

Una visión crítica del índice h: algunas consideraciones derivadas de su aplicación práctica

Por Rodrigo Costas y María Bordons

Resumen: La posible aplicación del índice h en la evaluación de investigadores individuales ha despertado un gran interés entre los gestores de la investigación y entre los propios autores, haciendo necesarios estudios sobre las ventajas y limitaciones de este nuevo indicador. En este trabajo se presentan algunos resultados de su aplicación al colectivo de científicos del área de Recursos Naturales del CSIC, estudiando su producción en el Web of Science durante 1994-2004. El índice h permite detectar a los investigadores más destacados dentro del área, pero presenta limitaciones para identificar a aquellos con estrategias de publicación “selectivas”, que no destacan por su número de documentos pero sí por el alto número de citas recibidas, por lo que este tipo de comportamiento puede verse penalizado en evaluaciones basadas sólo en este indicador. Se incide en la necesidad de combinar distintos indicadores en los procesos de evaluación para evitar las limitaciones propias de cada uno de ellos.

Palabras clave: Índice h, Análisis de citas, Indicadores bibliométricos, Evaluación científica.

Title: A critical view of the h-index: observations based on a practical application

Abstract: The possible use of h-index in the evaluation of individual researchers has raised great concern among policy-makers and scientists themselves. Research is needed to determine the advantages and limitations of this new indicator. A study of publications in the Web of Science, 1994-2004, applied the h-index to researchers working in the natural resources area at the Spanish Council for Scientific Research (CSIC). One of the findings presented is that the h-index detected the most outstanding scientists; however, scientists who follow a selective publication strategy, publishing a relatively low number of documents that are frequently cited, could be penalized for this approach in h-index based evaluations. The authors emphasize the need to avoid the specific limitations of each indicator by combining indicators in the evaluation process.

Keywords: h-index, Citation analysis, Bibliometric indicators, Research evaluation.

Costas, Rodrigo; Bordons, María. “Una visión crítica del índice h: algunas consideraciones derivadas de su aplicación práctica”. En: *El profesional de la información*, 2007, septiembre-octubre, v. 16, n. 5, pp. 427-432.

DOI: 10.3145/epi.2007.sep.04



Rodrigo Costas es diplomado en biblioteconomía y documentación por la Universidade da Coruña y licenciado en documentación por la Universidad de Extremadura. Actualmente completa su doctorado en la Universidad Carlos III de Madrid y forma parte del grupo de Análisis Cuantitativo de Ciencia y Tecnología (Acute) del Centro de Información y Documentación Científica (Cindoc) del CSIC.



María Bordons es investigadora científica en el Centro de Información y Documentación Científica (Cindoc) del CSIC y miembro del grupo de investigación de Análisis Cuantitativo de Ciencia y Tecnología (Acute). Su investigación se centra en los estudios cuantitativos de la ciencia y, en especial, en el uso de los indicadores bibliométricos para el estudio de la ciencia y la evaluación científica.

Introducción

Desde su aparición en 2005, el índice h propuesto por **Jorge Hirsch** (2005) ha causado un gran impacto internacional por su posible aplicación en la evaluación de la actividad de los investigadores (**Ball**, 2005). Su amplia difusión en la comunidad científica ha venido propiciada por dos de sus principales ventajas: el hecho de combinar una medida de cantidad e impacto en un

solo indicador superando los recuentos brutos de documentos y citas y, sobre todo, su sencillez de cálculo. Según **Hirsch**, un investigador tiene un índice h cuando h de sus documentos han recibido al menos h citas cada uno, y el resto tiene no más de h citas por documento. Su cálculo es sencillo, ya que sólo requiere ordenar los documentos de un investigador en orden descendente del número de citas recibidas, numerándolos, e identificar el punto en que el número de orden coincide con

Artículo recibido el 19-03-07

Aceptación definitiva: 21-05-07

el de citas recibidas por el documento. Este número constituye el *índice h* –véase la metodología de cálculo en **Hirsch** (2005) o un ejemplo detallado en *Grupo Scimago* (2006)–. No obstante, también presenta limitaciones, algunas de las cuales mencionamos a continuación brevemente.

– Es inadecuado para comparar investigadores de diferentes áreas científicas, lo que se explica por los distintos hábitos de publicación y citación según el campo. Un ejemplo de ello es que los investigadores más relevantes en biología doblan en sus valores de *índice h* a los de otras áreas como la física, y éstos a su vez se sitúan muy por encima de las matemáticas. Como respuesta a esta limitación han surgido distintas modificaciones del *índice h* orientadas a facilitar las comparaciones entre áreas (**Iglesias; Pecharroman**, 2006; **Batista**, et al., 2006).

– Muestra una alta correlación positiva con el número total de citas y de documentos de los investigadores (**van Raan**, 2006), por lo que tiende a favorecer a los que cuentan con carreras científicas más dilatadas (**Kelly; Jenions**, 2006) y tiene menos validez en los más jóvenes con un bajo número de publicaciones. Para comparar investigadores en distintas fases de su carrera investigadora **Hirsch** introdujo el *valor m*, que divide *h* entre la edad científica de los investigadores (número de años desde la tesis doctoral o desde la primera publicación), aunque otra alternativa para comparar investigadores en distinta etapa de su carrera es calcular el *índice h* para un determinado período de tiempo, y no para toda la vida profesional de los investigadores.

– No tiene en cuenta la calidad de las revistas de publicación (**van Raan**, 2006), lo que constituye una limitación importante, dado que existen notables diferencias entre publicaciones en cuanto a los filtros de calidad impuestos a los documentos.

– La creciente popularidad del índice conlleva el riesgo de pretender reducir la evaluación a un solo indicador, a pesar de que la investigación es un proceso multidimensional que difícilmente se puede caracterizar unidimensionalmente (**Glänzel**, 2006; **van Raan**, 2006; **Costas; Bordons**, 2007).

– La conveniencia de incluir o no las autocitas en su cálculo ha sido objeto de debate. Algunos autores sugieren su exclusión por el riesgo de que el índice pueda incrementarse artificialmente si los investigadores identifican y citan sus documentos más próximos a entrar en el cálculo del *índice h* (**van Raan**, 2006; **Purvis**, 2006). Sin embargo, la investigación es un proceso acumulativo, por lo que cierto grado de autocitación es normal en el proceso de comunicación científica, y la exclusión de las autocitas en el cálculo de indicadores basados en citas no siempre estaría justificada (**Glänzel**, et al., 2006; **Garfield**, 1979). Aunque algunos tra-

bajos señalan que la influencia de las autocitas sobre el *índice h* es limitada (**Hirsch**, 2005), podría ser mayor en el caso de investigadores jóvenes que tuvieran valores bajos (**Schreiber**, 2007).

– Existen también limitaciones técnicas, ya que aunque las principales bases de datos que incluyen citas (como son el *Web of Science* o *Scopus*) proporcionan diversas herramientas para facilitar la identificación de los documentos de un autor e incluso permiten el cálculo automático del *índice h*, los problemas de homonimia, variantes de firma de los autores, errores tipográficos y falta de normalización siguen vigentes y dificultan la obtención correcta de este indicador de una forma automática y fiable. Si a esto se une que el *índice h* de un investigador puede variar según la base de datos utilizada para su cálculo (*WoS* o *Scopus*), se explica la recomendación de algunos autores de obtenerlo partiendo de una relación completa de las publicaciones de los investigadores, validada por ellos mismos, para aumentar la fiabilidad de los resultados (**Bornman; Daniel**, 2007).

De entre todas las limitaciones, hay dos que sin lugar a dudas son cruciales. Una es la dependencia entre el *índice h* de un investigador y su número total de documentos, de forma que el *índice h* nunca puede ser superior al número total de documentos. De este modo, una persona con 20 documentos nunca podrá tener un *índice h* superior a 20, independientemente del número de citas que reciban sus artículos. Por otro lado, hay que señalar que los *highly cited papers* o trabajos muy citados contribuyen al *índice h* de un determinado autor, pero el número de veces que son citados no influye sobre su valor (**Sidiropoulos**, et al., 2006; **Egghe**, 2006). O sea, alguien con 20 documentos con 20 citas cada uno tendría un *índice h* de 20, mientras que otro con 20 artículos de 40 citas cada uno lo tendría igual, y otro con 10 trabajos y 100 citas cada uno solamente lo tendría de 10.

“El índice h penaliza a los autores que priman calidad frente a cantidad: no publican un número de documentos muy elevado, pero sus contribuciones son muy relevantes”

Esto nos lleva a pensar que el *índice h* penaliza a los autores que presentan estrategias de publicación selectivas. Entendemos por investigadores selectivos (**Cole; Cole**, 1967; **Moed**, 2000; **Costas; Bordons**, 2005) aquellos que tienden a primar calidad frente a cantidad: no publican un número de documentos muy elevado, pero sus contribuciones son muy relevantes y

producen un alto impacto en la comunidad científica, es decir, publican menos pero publican mejor.

El índice h en el área de recursos naturales. Metodología

Con el fin de determinar hasta qué punto puede perjudicar a los científicos denominados selectivos, se ha analizado la producción de los 348 investigadores del CSIC en el área de recursos naturales. Sus publicaciones (incluido todo tipo documental) en revistas recogidas en el *Web of Science* durante el período 1994-2004 han sido descargadas y analizadas, obteniéndose diferentes indicadores, entre ellos el número total de documentos, el número total de citas (autocitas no excluidas), el ratio de citas por documento y el ratio de documentos altamente citados o *highly cited papers* (HCP, que incluye los materiales con 15 o más citas, lo que corresponde al 20% de los artículos más citados de la población estudiada).

Los investigadores han sido clasificados en 4 clases inspiradas en la propuesta de Cole y Cole (1967), como se muestra en la tabla 1. Los límites para la clasificación de los científicos vienen dados por el valor de la mediana en dos indicadores: número total de documentos y ratio de citas por documento. Para incrementar la relevancia de los resultados de la clasificación, únicamente se han tenido en cuenta aquellos con una producción intermedia o alta (al menos 12 documentos en el período, lo que corresponde al percentil 25 de la distribución de la producción del área).

De este modo se distinguen cuatro tipos principales:

- Investigadores *top*: con una alta producción e impacto.
- Pequeños productores silentes: con baja producción e impacto.
- Grandes productores silentes: con alta producción de documentos pero bajo ratio de citas por documento.
- Selectivos: con baja producción pero alto ratio de citas por documentos.

A partir de esta clasificación, podemos plantearnos responder a nuestro interrogante inicial: ¿penaliza el índice h a los selectivos?

Principales resultados: tipos de investigadores

Durante el período 1994-2004 la producción de los investigadores del CSIC en recursos naturales ascendió a 6.093 documentos. La productividad osciló entre 0 y 162 trabajos, variando su impacto entre 0 y 41,87 citas por documento, y el índice h de los científicos entre 0 y 29. Se ha observado una correlación positiva entre el número de documentos y el número de citas ($R^2=0,790$), así como entre el índice h y las dos variables anteriores (figura 1).

Esta estrecha relación también ha sido observada por van Raan (2006) en el estudio de 147 grupos de investigación holandeses en el área de química. Aunque el estudio de Grupo Scimago (2007) no detecta correlación entre el índice h y el número de documentos, su análisis se limita a nueve investigadores top de áreas temáticas diferentes; de forma que el pequeño tamaño

Ratio de citas por documento	P50=8,79	Investigadores selectivos	Investigadores top
		N. documentos <= 26 y Cit/doc. > 8,79 (N = 54 investigadores)	N. documentos > 26 y Cit/doc. > 8,79 (N = 72 investigadores)
		Pequeños productores	Grandes productores
		N. documentos <= 26 y Cit/doc. <= 8,79 (N = 75 investigadores)	N. documentos > 26 y Cit/doc. <= 8,79 (N = 52 investigadores)
		P50=26 Número de documentos	

Tabla 1. Clasificación de los investigadores atendiendo a su estrategia de publicación

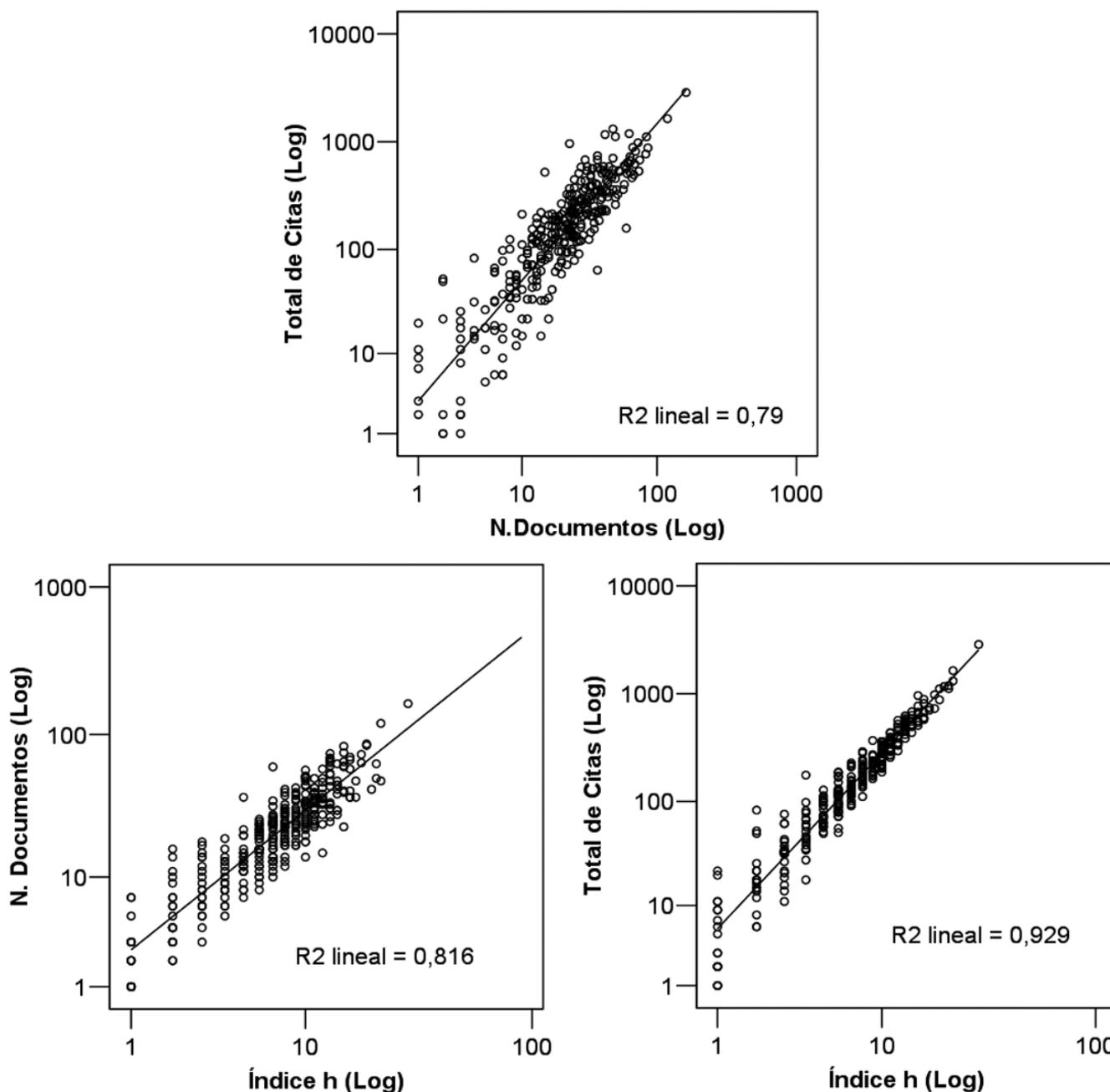


Figura 1. Correlaciones entre indicadores

de la muestra y su heterogeneidad temática reducen el valor de los resultados.

La figura 2 muestra la distribución del *índice h* para cada uno de los cuatro tipos de autores, considerando sólo aquellos con 12 o más documentos en el período. Se observa que los *top* presentan valores más elevados de *índice h* que los restantes, mientras que los selectivos y los grandes productores los tienen muy similares.

De la figura 2 se deduce que el *índice h* discrimina fácilmente a los pequeños productores y a los *top*, pero presenta limitaciones para discriminar entre los selectivos y los grandes productores. Este hallazgo corrobora la afirmación de otros autores (Jin, et al., 2007; Lehmann, et al., 2006) que indican la dificultad para discriminar entre científicos con niveles de actividad intermedios.

Los selectivos presentan más citas por documento, mayor porcentaje de documentos muy citados y mayor factor de impacto medio que los grandes productores (diferencias significativas, $p < 0,05$), y sin embargo obtienen menores o similares valores de *índice h*, porque éste viene limitado por su número de publicaciones. En la tabla 2 se muestran algunos ejemplos de casos reales para ilustrar este punto. Los investigadores B de la tabla obtienen un *índice h* igual o menor que sus correspondientes parejas A debido a su menor producción, y a pesar del mayor impacto observado de sus documentos (citas por documento y ratio de *HCP*, *highly cited papers*). Para los científicos selectivos es difícil obtener un alto *índice h*, dado que el valor máximo que se puede obtener es su número total de documentos, independientemente de la cantidad de citas que reciban

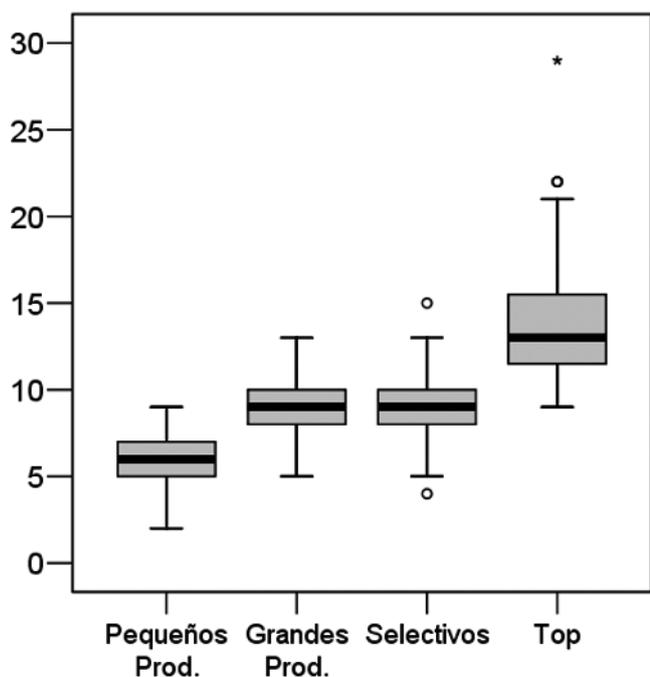


Figura 2. Índice h por tipo de investigador

sus publicaciones, y sin que además su mayor tasa de documentos muy citados contribuya a incrementar su índice. El último ejemplo de la tabla 2 es un caso extremo: se compara un gran productor con un investigador top. El autor A4 tiene el doble de producción que B4, y obtiene un mayor *índice h*, aunque B4 presenta mayor número de citas y citas por documento. Se observa así que en algunos casos también los top pueden verse penalizados frente a los grandes productores por seguir una estrategia selectiva.

Conclusiones

La principal ventaja del *índice h* es que combina en un solo indicador una medida de cantidad y otra del

“La principal ventaja del *índice h* es que combina en un solo indicador una medida de cantidad y otra del impacto de la producción”

impacto de la producción. Como se ha visto, es útil para detectar a los investigadores más destacados en un área (top) y a los pequeños productores más silenciosos, pero no discrimina entre los situados en niveles intermedios y en concreto entre los grandes productores y los selectivos. Debido a que el máximo valor que puede tomar para un investigador corresponde a su número total de documentos, los selectivos, que tienen un número intermedio de documentos pero de alto impacto, pueden verse penalizados en las evaluaciones basadas en este indicador. Esta limitación se mantiene en varias de las propuestas de mejora del *índice h* que han surgido hasta el momento (Batista, et al., 2006; Imperial; Rodríguez-Navarro, 2005; Sidiropoulos, et al., 2006).

Mención aparte merece el *índice g*, sugerido por Egghe (2006) ya que éste puede alcanzar valores superiores al número total de documentos de un autor y tiene en cuenta la cantidad de citas recibidas por los documentos incluidos en su cálculo. El principal inconveniente es que puede verse excesivamente influido por publicaciones ocasionales muy citadas, no representativas de la actividad científica del investigador (Jin, et al., 2007).

Recientes propuestas sugieren complementar el *índice h* con otros índices que tienen en cuenta las citas recibidas por los documentos del denominado *h-core*

Investigador	Nº docs.	Nº citas	Nº cit./doc.	HCP	Índice h
A1 (gran productor)	75	458	6,1	0,2	15
B1 (selectivo)	15	502	33,5	0,7	12
A2 (gran productor)	39	222	5,7	0,2	12
B2 (selectivo)	23	255	11,1	0,3	8
A3 (gran productor)	61	504	8,3	0,3	15
B3 (selectivo)	23	942	41,0	0,7	15
A4 (gran productor)	60	359	6,0	0,3	15
B4 (top)	34	568	16,7	0,3	12

Tabla 2. Ejemplos reales para la comparación de pares de investigadores

(los que determinan el *índice h*), lo que podría beneficiar a los autores selectivos. Así, se ha descrito el *índice A*, que calcula el número medio de citas recibidas por los artículos del *h-core*; el *índice R*, que mide la intensidad de citación de dicho núcleo, y el *índice AR*, que tiene en cuenta el año de publicación de los documentos (Jin, et al., 2007). No obstante, no son los sustitutos del *índice h*, sino que Jin sugiere el uso combinado de *AR* y *h* para superar algunas de las limitaciones de este último. De cualquier forma, todos ellos son todavía muy recientes y serán necesarios más estudios para analizar su validez y determinar las condiciones óptimas de su aplicación.

Creemos que es importante favorecer las buenas prácticas dentro del proceso de comunicación científica, fomentar la publicación de artículos de calidad y primar ésta sobre la cantidad de publicaciones. En este sentido consideramos necesario apoyar a los investigadores que muestran una estrategia de publicación selectiva que, aunque no tienen un número muy elevado de publicaciones, las concentran en revistas de alto prestigio y cuya investigación tiene un alto impacto sobre la comunidad científica (medido a través del número de citas recibidas). Sin embargo, el uso aislado del *índice h* puede penalizarles y, en consecuencia, disuadirles de seguir estos hábitos de publicación.

El *índice h* es un interesante indicador, pero presenta limitaciones y no debe utilizarse como único en la evaluación de la actividad investigadora de los científicos, sino que, tal y como se indica repetidamente en la literatura especializada (Glänzel, 2006; van Leeuwen, et al., 2003; Lewison, et al., 1999; Martin, 1996), es preferible combinar distintos indicadores y realizar planteamientos multidimensionales que contemplen la investigación científica en sus diferentes facetas.

Referencias

Ball, P. "Index aims for fair ranking of scientists". En: *Nature*, 2005, v. 436, n. 7.053, p. 900.

Batista, P. D.; Campiteli, M. G.; Kinouchi, O.; Martinez, A. S. "Is it possible to compare researchers with different scientific interests?". En: *Scientometrics*, 2006, v. 68, n. 1, pp. 179-189.

Bornmann, L.; Daniel, H. D. "What do we know about the h-index?". En: *Journal of the American Society for Information Science and Technology*, 2007 (en prensa).
DOI: 10.1002/asi.20609

Cole, S.; Cole, J. R. "Scientific output and recognition: a study in the operation of the reward system in science". En: *American sociological review*, 1967, v. 32, n. 3, pp. 377-390.

Costas, R.; Bordons, M. "Bibliometric indicators at the micro-level: some results in the area of natural resources at the spanish CSIC". En: *Research evaluation*, 2005, v. 14, n. 2, pp. 110-120.

Costas, R.; Bordons, M. "The h-index: advantages, limitations and its relation with other bibliometric indicators at the micro-level". En: *Journal of informetrics*, 2007 (en prensa).
DOI:10.106/j.joi.2007.02.001

Egghe, L. "Theory and practise of the g-index". En: *Scientometrics*, 2006, v. 69, n. 1, pp. 131-152.

Garfield, E. "Is citation analysis a legitimate evaluation tool?". En: *Scientometrics*, 1979, v. 1, n. 4, pp. 359-375.

Glänzel, W. "On the h-index – a mathematical approach to a new measure of publication activity and citation impact". En: *Scientometrics*, 2006, v. 67, n. 2, pp. 315-321.

Glänzel, W.; Debackere, K.; Thijs, B.; Schubert, A. "A concise review on the role of author self-citations in information science, bibliometrics and science policy". En: *Scientometrics*, 2006, v. 67, n. 2, pp. 263-277.

Grupo Scimago. "El índice h de Hirsch: aportaciones a un debate". En: *El profesional de la información*, 2006, v. 15, n. 4, pp. 304-306.

Grupo Scimago. "El índice h de Hirsch: su aplicación a algunos de los científicos españoles más destacados". En: *El profesional de la información*, 2007, v. 16, n. 1, pp. 47-49.

Hirsch, J. E. "An index to quantify an individual's scientific research output". En: *Proceedings of the National Academy of Sciences of the United States of America*, 2005, v. 102, n. 46, pp. 16.569-16.572.

Iglesias, J. E.; Pecharroman, C. "Scaling the h-index for different scientific ISI fields". En: *arXiv*, 2006. Consultado en: 12-03-07.
<http://arxiv.org/ftp/physics/papers/0607/0607224.pdf>

Imperial, J.; Rodríguez-Navarro, A. *Utilidad del índice h de Hirsch para evaluar la investigación en España*, 2005. Consultado en: 15-03-07.
http://www.bit.etsia.upm.es/Imperial_Rodriguez_Navarro.pdf

Jin, B.; Liang, L.; Rousseau, R.; Egghe, L. "The R- and AR-indices: complementing the h-index". En: *Chinese science bulletin*, 2007, v. 52, n. 6, pp. 855-863.

Kelly, C. D.; Jennions, M. D. "The h-index and career assessment by numbers". En: *Trends in ecology and evolution*, 2006, v. 21, n. 4, pp. 167-170.

Lehmann, S.; Jackson, A. D.; Lautrup, A. D. "Measures and mismeasures of scientific quality". En: *arXiv*, 2005. Consultado en: 08-05-07.
http://arxiv.org/PS_cache/physics/pdf/0512/0512238v2.pdf

Lewison, G.; Cottrell, R.; Dixon, D. "Bibliometric indicators to assist the peer review process in grant decisions". En: *Research evaluation*, 1999, v. 8, n. 1, pp. 47-52.

Martin, B. R. "The use of multiple indicators in the assessment of basic research". En: *Scientometrics*, 1996, v. 36, n. 3, pp. 343-362.

Moed, H. F. "Bibliometric indicators reflect publication and management strategies". En: *Scientometrics*, 2000, v. 47, n. 2, pp. 323-346.

Purvis, A. "The h index: playing the numbers game". En: *Trends in ecology and evolution*, v. 21, n. 8, p. 422.

Schreiber, M. "Self-citation corrections for the Hirsch index". En: *EPL*, 2007, v. 78, n. 30.002. Consultado en: 04-05-07.
<http://www.iop.org/EJ/abstract/0295-5075/78/3/30002>

Sidiropoulos, A.; Katsaros, D.; Manolopoulos, Y. "Generalized h-index for disclosing latent facts in citation networks". En: *Scientometrics*, 2007, v. 72, n. 2 (en prensa).

Van Leeuwen, T. N.; Visser, M. S.; Moed, H. F.; Nederhof, T. J.; Van Raan, A. F. J. "The holy grail of science policy: exploring and combining bibliometric tools in search of scientific excellence". En: *Scientometrics*, 2003, v. 57, n. 2, pp. 257-280.

Van Raan, A. F. J. "Comparisons of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups". En: *Scientometrics*, 2006, v. 67, n. 3, pp. 491-502.

Rodrigo Costas; María Bordons, Centro de Información y Documentación Científica (Cindoc), CSIC, Joaquín Costa 22, 28002 Madrid.
rodrigo.costas@cindoc.csic.es
mbordons@cindoc.csic.es