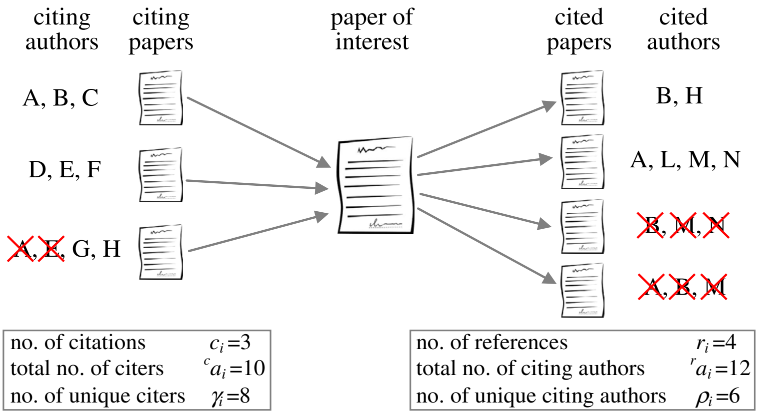

Fig. 2 Introduction of some indicators concerning the authors (represented by letters, e.g. A, B, C, etc.) of papers citing/cited by a fictitious paper of interest. Repeated authors, i.e. those authoring more than one of the citing/cited papers, are marked by *crosses*

$$\begin{cases} \text{score}_i = 1 & \text{when } \gamma_i \geq cCT_i \\ \text{score}_i = 0 & \text{when } \gamma_i < cCT_i \end{cases}, \tag{5}$$

where γ_i are the unique citers related to the *i*-th publication. The word "unique" means that repeated citers are counted only once. The *cs*-index is therefore given by:

$$cs\text{-index} = \sum_{i=1}^{P} \text{score}_i \tag{6}$$

Table 1 exemplifies the calculation of the *s*- and *cs*-index for a fictitious set of papers.

In analogy with CT_i, cCT_i is an estimate of the number of unique citers that articles homologous to that of interest are likely to "infect".

Similarly to CT_i, there are three basic steps when constructing the cCT_i relating to an *i*-th article of interest:

1. Selecting a sample of articles homologous to that interest.
2. Deciding whether to consider the distribution of (i) unique citers or (ii) unique cited authors, relating to the papers of the sample.
3. Defining cCT_i by an indicator of central tendency, applied to the distribution chosen at point (2).

For the purpose of example, a possible option for constructing cCT_i is using the mean value of the distribution of the number of unique (citing) authors of articles that cite a sample of articles, in the same ISI subject category of the article of interest.

The choice at point (2) is more delicate than in the case of the *s*-index. Intuitively, it may appear convenient to use the distribution of unique cited authors for the same reasons for which, in the case of the *s*-index, it was convenient to use the distribution of references. However, the link between unique citers and unique cited authors is not necessarily similar to that between r_i and c_i values; even in a model configuration of isolated fields:

Table 1 Calculation of the *s*- and *cs*-index for a fictitious set of papers

Paper no.	c_i	CT_i	γ_i	cCT_i	*s*-Elite	*cs*-Elite
1	115	20.3	297	60.1	✔	✔
2	86	21.2	187	71.0	✔	✔
3	17	14.5	31	44.8	✔	✗
4	15	20.4	68	72.4	✗	✗
5	12	11.8	30	29.2	✔	✔
6	9	15.7	12	61.9	✗	✗
					s-index = 4	*cs*-index = 3

$$\sum_{i=1}^{P} \gamma_i \text{ is not necessarily} = \sum_{i=1}^{P} \rho_i, \tag{7}$$

Being P the total number of papers in the isolated field; γ_i the number of unique citers of the i-th paper; ρ_i the number of unique authors cited by the i-th paper.

The reason for this lack of parallelism is twofold and will be examined later in the manuscript.

The rest of the paper is structured in three sections. The "General link between citers and cited authors" section investigates whether it is appropriate to construct the cCT_i by using the distribution of the number of unique authors cited by a sample of papers. The "Preliminary empirical analysis of the *cs*-index" section delves into the issue raised in the previous section, examining a large number of papers from different fields. After defining the cCT_i properly, it is studied the correlation between the *s*- and the *cs*-index. Finally, the "Further remarks" section summarizes the original contributions of the paper and the main advantages and disadvantages of the *cs*-index.

This paper is the extended version of the paper (Franceschini et al. 2013a, b), presented at ISSI'13 (14th International Society of Scientometrics and Informetrics Conference) in Vienna, Austria, July 2013.

General link between citers and cited authors

Even modelling a scientific field as isolated and considering the totality of the scientific production in it, there are two possible elements of diversity among citing and cited papers: (i) different average number of authors per paper, and (ii) different percentage of unique authors. Let us clarify this point with simple mathematical considerations. The quantity $\sum_{i=1}^{P} \gamma_i$ can be expressed as:

$$\sum_{i=1}^{P} \gamma_i = \left(\sum_{i=1}^{P} \gamma_i \middle/ \sum_{i=1}^{P} {}^c a_i \right) \cdot \left(\sum_{i=1}^{P} {}^c a_i \middle/ \sum_{i=1}^{P} c_i \right) \cdot \sum_{i=1}^{P} c_i = {}^c p \cdot {}^c app \cdot \sum_{i=1}^{P} c_i, \tag{8}$$

in which γ_i is the number of unique citers of the i-th paper in the isolated field; ${}^c a_i$ ($\geq \gamma_i$) is the total number of citers (even repeated, in the case that some citing papers are (co-)authored by the same individuals) related to the i-th paper; c_i is the number of citing papers (or the number of citations obtained) relating to the i-th paper; P is the total number of articles in the isolated field.

As shown in Eq. 8, the quantity $\sum_{i=1}^{P} \gamma_i$ can also be seen as the product of three terms: ${}^{c}p = \sum \gamma_i / \sum {}^{c}a_i$ (≤ 1) i.e. the percentage of unique citers; ${}^{c}app = \sum {}^{c}a_i / \sum c_i$ (≥ 1) i.e. the average number of authors per citing paper; $\sum_{i=1}^{P} c_i$ the total number of citations obtained.

A "decomposition" similar to that of Eq. 8 may apply to the quantity $\sum_{i=1}^{P} \rho_i$:

$$\sum_{i=1}^{P} \rho_i = \left(\sum_{i=1}^{P} \rho_i \Big/ \sum_{i=1}^{P} {}^{r}a_i \right) \cdot \left(\sum_{i=1}^{P} {}^{r}a_i \Big/ \sum_{i=1}^{P} r_i \right) \cdot \sum_{i=1}^{P} r_i = {}^{r}p \cdot {}^{r}app \cdot \sum_{i=1}^{P} r_i, \quad (9)$$

in which ρ_i is the number of unique authors cited by the i-th paper in the isolated field; ${}^{r}a_i$ ($\geq \rho_i$) is the total number of cited authors (even repeated, in the case that some cited papers are (co-)authored by the same individuals) related to the i-th paper; r_i is the number of papers cited (or the number of bibliographic references) relating to the i-th paper; P is the total number of articles in the isolated field.

Similarly to $\sum_{i=1}^{P} \gamma_i$, $\sum_{i=1}^{P} \rho_i$ can be seen as the product of three terms: ${}^{r}p = \sum \rho_i / \sum {}^{r}a_i$ (≤ 1) i.e. the percentage of unique cited authors; ${}^{r}app = \sum {}^{r}a_i / \sum r_i$ (≥ 1) i.e. the average number of authors per cited paper. $\sum_{i=1}^{P} r_i$ the total number of references given.

Combining Eqs. 8 and 9 with Eq. 3, it is obtained:

$$\sum_{i=1}^{P} \gamma_i = \left(\frac{{}^{c}p}{{}^{r}p} \cdot \frac{{}^{c}app}{{}^{r}app} \right) \cdot \sum_{i=1}^{P} \rho_i. \quad (10)$$

The "balanced" situation $\sum \gamma_i = \sum \rho_i$ can be achieved in the case the following two (sufficient but not necessary) conditions occur (also see the exemplification in Fig. 3):

$$\begin{aligned} {}^{c}p &= {}^{r}p \\ {}^{c}app &= {}^{r}app. \end{aligned} \quad (11)$$

that is to say, (i) equal average percentage of unique authors and (ii) equal average number of authors for the papers citing and being cited by the total P papers in the isolated field.

Equation 7 could also be met without necessarily satisfying the two conditions in Eq. 11, that is to say in the case the quantity in brackets in Eq. 10 was unitary. However, there is no practical reason that justifies the occurrence of this coincidence, which is purely conjectural. On the other hand, the two conditions of Eq. 11 seem reasonable for (citing and cited) papers within the same field. In any case, they will be tested empirically in the next section.

Preliminary empirical analysis of the *cs*-index

Data collection

A preliminary empirical analysis of the *cs*-index is performed by selecting some papers from a set of journals of seven different ISI subject categories (in brackets the total number of journals indexed by Thomson Scientific in each category): Biology (85), Analytical

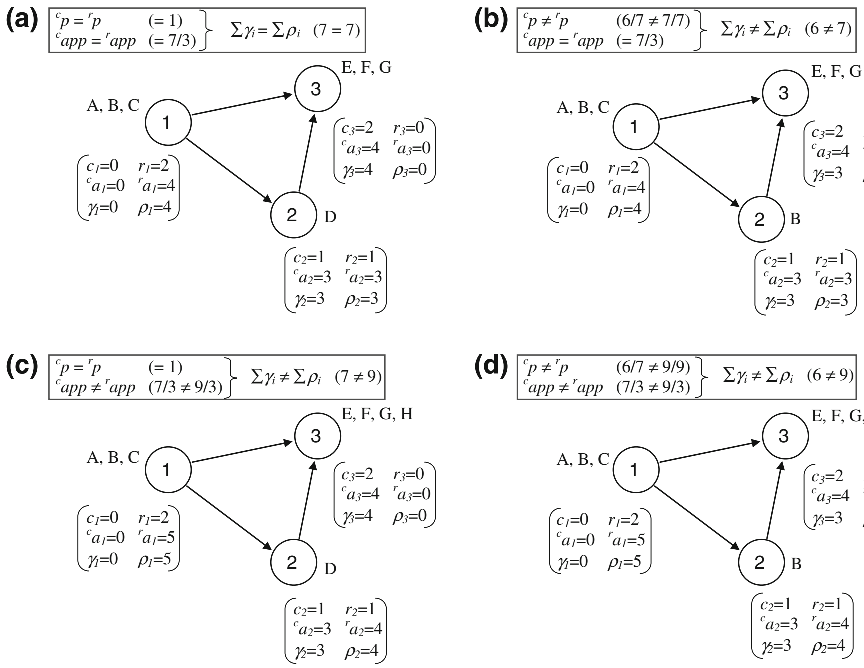

Fig. 3 Examples of isolated groups of three papers. Nodes represent the papers (1, 2 and 3), whose authors are A, B, C, D, etc.; *arrows* represent the citations given by one paper to another. For each paper, it is reported the number of citations obtained (c_i), the number of references given (r_i), the number of total citers (${}^{c}a_i$), the number of total cited authors (${}^{r}a_i$), the number of unique citers (γ_i) and the number of unique cited authors (ρ_i). The equality of Eq. 7 is satisfied in case (a) only, when ${}^{c}p = {}^{r}p$ and ${}^{c}app = {}^{r}app$

Chemistry (73), Manufacturing Engineering (37), Mathematics (289), General & Internal Medicine (155), Applied Physics (125), Psychology (75). For each discipline, we selected a random sample of three scientific journals. For each journal, we considered as articles of interest those produced in the 3-year period from 2008 to 2010, limiting the selection to research papers only (other document types, such as reviews, conference papers or letters, were excluded). Table 2 contains the journal titles and the number of articles examined for each year. Data are retrieved by querying the Web of Science[1] (WoS) database (Thomson Reuters 2012).

For each i-th article of interest, the following operations are performed.

1. Collection of the citation statistics, consisting of:

   c_i the number of citing papers published in 2011 and indexed by the database in use;
   ${}^{c}a_i$ the total number of authors of the (c_i) citing papers (even repeated, if different citing papers are (co-)authored by the same individuals);
   γ_i the total number of unique citers, obtained by performing the union of the (${}^{c}a_i$) total citers and removing those repeated.

[1] The WoS database configuration included the following resources: Citation Index Expanded (SCI-EXPANDED) from 1970 to present, Social Sciences Citation Index (SSCI) from 1970 to present, Arts & Humanities Citation Index (A&HCI) from 1975 to present, Conference Proceedings Citation Index-Science (CPCI-S) from 1990 to present, Conference Proceedings Citation Index-Social Science & Humanities (CPCI-SSH) from 1990 to present.

Table 2 List of journals analyzed within seven ISI subject categories (WoS)

Discipline (ISI subject category)	Journal	Abbreviation	No. of papers			
			2008	2009	2010	Total
Biology	Bio1	Bioscience	84	65	66	215
	Bio2	Biology Direct	46	41	65	152
	Bio3	Journal of Biosciences	60	65	52	177
Chemistry (analytical)	Che1	Analytical Sciences	264	238	209	711
	Che2	Journal of Chemometrics	83	68	76	227
	Che3	Microchemical Journal	85	114	151	350
Engineering (manufacturing)	Eng1	International Journal of Machine Tools & Manufacture	164	139	118	421
	Eng2	Robotics and Computer-Integrated Manufacturing	77	96	87	260
	Eng3	Journal of Intelligent Manufacturing	57	62	71	190
Mathematics	Mat1	Computational Complexity	20	20	21	61
	Mat2	Constructive Approximation	31	46	38	115
	Mat3	Advances in Mathematics	169	146	190	505
Medicine (general and internal)	Med1	American Journal of Medicine	112	98	119	329
	Med2	Mayo Clinic Proceedings	86	55	74	215
	Med3	Medicine	33	40	30	103
Physics (applied)	Phy1	Applied Physics Express	341	339	345	1,025
	Phy2	Current Applied Physics	177	430	436	1,043
	Phy3	Journal of Magnetic Resonance	230	214	241	685
Psychology	Psy1	Journal of Experimental Psychology: Learning Memory and Cognition	66	94	52	212
	Psy2	Cognitive Psychology	18	26	24	68
	Psy3	Health Psychology	125	90	73	288

For each journal, we considered the research papers issued in the three-year period from 2008 to 2010

The choice of a time window for citations accumulation of 1 year (2011) is to simplify the analysis.

2. Determination of an appropriate cCT_i, which takes into account the propensity to obtain citations from different authors. The construction of cCT_i is based on a sample of S articles that are issued in 2011 by the same journal of the (i-th) article of interest.

For each j-th of the articles of the sample, we determine:

r_j the number of cited papers that were published in the three-year period from 2008 to 2010 and are indexed by the database in use. These constraints were introduced to be consistent with the time window described at point (1) (Moed 2011);
${}^{r}a_j$ the total number of cited authors (even repeated, if different cited papers are authored by the same individuals);
ρ_j the total number of unique cited authors, obtained by the union of the (${}^{r}a_i$) total cited authors, removing those repeated.

Next, the distribution of the ρ_j values (relating to the papers of the sample) is constructed and the cCT_i is defined by an appropriate central tendency indicator—e.g. the

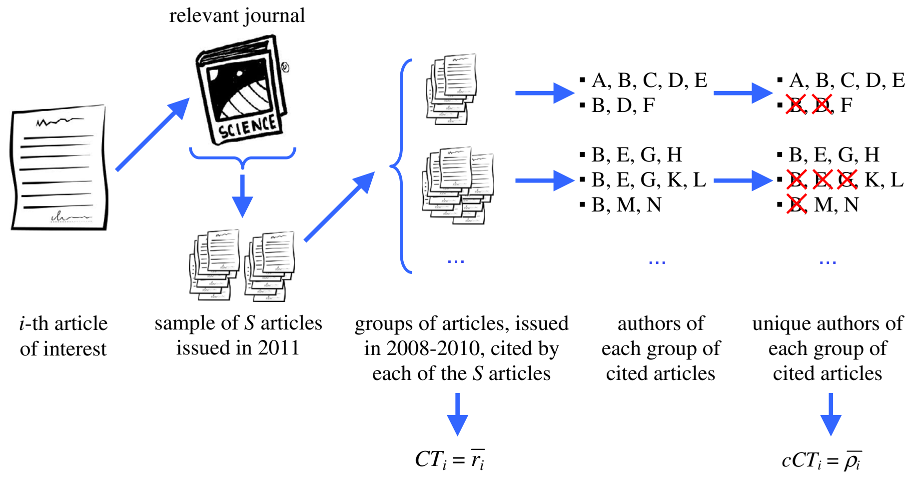

Fig. 4 Scheme of the construction of the CT_i and cCT_i values related to the articles of interest

mean ($\bar{\rho}$) or median ($\tilde{\rho}$). This construction (schematized in Fig. 4) is based on the assumption that, referring to the *i*-th article, the propensity to be cited by different authors is, on average, reasonably close to the propensity to cite different authors, referring to articles issued by the same journal. According to this construction, articles published in the same journal and in the same year will have the same cCT_i value. Probably, a more rigorous way to estimate the cCT_i—but also computationally more expensive—is to use the distribution of the ρ_j values relating to the articles that cite other articles, issued by the article of interest's journal. For further information about this point, please refer to (Franceschini et al. 2013a).

The *cs*-index related to the articles of each journal can be calculated using the cCT_i determined at point (2) (according to Eq. 5). The information at point (2) can also be used to determine the average number of authors (${}^{r}app$) and the percentage of unique authors (${}^{r}p$) of the articles cited by the (S) articles of the sample (see Eq. 9). Similarly, the information at point (1) can be used to determine the average number of authors (${}^{c}app$) and the percentage of unique authors (${}^{c}p$) of the articles that cite the (P) articles of interest (see Eq. 8).

The overall ${}^{c}app$, ${}^{r}app$, ${}^{c}p$ and ${}^{r}p$ values of the seven fields examined can be estimated by aggregating data related to the three journals considered in each discipline.

Information at point (1) can also be used to build other indicators: C (i.e., total number of citations), CPP (i.e., average citations per paper), h, ch and s. As regards the s-index, we will compare the (c_i) citations obtained by each (i-th) paper with a CT_i represented by the mean or median number of references ($\bar{r}$ and $\tilde{r}$ respectively) that are given by each (j-th) of the articles of the sample.

Conventionally, all indicators are constructed considering the citations obtained in 2011 and the references given to (cited) articles, issued from 2008 to 2010 and indexed by WoS.

Table 3 summarizes the name, meaning and the calculation method of the major indicators used in the empirical analysis. The purpose of this table is to ease the understanding of the remaining of the paper.

Table 3 List of the major indicators used in the empirical analysis

Abbr.	Name	Meaning	Calculation method
C	Total citations	Total number of citing papers (or citations) relating to the P total articles of interest	$\sum c_i$
$^{c}a_i$	Citers	Total number of citers (even repeated) related to the i-th paper of interest	For each i-th paper of interest, count the total authors (even repeated) of the citing papers, issued in a specific time window
^{c}app	Average citers	Average number of authors per citing paper, considering the P total articles of interest	$\sum {}^{c}a_i / \sum c_i$
cCT_i	Comparison term for *cs*-index	Comparison term, associated with the i-th paper of interest, to be compared with γ_i when determining *cs*	Several options, based on three steps: (i) selecting a sample of S articles homologous to that interest; (ii) deciding whether to consider the distribution of unique citers or unique cited authors, relating to the papers of the sample; (iii) defining cCT_i by an indicator of central tendency, applied to the distribution chosen at point (ii)
ch	*Citer-h*-index	Number (*ch*) such that, for a general group of papers, *ch* papers are cited by at least *ch* different citers while the other papers are cited by no more than *ch* different citers	Sort the P total papers of interest in decreasing order, according to γ_i, and then count them, stopping the tally at the breakeven point between γ_i and the number of examined papers
c_i	Citations	Number of citing papers (or citations) relating to the i–th paper of interest	For each i-th paper of interest, count the citations obtained from other papers, issued in a specific time-window
^{c}p	Percent of unique citers	Overall percentage of unique citers, related to the P total papers of interest	$\sum \gamma_i / \sum {}^{c}a_i$
CPP	Average citations per paper	Average number of citations per paper, related to the P article of interest	C/P
cs	*Citer-success*-index	Number of papers of interest, with a number of unique citers (γ_i) greater than or equal to cCT_i	Compare the unique citers (γ_i) related to each of the P papers of interest, with the comparison term (cCT_i); count the papers for which $\gamma_i \geq cCT_i$

Table 3 continued

Abbr.	Name	Meaning	Calculation method
CT_i	Comparison term for *s*-index	Comparison term for determining the *s*-index, associated with the *i*-th paper of interest	Several options, based on three steps: (i) selecting a sample of *S* articles homologous to that interest; (ii) deciding whether to consider the distribution of references given or citations obtained, relating to the papers of the sample; (iii) defining CT_i by an indicator of central tendency, applied to the distribution chosen at point (ii)
γ_i	Unique citers	Number of unique authors that cite the *i*-th paper of interest	For each *i*-th paper of interest, count the unique citing authors of papers issued in a specific time-window (i.e., a repeated citer counts one)
h	Hirsch's index	Number (h) such that, for a general group of papers, h papers obtained at least h citations while the other papers no more than h citations	Sort the P total papers of interest in decreasing order, according to c_i, and then count them, stopping the tally at the breakeven point between c_i and the number of examined papers
P	Total papers of interest	Total number of papers of interest, to be evaluated by several performance indicators	Count the papers of interest
R	Total references	Total number of references given by the totality of the S papers in the sample	$\sum r_i$
$^r a_j$	Cited authors	Total number of cited authors (even repeated) related to the *j*–th paper of the sample	For each *j*-th paper of the sample, count the total authors (even repeated) of the cited papers, issued in a specific time window and indexed by the database in use
$^r app$	Average cited authors	Average number of authors per cited paper, considering the S papers in the sample	$\sum {}^r a_i / \sum r_i$
r_j	References	Number of references given by the *j*–th paper of the sample	For each *j*-th paper of the sample, count the references to papers issued in a specific time-window and indexed by the database in use.
$\bar{r}$	Average references	Mean value of the distribution of the r_j values, relating to the S papers in the sample	Apply the average operator to the r_j values
$\tilde{r}$	Median references	Median value of the distribution of the r_j values, relating to the S papers in the sample	Apply the median operator to the r_j values

Table 3 continued

Abbr.	Name	Meaning	Calculation method
${}^{r}p$	Percent of unique cited authors	Overall percentage of unique cited authors, related to the S total papers in the sample	$\sum \rho_j / \sum {}^{r}a_j$
ρ_j	Unique cited authors	Number of unique authors cited by the j-th paper of the sample	For each j-th paper of the sample, count the unique authors of the cited papers, issued in a specific time window and indexed by the database in use (i.e., a repeated cited author counts one)
$\bar{\rho}$	Average unique cited authors	Mean value of the distribution of the ρ_j values, relating to the S papers in the sample	Apply the average operator to the ρ_j values
$\tilde{\rho}$	Median unique cited authors	Median value of the distribution of the ρ_j values, relating to the S papers in the sample	Apply the median operator to the ρ_j values
S	Total papers of the sample	Total number of papers of the sample, used for estimating cCT_i and CT_i	Count the papers of the sample
s	*Success*-index	Number of papers with a number of citations greater than or equal to CT_i	Compare the citations (c_i) obtained by each of the P papers of interest, with the comparison term (CT_i); count the papers for which $c_i \geq CT_i$

Indicators are sorted alphabetically according to their abbreviations

Data analysis

Table 4 summarises the results of the empirical analysis. For each journal, the $C = \sum c_i$ total citing papers are those citing each (i-th) of the P papers of interest, and the $R = \sum r_i$ total cited papers are the ones cited by each (j-th) of the S articles of the sample. All statistics were constructed considering the aforementioned time windows and the papers indexed by WoS.

For a specific journal, there are marginal differences between citing and cited authors, as regards (i) the average number of authors per paper (i.e. ${}^{c}app$ and ${}^{r}app$ values) and (ii) the percentage of unique authors (i.e. ${}^{c}p$ and ${}^{r}p$ values).

Besides, there are relatively small variations among the three journals in a specific field. For this reason, it seems appropriate to calculate some aggregated indicators for the whole disciplines (see "overall" indicators in Table 4). The determination of the overall indicators—by joining the data related to the three journals in each discipline—is extended to all the indicators presented in Table 4. In the case of the *cs*-index and *s*-index, overall indicators are constructed using cCT_i and CT_i values determined on the basis of macro-samples obtained by joining the articles issued in 2011 by the three journals selected for each discipline.

Returning to the comparison between ${}^{c}app$ and ${}^{r}app$ values in each field, a simple way to visualize their similarity is through box-plots based on overall statistics. In particular, two distributions are considered; (i) that of the number of authors per paper relating to articles that cite the papers of interest, and (ii) that of the papers cited by the papers of the (macro-)sample (see Fig. 5).

It can be seen that, for each discipline, the notches of the two box-plots (respectively for citing and cited papers) almost completely overlap, supporting the view of absence of systematic differences between the two distributions. The same hypothesis can be tested by more rigorous statistical tests, albeit introducing additional assumptions about distributions. On the contrary, when comparing different fields there are systematic differences, confirming what observed in other studies (Glänzel 2002). For example, let us consider the comparison between the notches relating to Mathematics and Physiscs.

As regards the comparison between ${}^{c}p$ and ${}^{r}p$ values, the question is a bit more complicated: the overall percentages of different authors (respectively citing or cited) can be seen as weighted averages of the same percentages, at the level of individual papers:

$$\begin{aligned} {}^{c}p &= \left(\sum_{i=1}^{P} \gamma_i\right) \Big/ \left(\sum_{i=1}^{P} {}^{c}a_i\right) = \left(\sum_{i=1}^{P} {}^{c}p_i \cdot {}^{c}a_i\right) \Big/ \left(\sum_{i=1}^{P} {}^{c}a_i\right) \\ {}^{r}p &= \left(\sum_{j=1}^{S} \rho_j\right) \Big/ \left(\sum_{j=1}^{S} {}^{c}a_j\right) = \left(\sum_{j=1}^{S} {}^{r}p_j \cdot {}^{r}a_j\right) \Big/ \left(\sum_{j=1}^{S} {}^{r}a_j\right), \end{aligned} \tag{13}$$

being ${}^{c}p_i$ the percentage of unique citers relating to the i-th of the P papers of interest; ${}^{c}a_i$ the "weight" of ${}^{c}p_i$, i.e. the number of authors (even repeated) citing the i-th paper; ${}^{r}p_j$ the percentage of unique authors cited by the j-th of the S papers of the sample; ${}^{r}a_j$ the "weight" of ${}^{r}p_j$, i.e. the number of authors (even repeated) cited by the j-th paper.

Being ${}^{c}p$ and ${}^{r}p$ weighted quantities, one can represent the distributions of ${}^{c}p_i$ and ${}^{r}p_j$ values by special box-plots based on *weighted quartiles*, defined as:

- ${}^{c}Q_w^{(1)}$, ${}^{c}Q_w^{(2)}$ and ${}^{c}Q_w^{(3)}$, i.e. the weighted first, second (or weighted median) and third quartile of the ${}^{c}p_i$ values. These indicators are obtained by ordering in ascending order

Table 4 Summary of the analysis results. For each of the journals (in Table 2), we report the indicators illustrated in the "Data collection" sub-section

Field	Journ.	${}^{c}app$	${}^{r}app$	${}^{c}p$	${}^{r}p$	P	C	CPP	h	ch	S	R	cCT_i		cs-index		CT_i		s-index	
													$\bar{\rho}$	$\tilde{\rho}$	$(\bar{\rho})$	$(\tilde{\rho})$	$\bar{r}$	$\tilde{r}$	$(\bar{r})$	$(\tilde{r})$
Bio	Bio1	4.6	5.5	0.95	0.91	215	1,131	5.3	14	37	76	792	52.3	35.0	25	38	10.4	9.0	30	35
	Bio2	4.9	6.5	0.94	0.86	152	469	3.1	9	26	59	943	89.4	60.0	3	4	16.0	14.0	2	2
	Bio3	5.3	5.9	0.86	0.93	177	274	1.5	7	19	71	382	29.3	18.0	9	20	5.4	4.0	16	17
	overall	4.8	6.0	0.93	0.89	544	1,874	3.4	15	45	206	2,117	55.0	35.0	31	57	10.3	8.5	37	52
Che	Che1	4.4	4.5	0.89	0.83	711	905	1.3	7	20	191	1,076	21.1	17.0	14	30	5.6	5.0	14	14
	Che2	3.9	3.9	0.92	0.86	227	371	1.6	7	17	65	304	15.8	12.0	22	29	4.7	4.0	15	15
	Che3	4.3	4.3	0.92	0.88	350	948	2.7	9	28	185	1,274	25.9	22.0	35	50	6.9	5.0	29	51
	overall	4.3	4.3	0.91	0.86	1,288	2,224	1.7	10	30	441	2,654	22.4	17.0	71	128	6.0	5.0	44	78
Eng	Eng1	3.6	3.3	0.86	0.84	421	1,148	2.7	9	23	98	392	11.3	9.0	115	142	4.0	3.0	78	126
	Eng2	3.2	3.1	0.93	0.88	260	374	1.4	6	15	101	229	6.2	5.0	74	86	2.3	2.0	57	57
	Eng3	3.0	2.8	0.90	0.93	190	191	1.0	6	10	78	140	4.6	3.0	41	54	1.8	1.0	43	43
	overall	3.4	3.2	0.88	0.87	871	1,713	2.0	10	24	277	761	7.6	5.0	261	341	2.7	2.0	266	266
Mat	Mat1	2.2	2.4	0.92	0.86	61	39	0.6	2	6	19	25	2.7	1.0	11	17	1.3	1.0	11	11
	Mat2	2.5	2.1	0.88	0.80	115	178	1.5	4	8	36	87	4.0	3.0	18	26	2.4	1.0	17	31
	Mat3	1.9	2.0	0.88	0.77	687	912	1.3	7	11	290	819	4.3	3.0	113	157	2.8	2.0	126	126
	overall	2.0	2.0	0.88	0.77	863	1,129	1.3	7	13	345	931	4.2	3.0	138	190	2.7	2.0	145	145
Med	Med1	5.3	7.5	0.93	0.91	329	533	1.6	6	25	125	946	51.4	36.0	1	7	7.6	6.0	1	4
	Med2	5.3	6.8	0.92	0.89	215	996	4.6	14	37	75	833	66.8	42.0	12	31	11.1	8.0	18	27
	Med3	5.6	7.7	0.92	0.91	103	489	4.7	10	29	48	424	61.8	45.5	7	12	8.8	7.0	17	20
	overall	5.4	7.3	0.92	0.90	647	2,018	3.1	15	44	248	2,203	58.1	40.0	26	56	8.9	6.0	45	82
Phy	Phy1	5.8	6.1	0.82	0.81	1,025	2,919	2.8	17	50	418	2,483	29.1	24.0	122	147	5.9	5.0	149	149
	Phy2	4.5	4.8	0.89	0.85	1,043	1,939	1.9	12	34	526	2,573	20.1	14.0	99	160	4.9	4.0	111	111
	Phy3	4.4	4.5	0.87	0.79	685	1,579	2.3	11	31	243	1,671	24.1	19.0	53	80	6.9	6.0	37	37
	overall	5.1	5.2	0.85	0.82	2,753	6,437	2.3	17	55	1,187	6,727	24.1	19.0	270	395	5.7	5.0	287	287

Table 4 continued

Field	Journ.	^{c}app	^{r}app	^{c}p	^{r}p	P	C	CPP	h	ch	S	R	cCT_i		cs-index		CT_i		s-index	
													$\bar{\rho}$	$\tilde{\rho}$	$(\bar{\rho})$	$(\tilde{\rho})$	$\bar{r}$	$\tilde{r}$	$(\bar{r})$	$(\tilde{r})$
Psy	Psy1	2.9	2.7	0.89	0.79	212	545	2.6	10	18	78	596	16.7	15.0	20	23	7.6	7.0	12	12
	Psy2	2.9	2.5	0.88	0.85	68	298	4.4	7	16	17	172	21.3	19.0	10	11	10.1	9.0	5	5
	Psy3	4.3	4.4	0.93	0.89	288	1,245	4.3	12	35	90	738	32.4	26.0	43	58	8.2	7.0	32	41
	overall	3.8	3.5	0.92	0.86	568	2,088	3.7	15	37	185	1,506	24.7	19.0	87	121	8.1	7.0	50	60

Overall indicators are obtained by aggregating the data relating to the three journals examined in each field

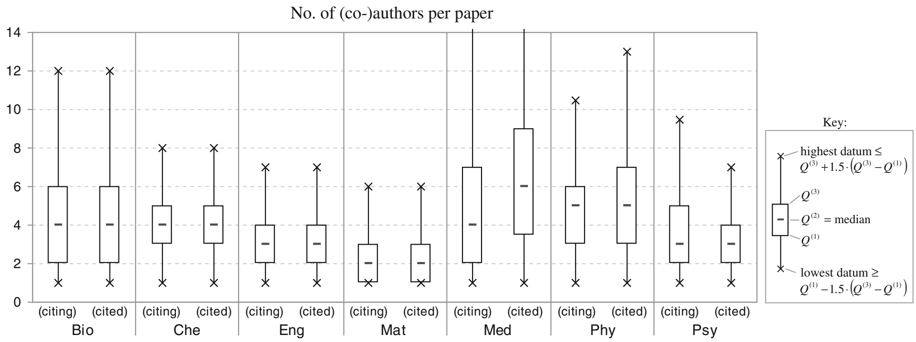

Fig. 5 Box-plot of the distribution of the number of (co-)authors relating to the citing and cited papers, concerning the seven fields examined. Citing papers are those that cite the *P* papers of interest while cited papers are those cited by the *S* papers of the macro-sample. $Q^{(1)}$, $Q^{(2)}$ and $Q^{(3)}$ are the first, second and the third quartile of the distributions of interest

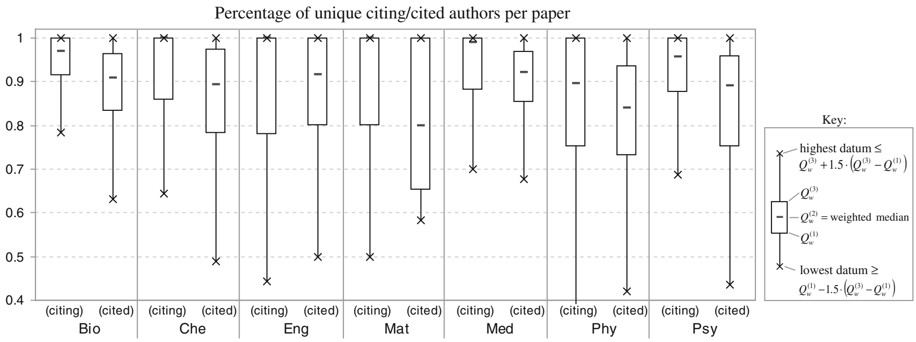

Fig. 6 "Weighted" box-plot of the percentage of unique citing (${}^{c}p_i$) and cited authors (${}^{r}p_i$), relating to the papers that cite the papers of interest and are cited by the papers of the macro-sample, in the seven fields examined. $Q_w^{(1)}$, $Q_w^{(2)}$ and $Q_w^{(3)}$ are the first, second and the third weighted quartile of the distributions of interest

the ${}^{c}p_i$ values of the articles of interest and considering the values for which the cumulative of weights is equal to respectively the 25, 50 and 75 % of their sum;

- ${}^{r}Q_w^{(1)}$, ${}^{r}Q_w^{(2)}$ and ${}^{r}Q_w^{(3)}$, i.e. the weighted first, second (i.e. the weighted median) and third quartile of the ${}^{r}p_i$ values.

The box-plots relating to weighted quartiles are represented in Fig. 6. The differences between the ${}^{c}p_i$ and ${}^{r}p_j$ distributions within the same field seem insignificant. We also note the absence of significant differences between fields.

Returning to Table 4, there are relatively little differences in terms of cCT_i values (i.e. estimators of the propensity to cite different authors), for journals of the same field. Some exceptions are: Bio2 for Biology and Eng1 for Engineering. This incomplete uniformity is probably due to the fact that some journals are influenced by publications of neighbouring fields, with different citation propensity. For a more rigorous estimate, it would probably be appropriate to define cCT_is using a larger sample of papers/journals.

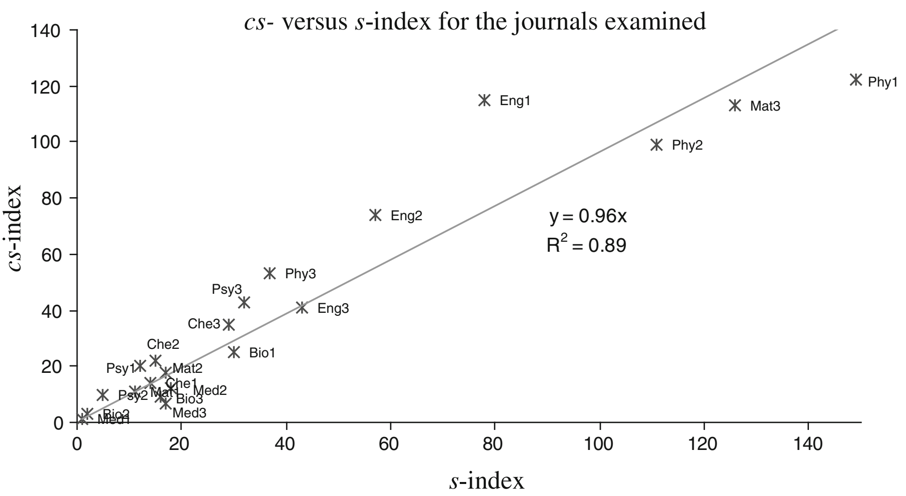

Fig. 7 Relationship between the *cs*- and *s*-index for the journals examined. Indicators are calculated considering respectively $cCT_i = \bar{\rho}$ and $CT_i = \bar{r}$ (see Table 4)

Table 5 Spearman's rank correlation coefficients between the impact indicators of the journals examined

	C	CPP	h	ch	$cs(\bar{\rho})$	$cs(\tilde{\rho})$	$s(\bar{r})$	$s(\tilde{r})$
C	1.00	0.47	0.83	0.82	0.59	0.64	0.57	0.55
CPP		1.00	0.75	0.74	−0.11	−0.13	−0.02	−0.04
h			1.00	0.90	0.31	0.35	0.36	0.31
ch				1.00	0.15	0.24	0.26	0.22
$cs(\bar{\rho})$					1.00	0.96	0.90	0.91
$cs(\tilde{\rho})$						1.00	0.94	0.94
$s(\bar{r})$							1.00	0.98
								1.00
Mean	832.5	2.5	8.9	23.6	40.3	56.3	39.0	44.5
Std. dev.	687.7	1.4	3.6	11.3	40.1	52.2	42.2	44.6

For each journal, in Table 4 are reported two different cCT_js: i.e. using $\bar{\rho}$ and $\tilde{\rho}$. In general, the resulting values are higher in the first case. This probably depends on the incidence of papers characterized by hyperauthorship—i.e. literally tens or even hundreds of authors (Cronin 2001)—which tends to "inflate" $\bar{\rho}$ but not $\tilde{\rho}$, as the latter indicator is only marginally sensitive to the right tail of the distribution of ρ_j values.

Another interesting aspect is the link between *cs*-index and *s*-index. The diagram in Fig. 7—which is constructed using $cCT_i = \bar{\rho}$ and $CT_i = \bar{r}$ (in Table 4)—shows a strong correlation ($R^2 \approx 89\ \%$), similar to that between *ch* and *h* (Franceschini et al. 2010; Egghe 2012). All the points of the graph—although resulting from articles of different scientific fields—tend to be distributed around the same trend line, which is very close to the bisector of the *cs*–*s* plane.

In the absence of "anomalies"—e.g. high incidence of self-citations or citations from recurrent citing authors—the *cs*-index and *s*-index should be very close. Therefore, the study of their difference can be useful to highlight abnormal situations.

In this specific case, there is no important difference between the journals analyzed, in terms of citations from self- or recurrent citers; this is also proven by the relatively similar ${}^{c}p$ values (in Table 4). The relatively important deviation of Eng1 from the tendency line is due to an abnormal citation transfer from external disciplines with different propensity to co-authorship. Precisely, it was observed that a relatively low portion (lower than 10 %) of the papers issued by Eng1 obtained several citations from journals in the Applied Chemistry field, in which co-authorship is relatively higher than that in the Engineering field. This is proven by the fact that the Eng1's ${}^{c}app$ value is "inflated" with respect to that ones of the other two journals in the same field (i.e. Eng2 and Eng3; see Table 4).

Table 5 shows the Spearman's rank correlation coefficients (Kendall 1970) relating to the indicators of impact in Table 4, at the level of single journal. Not surprisingly, most of the indicators are positively correlated. The only exception is the absence of correlation between the *CPP* and the *cs*- and *s*-index; the reason probably comes from the fact that the former, contrarily to the other ones, is size-dependent and non-field-normalized.

Final remarks

The first part of this study revealed that the comparison term (cCT_i) of the *cs*-index can be constructed using the distribution of the ρ_j values related to the papers of a sample. This is justified by the absence of systematic differences between (i) the average number of authors and (ii) the average percentage of unique authors, between citing and cited papers in a certain field. On the other hand, the analysis confirmed some systematic differences between fields, as regards the average number of authors per paper.

The empirical analysis is that the *cs*-index, although generally correlated with the *s*-index, can complement it, being only marginally affected by self-citations and citations from recurrent citers. Similarly to the *s*-index, the *cs*-index has an immediate meaning and is practical for normalizations aimed at obtaining the so-called size-independency, thanks to the ratio scale property (Franceschini et al. 2012). For example, scientific journals with a different number (*P*) of articles could be easily compared by means of the percentage of "successful" papers, i.e., *cs*-index/*P*.

Even if it was not shown directly in this paper, another advantage "inherited" by the *s*-index is that *cs*-index can be calculated for a set of multidisciplinary articles, thanks to the field-normalization that it achieves at the level of individual paper. For example, the *cs*-index can be used as a proxy for synthesizing the productivity and impact of (i) the whole publication output of scientists involved in multiple disciplines (e.g. mathematicians or computer scientists actively involved in bibliometrics), or (ii) that of entire multidisciplinary research institutions.

The major disadvantage of the *cs*-index is the computational complexity of the cCT_i values. E.g. our data collection and analysis—which was performed by an ad hoc application software able to query the WoS database automatically—took about twenty consecutive hours.

Another problem of the *cs*-index, as well as the totality of indicators based on the number of unique (citing/cited) authors, is author disambiguation (Jovanovic and Fritsche 2011). There are two typical error types. The first is represented by homonymous authors. Generally, authors with common names (e.g. Chinese family names) or identified by full surname and first name(s)' initial(s)—rather than full first name(s)—are subject to this kind of problem. The practical effect is that contributions of different homonym authors are erroneously added up. The second error type is that of failing to recognize a repeated author, e.g. due to multiple name spellings originated from omitted accents or omitted first names initials. The authors are aware that the only way to solve the disambiguation problem once and for all would be the use of a unified identifying system for scientific authors, i.e. a "universal registry" associating each author with a unique identifier (Dervos et al. 2006). Unfortunately, such a system is not yet available, although some attempts, such as the ResearcherID tool by Thomson Reuters, seem to go in this direction.

Finally, a potential drawback of *cs*-index is represented by hyperauthorship, which could lead to inflate cCT_i values. A partial solution to this problem is (i) to determine cCT_i by indicators that are insensitive to the right-hand tail of the distribution of ρ_j (e.g. $\tilde{\rho}$), or (ii) to apply some exclusion criteria, so as to curtail the count of the authors of a certain paper, according to a conventional threshold.

References

Ajiferuke, I., & Wolfram, D. (2010). Citer analysis as a measure of research impact: Library and information science as a case study. *Scientometrics, 83*(3), 623–638.

Bornmann, L. (2013). A better alternative to the *h* index. *Journal of Informetrics, 7*(1), 100.

Braun, T., Glänzel, W., & Schubert, A. (1985). *Scientometric Indicators: A 32-country comparative evaluation of publishing performance and citation impact*. Philadelphia: World Scientific.

Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology, 52*(7), 558–569.

Dervos, D. A., Samaras, N., Evangelidis, G., Hyvärinen, J. & Asmanidis, Y. (2006). The Universal Author Identifier System (UAI_Sys). *Proceedings of the 1st International Scientific Conference, eRA: The Contribution of Information Technology in Science, Economy, Society and Education*, Tripolis, Greece, 16–17 September 2006.

Dieks, D., & Chang, K. H. (1976). Differences in impact of scientific publications: Some indices derived from a citation analysis. *Social Studies of Science, 6*(2), 247–267.

Egghe, L. (2012). A rationale for the relation between the citer *h*-index and the classical *h*-index of a researcher. *Scientometrics*,. doi:10.1007/s11192-012-0770-1.

Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics: Quantitative methods in library, documentation and information science*. Amsterdam: Elsevier.

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., & Mastrogiacomo, L. (2013a). Evaluating research institutions: The potential of the success-index. *Scientometrics, 96*(1), 85–101.

Franceschini, F., Maisano, D., & Mastrogiacomo, L. (2013b). The citer-success-index: An indicator to select a subset of elite papers, based on citers. *Proceedings of the 14th International Society of Scientometrics and Informetrics (ISSI) Conference*, vol. I: 300–315, 15th-19th July 2013, Vienna, Austria, ISBN: 9783200031357.

Franceschini, F., Maisano, D., Perotti, A., & Proto, A. (2010). Analysis of the *ch*-index: An indicator to evaluate the diffusion of scientific research output by citers. *Scientometrics, 85*(1), 203–217.

Glänzel, W. (2002). Coauthorship patterns and trends in the sciences (1980–1998): A bibliometric study with implications for database indexing and search strategies. *Library Trends, 50*(3), 461–473.

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*, 16569–16572.

ISI Web of Knowledge. (2012). Essential Science Indicators. http://thomsonreuters.com. Accessed Nov 2012.

Jovanovic, M. & Fritsche, F. (2011). There goes another one: Introducing the NUCA-set of indicators. *Proceedings of the 13th International Society of Scientometrics and Informetrics (ISSI) Conference*, vol. I: 333–338, 4th–7th July 2011, Durban, South Africa, ISBN: 9789081752701.

Kendall, M.G. (1970). Rank correlation methods (4th ed.). London: Griffin. ISBN 978-0-852-6419-96.

Moed, H. F. (2011). The source-normalized impact per Paper (SNIP) is a valid and sophisticated indicator of journal citation impact. *Journal of the American Society for Information Science and Technology, 62*(1), 211–213.

Thomson Reuters. (2012). 2011 Journal Citation Reports—Science Edition, www.isiknowledge.com.

Vinkler, P. (2009). The π-index: A new indicator for assessing scientific impact. *Journal of Information Science, 35*(5), 602–612.

Vinkler, P. (2010). The π_v-index: A new indicator to characterize the impact of journals. *Scientometrics, 82*(3), 461–475.

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.

White, H. D. (2001). Authors as citers over time. *Journal of the American Society for Information Science and Technology, 52*(2), 87–108.