https://doi.org/10.26219/heal.aueb.6022

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

SCHOOL OF INFORMATION SCIENCES AND TECHNOLOGY

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

MEASURING SCIENTIFIC PUBLICATION IMPACT: ESSENTIALS AND LATEST TRENDS

by

Irini A. Arvaniti

A THESIS

Submitted to the Department of Statistics

of the Athens University of Economics and Business

in partial fulfilment of the requirements for

the degree of Master of Science in Statistics

Athens, Greece

January 2015

https://doi.org/10.26219/heal.aueb.6022

https://doi.org/10.26219/heal.aueb.6022

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

Μετρώντας τον Αντίκτυπο των Επιστημονικών Δημοσιεύσεων: Βασικές Αρχές και Τελευταίες Εξελίξεις

Ειρήνη Α. Αρβανίτη

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής

του Οικονομικού Πανεπιστημίου Αθηνών

ως μέρος των απαιτήσεων για την απόκτηση

Μεταπτυχιακού Διπλώματος

Ειδίκευσης στη Στατιστική

Αθήνα,
Ιανουάριος, 2015

https://doi.org/10.26219/heal.aueb.6022

DEDICATION

To my Parents and my Brother.

https://doi.org/10.26219/heal.aueb.6022

https://doi.org/10.26219/heal.aueb.6022

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Psarakis Stelios for his helpful and important suggestions during the writing of this thesis.

https://doi.org/10.26219/heal.aueb.6022

https://doi.org/10.26219/heal.aueb.6022

VITA

Eirini Arvaniti was born in 1981 in Athens, where she grew up with her family.

She finished high school in Lycée Franco-Hellénique St. Joseph.

She studied mathematics at the University of Athens, where she majored in Applied Mathematics with specialization in Teaching and Computational Mathematics.

She has acquired an MBA degree on Banking and Industry Finance and Management from the University Paris 13, where she was chosen with full scholarship.

She has taught mathematics and science for high school students, national admittance exams and mathematics for business and finance for university students.

Her research interests include number theory, probability theory and fractals.

She speaks English, French and Spanish.

Her personal interests include scouting, mountaineering, dancing tango and jewellery making.

https://doi.org/10.26219/heal.aueb.6022

 https://doi.org/10.26219/heal.aueb.6022

ABSTRACT

Eirini Antoniou Arvaniti

Title: Measuring Scientific Publication Impact: Essentials and Latest Trends

August 2014

The need for measurable in the field of higher education and research has prompted many governments, educational and research institutions to adopt the use of indicators (simple or not) to study the quality and importance of scientific research.

It has now gone more than half a century since the Eugene Garfield (1955) published his pioneering work "Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas". This work, according to many, laid the foundations of a scientific field called "Bibliometrics" or "Scientometrics" and the substance is intended to refer to features of scientific research and related publications.

The evolution of this scientific field continues even today, with the next important milestone in 2005, when Hirsch JE introduced the concept of the index h. Since this indicator, despite the many criticisms and several variations in its use, remains the main tool and measurement of the quality and importance of scientific publications. Essentially, quality measured through another term, the impact of these publications in the scientific community, through citations gaining each publication.

This thesis monitoring throughout evolution history, before the introduction of the h-index, until the last important contributions which have been published on this field.

https://doi.org/10.26219/heal.aueb.6022

ΠΕΡΙΛΗΨΗ

Ειρήνη Αντωνίου Αρβανίτη

Τίτλος: Μετρώντας τον Αντίκτυπο των Επιστημονικών Δημοσιεύσεων: Βασικές Αρχές και Νεώτερες Εξελίξεις.

Αύγουστος 2014

Η ανάγκη για μετρησιμότητα στον χώρο της ανώτατης εκπαίδευσης και της έρευνας έχει ωθήσει πάρα πολλές κυβερνήσεις, εκπαιδευτικούς και ερευνητικούς φορείς να υιοθετήσουν την χρήση δεικτών (απλών ή όχι) για να μελετήσουν την ποιότητα και την σημαντικότητα της επιστημονικής έρευνας.

Έχει περάσει πια πάνω από μισός αιώνας από τότε που ο Eugene Garfield (1955) δημοσίευσε την πρωτοπόρα εργασία του “Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas” . Η εργασία αυτή, σύμφωνα με πολλούς, έθεσε τα θεμέλια μιας επιστημονικής περιοχής η οποία ονομάζεται "Bibliometrics" ή "Scientometrics" και στην ουσία έχει στόχο να αναφερθεί σε χαρακτηριστικά της επιστημονικής έρευνας και των σχετικών δημοσιεύσεων.

Η εξέλιξη αυτού του επιστημονικού πεδίου συνεχίζεται ακόμα και σήμερα με επόμενο σημαντικό σταθμό το 2005 όταν ο Hirsch J.E. εισήγαγε την έννοια του δείκτη h. Από τότε ο δείκτης αυτός, παρά τις πολλές κριτικές και τις διάφορες παραλλαγές στην χρήση του, παραμένει ο βασικός δείκτης υπολογισμού και μέτρησης της ποιότητας και της σημαντικότητας των επιστημονικών δημοσιεύσεων. Ουσιαστικά μετράται η ποιότητα μέσω μιας άλλης έννοιας, του αντίκτυπου που έχουν οι δημοσιεύσεις αυτές στην επιστημονική κοινότητα, μέσω των αναφορών που κερδίζει κάθε δημοσίευση.

Η εργασία αυτή παρακολουθεί όλη την ιστορική διαδρομή, πριν από την εισαγωγή του h-index, μέχρι τις τελευταίες σημαντικές παρατηρήσεις οι οποίες έχουν δημοσιευθεί, σχετικά με τον χώρο αυτό.

 https://doi.org/10.26219/heal.aueb.6022

Table of Contents

https://doi.org/10.26219/heal.aueb.6022

https://doi.org/10.26219/heal.aueb.6022

https://doi.org/10.26219/heal.aueb.6022

LIST OF FIGURES

https://doi.org/10.26219/heal.aueb.6022

Chapter 1: Introduction

This postgraduate thesis aims to refer to examine and describe the main methods and techniques for measuring the performance (in terms of impact) of scientific research.

The studies based on bibliometric analysis increased continuously recent years in the international scientific community. The bibliometric indicators (used to say “indices”) of research performance, contribute a lot in the measurable image of Research, Technological Development and Innovation systems. As they used for the evaluation of research organizations, groups and researchers record the research fields in which active scientific community, reflecting new research emerging fields and scientific networks established for the implementation of common research goals.

Specifically, based on scientific publications is possible to identify characteristics and trends of research production entity-level, country or wider set of countries, the assessment the impact of scientific work, the evaluation of research activity and identify national and multinational networks scientists and disciplines.

Measurements of scientific publications with figures expressed with "bibliometric indicators." Such of these indicators, the number of Publication is the simplest indicator for recording production scientific work and therefore the research project by scientist, institution, discipline or country.

Apart from the number of publications, the most common bibliometric indicators used to assess the impact and originality scientific work, based on analysis of citations to publications by other scientific publications.

In the literature, very often are referred problems associated with the purification of the primary bibliographic data and the identification of publications and also some serious constraints must be taken into account in the calculation and use of number of reports, since the number of petitions for received publications are significantly influenced by parameters as the scientific

https://doi.org/10.26219/heal.aueb.6022

or research field, the time range of examining a publication and the type of the publication. The influence of such factors will be briefly described below.

The influence of research fields: there are significant variations in publication practice, in year depreciation of research results and the patterns of publications reference (citation patterns), depending on their disciplines, making it inappropriate to directly compare indicators between different research fields, and often between individual regions the same broader scientific field. For example, in medicine and molecular biology, there is very high productivity in terms of amount of scientific articles published per year.

Usually, the number of citations of these publications is increasing and reaches its maximum at short period after their publication. In contrary, publications rates in social sciences are much slower while recording reports for a long time after their publication. On humanities, the majority of scientific publications are books, monographs and articles published in journals nationwide.

In other academic areas (eg in certain sectors IT) is common practice to publish significant research results at conferences, without necessarily follow the same publication in scientific journals. In both cases, publications with high scientific value, as part of their petitions, generally they do not counted in the databases so bibliometric indicators in the respective academic areas not fully capture the reality.

About the time interval of reports analysis: the number of citations to a scientific work depends on the time elapsed after its publication (time of appearance in web). Usually, the older publications have more references, without always-analogous association with their impact on scientific community. To normalize the differences due to the normal growth of the number of references to earlier publications, specifically defined intervals used to measure reference to a publication.

The effect of the different types of scientific publications: the number of petitions affected also significantly by the type of publication. The review

https://doi.org/10.26219/heal.aueb.6022

articles (review) have more references than just scientific articles. For this reason, identifying different bibliometric indices depending on the type of publication is necessary in cases of scientific evaluation work of a research team or individual researchers. The above considerations do not negate the importance of bibliometric indicators as a valuable source of data, as it applies to the interpretation most indicators, can be overcome when the bibliometric indicators seen in the right context.

All these arguments simply highlight the fact that it is not feasible to describe the quality and the significance that is included in a scientific publication with an index, while this index being widely accepted. This finding is the main reason that has trigered the creation of scientific/ knowledge area called “scientometrics”. In subsequent chapters of this thesis the main tools for measuring the influences of a scientific publication will be presented. This presentation will be evolutionarily from the simplest to the most complex and improvement tools (indicators).

https://doi.org/10.26219/heal.aueb.6022

Chapter 2: Bibliographic Data Sources and Crude Indices

2.1 Introduction

In this chapter will refer to the existing most popular bibliographic data sources (bibliographic databases) and some of the most simple and crude (simplicity based on construction) bibliographic indices.

The purpose of this chapter, is that the reader of this postgraduate thesis to recall and equip with all the necessary background concepts and tools, for continuing at the next chapters.

2.2 Bibliographic Databases

Internationally, the most well-established databases that include bibliographic records of worldwide scientific publications and data references between them, are:

- Web of Science (the Company Thomson Reuters),
- Scopus (of Elsevier) and
- Google Scholar (the Google).

In addition, some other databases existing covering (wide or not) knowledge areas, such as PubMed, ScienseDirect.

2.2.1 Web of Science (WoS)

In the of Web of Science (WoS) system are indexed over 13,000 periodically subject to review judges (peer-review) and from 1990 was added and recording some conference proceedings. There are not indexed monographs and books. It is the oldest database of scientific publications with bibliographic records and reports dating for some disciplines since 1900.

The WoS system includes detailed metadata for articles, authors and research organizations. An important advantage of Web of Science is the considered reliability of the resultw of rigorous evaluation versions and mainly journals into the system, according to specific criteria, including their scientific impact.

Disadvantages in particular: the uneven coverage of scientific publications, in terms of geography where the vast majority of the material comes from English-speaking countries, especially the U.S., as well as thematic-lag mainly areas of scientific fields "Engineering and Technology", "Social Sciences "and" Humanities ", while outweigh other field" Natural Sciences ".

2.2.2 Scopus

In the Scopus system, are indexed about 19.000 journals, proceedings conferences and books. As in the system WoS, metadata-containing details of articles, authors and institutions, while the inclusion of publishers or journals is being after qualitative evaluation based on specific criteria.

The Scopus system provides, in relation to the system WoS, more balanced geographical coverage, but time lags: the earlier literature Records beginning in 1966, while references to publications listed after 1995. Moreover, as in WoS, shows uneven coverage certain scientific areas-lagging areas of scientific Field "Social Sciences" and "Humanities", while greater than those of the "Medical Sciences ".

2.2.3 Google Scholar

The system Google Scholar includes a huge number of sources, many of

https://doi.org/10.26219/heal.aueb.6022

which concern reports 'gray or grizzly literature' (eg academic papers, theses, technical reports, research projects). Despite the wide range of content, metadata records are limited and missing critical information on geographical origin or address of organizations to which the authors are also limited.

Moreover, the criteria for importing sources of content are limited and the is not provided available accurate information on geographical or thematic coverage of the scheme. Main advantage of this system is the freely availability.

2.2.4 Medline

Medline (Medical Literature Analysis and Retrieval System Online, or MedlarS Online) is a bibliographic database of life sciences and biomedical information. It includes bibliographic information for articles from academic journals covering medicine, nursing, pharmacy, dentistry, veterinary medicine, and health care. Medline also covers much of the literature in biology and biochemistry, as well as fields such as molecular evolution.

Compiled by the United States National Library of Medicine (NLM), Medline is freely available on the Internet and searchable via PubMed and NLM's National Center for Biotechnology Information's Entrez system.

Medlars is a computerized biomedical bibliographic retrieval system. It was launched by the National Library of Medicine in 1964 and was the first large scale, computer based, retrospective search service available to the public.

The database contains more than 21.6 million records from 5,639 selected publications, covering biomedicine and health from 1950 to the present. Originally, the database covered articles starting from 1965, but this has been enhanced, and records as far back as 1950/51 are now available within the main index.

2.2.5 PubMed

PubMed is a free search engine accessing primarily the Medline database of references and abstracts on life sciences and biomedical topics. The United States National Library of Medicine (NLM) at the National Institutes of Health maintains the database as part of the Entrez system of information retrieval.

From 1971 to 1997, Medline online access to the Medlars computerized database had been primarily through institutional facilities such as university libraries. PubMed, first released in January 1996, ushered in the era of private, free, home and office-based Medline searching. The PubMed system was offered free to the public in June 1997.

2.2.6 ScienceDirect

The ScienceDirect system, is managed by the German/Enflish publisher Elsevier. The ScienceDirect system includes (as in 2013) about 11 million articles from 2,500 journals and over 25,000 electronic books, work references, handbooks and book series.

The articles are categorized into four areas: Physical Sciences and Engineering, Life Sciences, Health Sciences, social sciences and Humanities. Most articles on the site are summaries, which are freely available to the user. Access to the full text of the article (in PDF and HTML for the newer versions) require registration or purchase process with pay-per-view method.

2.3 Crude bibliographic indices

2.3.1 Introduction

In the paragraphs below, some crude or simple indices are described. These indices are mainly simple to be calculated, and the main usage of them may be the relative inference of the publications of different individuals (such researchers or organizations) in the same topic.

2.3.2 Number of publications

The simplest of all indices. It measures the overall publications and it represents the whole volume of the scientific work.

2.3.3 Share (%) publications

It is calculated as the percentage (%) of the number of scientific publications of a researcher as part of the total publications of an institute.

2.3.4 Percentage (%) publications receive reports (% cited papers)

It is calculated as the percentage (%) of publications that have received at least one reference. The calculation is performed in overlapping intervals five years. It is a recognition evidence of the scientific publications.

2.3.5 Number of references to publications (Number of citations)

It is the number of reports that received by the publications during the last five years. It is an indicator of the visibility and the influence of scientific publications.

2.3.6 Share (%) reporting (Share of citations)

Calculated as the percentage (%) of number of reports (time intervals of five years) who received the scientific publications of an individual as a member of a whole: Indicates involvement as regards visibility and influence scientific publications of an individual member.

2.3.7 Impact Index (Citation impact)

The impact index is the average citations per publication and calculated as the ratio of the number the references listed in certain period to total number of publications in the same period. This Indicates the response of the corresponding scientific community to the publications.

2.3.8 Relative impact index (Relative citation impact)

The relative impact index compares influential publications of an individual than influential publications references within a defined set of members. When the value of the impact index is greater than 1, publications of the examined individual have greater appeal than the average of all reference members. This indicator does not take into account the variations in practices references to different academic areas.

2.3.9 Relative impact indicator - normalized based on scientific field (Field normalized citation score)

It comes from the previous examined impact indicator, after normalization according to the different scientific thematic areas. It compares the impact of a publication relative to the resonance having the publications worldwide in same scientific area. It is also known as Index Field Normalized Citation Score, which is calculated based on the distribution of publications on 253 thematic areas of the bases NSI and InCites through specialized software developed by the ECB. When the value of the relevant index resonance is greater than 1, the examined individual publications have greater appeal than the global average.

2.3.10 Number of publications with high impact (P Top X%)

This is the number of scientific publications, in globally, per year rank high in percentage classification, of publications in respective scientific field. The Classification is based on number of citations. The index calculated for intervals five years, concerns the percentiles 1 5, 10, 25 and 50 (ie the number of publications classified Awards at 1%, 5%, 10%, 25% and 50% of the publications higher resonance).

2.3.11 Percentage (%) publications in high impact (Top X%)

It is calculated as the percentage (%) of the number of an individual publications per year which are characterized as publications with high impact compared on the total individual publications of the same year. When the percentage distribution of publications in high impact of the examined individual, approaches or exceeds the global allocating 1%, 5%, 10%, 25% and 50%, the individual is considered to approximate or exceed the corresponding global average. The index is calculated for intervals of five years.

https://doi.org/10.26219/heal.aueb.6022

https://doi.org/10.26219/heal.aueb.6022

Chapter 3: The construction of the h-index

3.1 Introduction

In previous chapters showed some simplified methods and indices, which are useful to make inference about the volume and the quality of the scientific research of an entity.

As a continuation, in this chapter, the novel introduction of the h-index will be presented. Such presentation will include all the initial thoughts and assumptions than made from the introducer of h-index, the J. E. Hirsch, in 2005.

3.2 The h-index at a glance

The h-index (Hirsch index, Hirsch factor also) is a bibliometric measure, based on citations of publications of an author based at a time. A high h-index speaks for major scientific impact of the author. The index h of the author cannot decrease over time.

A scientist has a Hirsch index h if h of his total N publications have at least h times, the remaining (N - h) publications have been cited more than h times. For the determination you can string descending all publications the author's citation frequencies. It now counts down until the r-th publication has less than r h citations is then r - 1.

The h-index to evaluate scientific achievement in 2005 by the Argentine physicist Jorge E. Hirsch developed and in the Proceedings of the National Academy of Sciences published. [1]

A database needed for the calculation of this factor, a variety of data sources are conceivable. However, Hirsch himself has suggested significant data from the Web of Science to use, as these are currently the most reliable and

https://doi.org/10.26219/heal.aueb.6022

comprehensive data base.

The index is based on the distribution of citations received by the publications of a specific researcher. Hirsch writes: “A scientist has index h, if the h/Np papers have at least h references each, and the other (Np - h) papers have no more than h each references”.

In other words, a scholar with an index of h, has published h each of which has been reported in other studies least twice h. Thus, the h index reflects both the number of publications and number of citations per publication.

The index is designed to improve simpler measures such as the total number of citations or publications. The index works properly only for comparing scientists working in the same field; citation conventions differ widely between sectors.

The h index serves as an alternative to more traditional measurements factor of influence for journal impact assessment project of a specific researcher.

Because only the most highly cited articles contribute to the h-index, its determination is a relatively simple process. Hirsch has demonstrated that it has a high predictive value for whether a scientist has won honors as National Academy membership.. The index h grows as reports accumulate and thus depends on the "academic age 'of the researcher.

3.3 Foundation of h-index (the story from the beginning)

Hirsch in 2005 initialy has set the question: how does one researcher quantify the cumulative impact and relevance of the scientific research output of an individual?

In a world of limited resources, such quantification is often necessary for the evaluation and comparison purposes. The publication of a single record and the citation record are pure data that contains very useful information.

This information includes such parameters as:

- the number (NP) of papers that have published over n years,
- the number of citations (Nc j) for each paper (j),
- the journals where the papers were published,
- the effects of parameters, etc.

This is great deal of information: different criteria are evaluated by different people. Here it obvious the need for a single number, the" h-index" as a particularly simple and useful way to propose characterize the scientific output of a researcher.

A scientist has index h if h of his Np papers have at least h citations each and the other (Np - h) papers have each less than h. The research reported by Hirsch in his famous work, focuses on the community of physicist; but he suggested that the h-index should also be able to work for other scientific disciplines as well.

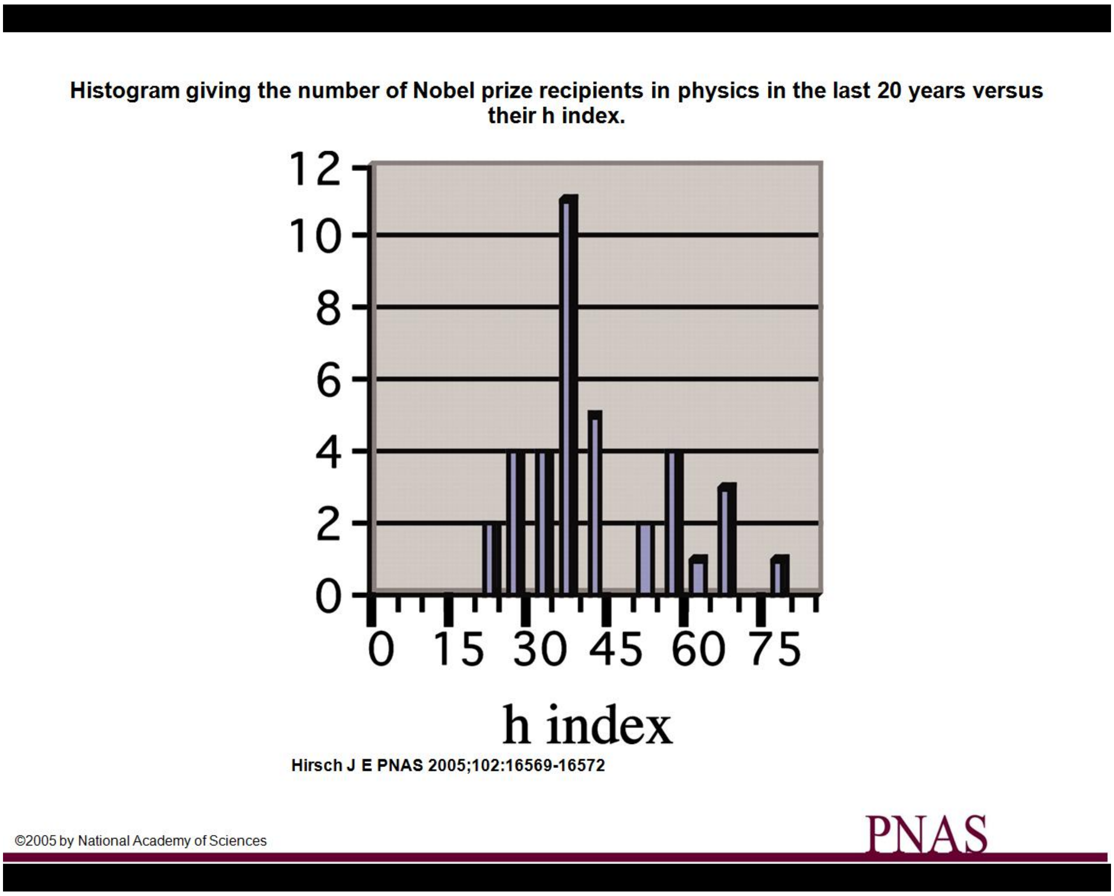

Figure 1: Histogram giving the number of Nobel prize recipients in physics in the last 20 years versus their h index. The peak is at the h index between 35 and 39.

https://doi.org/10.26219/heal.aueb.6022

Given the data had collected until 2004, Hirsch commented that, the highest h among physicists seems E. Witten h be the 110. That is, Witten has 110 papers with at least 110 citations written each. That's a bottom to the total number of citations to Witten papers bound $(h)^2 = 12{,}100$. Of course, the total number of citations is ($N_{c,tot}$) is usually much larger than h^2, and that because both underestimated the total number of citations of the most cited papers h and ignored the papers with less than h citations.

The relationship between $N_{c,tot}$ and h is on the detailed form of the special distribution (see scheme below) the come from, and it is useful to define the proportionality constant a as $N_{c,tot}$ - ah^2.

Hirsch found empirically that a in the range 3 to 5.

I argue that h is better, other single-number criteria commonly used to evaluate scientific output of a researcher, as follows:

According to Hirsch, the use of h-index is preferable than simplest constructions, (such constructions presented in chapter 2 of this thesis), due to the next reasons, in comparison:

A. In comparison with Total number of papers (Np). Advantage: measures productivity. Disadvantage: does not measure importance or impact of papers.

B. In comparison with Total number of citations (Nc,tot). Advantage: measures total impact. Disadvantage: hard to find and may be inflated by a small number of ‘‘big hits,’’ which may not be representative of the individual if he or she is a coauthor with many others on those papers. Another disadvantage is that Nc,tot gives undue weight to highly cited review articles versus original research contributions.

C. In comparison with Citations per paper. Advantage: allows comparison of scientists of different ages. Disadvantage: hard to find, rewards low productivity, and penalizes high productivity.

https://doi.org/10.26219/heal.aueb.6022

D. In comparison with Number of ‘‘significant papers,’’ defined as the number of papers with more than y citations. Advantage: eliminates the disadvantages of criteria A, B, and C and gives an sense of broad and sustained impact. Disadvantage: y is arbitrary and will randomly favor or disfavor individuals, and y needs to be adjusted for different levels of seniority.

E. In comparison with “Number of citations to each of the q most-cited papers”. Advantage: overcomes many of the disadvantages of the criteria above. Disadvantage: It is not a single number, making it more difficult to obtain and compare. Also, q is arbitrary and will randomly favor and disfavor individuals.

Instead, the proposed h index measures the broad impact of an individual’s work, avoids all of the disadvantages of the criteria listed above, usually can be found very easily by ordering papers by ‘‘times cited’’.

3.4 Calculation and properties of the h index.

The h-index can be determined manually with quote databases or by automatic tools. Subscription-based databases such as Scopus and Web of Science provide automated computer. Several programs calculate the h-index on Google Scholar entries.

Each database is likely to produce another value for the same publications, because of the different cover and the different records per publication. In addition, certain databases, such as the Inspire Hep database can automatically calculate the h-index for researchers. In some fields, some engines may have a strong coverage of journal publications, but poor coverage of high impact conferences.

Scopus has better coverage of conferences, but poor coverage of publications prior to 1996.

Google Scholar has the best coverage of conferences and journals, most (but not all), but like Scopus limited coverage of pre-1990 publications. The exclusion of conference proceedings papers, is a particular problem for researchers in computer science. In the conference proceedings, will be an important part of the literature. It has been suggested that in order to meet measured for an academic about the possible citation databases with the sometimes large variation, one should assume that false negative in the databases are more problematic than false positives and take the maximum hours for a measured academically.

The proposed h index measures the broad impact of an researcher work, rejects all of the disadvantages of the criteria listed in previous paragraph, and usually can be calculated very easily by ordering papers with key the times cited in a Web database and gives a ballpark estimate of the total number of citations.

Therefore, one can argue that two researchers with similar h values are comparable in terms of their total scientific impact, even if their total number (sum) of papers or their total number of citations is not the same and varies. Conversely, comparing two researchers (in condition they are of the same scientific age; this is a limitation of the h index) with a comparable number of total papers or of total citation number and different h values, the one with the higher h value is likely to be the more significant scientist.

For a given researcher, one may expects that h value should increase almost linearly with time. In the simplest possible model, let assume that the researcher publishes p papers per year and that each appearance paper gets c new citations per year every next year. The total number of citations after n _ 1 years is then given by the equation:

$$N_{c,tot} = \sum_{j=1}^{n} pcj = \frac{pcn(n+1)}{2}.$$

It makes sense to assume that all papers up to year y contribute to the index h, we have:

https://doi.org/10.26219/heal.aueb.6022

$$(n - y)c = h$$

$$py = h.$$

Then, the left side of the first equation is the number of citations to the most recent of the papers contributing to h; the left side of second equation is the total number of papers contributing to h. Consequently, we are resulting:

$$h = \frac{c}{1 + c/p} n.$$

Then, the total number of citations is approximately:

$$N_{c,tot} \sim \frac{(1 + c/p)^2}{2c/p} h^2$$

This result is an approximation under the assumption that n remains small. This is a limitation of the h index, as it will be discussed in the next paragraph.

Comparing this form with the previous showed:

$$N_{c,tot} = ah^2 .$$

We get that the constant a is the quantity:

$$\frac{(1 + c/p)^2}{2c/p}$$

The constant a is related to the number of papers and the number of citations per paper earned per year. Hirsch (2005) empirivaly found that:

$$a \approx 3\text{–}5$$

The linear association:

$$h \sim mn$$

Occurs for the cases that researchers who produce papers of similar quality at a steady rate over the course of their careers; of course, m will vary widely among different researchers.

The linear model defined above corresponds to the distribution:

$$N_c(y) = N_0 - \left(\frac{N_0}{h} - 1\right)y$$

Where Ncy is the number of of citations to the y^{th} paper (ordered from most cited to least cited) and N_0 is the number of citations of the most highly cited paper. The total number of papers y_m is given by $N_c(y_m) = 0$; hence:

$$y_m = \frac{N_0 h}{N_0 - h}.$$

In a more realistic model, $N_c(y)$ be a non linear function of y. it is denoted that a = 2 can safely be hypothesized to be a lower bound in generally, because a smaller value of a would require the second derivative $\theta^2 N_C/\theta y^2$ to be negative over large regions of y, which is not realistic. The total number of citations is given by the area under the $N_c(y)$ curve that passes through the point $N_c(h) = h$. In the linear model, the lowest a = 2 corresponds to the line of slope _1, as shown in Fig. 2.

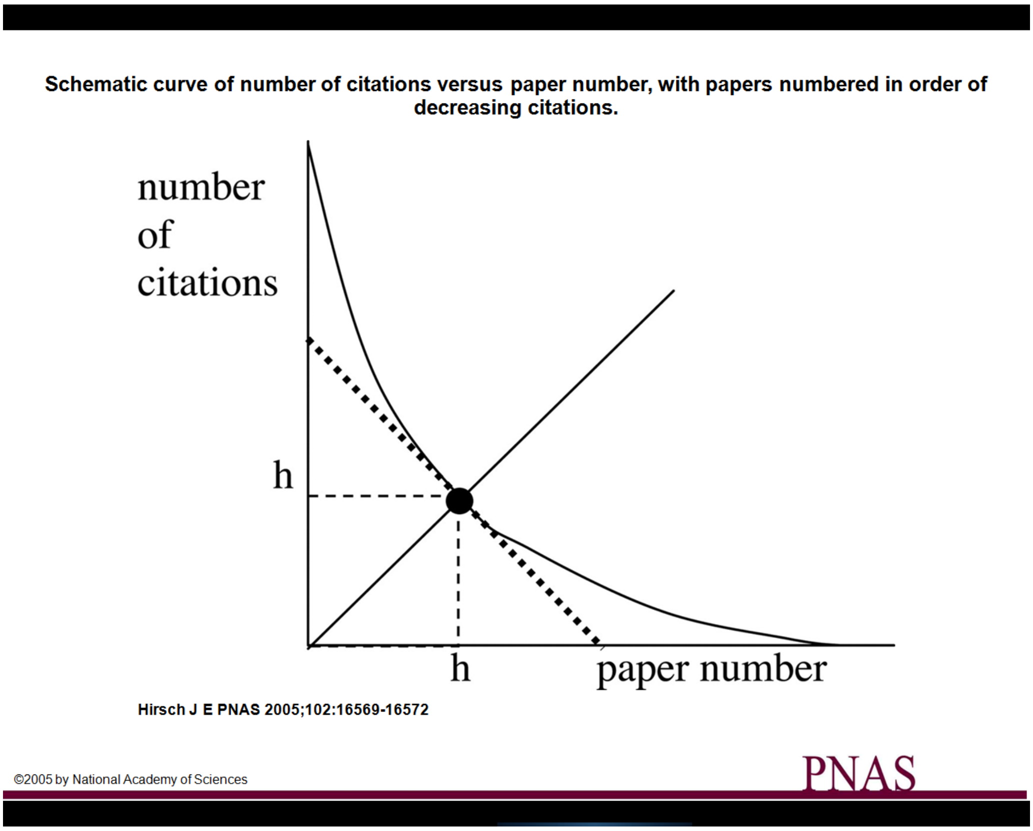

Figure 2: Schematic curve of number of citations versus paper number, with papers numbered in order of decreasing citations. The intersection of the 45° line with the curve gives h. The total number of citations is the area under the curve. Assuming the second derivative is nonnegative everywhere, the minimum area is given by the distribution indicated by the dotted line, yielding a = 2 in Eq. 1.

A more realistic model would be a stretched exponential such as the model:

$$N_c(y) = N_0 e^{-\left(\frac{y}{y_0}\right)^{\beta}}.$$

https://doi.org/10.26219/heal.aueb.6022

where β<=1,N''(y)>0,for all y.

Then, the distribution of N can be written in terms of h and a, in the form:

$$N_c(y) = \frac{a}{\alpha I(\beta)} h e^{-\left(\frac{y}{h\alpha}\right)^{\beta}}$$

the quantity I(β) is an integral:

$$I(\beta) = \int_0^\infty dz e^{-z^\beta}$$

and the parameter a can be found via the equation:

$$\alpha e^{\alpha^{-\beta}} = \frac{a}{I(\beta)}.$$

Then, the most cited article will have N_0 citations:

$$N_0 = \frac{a}{\alpha I(\beta)} h,$$

Consequently, the total number of papers of an individual, with at least one time cited, can be founded setting N(ym)=1, and it comes as:

$$y_m = h[1 + \alpha^\beta \ln(h)]^{1/\beta}.$$

The linear association between h and n will of course collapse when the

https://doi.org/10.26219/heal.aueb.6022

scientist slows down in paper publishing. There is a time lag between the appearances of the two events. In the linear model, if the researcher stops

$$n_{lag} = \frac{h}{c} = \frac{1}{1 + c/p} n_{\text{stop}}$$

publishing after n stop years, h continues to increase at the same rate for a time:

After of such period of increment, stays constant because all papers contributing to h.

3.5 Some points of criticism in h index

There are a large number of cases in which h may provide misleading information about the output of a scientific researcher (However, most of the cases are not exclusive to the h-index.). Below, some of those cases are listed:

The h-index does not take into account the number of authors of a paper. In the original article (as already showed), Hirsch (2005) suggested citations separation between the co-authors. Even in the absence of explicit game, the h index and similar indices tend to merge fields with larger groups, experimentally and theoretically in example.

The h-index does not reflect the typical number of citations in different scientific fields. Different scientific areas, or publishing means, traditionally use different numbers of citations.

The h-index deletes the information, which contained in the author's placement in the list of authors, which in some scientific fields is important.

The total number of publications limits the h-index. This means that scientists with a short career (in example when they begin it) are at an inherent disadvantage, regardless of the possible importance of their discoveries. For example, Albert Einstein had died after the publication of his Annus Mirabilis revolutionary four documents in 1905, his h-index is stuck at 4 or 5 maximum. It is also a problem for any measure that is based on the number of publications (n). However, as Hirsch (2005) said in the original document, “the index is designed as a tool to evaluate researchers in the same stage of their careers”. It is not intended as a tool for historical comparisons.

The h-index does not take into account the context of citations. For example, citations in a document are often simply to flesh out an introduction that was also no other meaning to the work. h also does not resolve other cases of context. It is also a problem for regular citation counts.

The h-index assigning the same number of books as articles, making the comparison of research in areas that are not book oriented, like the humanities, difficult.

The h-index does not take into account confounding factors such as "free paternity," the so-called “Matthew effect”[1] (Merton, 1968)and the favorable citation bias associated with review articles. Again, this is a problem for all other parameters using publications or citations.

The h-index was found to have predictive accuracy and slightly less accuracy than the simple measure of average citations per paper.

The h-index is a natural number, which reduces its power of discrimination.

The h-index can be affected by self-citations, and if it is based on Google Scholar, output, even computer-generated documents can be used for this purpose.

---

[1] In the so called field, sociology of science, "Matthew effect" was a term founded by Robert K. Merton to refer how, among other things, eminent scientists will often get more measurable credit than a corresponding unknown researcher, even if their work is similar. In addition, it means that credit will usually be given to researchers who are already famous. In example, a prize is possible to be awarded to the most senior researcher involved in a project, even if all the work was done by a graduate student.

https://doi.org/10.26219/heal.aueb.6022

Chapter 4: Disadvantages and Variants of h-index

4.1 Introduction

This chapter is a conceptual continuation of the previous chapter. In the previous chapter the original construction and the properties of the h-index, was presented. Continuing, the current chapter attempts to present the main disadvantages of h-index. Also attempts to address the disadvantages of the indicator as e trigger resulting the creation of many other indices, members of h-index family. These indices are derivatives or variants of h index..

Also, in this chapter, are being presenting topics related with some critical issues, such as the self citation issue, the time growth issue, the h index of a journal, etc.

4.2 Disadvantages of the h-index

The grade popularity and the extended use of the h-index has raised a lot of criticism in several concerns. As it has already mentioned in the previous chapter, an enough notable and right documented of critical theorizes on the h-index is a report by the joint Committee on Quantitative Assessment of Research (Adler, 2008). According to this report, the authors argued detailed versus the use (or misuse) of citation metrics alone as a tool for assessing quality of research, and suggested the use of more complex methods for judging researchers, journals or disciplines, that should combine citation metrics and other criteria such as memberships on editorial boards, awards, invitations or peer reviews

Together with the above-referred work, there are many more publications

https://doi.org/10.26219/heal.aueb.6022

referring to disadvantages or some type of criticism on the h-index.

The h-index is ,as mentioned in the previous chapter, bounded by the total number of publications. That means that researchers with a career in the beginning are at an obvious disadvantage, although of the possible importance of their discoveries. The h-index puts newcomers at a disadvantage position since both publication output and citation rates will be relatively low for them. Some authors have also discussed that the h-index is negative affected by self-citations (Hirsch J. E., 2007). Many self-citations would lead a false impression that the researchers' publications are widely accepted by the scientific community. In these terms, the emerging problem is that researchers with many co-operating partners may receive many self-citations, in contrast to researchers that make publications alone. The h-index, under these conditions, has slightly less predictive accuracy and precision than the simpler index of mean citations per publication (Schreiber, 2007). Another problem is that the h-index rank small (numbered) but highly-cited scientific outputs in a disadvantage. While the h-index non-giving emphasis at singular successful publications in favor of stable productivity, it may do this too strongly (Anderson, 2008). Two researchers may have the same h-index, say, h = 40, i.e., they both have 40 publications with at least 40 citations each. However, one may have 30 of these publications that have been cited more than 1000 times and the other may have all of his/hers h type publications receiving just above 40 citations each. It is evident that the scientific work of the former scientist is more influential. Many methods to correct for this have been introduced, but none has gained universal support.

The existence of limitations and differences of the citation data bases may also affect the h-index. An automated searching machine finds citations to publications going back many years, while others find only recent publications or citations. Another database related problem often happens with a serious effect on the correct computation of the h-index, is that of name similarities between researchers. It is almost impossible to find a scientist with a unique combination of family name and initials while searching the most known citation databases (Kosmulski (2006) in Panaretos et al (2012)). Resulting of this, in many circumstances the h-index will be strongly overestimated, since in its

https://doi.org/10.26219/heal.aueb.6022

computation the works of more than one researcher are summarized.

It seems that the h-index may-not be used for comparing researchers working in different scientific fields. It has been showed that average citation numbers differ widely among different scientific areas (Hirsch J. E., 2005).

In addition, several types of problems associated with any bibliometric index, based on the necessity to account the scientific impact by a single number, apply here as well. Since the h-index is just one ‘index’ of scientific productivity, some object to the practice of taking a human activities as complex activities as the formal acquisition of knowledge and condense it to a single number (Kelly & Jennions, (2006) in Panaretos et al (2012)). Two crucial risks of this fact have been noted: The first is that career progression and other aspects related to human’s life may be damaged by the use of a simple index in a decision-making procedure by someone who has neither the time nor the intelligence to consider more appropriate decision indexes. The second risk is that researchers may reflect to this by attending to maximize their h-index to the detriment of embedding more quality in their work. The effect of using simple indices for making decisions has often been observed to be an non-intended consequence of index-based decision making; for example, governments routinely aply policies designed to minimize crime indexes and not crime itself (Panaretos & Malesios, 2009).

4.3 Some variants/modifications of the h-index

Very soon after the introduction of the h-index by Hirsch (2005), various criticisms, modifications and generalizations on h index, have appeared in the literature. Most of those publications are focusing to the non efficient points or insufficiencies and discussing techniques to correct the insufficiencies of the h-index. During this dissertation, many of these articles have already described in the previous sections. There are decades of several types of indices. Some of them are close to h-theme some other may not.

In the following paragraphs, will be an attempting to define and review

https://doi.org/10.26219/heal.aueb.6022

some of the modifications of the h-index.

4.3.1 The g-index

The h-index is a robust index in terms of insensitive to an accidental set of unfitted (or lowly cited) publications and to one, or several, high-cited publications (let say extreme values) (Panaretos & Malesios, Assessing scientific research performance and impact with single indices, 2009) and (Ruscio, Seaman, Orriano, Stremlo, & Mahalchik, 2012). This topic, the strength of g-index in extreme values, will be covered intensively at the last chapter of ths dissertation. However, g-index is not sensitive to the level of the highly cited articles. Indeed, suppose a researcher having an h-index of value let say 10. Then for an article belonging to the h-core[2] of this scientist, it is not important whether this paper has 10 or 100, or even 10000 citations (supposing it still remains in the core). Therefore, there is a type of tolerance loss. In order to overcome this loss, Egghe(2006a,b,c) introduced the g-index.

Definition of the g-index: The g-index is the highest number g of publications that together received at least g^2 citations. It is clear that:

$$g-index \geq h-index$$

By its definition, this g-index is increased with a skewed frequency distribution of the citations. That is the higher the number of the citations in the top range, the higher the g-index Egghe (2006b). The g-index clearly overcomes the h-index's insufficiency in terms that it is sensitive in the internal changes of the Hirsch core. Furthermore, it requires drawing a longer list than the necessary one for calculating the h-index, hence it increases the precision problem.

[2] H-core is the square matrix nxn, where n is the h index of the target scientist.

https://doi.org/10.26219/heal.aueb.6022

4.3.2 The A-index

The A index introduced by Jin's(2006). The A-index shared the same goal as the g-index, namely correcting for the fact that the original h-index does not take into account the exact number of citations when they belonging in the h-core.

Definition: A-index defined as the average number of citations received by the publications included in the Hirsch core.

The definition of A-index given in formula is as follows:

$$A - index \equiv \frac{1}{h} \sum_{j=1}^{h} citation_j$$

In the above formula, the numbers of citations must firstly be ranked in decreasing order. The A-index uses the same data as the h-index. It is obvious that always:

$$h - index \leq A - index$$

A weak point of A-index is that suffers from the problem of punishing among researchers with the same A-index the one with higher h-index, since the A-index comes by a division of the shown integral by h.

https://doi.org/10.26219/heal.aueb.6022

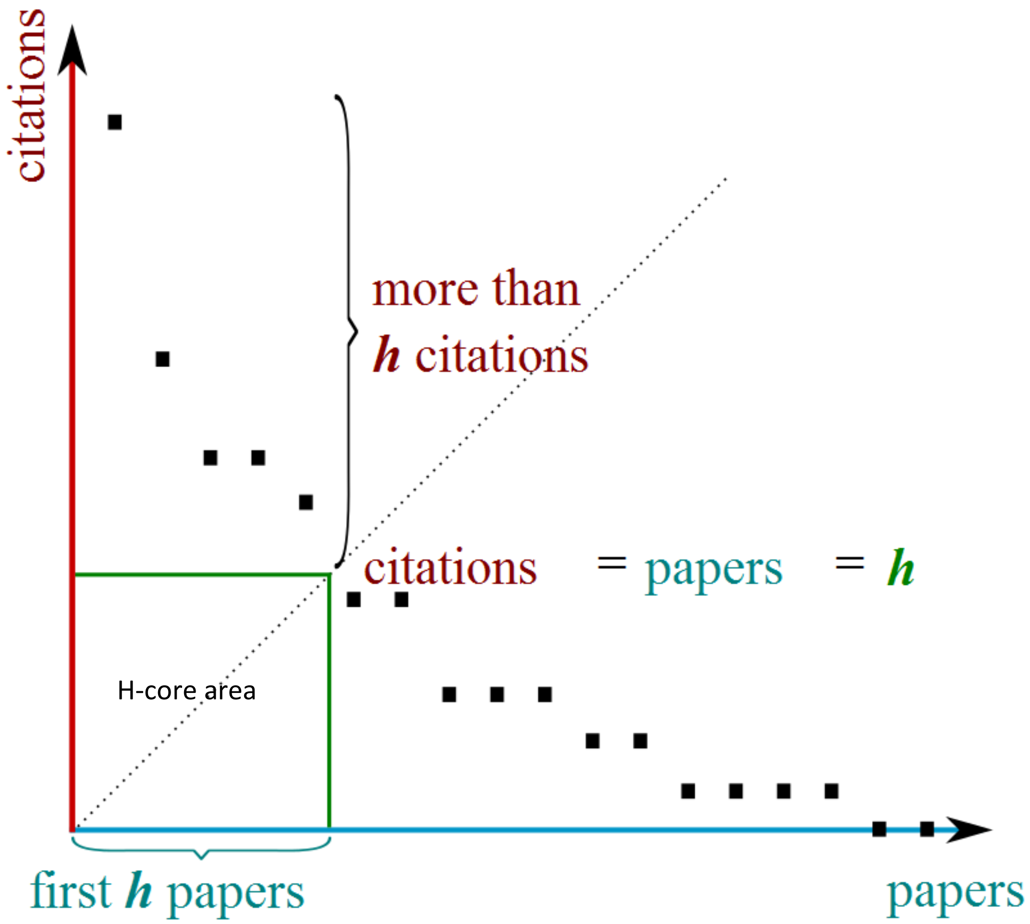

Figure 3: A typical figure of the h-core area.

4.3.3 The R-index

The R-index introduced by Jin & Al. [2007], which according to the authors arguments, eliminates some of the crucial disadvantages of the h-index. The R-index is a kind of improvement of the A-index. Specifically, the R-index tries to eliminate the main disadvantage of the A-index (that of punishing among researchers with the same A-index the one with higher h-index, since the A-index comes by a division of the shown integral by h), by calculating the square root of the summary of the h citations bounded in the Hirsch core.

The corresponding formula is the next one:

https://doi.org/10.26219/heal.aueb.6022

$$R-index = \sqrt{\sum_{j=1}^{h} citation_j}$$

It is clear that occurs:

$$R = \sqrt{h * \frac{1}{h} * \sum_{j=1}^{h} citation_j} \Rightarrow$$
$$R = \sqrt{h * A\_index}$$

And therefore occurs:

$$h\_index \leq R\_index$$

4.3.4 The hw-index

The hw index introduced by Egghe & Rousseau [2007]. The hw index aims to be more sensitive to variations in the h-core. The authors defined the hw-index in a discrete version and a continuous version, and gave a series of theoretical properties of the hw-index in both versions. In order to construct the hw-index in the discrete case (obviously the most practical of the two versions) one has to compute the weighted ranks as it follows:

$$r_W(j) = \sum_{i=1}^{h} \frac{C_j}{h}$$

Where Cj denotes the number of citations gained by the jth publication, and h is the h index. Then, someone has to substitute with this weighted ranking the journal ranking according to the citations received used for the computation of the h-index. The new index wich produced is given by:

$$h_w = \sqrt{\sum_{i=1}^{r_0} C_i}$$

Where r_0 is the largest r_w-value such that:

$$r_w(j) \leq C_j$$

By applying this weighted sum of ranks to the citations, Egghe & Rousseau (2007) introduced a measure that accounting the overall number of h-core citations as well as the distribution of the citations in the h-core.

4.3.5 The h(2) index

The $h^{(2)}$ index introduced by Kosmulski (2006), in an try to avoid and overcome the problems coming of name similarities between researchers. Those type of problems reduces the precision in the computation of h-index.

The $h^{(2)}$-index, defined as follows: A researcher has $h^{(2)}$-index, let say k, if k of his Np publications have at least k^2 citation, and the other (Np–h) publications have at least h^2 citations. Obviously:

$$h^{(2)} \leq h \quad \text{for any scientist.}$$

According to authors (Kosmulski 2006), $h^{(2)}$ is strongly correlated with the total number of citations received by a researcher. The $h^{(2)}$ index lacks in tolerance ability when used to assign ranks to different classes of books and the loans on them, when compared to the other two Hirsch-type indices, the h-index and g-index.

4.3.6 The w-index

The w-index introduced by Wu (2008). Definition of the w-index: A researcher has a value of w-index if w of his publications have at least 10*w citations each, and the remaining publications have fewer than 10*(w+1) citations. According to the author, the w-index seems to be very similar to the h-index, so it better behaved to the influence of a scientist’s top publications. Author examines the accuracy and properties of the w-index through an

https://doi.org/10.26219/heal.aueb.6022

empirical analysis (analogous to Ruscio et al (2012)) using bibliometric data. By practical implementations, the author found that it is:

$$w \cong \frac{h}{4}$$

4.3.7 The i×ci index (maxprod index)

In 2007, Kosmulski (2007) introduced the context of h-type indices, the maxprod index. Maxprod index defined as: “the highest number among values i×ci, where i denotes the ith publication and ci is the number of citations gained by the ith article”.

Maxprod is related to h-index, as to the next argument:

$$\max prod \geq h * c_h \geq h^2$$

According to the introducer (Kosmulski (2007)), the maxprod index has an advantage over the h-index, since it can be used as a selective tool for identifying researchers of outstanding achievements filtered them from the vast majority of researchers of “typical” scientific behavior.

4.3.8 The t and f indices

Two new modifications of the h-theme introduced in 2007 by Tol (2007), in an effort to remove some of h disadvantages. The two variants of h, are close to the g-index, although they are based on harmonic and geometric averages instead of conventional averages.

The f-index is computed by solving with respect to f the inequality:

$$\max_f \frac{1}{\frac{1}{f}\sum_{i=1}^{f}\frac{1}{c_i}} \geq f$$

 https://doi.org/10.26219/heal.aueb.6022

Where ci denotes the number of citations gained by publication i (i=1,2,…,n). Similarly, the t-index is computed by solving the following expression:

$$\max_t \exp\left[\frac{1}{t}\sum_{i=1}^{t}\ln(c_i)\right] \geq t$$

It is always:

$$h \leq f \leq t \leq g$$

The f- and t-indices have more tolerance power in comparison to h- and g-indices.

A real-data application, obtained from bibliographic databases, of these four indices (Kosmulski, 2007) showed the existence of strong correlations between them; however, the application of the new indices did not change enough the ranking of the researchers (Panaretos & Malesios, 2009).

4.3.9 Contemporary h index and trend h index

The contemporary h-index and some variants as the trend h-index, normalized h-index. (Sidiropoulos et al. 2006) treat some of the disadvantages of the h-index, and propose a series of generalizations of the specific index.

Sidiropoulos et al (2006) introduce two generalizations of the h-index, the contemporary h-index and the trend h-index, which are modifications of the h-index in terms to reveal the genius young researchers and trendsetters, respectively.

Furthermore, they define a normalized h-index, which attends to give advantage to the researchers with few but largely cited publications. As mentioned already, the h-index does not take into account the age of the publication. This may be the case that a scientist has published in the past a

https://doi.org/10.26219/heal.aueb.6022

number of significant publications that resulted a large h-index value, but now is rather inactive. In contrary, another scientist may still producing significant publications. To detect these differences in terms of time , Sidiropoulos & al. (2006) define the contemporary h-index hc as follows:

A researcher has contemporary h-index, hc , if hc of its Np publications get a score of S^c(i)>hc each, and the rest (Np-hc) publications get a score of S^c(i)<hc, where S^c(i) is defined as:

$$S^c(\mathrm{i}) = \gamma \times (\mathrm{Y(now)} - \mathrm{Y(i+1)})^{-\delta} \times C(i)$$

Y(i) denoting the publication year of article i, and C(i) the publications citing the i-th article. If δ=1, Sc(i) is the number of citations the publication i has received divided by the “age” of the publication. The coefficient γ is used to correct the small value of the derived index, and is suggested to take the value of 4.

4.3.10 The AR-index

The AR-index introduced by Jin et al (2007) is an index which is depended from the age of publication, built in order to overcome the problem that the h-index always increases its value even in the case where a scientist stops to publish new work or even he dies. Assuming that aj is the age of the publication j, then the A index defined as:

$$AR = \sqrt{\sum_{j=1}^{h} \frac{citation_j}{a_j}}$$

The main advantage of the AR-index is that it includes in its computation the age of the publications, thus decreasing the contribution of the old publications as they becoming older and older.

In this way the h-index is complemented by an index that can actually decrease as time pass. The AR-index is based on the h-index as it uses of the h-core, as it seems from its formula.

https://doi.org/10.26219/heal.aueb.6022

4.3.11 Correcting the h-index for different fields of research

As Hirsch (2005) denoted, researchers in humanities tend to achieve much higher h-values when compared to researchers in physics or mathematicss etc. Therefore, prior to comparisons of the h-index, in such scientific areas, some kind of purification of the h-indices from that effect is required. Following this direction, Iglesias & Pecharromán [2007a,b] introduced a variant of the h-index in order to take account for the different scientific areas of researchers, introducing a stochastic model for the distribution of citations . This which leads to the following expression for the theoretical h-index:

$$h = \sqrt[3]{\frac{Np}{4}\chi^{(2/3)}}$$

Where Np denotes the total number of publications published and χ is the average number of citations per publication for the researcher.

4.3.12 Other variants of the h index

In the previous paragraphs described a small number of different types and variants of scientific impact indices. As it is already understood the majority of them concerns variations on h index, and many of them uses the h-core, the basic idea of h index.

There are many others, totally may numbered in some decades, variants of such indices (Panaretos & Malesios, 2009), (Anderson, 2008), (Alonso, 2009), etc.

Trying to classify in categories the whole population of such variations, ones may find this classification:

https://doi.org/10.26219/heal.aueb.6022

- Variants seeking to take account the age of publications
- Variants seeking to adjust for self citations
- Co-authorship adjustable variations
- h-variants for entities such journals or institutions
- variants seeking to attach tolerance for the area into the h-core
- variants to take account of the several scientific areas

https://doi.org/10.26219/heal.aueb.6022

Chapter 5: Latest Trends on h-index

5.1 Introduction

This chapter seeks to introduce the latest contributions to the examined field. The discussion about the disadvantages and the problems occurring while using the h index and its derivatives still continues. The whole discussion mainly represented and concentrated at the contribution of two papers: the first is the evolutionary work of Ruscio et all (2012) and the second one is the commentary on this work of Panaretos et al (2013). Ruscio et al (2012) evaluated 22 metrics, including conventional measures, the h index, and many variations on the h theme. The criteria they used for this evaluation criteria encompass conceptual, empirical, and practical issues such as: ease of understanding, accuracy of calculation, effects on incentives, influence of extreme scores, and validity. Based on this criteria, they proceed to the evaluation work of several bibliometric indices.

5.2 Criteria that can be used in order to asses and evaluate the behavior of an index

In terms of research, it may be fewer feasible to gain peer reviews or to collect data on other qualitative indicators and incorporate it into a data analyses. Because the value of well-chosen metrics can be substantial in either context, it is worth to assess the merits of quantitative measures using specific evaluation criteria. This will enhance to solve serious empirical and practical issues. Ruscio et all used five main criteria in order to examine the indices. More over two additional criteria were examined: Arithmetic stability and reliability. Table 1 shows the indices were compares, and the relative definition of each index.

Measure	*Definition*
N_a	number of articles
C	$\sum_{i=1}^{N_a} c_i$, where c_i is the citation count for article i
sqrt-C	$\sqrt{C}$
M_c	C/N_a
Mdn_c	if N_a is odd: c_j, where $j = (1 + N_a)/2$; if N_a is even: $(c_j + c_{j+1})/2$, where $j = N_a/2$
h	largest value of i such that $c_i \geq i$
h_t	$\sum_{i=1}^{N_1}\sum_{j=1}^{c_i} f_{ij}$, where N_1 is the number of articles cited at least once and $f_{ij} = 1/(2 \times \max(i, j) - 1)$
f	largest value of j such that $\sqrt[j]{\prod_{i=1}^{j} c_i} \geq j$
t	largest value of j such that $\left(1 \Big/ \sum_{i=1}^{j} \frac{1}{c_i}\right) \geq j$
g	largest value of j such that $\left(\sum_{i=1}^{j} c_i \Big/ j\right) \geq j$
hg	$\sqrt{h \times g}$
a	$\sum_{i=1}^{h} c_i \Big/ h$
m_i	if h is odd: c_j, where $j = (1 + h)/2$ if h is even: $(c_j + c_{j+1}) / 2$, where $j = h/2$
r	$\sqrt{\sum_{i=1}^{h} c_i}$
h_w	$\sqrt{\sum_{i=1}^{r_0} c_i}$, where r_0 = largest value of j such that $\left(\sum_{i=1}^{j} c_i \Big/ h\right) \leq c_j$
q^2	$\sqrt{h \times m_i}$
e	$\sqrt{\left(\sum_{i=1}^{h} c_i\right) - h^2}$
mp	$\max(i \times c_i)$ for i from 1 to N_a
sqrt-mp	$\sqrt{mp}$
$h^{(2)}$	largest value of i such that $c_i \geq i^2$
m_q	h/A, where A is publishing age (current year − year first article was published)
m_{qt}	h_t/A

Figure 4: Publication Impact Measures used in a comparable study of Ruschio et all (2012). Index I refers to a specific article. In addition, the referred I publications hypothesized to be ordered by citation from highest to lowest.

https://doi.org/10.26219/heal.aueb.6022

5.2.1 Criterion 1. Ease of understanding.

It makes sense that simpler measures should be more preferrable than those that are more complex. A well-chosen and understandable measure should be easy to explain and, if researchers with little or no training in measurement or other quantitative methods can use it. In addition, the ease of understanding has an extra meaning: measure must be meaningful.

5.2.2 Criterion 2. Accuracy of calculation.

It is assumed that measures that are more easy to calculate accurately should be more preferable than those that are more error generating. Tthe complexity of a calculation is of little concern in research settings, in practice it may be desirable to an individual to use metrics that can be calculated without any special-purpose algorithms or software. In addition, more important than the complexity of a measure is the availability of the data that one must retrieve to calculate a index. Indexes that require less information, or they are more robust to missing data, should be more preferable.

5.2.3 Criterion 3. Effects on incentives.

Ruscio et al(2012) recalling Haslam & Laham, (2010), and introducing a qualitative criterion: How an individual who measures a publication impact can reinforce the perception that a selected research strategy focused on the quantity, quality, or balance between quantity and quality of publications will be rewarded. The users of the indexes should carefully select the apropriate metrics that are consistent with the desired incentives, in order to avoid influencing researchers' strategies in non-desired ways.

5.2.4 *Criterion 4. Influence of extreme scores.*

The distributions of citations seem to be heavy-tailed. Clauset, Shalizi, and Newman (2009) found that citation numbers can be modeled good, using a power law distribution, Even if a lot of extreme scores (of Ruscio et al data) correspond to genuinely influential publications that merit a strong influence on measures of publication impact, others may stem from well fortune, unique insights leading to unrepresentative collaborative efforts (e.g., when a post-graduate student becomes coauthor on a faculty mentor's research).

Individual user of indexes must decide whether they prefer metrics that are robust or sensitive to the influence of outliers (extreme scores). Let say, for example, if one's aim to measure sustained or programmatic impact, robustness to extreme scores should be preferred.

5.2.5 *Criterion 5. Validity.*

It is common sense, that a more valid measure is preferable to one that has less validity. Ruscio et al place special emphasis on validity when discussing the selection of metrics for applications.

5.3 Statistical methods and data

The statistical methods used at Ruscio et al (2012) varying across several methods. Initially they collected data making two different samples that contained values of several indices for specific researchers.

A large sample of data representative of professors working at U. S. universities obtained. The score assigned to each department is recorded by the United States National Research Council (NRC), which ranged from 29 to 72 in an approximately normal distribution. For each faculty member, made a record

https://doi.org/10.26219/heal.aueb.6022

of his or her rank. The exact sample sise was: n = 450 assistant professors, n = 471 associate professors and n = 829 full professors. From the whole procedure were excluded the adjunct, affiliated, and emeritus faculty; in case of distinguished professors Ruscio et all assumed to have achieved the same rank as a full professor.

The researcher's names were entered into advanced searches on PsycINFO Engine. The results obtained, restricted to peer-reviewed journal publications. Individuals were excluded from the account if their names did not appear in PsycINFO search engine or if it was impossible to be separated from authors with same or nearly identical names. Several more measures were taken in order to purify the citations.

The number of citations and the age of publication were recorded for 1,750 individuals who published 48,692 peer-reviewed journal publications that were cited 919,883 times.

Methods such as repeated sampling, extreme values handling, split-half reliability, bootstrapping, several types of correlations, reliability analysis, were used.

5.4 Evaluating Indices

Bellow will be representing the main methods and results of examining the indices into the specific criteria.

5.4.1 Ease of understanding

The need of ease of understanding of a measure is obvious. Such a measure must be as easily understandable can be. Simultaneously, this measure must to include as much information it can be. In this terms, possibly the simplest index to understand is the total number of publications (Na). In the same category of understandable and information carrying is the total number of

 https://doi.org/10.26219/heal.aueb.6022

citations (C), followed closely (in terms of those two criteria were given) by the mean number of citations per publication (Mc). As an alternative of the Mc, the median number of citations per article (Mdnc), can be used, especially under the case of non symmetric distributions of C. In such list of ease of use and understanding, the square root of total citations (sqrt-C), must follows.

The large popularity of the h index shows how effectively its simple definition lends itself to arithmetical and graphical representations. The other variations on the h index introduce ideas and elements that are somewhat to considerably more complex.

5.4.2 Accuracy of calculation

Two main criteria were cited in order to approximate the accuracy of calculation:

a) the complexity of a measure's algorithm and

b) its robustness to missing information (e.g., publications that are not retrieved due to data entry errors in the database).

Under those two criteria the extraction of the total number of published publications (Na) is easy to calculate. The citation databases search results will indicate the number of publications identified. The same accuracy and ace of calculating occurs with summing citation counts to yield the C index. Examining the sqrt-C index, ones find that is simply the square root of C, Mc is just C divided by Na, all must assumed simply calculated. The calculation of the median Mdnc requires a bit more complex algorithm: firstly ranking the citation counts and locating the middle value (or, for an even Na, calculating the mean of the two middle values).

In conclusion, these conventional measures are all easy to calculate. The main disadvantage (in terms of accuracy) is that they are at least somewhat sensitive to missing information. Total number of publication (Na) is the most

https://doi.org/10.26219/heal.aueb.6022

straightforwardly affected by missing publications. Total number of citations C and sqrt-C will be affected to varying degrees depending on the frequency with which the missing publications were cited and the citation intense of the available publications. The average-citation indexes, of the available Mc and Mdnc can worse the influence of missing data, especially for individuals with few number of published publications. Comparing the modern citation-based indices (except the conventional), h is arguably the easiest to obtain because it requires little more than ranking the citation counts. It simply needs to, among the ranked values, locate the last one publication that equals or exceeds its position in the list. Since $h \leq Na$, h is less affected by missing data. For example (Ruscio, Seaman, Orriano, Stremlo, & Mahalchik, 2012) , Professor X published 18 publications, with h = 5. Whereas Na would be reduced by 1 unit for each missing article, h would not be reduced at all unless there was so much missing data that there were no longer at least 5 publications cited at least five times apiece. This could not happen with only 1 or 2 missing publications, because there were originally 7 cited at least five times. There is only a 4% chance that Professor X's h index would be affected with 3 missing publications, and the chance of this happening surpasses 50% only when 7 (of the total of 18) publications are missing. This can be interpreted as an impressive robustness to randomly missing (not collected) data. Consequently, all of the modern citation-based indices (derivatives of h, or not) should be relatively less weak to missing data.

The mq index is comes from h divided by publishing age. That means it's too easy to calculate and inherit its robustness to missing data. The $h^{(2)}$ measure is being calculating in much the same manner as h, but in addition with a much more stringent added criterion at each serial position among the citation times counts. This addition can be hypothesized as a minor increase in the calculation complexity, and its stringency makes it even more robust to missing data.

Many other measures require several (iterative or no) series of computations. Calculating f, t, or g indices is similar to counting one's way to h, but rather than assessing only to the citation number at serial position i, one calculates an average (in terms of harmonic, geometric, or arithmetic mean,

 https://doi.org/10.26219/heal.aueb.6022

respectively) of the i mostly cited publications. Then, asks whether this average value of publications citations, equals or exceeds i. This has a burden to do by hand. The hg index is the geometric mean of h and g, and thus it requires calculating both indices h and g.

There are four measures, which follow a different kind of two-step procedure. First, h is produced to determine how many publications are in the "Hirsch core," and then the a, mi, and r indexes are calculated as the mean, median, and square root, respectively, of the sum of these h citation numbers. The e index is calculating as described before. The q2 index is the geometric average of h and mi, so it requires calculating both h and mi. The hw index needs a two-step algorithm, that is very similar to that used to calculate the r index, but more burden needing at the first step of its calculation.

The remaining indices involve computations that can be cumbersome with even a modest number of citations. The mp index needs the asumption that one rank-order citation counts; and then multiplies these citation counts by their serial positions (rank). The mp index equals the largest of these products, and the sqrt-mp comes simply as the square root of mp. The ht index requires summing fractions as shown earlier. It is clear now that all the modern citation-based indexes must be robust to missing data, providing more accurate results, despite imperfections in data sources and retrieval algorithms.

5.4.3 Effects on incentives

Using Na (number of published publications) as a metric can exert a strong affect on the incentives that shape researchers' activities by giving rewards only in terms of the quantity of research and not in terms of its quality. Because a large influential publication receives no more credit than any other does, researchers might be tempted to publish as much publications, or as quickly as it is possible, even if the results of their work is not more important or making publications less carefully. This problem can be solved introducing the

https://doi.org/10.26219/heal.aueb.6022

use of C or sqrt-C as a metric, but it might lead to others problems. While no existing explicit reward for the quantity of producing publications, the completion of the publication in a timely manner, seems less urgent. Therefore, the effects could be beneficial; for example, investigators may address more important problems even if they require greater time and effort.

The use of Mc or Mdnc indices must attain some kind of balance between rewarding for quantity and quality of research. But there remains the problem that this could affect the reduction of the incentive to produce new publications among those who have gained a good score on an average-based metric. Any article of an individual with low citation probability may pull down the average. Therefore, for researchers who have attained a high citation average score, publishing anything new becomes a type of gamble: The higher one's current average score, the less likely it is that a new article will be cited enough times to maintain or exceed the existing average. In economic terms, that called that the marginal product comes smaller. In contrast to conventional indices, the h index establishes and maintains stimulus to take position in a balance between the quantity of published publications and the included quality of them. In contrary, Focusing for quantity only, is not likely to result a large h index, as it is obvious that a high volume of infrequently cited publications corresponds to a long, thin array of cited publications into which a large square will not fit.

In addition, focusing for quality only is not likely to gain a large h index, as a low number of highly cited publications, corresponding to a tall, narrow array of citations into which a large square table will not fit. **Thus, achieving a high score on h requires a large quantity of high-quality publications**.

As introduced in Bartolucci (2013) a decomposition of h into two main components so called "impact" and "concentration", which remark (that also noted in the previous lines) the existing need for publication to be frequently cited and to have these citations spread to a large number of discrete publications in order to gain a high h index. The h(2) index maintains an emphasis on balance between quality and quantity, similar to that of the h index, but it introduting a enough stringent criterion, as already mentioned above. Some more variations on the h family tend to focus in quality more than quantity. This includes all of the

https://doi.org/10.26219/heal.aueb.6022

indices that award credit for only those overflow citations that fall above the h square, called f , t, g, hg, a, mi, r, q2, and e. Since ht, mp, and sqrt-mp indexes provide the same amount of rewards for excess citations falling above or to the right of the h square, there is no built-in reward for quantity or quality of publications. These indices are flexible and usefulness in that they provide rewards for either research strategy.

5.4.4 Influence of extreme scores

5.4.4.1 Stability

It makes sense for someone to expect that the index, which is more robust to extreme scores, must be more stable under repeated sampling techniques from a population than an index that is more sensitive to extreme scores. Ruscio et all (2012) studied the stability of each index using two ways. According to first way, 100 bootstrap samples (Efron & Tibshirani, 1993) were drawn from the initial collected data. For each of the 1,750 researchers, the observed distribution of citation counts was handled as a population and a random sample of equal size was created with replacement (that holds the probability constant). For each bootstrap sample, all 22 indices were computed for all 1,750 researchers. Then, scores on each index correlated across all pairwise combinations of the 100 bootstrap samples were collected. The mean of these 4,950 correlations for each index is shown in the table below:

https://doi.org/10.26219/heal.aueb.6022

Index	*Mean r for 100 Bootstrap Samples*	*CV for 10,000 Randomly Sampled Scores*
N_a	N/A	N/A
C	.950	.494
sqrt-C	.963	.222
M_c	.791	.494
Mdn_c	.646	.457
h	.966	.154
h_t	.976	.138
f	.967	.142
t	.976	.141
g	.971	.188
hg	.975	.150
a	.839	.500
m_i	.700	.340
r	.951	.245
h_w	.944	.256
q^2	.929	.196
e	.928	.303
mp	.885	.697
sqrt-mp	.925	.277
$h^{(2)}$	.920	.155
m_q	.919	.154
m_{qt}	.940	.138

Figure 5: Average values of the Pearson Product Moment Correlation and the Coefficient of variation, for 100 Bootstrap samples.

A partition of 6 indices exhibited the largest correlations ($r \geq .966$): h, ht, f , t, g, and hg. Most of the remaining indices were also highly stable, with only 3 citation-based indices (a, mi, and mp), Mc, and Mdnc exhibiting $r < .90$.

The second way that Ruscio et all were checked stability of the indexes was to randomly sample citation counts for 28 publications; this number chosen

https://doi.org/10.26219/heal.aueb.6022

because it represents the mean number of publications per researcher, from the distribution of all 48,692 available publications. This resulted a set of scores on all indices, and the whole process was repeated 10,000 times. In order to calculate m_q and m_{qt}, which refer to career stage, publishing age was held constant at a value of 21 years; this age was selected because it represents the average value of the selected researchers. To examine the stability under repeated sampling, the coefficient of variation was computed for each index; these results are shown in the Table above.

Again, the same partition of citation-based indices, joined by q^2, $h^{(2)}$, m_q, and m_{qt}, exhibited the greatest stability of all. Conventional citation indices such as C,Mc, and Mdnc were considerably less stable, as were indices a and m_p.

5.4.4.2 Adding or Removing Extreme Scores

In a way, an index must appearance a kind of robustness in the existence of extreme scores. According to Ruscio et all, four type of extreme score analyses were performed. The Type I extreme score was constructed by multiplying by 2 the largest citation count. The Type II extreme score was performed by adding a preceding digit of 1 to the largest citation number (e.g., if largest count was 25, an extreme score of 125 was added). The median percent increase in each metric was computed when adding a Type I extreme score, a Type II extreme score, or one of each. The Type III extreme score analysis coming as the largest citation count was dropped out of the sample for each researcher and the median percent decrease in each metric was computed. Results for these extreme score analyses are shown in Table below (four labeled columns). Between the most strength and robust indices were Na, h, and mq, which increase as increases the number of extreme scores but not with their magnitude, and $h^{(2)}$, which does not necessarily increase at all due to its stringent criterion as this told before. The most sensitive indices to the influence of extreme scores were C, Mc, a, and mp.

https://doi.org/10.26219/heal.aueb.6022

Extreme Score Analyses: Median % Change

Type I Extreme Score Added	*Type II Extreme Score Added*	*Type I + II Extreme Scores Added*	*Largest Value Dropped*	*Split-Half Reliability*
5.3	5.3	10.5	−5.3	N/A
53.4	118.5	179.5	−27.4	.954
23.9	47.8	67.2	−14.8	.962
44.8	102.7	140.6	−22.0	.708
9.1	9.1	20.0	−8.3	.586
7.7	7.7	16.7	−8.3	.980
13.6	16.8	29.9	−11.0	.984
10.0	10.0	20.0	−9.1	.977
12.5	15.8	28.6	−11.1	.986
19.1	31.2	46.1	−14.3	.982
13.4	20.3	30.9	−11.7	.986
46.0	109.2	142.3	−22.5	.775
13.6	13.6	35.7	−11.7	.648
26.6	54.0	74.7	−16.3	.954
26.5	51.9	77.5	−15.5	.949
11.8	11.8	27.6	−11.0	.950
36.1	73.2	97.4	−20.6	.933
33.3	138.5	190.2	−20.0	.919
15.5	54.4	70.3	−10.6	.934
0.0	0.0	20.0	0.0	.945
7.7	7.7	16.7	−8.3	.941
13.6	16.8	29.9	−11.0	.946

Figure 6: Resulting from the extreme scores analysis. The values are the medians of the scores produced under several specific types of handling in order to check the strength of indices against extreme scores.

5.4.4.3 Reliability

A final test in order to check the influence of extreme scores was performed (Ruscio, Seaman, Orriano, Stremlo, & Mahalchik, 2012), and this is the estimation of each index's reliability. A known psychometric tool used for this: one can estimate the reliability of scores on a test by using the split-half method. Data used, comes from PsycINFO which lists publications chronologically, and computed each index's separately using odd- and even-numbered publications; researchers who had one article were dropped out from

https://doi.org/10.26219/heal.aueb.6022

these analyses data set. This method is analogous to splitting a test into halves using odd- and even-numbered items. In addition the usual Spearman-Brown correction of applied:

$$r_{xx} = \frac{2 * r_{12}}{1 + r_{12}}$$

where r_{12} is the correlation between halves and r_{xx} is the final corrected estimation of reliability. Since extreme scores between citation counts must contribute with differentially degrees to the scores for odd- and even-numbered publications, reliability estimations must be larger for indices that are more robust to the influence of extreme citation scores. The Reliability estimations are shown in the table above, in the last column. Split-half reliabilities exceeded the value 0.90 for most indices, but reliability was especially high (r_{xx} >0.97) for the six citation-based indices that segmented together in each type of analysis, as has already mentioned before. The extracted Reliability was some kind lower (r_{xx} < 0.80) for four indices (Mc, Mdnc, a, and mi) .

5.4.5 Validity

As mentioned in (Panaretos & Malesios, 2012) and (Haslam, 2010) validity is the most important from all the related criteria. A standardized index of effect size (ω2) computed, for the mean difference across researcher ranks (Ruscio, Seaman, Orriano, Stremlo, & Mahalchik, 2012). All else being equal, one would probably expect a valid metric to score the greater bibliographic impact of researchers at higher ranks. Since that there are many other factors involved in attaining promotion and achieving scholarly impact, someone must not be so optimist to expect especially large values of ω^2. This comparison affords some insight into metrics' relative validities.

For all indices, except mq and mqt, which are designed to control for career level, means followed the expected order of full researchers --> associate researchers --> assistant researchers. The magnitude of ω^2 served as the first

https://doi.org/10.26219/heal.aueb.6022

index of validity. Together with the conventional index Na, the citation-based indices h, ht, f , t, g, and hg showed serious differences between researcher ranks ($0.178 \leq \omega^2 \leq 0.202$ for these metrics). Most of the other conventional indices scored comparatively poorly with ω values: $\omega^2 < .100$ for C, Mc, and Mdnc; but $\omega^2 = 0.171$ for sqrt-C), as did two of the citation-based indices (ω2 = 0.064 for a and .055 for mi). The extent to which the two metrics designed to control for career level attained this goal is reflected in their very small differences across researcher ranks ($\omega^2 = 0.007$ and 0.008 for mq and mqt, respectively). The remaining seven citation-based indices attained lowest validity ($0.105 \leq \omega^2 \leq 0.159$) in differentiating researcher ranks.

As an extra test of validity, correlations between each metric and NRC

Index	*Effect size (ω^2) across Professor Ranks*	*Correlation with NRC School Scores*
N_a	.200	.226
C	.097	.258
sqrt-C	.171	.346
M_c	.017	.249
Mdn_c	.004	.126
h	.202	.335
h_t	.195	.339
f	.197	.333
t	.194	.340
g	.178	.342
hg	.192	.341
a	.064	.306
m_i	.055	.291
r	.150	.354
h_w	.147	.354
q^2	.156	.355
e	.120	.352
mp	.105	.281
sqrt-mp	.156	.351
$h^{(2)}$	.159	.366
m_q	.007	.269
m_{qt}	.008	.278

Figure 7: Validity coefficient for bibliographic impact measures

school scores computed. The whole results shown in the table below:

5.5 Adjusting for self-citations and shared authorship

The self-citation theme is a well-known problem in the scientific community. It is also known that everybody interests in results cleared from the sef-citation effects. In the pioneering comparative study of Ruschio et all, indices of scholarly impact in their original form, unadjusted for self-citations or shared authorship, were assessed. As it is denoted by Schreiber (2007); Zhivotovsky & Krutovsky (2008), many investigators may argue that self-citations must be identified and removed or that corrections must be made for shared authorship and that failing to do so compromises the usability of these indices.

Of course, there are many logical and understandable reasons why the publications cited in an article might include one or more of its authors. self-citations can be appropriate when for example citing earlier stages in a program of research, or being convenient when authors might be more familiar with their own work than that of others etc. The identification of self-citations requires considerable serious burden. It is needed to be defined what qualifies as a self-citation and inspect every published citation to every article published by the main author to determine whether it meets the definition of self citation.

In the test, that Ruscio et all (2012) performed, self-citations defined in five ways:

1. The first citing author matched the first cited author,
2. The first citing author matched any cited author,
3. The target author matched any cited author,
4. Any citing author matched the first cited author, and
5. Any citing author matched any cited author.

https://doi.org/10.26219/heal.aueb.6022

According to these definitions, five variants of each metric were constructed by removing self-citations according to each of these five criteria. For shared authorship check, they coded two variables—the number of author and the main author's position on the list of authors—for each article and adjusted citations counts in four different ways. Again, four variants of each metric were constructed by sorting citation counts in each of these four ways. In a sample that participated across 156 publications authors, the percent of citations coded as self-citations ranged from Mdn = 5.23% (IQR = 3.04% to 9.01%) for the most stringent criterion to Mdn = 15.79% (IQR = 11.20% to 22.04%) for the most relaxed defined criterion.

All the calculated rank correlations between the non self-citation adjusted version of each metric and its five self-citation adjusted variants exceeded 0.950, and most (89%) exceeded 0.980; Mdn = 0.992. Under the fact of these extremely high correlations, it should not be surprising that split-half reliabilities changed very little after the adjustments were introduced. In most cases, there was a decrement in reliability, though all changes were quite small: No one reliability changed by more than ±0.03, and most reliabilities did not changed by more than ±.01.

5.6 Final conclusions- Selecting indices for applications in research. Future trends.

It is clear and totally accepted that citation based type of indices of scholarly impact have much credits to offer as accounting tools and as supplements to quantitative but major in qualitative information in many important contexts. A question may remain after all the previous chapter discussion: which index to choose someone in a specific situation? Is there an ierarcy on the criteria introduced in the previous paragraphs? Ruscio (2012) and Panaretos & Malesios (2012) considering that validity must be the most important criterion of all. According to this criterion (validity), among the conventional indices, only Na achieved a respectable degree of validity. Na showed a strong difference across researcher ranks, and a comparatively small correlation with NRC school scores. The classical indices C, Mc, and Mdnc behaved more poorly, especially across researchers ranks. The index, sqrt-C behaved well; probably because the square root transformation reduced the influence of extreme scores, reducing with this way the intra-sample variance.

In contrary, some of the citation based indices, exhibited good validity in both tests. Among the most valid were the h index and some close to h variants, including ht, f , t, g, and hg. These indices also were very highly correlated; with Pearson or Spearman correlation among these indices being more than 0.97. That means that they would resulting very similar results in research. Conclusively, it is denoted that two indexes are coming totally validity scored; the Na as a conventional and the h as a modern one.

According to Ruscio et al (2012) & Panaretos et al (2012) some practical directions in order to choose the appropriate index depending to a given situation are the following.

When the audience consists of no specialists such as academic administrators or substantive researchers more interested in findings than methods, the easiest of the valid indices to understand and calculate is the h index.

In case that the frequency of tied scores is a concern, the ht index usually avoids this problem.

If fine distinctions are important, an individual might use ht in place of h or use scores on ht to break ties between researchers with equal scores on h. These indices' validities were same valued, but ht is less easy to understand or calculate accurately than h.

Another potentially serious factor to take considering is that some indices giving different reward in the quantity or the quality of publications. The h index giving rewards to a balance between quantity and quality (this explained detailly in the paragraph 5.4.3) because it corresponds to the length of the largest square that fits within a citation plot array. Neither a few highly influential papers nor a

https://doi.org/10.26219/heal.aueb.6022

large number of poorly-cited papers will produce a large h.

In case an individual prefers to reward quality more than quantity, one could use the g, ht, or sqrt-mp indices. The g index focus on crediting for excess citations above the h square and is more sensitive to especially highly-cited papers than h .

The ht index also look to awards credit for excess citations above the h square, and it showed of comparable validity to indices h and g.

The sqrt-mp index corresponding to the square-root of the largest rectangle that fits normally within the array of citations X publications. This rectangle can be large if there are a lot of publications with very large citation counts.

In case of an individual who prefers to reward quantity more than quality, the ht or sqrt-mp indices may be used, which awarding credit for excess citations to the right of the h square.

Finally, in case of someone might prefer not to favor any specific research strategy. The ht and sqrt-mp indices can accept as flexible variations on h in that they neither rigidly reward a balance between quantity and quality, as does h, nor favor quantity over quality.

However, they did not appear between the most valid indices, the two metrics designed to control for career level, mq and mqt, performed enough well. Their correlations with NRC school scores were satisfying strong. Because they exhibited similar validities, the choice between them could be made in much the same way as the choice between h and ht. On haves to recall that mq and mqt are simply h and ht divided by publishing age and the parallelization become clear. Whereas mq and h reward a balance between quantity and quality of publications, mqt and ht can flexibly rewarding either a greater quantity or a higher quality of publications.

The selection of one or more metrics based on the available citation counts, as opposed to more laborious coding of inconsistently available qualitative data, must greatly facilitate research on scholarly impact (Ruscio, Seaman, Orriano, Stremlo, & Mahalchik, 2012). In practice and depending on the decision given, one might hypothesize qualitative indicators such as candidates' editing experience, membership on editorial boards, awards for scholarly achievement, fellow status or leadership in professional organizations, invited talks, grant support, peer reviews, or other pertinent information sources (Panaretos & Malesios, Commentary on Ruscio et al.: "Measuring Scholarly Impact Using Modern Citation-Based Indices", 2012).

In case that the individual is not the element unit of analysis, then one can use either the h index or a variation of h, that adjusts for the number of publications published by each unit of interesting

There are a number of situations, however, in which the h index may provide misleading information about a scientist's output. There are also a lot of voices arguing that scientific impact is a multidimensional notion that cannot be effectively reduced to a single indicator (Bollen, Van de Sompel, Hagberg, 2009: at (Panaretos & Malesios, Commentary on Ruscio et al.: "Measuring Scholarly Impact Using Modern Citation-Based Indices", 2012)).

In conclusion (Ruscio, Seaman, Orriano, Stremlo, & Mahalchik, 2012), to evaluate researchers' scholarly impact it is suggested that in most cases Hirsch's (2005) h index must be seriously considered.

There is a huge volume of publications in recent years, presenting more and more h-type indices seemed to improve on the typical h index. These indices are found to be highly correlated to the h index and to each other. This direction of research does not seem to be the one necessary for setting up basic standards for the assessment of research impact. Since most of the indices involved in the current and previous studies are directly or indirectly related to each other (Panaretos & Malesios, 2012). More intensive theoretical work needs to be for the analysis of their interdependence, than in direction of just relying on a hypothetic linear relation, based on their correlations. Consequently, future investigations toward methodologies of accounting scientific performance based

https://doi.org/10.26219/heal.aueb.6022

on family of h-type indices proposed in the literature with older bibliometric indices should provide a more accurate picture of a scientist's productivity.

Bibliography

Abbott, A. C. (2010). Do metrics matter? *Nature* , pp. 860-862.

Adler, R. E. (2008). *Citation Statistics.* Retrieved from Comitte of Quantitative Assessment of Research.: http://www.mathunion.org/fileadmin/IMU /Report/CitationStatistics.pdf

Alonso, S. C.-V. (2009). h-index: A review focused in its variants,computation and standardization for different scientific fields. *Journal of Informetrics, , 3*, pp. 273–289.

Anderson, T. H. (2008). Beyond the Durfee square: Enhancing the h-index to score total publication output. *Scientometrics , 76*, pp. 577–588.

Bartolucci, F. (2013). On a possible decomposition of the h-index. *Journal of the American Society for Information* .

Bornmann, L. M.-D. (2009). Are there better indices for evaluation purposes than the h index? A comparison of nine different variants on the h index using data from biomedicine. *Journal of the American Society for Information for Information Science and Technology. , 59*, pp. 830–837.

Casella, G., & Berger, R. (2001). *Statistical Inference - Second edition.* Duxbury Advanced Series.

Clauset, A. S. (2009). Power-law distributions in empirical data. *SIAM Review , 51*, pp. 661-703.

Efron, B., & Tibshirani, R. (1993). *An Introduction to Bootstrap.* New York: Chapman Hall, Inc.

Egghe, L. (2006c). Theory and practice of the g-index. *Scientometrics* , pp. 69 (1) : 131–152.

Egghe, L. (2006b). An improvement of the H-index: the G-index. *ISSI Newsletter* , pp. 2 (1) : 8–9.

Egghe, L. (2006a). How to improve the h-index. *The Scientist* , p. 20 (3) : 14.

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics , 69*, pp. 131–152.

  https://doi.org/10.26219/heal.aueb.6022

Garfield, E. (1955). Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas. *Science , 122*, pp. 108-111.

Haslam, N. &. (2010). Quality, quantity, and impact in academic publication. *European Journal of Social Psychology , 40*, pp. 216–220.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, USA* , pp. 16569–16572.

Hirsch, J. E. (2007). Does the h-index have predictive power? *Proceedings of the National Academy of Sciences , 104*, pp. 19193–19198.

Iglesias, J. E. (2007). Scaling the h-index for different scientific ISI field. *Scientometrics* , pp. 73 (3) : 303–320.

Jin, B.-H. (2006). An evaluation indicator proposed by scientist. *Science Focus* , pp. 1 (1) : 8–9.

Jin, B.-H. L. (n.d.). The R- and AR- indices: Complementing the h-index. *Chinese Science Bulletin* , pp. 52 : 855–863.

Kosmulski, M. (2007). MAXPROD – A new index for assessment of the scientific output of an individual, and a comparison with the h-index. *International Journal of Scientometrics and Bibliometrics* , p. 11 (1).

Merton, R. K. (1968). The Matthew Effect in Science . *Science , 159*, pp. 56–63.

Panaretos, J., & Malesios, C. (2009). Assessing scientific research performance and impact with single indices. *Scientometrics , Vol. 81*, pp. 635-670.

Panaretos, J., & Malesios, C. (2012). Commentary on Ruscio et al.: "Measuring Scholarly Impact Using Modern Citation-Based Indices". *Measurement , 10*, pp. 164–166.

Ruscio, J., Seaman, F., Orriano, C., Stremlo, E., & Mahalchik, K. (2012, Sep). Measuring Scholarly Impact Using Modern Citation-Based Indices. *Measurement:Interdisciplinary Research and Perspectives , 10:3*, pp. 123-146.

Schreiber, M. (2007). Self-citation corrections for the Hirsch index. *Euro-Physics Letters , 78, 30002*.

Sidiropoulos, A. K. (2006). Generalized h-index for disclosing latent facts in citation networks. *Scientometrics* , pp. 72 (2) : 253–280.

Zhivotovsky, L. A. (2008). Self-citation can inflate h-index. *Scientometrics* , *77*, pp. 373–375.