Scientific metrics

Oleh Shkalikov Matteo Zannini Qader Qaribiyan

TU Dresden, Computer Science Faculty

December, 2022

Agenda

1. Introduction
   - Bibliographic Databases
   - Scientific Metrics Types
   - Altmetric Attention Score
2. Bibliometrics
   - Impact Factor
   - CiteScore
   - SNIP
   - SJR
3. Researcher's metrics
   - h-index
   - g-index
   - c-index
   - l-index

Problem

Assessing the quality and impact of research outputs is necessary

Every Metric has its limitations

No easy way exists to measure scientific performance

Problem

The original purpose of scientific publishing was to enable the global sharing of scientific results, ideas, and discussions within the academic society for more effective scientific achievements.

Influence of a publication is used for :

- allocation of funding resources
- industrial and economic growth priorities
- education policies
- the hiring of personnel academics

Bibliographic Databases

Scopus®

Google Scholar

Bibliographic Databases

What is a bibliographic database?

A bibliographic database is a repository of bibliographic or publication records.

They are built on the fact that citations in science serve as linkages between similar research items, and lead to related scientific literature

Frequently used biomedical databases include

- Web of Science
- Scopus
- Google Scholar

Advantages and disadvantages of databases

Pros	Cons
✓ Easy access to publication metadata and citation metrics	

Advantages and disadvantages of databases

Pros	Cons
✓ Easy access to publication metadata and citation metrics	
✓ Rigorous quality control	

Advantages and disadvantages of databases

Pros

- ✓ Easy access to publication metadata and citation metrics
- ✓ Rigorous quality control

Cons

- ✖ Databases may not contain the most recent references

Advantages and disadvantages of databases

Pros

- ✓ Easy access to publication metadata and citation metrics
- ✓ Rigorous quality control

Cons

- ✖ Databases may not contain the most recent references
- ✖ Most databases only include published articles

Advantages and disadvantages of databases

Pros

- ✓ Easy access to publication metadata and citation metrics
- ✓ Rigorous quality control

Cons

- ✖ Databases may not contain the most recent references
- ✖ Most databases only include published articles
- ✖ Bias towards English

Advantages and disadvantages of databases

Pros	Cons
✓ Easy access to publication metadata and citation metrics	✖ Databases may not contain the most recent references
✓ Rigorous quality control	✖ Most databases only include published articles
	✖ Bias towards English
	✖ Limited to specific fields

Scientific Metrics Types

Journal-level metrics are used to determine the impact a journal has on the scientific community

- Impact Factor

Article Metrics and Altmetrics are used to quantify the impact of published articles-how published papers are being discussed and shared.

- Citation Counts
- Attention Score

Author-level metrics assess the impact an author makes on the scientific community or field of the study.

- H-Index
- G-Index

Bibliometric Incentive

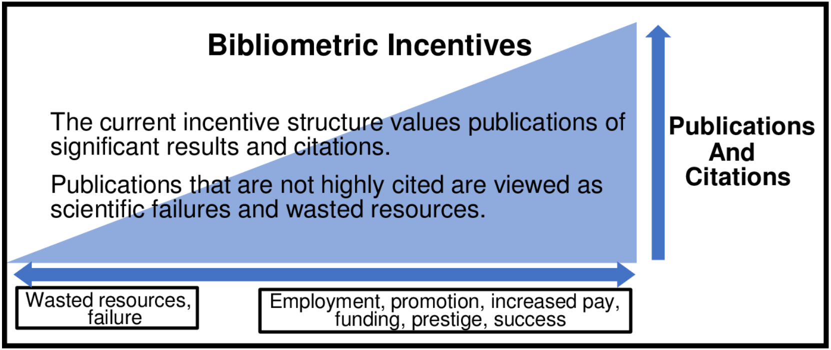


Figure 2: Bibliometric incentives model[1]

[1] https://doi.org/10.1371/journal.pone.0195321.g001

Altmetric Attention Score

What is Altmetric Attention Score?

The Altmetric Attention Score is an automatically calculated, weighted count of all of the attention a research output has received

It is calculated using data collected from research articles such as mentions on social media, news outlets, blogs, patents, etc.

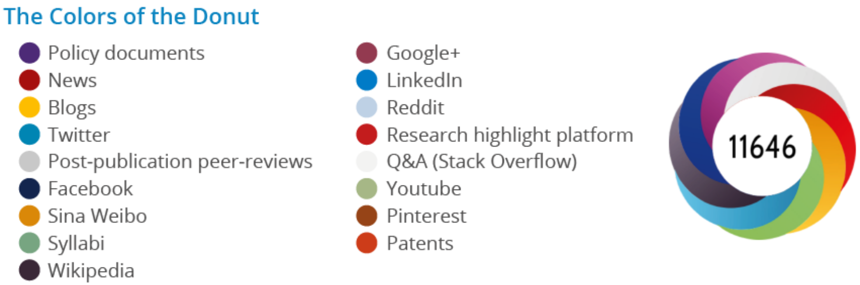


Figure 3: The colors of different sources of attention

Advantages to using altmetrics and limitations

Pros

✓ Quicker to accumulate than citation-based metrics

Cons

Advantages to using altmetrics and limitations

Pros	Cons
✓ Quicker to accumulate than citation-based metrics	
✓ Capture more diverse impacts than citation-based metrics	

Advantages to using altmetrics and limitations

Pros

- ✓ Quicker to accumulate than citation-based metrics
- ✓ Capture more diverse impacts than citation-based metrics
- ✓ Actively engage with comments and conversation around your work

Cons

Advantages to using altmetrics and limitations

Pros

- ✓ Quicker to accumulate than citation-based metrics
- ✓ Capture more diverse impacts than citation-based metrics
- ✓ Actively engage with comments and conversation around your work

Cons

- ✖ Altmetrics don't tell the whole story

Advantages to using altmetrics and limitations

Pros

- ✓ Quicker to accumulate than citation-based metrics
- ✓ Capture more diverse impacts than citation-based metrics
- ✓ Actively engage with comments and conversation around your work

Cons

- ✖ Altmetrics don't tell the whole story
- ✖ People can artificially inflate the altmetrics for their research

Advantages to using altmetrics and limitations

Pros

- ✓ Quicker to accumulate than citation-based metrics
- ✓ Capture more diverse impacts than citation-based metrics
- ✓ Actively engage with comments and conversation around your work

Cons

- ✖ Altmetrics don't tell the whole story
- ✖ People can artificially inflate the altmetrics for their research
- ✖ Altmetrics are relatively new

Bibliography metrics

IF (Impact Factor)

What is Impact Factor?

Impact Factor[Garfield, 1972] — a per-year metric based on *Web of Science* database which reflect the mean number of citations to the papers published within prior 2 year.

The Impact Factor is calculated by the following formula:

$$\mathrm{IF}(y) = \frac{C(y)}{D(y-1) + D(y-2)}$$

where

- $C(y)$ – number of citations to the articles published in $[y-2, y-1]$ years
- $D(y)$ – number of published articles in year y

IF: advantages and disadvantages

Pros	**Cons**
✓ Easily calculated and interpreted	

IF: advantages and disadvantages

Pros

✓ Easily calculated and interpreted

Cons

✖ Limited availability

IF: advantages and disadvantages

Pros

✓ Easily calculated and interpreted

Cons

✖ Limited availability

✖ Lack of reproducibility

IF: advantages and disadvantages

Pros

✓ Easily calculated and interpreted

Cons

✖ Limited availability

✖ Lack of reproducibility

✖ Doesn't adjust for the distribution of citations

IF: advantages and disadvantages

Pros

- ✓ Easily calculated and interpreted

Cons

- ✖ Limited availability
- ✖ Lack of reproducibility
- ✖ Doesn't adjust for the distribution of citations
- ✖ Windowed approach (especially 2 year window)

IF: advantages and disadvantages

Pros

✓ Easily calculated and interpreted

Cons

✖ Limited availability

✖ Lack of reproducibility

✖ Doesn't adjust for the distribution of citations

✖ Windowed approach (especially 2 year window)

✖ Doesn't distinguish between citations made to articles, reviews, or editorials

IF: advantages and disadvantages

Pros

- ✓ Easily calculated and interpreted

Cons

- ✖ Limited availability
- ✖ Lack of reproducibility
- ✖ Doesn't adjust for the distribution of citations
- ✖ Windowed approach (especially 2 year window)
- ✖ Doesn't distinguish between citations made to articles, reviews, or editorials
- ✖ Unstable

IF: advantages and disadvantages

Pros

✓ Easily calculated and interpreted

Cons

✖ Limited availability

✖ Lack of reproducibility

✖ Doesn't adjust for the distribution of citations

✖ Windowed approach (especially 2 year window)

✖ Doesn't distinguish between citations made to articles, reviews, or editorials

✖ Unstable

✖ Doesn't allow to compare titles from different fields

CiteScore

What is CiteScore?

CiteScore[Zijlstra and McCullough, 2016] — a per-year metric provided by *Scopus* which measure rate between the received citations to **all peer-reviewed documents** of a title published within 4 years.

The CiteScore is calculated by the following formula:

$$\text{CiteScore}(y) = \frac{CD_{y-3}(y)}{D(y) + D(y-1) + D(y-2) + D(y-3)}$$

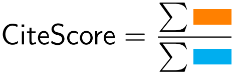
where

- $CD_x(y)$ – the number of citations in year range from x to y to all documents published within those years.
- $D(y)$ – the number of all **peer-reviewed** documents published in year y

CiteScore 2016 vs CiteScore 2019

CiteScore 2019-2021 vs CiteScore 2016-2018

Previously CiteScore has been calculated within 3 year range, and count citation received only in year *y*.



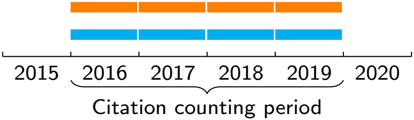


Figure 4: CiteScore 2019

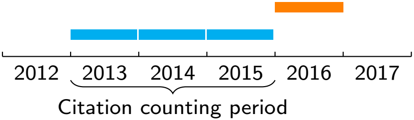


Figure 5: CiteScore 2016

– number of citation received by publications published in citation counting period

– number of documents (in CiteScore 2019 only peer-reviewed) published in corresponding year

CiteScore: advantages and disadvantages

Pros	**Cons**
✓ Transparency	

CiteScore: advantages and disadvantages

Pros

✓ Transparency

✓ Only citation from peer-reviewed content counted

Cons

CiteScore: advantages and disadvantages

Pros

✓ Transparency

✓ Only citation from peer-reviewed content counted

✓ Covers more titles than IF

Cons

CiteScore: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Easily calculated and interpreted

Cons

CiteScore: advantages and disadvantages

Pros

✓ Transparency

✓ Only citation from peer-reviewed content counted

✓ Covers more titles than IF

✓ Easily calculated and interpreted

✓ More robust and stable than IF

Cons

CiteScore: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Easily calculated and interpreted
- ✓ More robust and stable than IF

Cons

- ✖ Bounded to the *Scopus* database

CiteScore: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Easily calculated and interpreted
- ✓ More robust and stable than IF

Cons

- ✖ Bounded to the *Scopus* database
- ✖ Doesn't adjust for the distribution of citations

CiteScore: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Easily calculated and interpreted
- ✓ More robust and stable than IF

Cons

- ✖ Bounded to the *Scopus* database
- ✖ Doesn't adjust for the distribution of citations
- ✖ Windowed approach

CiteScore: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Easily calculated and interpreted
- ✓ More robust and stable than IF

Cons

- ✖ Bounded to the *Scopus* database
- ✖ Doesn't adjust for the distribution of citations
- ✖ Windowed approach
- ✖ Doesn't allow to compare titles from different fields

SNIP (Source Normalized Impact per Paper)

What is SNIP?

SNIP[Moed, 2010] measures a journal's contextual citation impact, taking into account subject field with use of, so called, citation potential of the field of science. The SNIP score is calculated based on *Scopus* database.

The SNIP is calculated by the following formula:

$$\mathrm{SNIP}(y) = \frac{RIP(y)}{RDCP(y)}$$

where

- $RIP(y)$ – raw impact per paper in year y (IPP / CiteScore 2016)
- $RDCP(y)$ – relative database citation potential in a journal's subject field in year y

RDCP

$$RDCP^j(y) = \frac{DCP^j(y)}{\operatorname*{med}_i\ DCP^i(y)}$$

where DCP – database citation potential in a journal's subject field:

$$DCP^j(y) = \frac{\sum\limits_{p \in CP(y)} DCP^j_p(y)}{\#CP^j(y)}$$

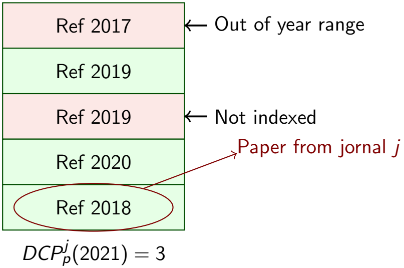


Figure 6: Example of DCP^j_p calculation

where

- $CP^j(y)$ – papers citing the journal j and **published in journals processed for the database** in the range $[y-3, y-1]$ years
- $DCP^j_p(y)$ – number of references to the papers published in those years from paper $p \in CP^j(y)$

SNIP: advantages and disadvantages

Pros	Cons
✓ Transparency	

SNIP: advantages and disadvantages

Pros	Cons
✓ Transparency	
✓ Only citation from peer-reviewed content counted	

SNIP: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF

Cons

SNIP: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ More robust and stable than IF

Cons

SNIP: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ More robust and stable than IF
- ✓ Allow to compare titles from different fields

Cons

SNIP: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ More robust and stable than IF
- ✓ Allow to compare titles from different fields

Cons

- ✖ Bounded to the *Scopus* database

SNIP: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ More robust and stable than IF
- ✓ Allow to compare titles from different fields

Cons

- ✖ Bounded to the *Scopus* database
- ✖ Hard to interpret

SNIP: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ More robust and stable than IF
- ✓ Allow to compare titles from different fields

Cons

- ✖ Bounded to the *Scopus* database
- ✖ Hard to interpret
- ✖ Calculation is more complicated

SNIP: advantages and disadvantages

Pros

✓ Transparency

✓ Only citation from peer-reviewed content counted

✓ Covers more titles than IF

✓ More robust and stable than IF

✓ Allow to compare titles from different fields

Cons

✖ Bounded to the *Scopus* database

✖ Hard to interpret

✖ Calculation is more complicated

✖ Citation potential tends to be highest for topical journals

SJR (SCImago Journal Rank)

What is SJR?

SJR[González-Pereira et al., 2009] is a metric provided by *Scopus* which is based on the number of citations and the prestige of the titles which cite this journal.

Key points of SJR:

- Based on PageRank[Page et al., 1999] algorithm to calculate the prestige of the journal: citations issued by more important journals will be more valuable than those issued by less important ones
- Self-citation limited to 33% of the total references
- Citation time frame is 3 year

SJR: methodology

The SJR indicator is computed in two phases:

1. Prestige SJR (PSJR) – size- dependent measure that reflects the overall journal prestige
2. normalization of this measure to give a size-independent metric

The formula for the SJR metric (2 stage) is the following:

$$\mathrm{SJR}_i(y) = c \cdot \frac{PSJR_i(y)}{Art_i(y)}$$

where

- $Art_i(y)$ – number of primary items (articles, reviews, and conference papers) of journal i published in three previous years [y-3, y-1].
- c – common constant to prettify values of SJR

Prestige SJR

The PSJR calculation is an iterative process which started with **identical amount** of prestige to each journal until the differences between journal prestige in consecutive iterations is under threshold.

$$PSJR_i = \overbrace{\frac{1-d-e}{N}}^{\text{min value}} + \overbrace{e \cdot \frac{Art_i}{\sum_{j=1}^{N} Art_j}}^{\text{number of papers}} + d \overbrace{\left[\sum_{j=1}^{N} C_{ji} \cdot \frac{PSJR_i}{C_j} \cdot CF + \frac{Art_i}{\sum_{j=1}^{N} Art_j} \cdot \sum_{k \in DN} PSJR_k \right]}^{\text{number and ``importance'' of the citations from other journals}}$$

where

- C_j – number of references of journal j
- C_{ji} – references from journal j to journal i
- d – constant: 0.9
- N – number of journals in the database
- DN – journals that do not cite other journals
- e – constant: 0.0999

Prestige SJR

The PSJR calculation is an iterative process which started with **identical amount** of prestige to each journal until the differences between journal prestige in consecutive iterations is under threshold.

$$PSJR_i = \overbrace{\frac{1-d-e}{N}}^{\text{min value}} + \overbrace{e \cdot \frac{Art_i}{\sum\limits_{j=1}^{N} Art_j}}^{\text{number of papers}} + d \overbrace{\left[\sum_{j=1}^{N} C_{ji} \cdot \frac{PSJR_i}{C_j} \cdot CF + \frac{Art_i}{\sum\limits_{j=1}^{N} Art_j} \cdot \sum_{k \in DN} PSJR_k \right]}^{\text{number and ``importance'' of the citations from other journals}}$$

where

- CF(Correction Factor) is introduced that spreads the undistributed prestige over all the journals proportionally to their accumulated prestige

$$CF = \frac{1 - \sum\limits_{k \in DN} PSJR_k}{\sum\limits_{h=1}^{N} \sum\limits_{l=1}^{N} C_{lh} \cdot \dfrac{PSJR_l}{C_l}}$$

SJR: advantages and disadvantages

Pros

✓ Transparency

Cons

SJR: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted

Cons

SJR: advantages and disadvantages

Pros

✓ Transparency

✓ Only citation from peer-reviewed content counted

✓ Covers more titles than IF

Cons

SJR: advantages and disadvantages

Pros

✓ Transparency

✓ Only citation from peer-reviewed content counted

✓ Covers more titles than IF

✓ Robust and stable

Cons

SJR: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Robust and stable
- ✓ Self-citation is limited

Cons

SJR: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Robust and stable
- ✓ Self-citation is limited
- ✓ Allow to compare titles from different fields

Cons

SJR: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Robust and stable
- ✓ Self-citation is limited
- ✓ Allow to compare titles from different fields

Cons

- ✖ Bounded to the *Scopus* database

SJR: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Robust and stable
- ✓ Self-citation is limited
- ✓ Allow to compare titles from different fields

Cons

- ✖ Bounded to the *Scopus* database
- ✖ Hard to interpret and calculate

SJR: advantages and disadvantages

Pros

- ✓ Transparency
- ✓ Only citation from peer-reviewed content counted
- ✓ Covers more titles than IF
- ✓ Robust and stable
- ✓ Self-citation is limited
- ✓ Allow to compare titles from different fields

Cons

- ✖ Bounded to the *Scopus* database
- ✖ Hard to interpret and calculate
- ✖ Still windowed approach

Author-level metrics

What is the h-index?

The **h-index**[Hirsch, 2005] is a measure used to quantify the cumulative impact and relevance of an individual's scientific research output.

First introduced by Professor Jorge E. Hirsch in 2005.

h-index formula:

$$h = \max i \in N : f(i) \geq i$$

where

- i is the published paper
- $f(i)$ is the number of citations for i

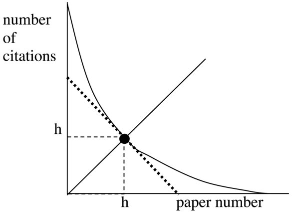


Figure 7: Graphical representation of the h-index

h-index

Example

4 publications: A,B,C,D citations per paper: 7,5,4,3

h-index=3

Pros

✓ It avoids previous indexes' problems

Cons

h-index

Example

4 publications: A,B,C,D   citations per paper: 7,5,4,3

h-index=3

Pros

- ✓ It avoids previous indexes' problems
- ✓ It measures both productivity and impact

Cons

h-index

Example

4 publications: A,B,C,D citations per paper: 7,5,4,3

h-index=3

Pros

- ✓ It avoids previous indexes' problems
- ✓ It measures both productivity and impact
- ✓ High h-index = high number of papers and citations

Cons

h-index

Example

4 publications: A,B,C,D citations per paper: 7,5,4,3

h-index=3

Pros

✓ It avoids previous indexes' problems

✓ It measures both productivity and impact

✓ High h-index = high number of papers and citations

Cons

✖ Field of research is ignored

h-index

Example

4 publications: A,B,C,D   citations per paper: 7,5,4,3

h-index=3

Pros

- ✓ It avoids previous indexes' problems
- ✓ It measures both productivity and impact
- ✓ High h-index = high number of papers and citations

Cons

- ✖ Field of research is ignored
- ✖ Order of authors' contribution is ignored

h-index

Example

4 publications: A,B,C,D citations per paper: 7,5,4,3

h-index=3

Pros

- ✓ It avoids previous indexes' problems
- ✓ It measures both productivity and impact
- ✓ High h-index = high number of papers and citations

Cons

- ✖ Field of research is ignored
- ✖ Order of authors' contribution is ignored
- ✖ Insensitive to highly cited publications

h-index

Example

4 publications: A,B,C,D citations per paper: 7,5,4,3

h-index=3

Pros

- ✓ It avoids previous indexes' problems
- ✓ It measures both productivity and impact
- ✓ High h-index = high number of papers and citations

Cons

- ✖ Field of research is ignored
- ✖ Order of authors' contribution is ignored
- ✖ Insensitive to highly cited publications
- ✖ Exposed to manipulation

h-index

Example

4 publications: A,B,C,D citations per paper: 7,5,4,3

h-index=3

Pros

- ✓ It avoids previous indexes' problems
- ✓ It measures both productivity and impact
- ✓ High h-index = high number of papers and citations

Cons

- ✖ Field of research is ignored
- ✖ Order of authors' contribution is ignored
- ✖ Insensitive to highly cited publications
- ✖ Exposed to manipulation
- ✖ Time-dependent

h-index variations

m quotient

It allows comparisons between scientists with different career lengths.

m quotient formula:

$$m = \frac{h}{n}$$

where

- n the number of years of academic activity

normalized h-index

It reduces the discipline bias.

normalized h-index formula:

$$|h| = \frac{h}{d}$$

where

- d the average number of academics in the field

individual h-index or hI

It reduces the effects of co-authorship.

hI formula:

$$hI = \frac{h}{\sum_{i=1}^{n} a_i}$$

where

- h – the h-index
- n – the number of papers and
- a_i – average number of authors

g-index

What is the g-index?

The **g-index**[Egghe, 2006] is a modified version of the h-index

First introduced by Leo Egghe in 2006.

How is it calculated?

Decreasingly order the papers based on the number of citations for each.
The index is the largest number s.t. the top g articles received together at least g^2 citations.

g-index

g-index formula:

$$g^2 = \sum_{i=1}^{g} c_i$$

where

- c_i – number of citations for each i paper

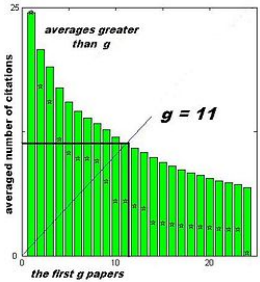


Figure 8: Graphical representation of the g-index

g-index

Example

Top publications: 5
Citations: 25
g-index=5

g is always greater or equal to h.

Pros

✓ All h-index advantages

Cons

g-index

Example

Top publications: 5
Citations: 25
g-index=5

g is always greater or equal to h.

Pros

- ✓ All h-index advantages
- ✓ It rewards papers with many citations

Cons

g-index

Example

Top publications: 5
Citations: 25
g-index=5

g is always greater or equal to h.

Pros

- ✓ All h-index advantages
- ✓ It rewards papers with many citations

Cons

- ✖ Saturation value

g-index

Example

Top publications: 5
Citations: 25
g-index=5

g is always greater or equal to h.

Pros

- ✓ All h-index advantages
- ✓ It rewards papers with many citations

Cons

- ✖ Saturation value
- ✖ It lacks discriminatory power

g-index

Example

Top publications: 5
Citations: 25
g-index=5

g is always greater or equal to h.

Pros

- ✓ All h-index advantages
- ✓ It rewards papers with many citations

Cons

- ✖ Saturation value
- ✖ It lacks discriminatory power
- ✖ It only favours academics with many publications

c-indexes

What are c-indexes?

c-indexes[Post et al., 2018] are the result of the augmentation of the h-index based on the citation order.

Introduced in 2018 in the magazine Cureus.

c-indexes formulas:

$$c_p = h + i_p \qquad c_s = h + i_s \qquad c_o = h + i_p + i_s$$

The augmentations are given by the following formulas:

$$i_p = n_f + 0.5 n_s$$
$$i_s = n_l + 0.5 n_s$$

c-indexes

Pros

✓ All the advantages of the h-index

Cons

c-indexes

Pros

- ✓ All the advantages of the h-index
- ✓ More accurate

Cons

c-indexes

Pros

✓ All the advantages of the h-index

✓ More accurate

✓ Takes citation order into account

Cons

c-indexes

Pros

- ✓ All the advantages of the h-index
- ✓ More accurate
- ✓ Takes citation order into account

Cons

- ✖ Time-dependent

c-indexes

Pros

- ✓ All the advantages of the h-index
- ✓ More accurate
- ✓ Takes citation order into account

Cons

- ✖ Time-dependent
- ✖ Field of research is ignored

l-index

What is the l-index?

The **l-index**[Lando and Bertoli-Barsotti, 2014] is a bibliometric based measure derived from the h-index which enhances highly cited papers.

Introduced in 2014.

l-index formula:

$$L(x) = \sum_{i=1}^{n} \ln\left(1 + \frac{x_i}{x_i^*}\right) x_i^*$$

After normalization:

$$l(x) = \sqrt{\frac{L(x)}{\ln 2}}$$

where

- x – real citation distribution
- x^* – ideal citation distribution

l-index

Pros

✓ All the advantages of the h-index

Cons

I-index

Pros

✓ All the advantages of the h-index

✓ It rewards papers with many citations

Cons

I-index

Pros	Cons
✓ All the advantages of the h-index ✓ It rewards papers with many citations ✓ Regular and robust	

I-index

Pros

- ✓ All the advantages of the h-index
- ✓ It rewards papers with many citations
- ✓ Regular and robust

Cons

- ✖ Time-dependent

I-index

Pros

- ✓ All the advantages of the h-index
- ✓ It rewards papers with many citations
- ✓ Regular and robust

Cons

- ✖ Time-dependent
- ✖ Order of authors' contribution is ignored

I-index

Pros

- ✓ All the advantages of the h-index
- ✓ It rewards papers with many citations
- ✓ Regular and robust

Cons

- ✖ Time-dependent
- ✖ Order of authors' contribution is ignored
- ✖ Field of research is ignored

I-index

Pros	Cons
✓ All the advantages of the h-index	✖ Time-dependent
✓ It rewards papers with many citations	✖ Order of authors' contribution is ignored
✓ Regular and robust	✖ Field of research is ignored
	✖ Complex calculation

CONCLUSION

- Simple metrics like number of publications, number of citations or impact factor are often inappropriate to quantify the impact of a journal or an academic.
- More complex metrics must be employed.
- Deciding which metric to use depends on its specific properties and on the academic context.

[Egghe, 2006] Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1):131–152.

[Garfield, 1972] Garfield, E. (1972). Citation analysis as a tool in journal evaluation: Journals can be ranked by frequency and impact of citations for science policy studies. *Science*, 178(4060):471–479.

[González-Pereira et al., 2009] González-Pereira, B., Guerrero-Bote, V., and Moya-Anegon, F. (2009). The sjr indicator: A new indicator of journals' scientific prestige. *arXiv preprint arXiv:0912.4141*.

[Hirsch, 2005] Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences*, 102(46):16569–16572.

References II

[Lando and Bertoli-Barsotti, 2014] Lando, T. and Bertoli-Barsotti, L. (2014). A new bibliometric index based on the shape of the citation distribution. *PLoS One*, 9(12):e115962.

[Moed, 2010] Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of informetrics*, 4(3):265–277.

[Page et al., 1999] Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab.

[Post et al., 2018] Post, A., Li, A. Y., Dai, J. B., Maniya, A. Y., Haider, S., Sobotka, S., and Choudhri, T. F. (2018). c-index and subindices of the h-index: New variants of the h-index to account for variations in author contribution. *Cureus*, 10(5).

References III

[Zijlstra and McCullough, 2016] Zijlstra, H. and McCullough, R. (2016).
Citescore: a new metric to help you track journal performance and make decisions.
https://www.elsevier.com/connect/editors-update/citescore-a-new-metric-to-help-you-choose-the-right-journal.
Accessed: 26.11.2022.