

Factorul G și schema de clasificare SPIRES --- completări la indexul Hirsch (G-factor and SPIRES Classification Scheme --- complements to H-index)

DORIN N. POENARU

Institutul Național de Fizică și Inginerie Nucleară “Horia Hulubei”, Măgurele-București

The G-index introduced by Leo Egghe in 2006, as an improvement of Hirsch index, is briefly presented together with SPIRES classification scheme of publications based on the number of citations. Unlike H-index, G-index is sensitive to the level of highly cited articles. Two figures illustrate the difference between G and H. A short FORTRAN 77 program, allowing to compute the G-index, is provided.

1. Introducere

Societatea Europeană de Fizică, având ca președinte pe Doamna Luisa Cifarelli, fiind interesată de promovarea fizicienilor și fizicii, a publicat recent [1] un comunicat referitor la utilizarea tot mai frecventă a indicilor bibliometrici pentru evaluarea cercetătorilor, colectivelor și proiectelor acestora. Se subliniază (pentru a câta oară!) că nu trebuie absolutizat un sistem pur numeric, ci se insistă pe necesitatea completării cu “peer review”. “*Rigorous but broader assessment procedures - says EPS President Luisa Cifarelli - must take into account the research environment and particularities of the domain in the evaluation of the scientific content.*” În ce mă privește, subscriu la acest deziderat, cu condiția să existe cercetători calificați având înalte standarde etice în domeniul analizat. În cazul în care expertiza și obiectivitatea celor care fac evaluarea poate fi pusă la îndoială, metoda scientometrică este o foarte bună alternativă și, în toate cazurile, reprezintă un bun punct de plecare.

În ultimii ani, remarcăm, în acest sens, o bună practică în România introdusă, de exemplu, de către UEFISCDI (Unitatea Executivă pentru Finanțarea Învățământului Superior, a Cercetării, Dezvoltării și Inovării). Când depunem propuneri de proiecte ne-am obișnuit să ni se ceară factorul de impact și, mai nou, scorul de influență al revistelor ISI în care publicăm. Din păcate, la noi încă nu este bine apreciat numărul citărilor individuale.

Schema de clasificare SPIRES (Stanford Physics Information Retrieval System) împarte producția științifică a unui autor în mai multe categorii, în

funcție de numărul de citări. SPIRES este o bază de date de lucrări (articole, preprinturi, cărți) în domeniul fizicii energiilor înalte, creată la accelerorul liniar de la Universitatea Stanford din California, SUA.

În cazul unor autori mult citați, această clasificare poate completa foarte bine indexul Hirsch [2,3,4].

În lucrarea de față, vom prezenta pe scurt această clasificare, precum și indicele sau factorul G introdus de către Leo Egghe [5]. Nu credem că acest factor este mai bun sau mai rău decât indexul H [6] ci, mai curând, că el poate completa foarte bine ceilalți indici scientometrici, precum și schema de clasificare SPIRES. Cert este că atât H cât și G au avantajul de a fi numere unice.

2. Schema de clasificare SPIRES

Conform criteriilor SPIRES, există următoarele categorii de publicații:

- renumite “*renowned*” (peste 500 citări)
- faimoase “*famous*” (250-499 citări)
- foarte bine cunoscute “*very well known*” (100-249 citări)
- bine cunoscute “*well known*” (50-99 citări)
- cunoscute “*known*” (10-49 citări)
- mai puțin cunoscute “*less known*” (1-9 citări)
- necunoscute “*unknown*” (0 citări).

Un exemplu de aplicare a schemei de clasificare SPIRES și de determinare a indicelui Hirsch se dă în Figura 1. Într-o primă etapă, se face lista articolelor în ordinea descrescătoare a numărului total de citări.

În etapa următoare, reprezentăm grafic numărul de citări, în funcție de numărul de ordine al publicației.

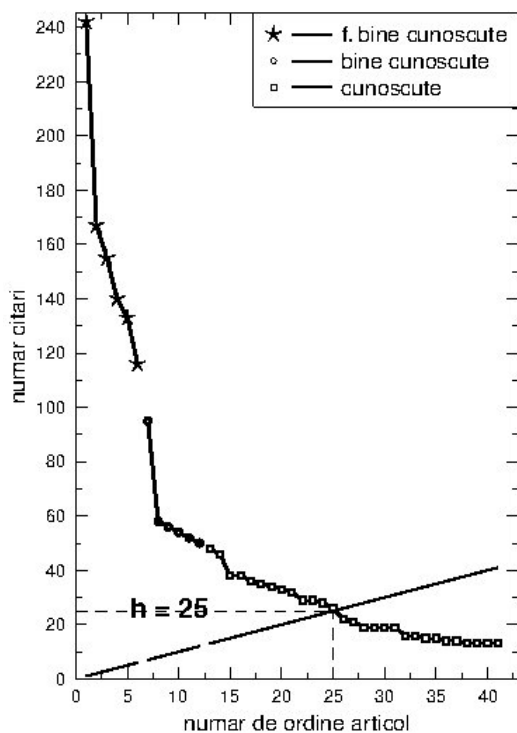


Fig. 1. Un exemplu de aplicare a schemei de clasificare SPIRES și de determinare a indicelui $H=25$.

Figura 1 se referă la producția științifică a unui cercetător având 6 publicații foarte bine cunoscute, 6 publicații bine cunoscute și 29 publicații cunoscute. De asemenea, se observă că indexul Hirsch al acestui cercetător este 25. Grafic, indicele Hirsch se obține foarte ușor comparând curba care trece prin punctele reprezentând numărul de citări ale fiecărui articol cu dreapta care trece prin zero și orice punct care are ordonată egală cu abscisa, de exemplu (25, 25).

3. Indexul G

Întrucât indexul H este corelat cu cel mai mare număr de citări al câtorva articole și este limitat de către numărul total de articole citate, Leo Egghe [5] a dorit să facă și o clasificare bazată pe numărul total de citări. Ca și în cazul precedent se pleacă de la o ordonare a numărului de citări, c_i în ordine descrescătoare; dar, în loc să se reprezinte numărul de citări în funcție de numărul de ordine al publicației, n și să se compare cu o dreaptă $y = n$, ca în Figura 1, se reprezintă numărul cumulat al citărilor, Σc_i , în funcție

de n . A se vedea Figura 2. De această dată, comparația se face cu o parabolă $y = n^2$.

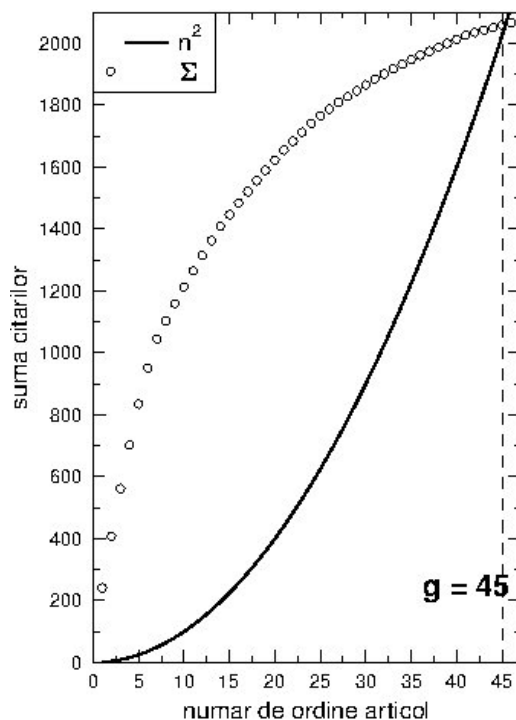


Fig. 2. Un exemplu de determinare a indicelui $G=45$ al unui cercetător care are indicele $H=25$ determinat în Fig. 1.

În Figura 2 se arată, pentru același cercetător, care are un indice Hirsch, $H = 25$, determinat în Figura 1, cum se obține indicele $G = 45$. În general $G > H$.

Indicele G este definit ca cel mai mare număr (unic) de publicații pentru care primele G articole au obținut împreună un număr de cel puțin G^2 citări:

$$G^2 \text{ mai mic sau egal cu } \Sigma c_i$$

suma fiind luată după toți indicii $i < G$.

O analiză axiomatică a indicelui G a fost făcută de către Gerhard J. Woeginger [7], iar R. S. J. Tol [8] a propus un indice G colectiv.

Indicele G își propune să completeze indicele H dând pondere mai mare articolelor mult citate. Într-adevăr, se poate afirma că la fel cu indicele H, indicele G este un număr unic, nu mai multe numere, ca în cazul clasificării SPIRES. Faptul că indicele H este insensibil la un set de articole slab citate sau necitate este un avantaj al acestuia. Un dezavantaj îl constituie însă lipsa de sensibilitate la nivelul citărilor articolelor mult citate. În cazul prezentat în Figura 1, indicele H rămâne egal cu 25, chiar dacă

autorul corespunzător are câteva articole citate de 1000 ori. În consecință, indicele G s-a introdus pentru a măsura mai bine calitatea globală a cercetătorului. Se observă că, în exemplele date, raportul $G/H=1,80$. Dacă oricare dintre primele 25 articole vor primi pe viitor mai multe citări, H nu se va modifica; dar G va crește astfel că G poate fi mai mare decât dublul lui H. În cazul în care numărul total de articole citate este insuficient de mare, se pot include în calcul unul sau mai multe articole fictive cu zero citări.

Un mic program FORTRAN77, prezentat mai jos, poate fi utilizat pentru calculul indicelui G, fără a face graficul din Figura 2.

```
c calculul factorului G pornind de la lista citărilor
c citări.inp
c aranjate în ordinea descrescătoare și având nc=0
c pe ultimul rând
c se compară o parabolă  $i^2$  cu suma citărilor
c rezultatul se scrie în fișierul gindex.out
c
OPEN(5,FILE='citari.inp',STATUS='OLD')
OPEN(12, FILE='gindex.out', STATUS='OLD')
rewind 12
i=0
isc=0
1 read(5,*)nc,ireal
  if(nc.eq.0)stop
  i=i+1
  i2=i*i
  isc=isc+ireal
  write(12,12)i,ireal,i2,isc
  ig=i-1
  if(i2.le.isc) goto 1
  write(12,14)ig
12 format(1x,4i7)
14 format(1x,'G-index=',i5)
  stop
  end
```

Primele patru și ultimele 2 linii ale fișierului de intrare '*citări.inp*', utilizat pentru Figurile 1 și 2 sunt:

```
1 242
2 167
3 155
4 140
...
46 8
0 0
```

În referința [6] se analizează în detaliu performanțele științifice ale unui mare număr de autori clasificați în patru categorii: (1) "*top scientists*",

(2) "*big producers*", (3) "*selective scientists*" și (4) "*low producers*" sau "*silent scientists*", arătând că cei din categoria a 2-a sunt avantajați de indicele H, iar cei din (3) de indicele G. Se subliniază din nou să nu se compare numeric domenii diferite.

În concluzie, schema de clasificare SPIRES și indicele G completează foarte bine indicele H, care împreună cu alte mărimi scientometrice și "*peer review*" cu referenți, care au expertiză în domeniu și sunt demni de încredere pentru obiectivitatea lor, constituie instrumentele de bază pentru evaluarea cercetătorilor, departamentelor, institutelor, precum și a propunerilor de proiecte ale acestora. Generalizarea aplicării acestor instrumente pentru judecări de valoare va îmbunătăți climatul cercetării pentru a stimula pe cei harnici și eficienți.

Bibliografie

- [1] The European Physical Society Statement: On the use of bibliometric indices during assessment, 11 June 2012, http://www.eps.org/resource/resmgr/policy/eps_statement_june2012.pdf
- [2] J. E. Hirsch, *An index to quantify an individual's scientific research output*, Proc. of the Nat. Acad. of Sci. of the USA, **102**, 16569-16572, 2005.
- [3] T. Braun, *The Hirsch index as a result of the frustration of a scientist*, Revista de politica științei și scientometrie - Serie nouă, **1**(1), 79 (2012).
- [4] P. T. Frangopol, *The Hirsch-Index - a new scientometric indicator for the evaluation of a scientist*, Revista de politica științei și scientometrie - Serie nouă, **1**(1), 75 (2012).
- [5] Leo Egghe, *Theory and practise of the g-index*, Scientometrics, **69**, nr. 1, 131-152, 2006. doi:10.1007/s11192-006-0144-7.
- [6] R. Costas, M. Bordons, *Is g-index better than h-index? An exploratory study at the individual level*, Scientometrics, **77**, 267 (2008). doi:10.1007/s11192-007-1997-0.
- [7] G. J. Woeginger, *An axiomatic analysis of Egghe's g-index*, Journal of Informetrics, **1**, 364 (2008). doi: 10.1016/j.joi.2008.05.002.
- [8] R. S. J. Tol, *A Rational, Successive G-Index Applied To Economics Departments In Ireland*, Journal of Informetrics, **2**, 149 (2008).