

Correlaciones entre conteos de uso y citas: aplicación en una muestra de documentos altamente citados sobre emprendimiento

Correlations between usage counts and citations in a sample of high impact documents about entrepreneurship

L. Javier Cabeza-Ramírez; Sandra Sánchez-Cañizares; Fernando J. Fuentes-García

Cómo citar este artículo:

Cabeza-Ramírez, L. Javier; Sánchez-Cañizares, Sandra; Fuentes-García, Fernando J. (2019). "Correlaciones entre conteos de uso y citas: aplicación en una muestra de documentos altamente citados sobre emprendimiento". *El profesional de la información*, v. 28, n. 4, e280418.

<https://doi.org/10.3145/epi.2019.jul.18>

Artículo recibido el 25-03-2019
Aceptación definitiva: 18-07-2019



L. Javier Cabeza-Ramírez

<https://orcid.org/0000-0002-5603-2365>

Universidad de Córdoba
Facultad de Derecho y Ciencias Económicas y Empresariales
Plaza de Puerta Nueva, s/n.
14002 Córdoba, España
r62caral@uco.es



Sandra Sánchez-Cañizares ✉

<https://orcid.org/0000-0003-0705-3327>

Universidad de Córdoba
Facultad de Derecho y Ciencias Económicas y Empresariales
Plaza de Puerta Nueva, s/n.
14002 Córdoba, España
sandra.sanchez@uco.es



Fernando J. Fuentes-García

<https://orcid.org/0000-0002-6477-3630>

Universidad de Córdoba
Facultad de Derecho y Ciencias Económicas y Empresariales
Plaza de Puerta Nueva, s/n.
14002 Córdoba, España
fernando.fuentes@uco.es

Resumen

Las métricas de uso están ganando importancia desde que las principales fuentes de información científica las facilitan. En este trabajo se recopilaron los períodos de exposición, citas y recuentos de uso de la *Web of Science Core Collection* (WoS) en una muestra multidisciplinar de documentos con altas tasas de citación sobre un tema específico: emprendimiento. Se aplicó la metodología h-classic para obtener un grupo de documentos representativo y se estudiaron las relaciones estadísticas que se establecen entre las tres variables. Los resultados mostraron cómo los años de exposición implican un descenso en los conteos de uso, pero no influyen en las citas recibidas a largo plazo. Por otro lado, citas y recuentos de uso poseen una clara correlación directa y positiva que permite realizar predicciones de cualquiera de las dos variables a partir de la otra.

Palabras clave

Emprendimiento; Bibliometría; Conteo de uso; Citas; Uso de artículos; Métricas de uso; Uso de datos; Descargas; *Web of Science*; H-classic.

Abstract

Metrics usage are becoming more important since the main sources of scientific information facilitate them. In this work, exposure years, citations and usage counts of the *Web of Science Core Collection (WoS)* were collected for a sample of multidisciplinary documents with high citation rates on a specific topic: entrepreneurship. The h-classic methodology was applied to obtain a representative group of documents and the statistical relationships established between the three variables were studied. The results showed how the years of exposure imply a decrease in usage counts, but do not influence the citations received in the long term. Citations and usage counts have a clearly direct and positive correlation that allows to predict either of the two variables from the other.

Keywords

Entrepreneurship; Bibliometrics; Usage counts; Citations; Article usage; Usage metrics; Usage data; Downloads; *Web of Science*; H-classic.

1. Introducción

En la actualidad, la manera de investigar está completamente adaptada a las mejoras tecnológicas que los científicos tienen a su disposición. Si hace no tanto la investigación pasaba por revisar copias impresas de los documentos más interesantes para su trabajo, hoy día lo habitual es consultar fuentes electrónicas (Kurtz; Bollen, 2010; Wang; Wang; Xu, 2013) y almacenar documentos en gestores bibliográficos personalizados para su posterior uso. El imparable avance de internet y la digitalización de las publicaciones ha supuesto una auténtica revolución, que en bibliometría se traduce en nuevos enfoques vinculados a la denominada *cienciometría 2.0* que incluye métricas más amplias, rápidas y abiertas (Priem; Hemminger, 2010; Thelwall; Kousha, 2015). Uno de esos enfoques se ha centrado en el estudio de los conteos de uso y los hábitos de descarga de los investigadores. Los conteos de uso no son recientes, los bibliotecarios llevan utilizándolos desde el comienzo de su profesión y son incluso más antiguos que las citas (Glänzel; Gorraiz, 2015).

Los recuentos de uso y descargas a nivel de artículo empiezan a estar disponibles en dos de las principales fuentes de información científica:

- a partir de 2012 en *Springer* (Chen, 2018);
- desde septiembre de 2015 en la *Web of Science* (Wang; Fang; Sun, 2016).

En este contexto la medida está ganando importancia y captando el interés fuera del ámbito de las bibliotecas. Henneken y Kurtz, (2017) la sitúan como un cuarto indicador para medir la cantidad y calidad de la investigación académica junto con

“fondos gastados, número de artículos publicados y número de citas de los documentos y medidas derivadas de estas citas”,

aunque señalan que su aplicación, más allá de la toma de decisiones de compra en las bibliotecas, apenas se encuentra en sus inicios.

Recientemente está aumentando el número de artículos que estudian sus posibilidades. Chen (2018) y Wang, Fang y Sun (2016) las dividen en temas o líneas a explorar:

- obsolescencia de artículos y documentos (Kurtz; Bollen, 2010; Wang *et al.*, 2014a; 2014b; Moed; Halevi, 2016);
- relación entre los conteos de uso y métricas tradicionales (Schloegl; Gorraiz, 2010; Subotic; Mukherjee, 2014; Chi; Glänzel, 2017; Xing, 2017; Markusova; Bogorov; Libkind, 2018);
- alternativa o complemento a las citas y otras métricas (Bollen *et al.*, 2005; Duy; Vaughan, 2006; Taylor, 2013);
- evaluación del desempeño (De-Sordi; Conejero; Meireles, 2016; Henneken; Kurtz, 2017);
- identificación de tendencias y frentes de investigación (Wang; Wang; Xu, 2013);
- comportamiento de los usuarios (Davis; Price, 2006; Khan; Younas, 2017).

La relación existente entre descargas/usos y citas es controvertida (Line; Sandison, 1975) y exige más investigación. Son métricas correlacionadas aunque no puede asumirse ninguna causalidad entre ellas, no existe una dirección específica desde las citas al uso o viceversa (Chi; Glänzel, 2018), ambas se retroalimentan y representan diferentes formas de comunicación (Moed; Halevi, 2016). A pesar de las influencias o múltiples motivaciones para citar (Borgman; Furner, 2002), los recuentos de citas están bien establecidos como medida de impacto (Brody; Harnad; Carr, 2006), se consideran una forma de uso, real o esencial (Chen, 2018). Las citas son producto de la reflexión en el proceso de investigación, que lleva a reflejarlas como significativas dentro del documento. En cambio, la definición de conteo de uso puede ser muy amplia, es una medida temprana de la conciencia o atención prestada hacia un documento (Moed; Halevi, 2016; Kurtz; Bollen, 2010), pueden referirse a vistas html o descargas pdf. En la *Web of Science* el acto de uso se realiza cuando el usuario intenta descargar el texto completo de un documento o guarda el registro del mismo en un administrador de referencias o en algún otro formato descargable (Markusova; Bogorov; Libkind, 2018). Las descargas/vistas no reflejan uso real sino potencial, ya que no existe evidencia de que los usuarios realmente vayan a usar el documento o finalmente acaben convirtiéndose en citas.

Aunque existen estudios anteriores que arrojan luz sobre la dinámica de la correlación entre citas y datos de uso (O'Leary, 2008; Watson, 2009; Wan *et al.*, 2010; Schlögl; Gorraiz, 2010; Schlögl *et al.*, 2014), la mayor parte de los análisis realizados se fundamentan en datos de uso y citas según disciplina, revistas o repositorios individuales como *Elsevier*, *Springer*, *PloS*, *Nature*, y *arXiv.org* (Wang; Fang; Sun, 2016). Por ejemplo, Nieder, Dalhaug y Aandahl (2013) analizaron los 50 artículos más vistos en 5 revistas de oncología y concluyeron correlaciones limitadas entre número de accesos y citas, o Moed y Halevi (2016) que compararon sus resultados con los obtenidos por Kurtz *et al.* (2005) y llegaron a conclusiones contradictorias sobre la capacidad predictora entre ambas medidas. Sin embargo, en la investigación moderna cada vez es más frecuente trabajar con muestras de documentos altamente citados, multidisciplinares, con distintos formatos y no limitados a una única revista, tipo de documento o disciplina, pero de un tema en particular.

De acuerdo con lo anterior, el objetivo de este trabajo consiste en plantear un enfoque que contribuya a profundizar en nuestro conocimiento sobre las correlaciones entre citas y conteos de uso en muestras multidisciplinares de documentos altamente citados. Para ello se abordan las siguientes cuestiones:

- Cómo afectan los años de exposición a las citas y a los conteos de uso en este tipo de muestras.
- Cómo se establece la correlación entre citas y conteos de uso o viceversa cuando los documentos muestran estas características.
- Determinar la posibilidad de desarrollar predicciones adecuadas de citas a partir de los conteos de uso.

2. Antecedentes de la investigación

Dilucidar las relaciones existentes entre indicadores bibliométricos se ha convertido en una cuestión determinante para explotar las posibilidades de las nuevas métricas que van surgiendo. Trabajos como los de Van-Raan (2006) ofrecen una mejor comprensión de las propiedades estadísticas de estos instrumentos de medición. Más recientemente estudios como Ortega (2015); Costas, Zahedi y Wouters (2015); Torres-Salinas, Cabezas-Clavijo y Jiménez-Contreras (2013) han puesto el foco de atención en las denominadas métricas alternativas. Especialmente en la pasada década se incrementó el número de análisis que abordan de manera específica la relación entre métricas de uso y citas.

A nivel de artículo o documento, Perneger (2004) presentó un breve análisis sobre 153 artículos publicados en la revista *British medical journal* que revelaba cómo los recuentos de visitas o número de veces que se accedió a un artículo online (*hit counts*) recogidos una semana después de su publicación, se relacionaban con el número de citas obtenidas en años sucesivos (aproximadamente 5 años después), mostrando cómo el recuento de visitas era una medida potencialmente útil del valor científico de un trabajo de investigación. Moed (2005) examinó relaciones estadísticas entre descargas y citas en una muestra más amplia de 1.190 artículos en la revista *Tetrahedron letters*, y pudo determinar que durante los primeros 3 meses después de la publicación de un artículo citado, su número de descargas aumentaba un 25% en comparación con lo que se habría esperado si ese artículo no hubiera sido citado, evidenciando que las descargas de los documentos citados podrían conducir a más descargas y más posibilidad de cita. Los documentos fueron estudiados durante 2 años tras su publicación, mostrando en su caso concreto que se producía aproximadamente 1 cita por cada 100 descargas, con coeficientes de correlación de Spearman más elevados cuando se descartaban las descargas realizadas en los tres primeros meses. Estos hallazgos evidenciaron cómo descargas o uso y citas iniciales se relacionan con las distintas fases del proceso de recopilación de información científica en la construcción de nuevos artículos.

Brody, Harnad y Carr (2006) utilizaron la base de datos *arXiv.org* para analizar las mismas relaciones, observando correlaciones significativas considerables entre el impacto de la citación y la descarga de artículos en áreas específicas: física, matemáticas, astrofísica y materia condensada. Comprobaron si el impacto del uso temprano podía predecir el impacto en citas posteriores, utilizando muestras aleatorias de 500 artículos en periodos de tiempo de 2 años acumulativos de citas y descargas. La mayor parte de los efectos se alcanzaron incluso en intervalos más cortos (6 meses).

Por su parte Chu y Krichel (2007) observaron la relación en un estudio transversal realizado en 2006 sobre los 200 documentos más descargados en *RePec* (*Research papers in economics*), revelando que un documento descargado en dicha base de datos recibía en promedio el doble de citas en *Google Scholar* que en *SSCI*. Según sus datos, las descargas tenían una relación moderada con las citas y advirtieron cómo la reputación del autor, el tipo de lector o el contenido temático del documento afectaban de manera positiva tanto las frecuencias de descarga como de cita.

En estudios de documentos altamente descargados, O'Leary (2008) observó los 25 documentos más usados de la revista *International journal of accounting information systems* y encontró relación entre la cantidad de citas y el número de veces que los artículos se encontraban en la lista de más descargados, presentando estos documentos un número de citas desproporcionado.

Ya en la década actual Cabezas-Clavijo y Torres-Salinas (2010) realizaron una investigación transversal sobre las métricas generadas a partir de la web 2.0 científica, en un análisis de caso más a partir de 8.945 artículos publicados en la revista *Plos one* entre finales de 2006 y principios de 2010. Calculando estadísticas básicas y correlaciones, apareció como destacable la correlación media entre citas y descargas.

Por su parte, Schlögl y Gorraiz (2010) examinaron diferencias entre los datos de uso y citas en todas las revistas de oncología facilitadas por *Science direct* desde 2001 hasta 2006. La investigación mostró cómo descargas y citas presen-

taban patrones de obsolescencia diferentes, con promedios de vida en las citas de 5,6 años y 1,7 años en las descargas; identificando también fuertes correlaciones entre frecuencias de citas y número de descargas a nivel revista y relaciones menores artículo por artículo.

Por otro lado, **Liu, Fang y Wang** (2011) observaron las frecuencias acumuladas de citas (2000-2009) y descargas (2005 a 2009) en artículos publicados en cinco revistas oftalmológicas chinas, encontrando correlaciones positivas entre ambas frecuencias, con una variación en el tiempo en que alcanzan los picos máximos, entre el séptimo y octavo año para las citas y a partir del segundo año para las descargas.

La lista de análisis dedicados a la mejor comprensión de la dinámica cita/uso o descarga, sin ánimo de ser exhaustiva, va incrementando sobre todo en los últimos años y se extiende entre otros trabajos a:

- los 25 artículos más descargados de una revista concreta (**Gregory; Denniss**, 2016);
- la descarga y distribución de citas en las revistas científicas (**Moed; Halevi**, 2016);
- comparativas entre distintos patrones y bases de datos (**Naude**, 2017; **Chen**, 2018);
- diferencia en los patrones según disciplina o revista (**Vaughan; Tang; Yang**, 2017; **Chi; Glänzel**, 2018);
- correlaciones entre citas y lectores de gestores bibliográficos (**Theilwall**, 2018);
- relaciones entre métricas de uso y tradicionales a nivel país (**Markusova; Bogorov; Libkind**, 2018).

En su mayoría, los antecedentes de la investigación referenciados presentan un denominador común: son estudios de caso en muestras muy distintas de documentos, revistas concretas, áreas o disciplinas específicas, etc. Todos los análisis mencionados evidencian que existe una clara retroalimentación entre citas y uso. Sin embargo, la dinámica actual de la investigación es cada vez más multidisciplinar y, por tanto, dentro de una temática concreta será habitual estudiar documentos procedentes de varias disciplinas, revistas o años de publicación. También resulta habitual para profundizar en ciertos conocimientos consolidados y nuevas tendencias, visualizar, descargar o citar aquellos documentos que son un referente claro, buscándolos entre los más citados o descargados.

3. Material y métodos

3.1. Selección de la muestra

La selección de la muestra de documentos se realizó utilizando una adaptación de la propuesta metodológica de **Martínez-Sánchez et al.** (2014), que había sido aplicada previamente por **Martínez-Sánchez et al.**, (2015): h-classic. El punto de partida es la definición de bibliografía clásica o altamente citada de **Garfield** (1977), es decir, aquellos documentos que han dejado marca en una determinada disciplina o área y tienen mayor visibilidad en cuanto a número de citas acumuladas. Se consideró la elección de este enfoque de acuerdo con los objetivos de la investigación. En este sentido resulta una metodología adecuada, no arbitraria, que ayuda a recopilar los documentos con mayor impacto dentro de un universo documental amplio, evitando la elección de un umbral aleatorio. **Rousseau** (2006) profundizó en el concepto de h-core de **Hirsch** (2005) para trabajar con grupos de publicaciones de alto rendimiento. Este indicador ha sido aplicado por ejemplo a revistas específicas (**Li et al.**, 2012) y se encuentra documentado entre otros en **Sanz-Casado, García-Zorita y Rousseau** (2016) o **Zhang, Glänzel y Ye** (2016):

1. La recogida de documentos se realiza en la *Core Collection* de la *Web of Science*.
2. Se establece un área de investigación multidisciplinar, en este caso *entrepreneurship*. Búsqueda de documentos mediante la ecuación topic: (entrep; timespan: All years; Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC).
3. Se calcula el índice h de los documentos, estableciendo una clasificación acorde a sus citas. En la práctica se calcula ordenando las publicaciones más citadas de manera descendente y examinando la lista para determinar el punto en el que el número de orden coincide con el número de citas recibidas. En junio de 2016 el resultado fue de 201.
4. Se recuperan los documentos altamente citados que constituyen el h-core de la muestra (201 documentos más citados).
5. Se extrae la información necesaria para realizar los análisis posteriores de acuerdo con la figura 1: año de publicación, citas acumuladas en junio de 2016, citas acumuladas en diciembre de 2018, conteo de uso desde 2013 recogido el 15 de diciembre de 2018, y conteo de uso desde 2013 anotado el 15 de marzo de 2019.

La selección de estos umbrales temporales se hizo de acuerdo con las características específicas de la muestra, buscando un equilibrio entre largo y corto plazo. Tal como señalaron **Brody, Harnad y Carr** (2006), el tiempo desde que un artículo se envía, revisa por pares, se publica, lo leen otros autores, lo citan en un nuevo artículo y se publica haciendo visible la cita, puede variar desde los 3 meses hasta más de dos años, motivo por el que se consideró adecuado recoger las citas tras 2 años y medio, aunque también se introdujo el dato de las citas en marzo de 2019 (tres meses después de la última recogida, efecto a corto).

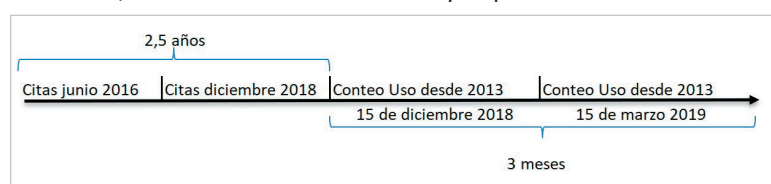


Figura 1. Información de las variables recogidas de la muestra

Para los conteos de uso la estrategia fue distinta: Se optó por anotar la medida cada tres meses, dado que la muestra estaba compuesta por documentos altamente citados, cuyo año más reciente de publicación era 2011. En este caso se pretende captar la correlación con el uso de este tipo de documentos concreto. De acuerdo con Wang, Fang y Sun (2016), los documentos que acaban de publicarse captan más atención en forma de uso que los más antiguos, debido a que los investigadores podrían querer usar publicaciones más recientes. Por tanto, nuestra muestra ya tendría acumulado ese uso más elevado inicial y, en consecuencia, el interés radica precisamente en verificar las correlaciones con documentos de alto impacto, con lo que un seguimiento cada 3 meses era una opción idónea; precisamente la *Web of Science* ofrece datos a este respecto en la doble ventana, acumulado desde 2013 (que fue el que se anotó) y cada 180 días.

3.2. Métodos de análisis

Los datos recopilados de la muestra de documentos seleccionada se han introducido en una base de datos de SPSS 22.0 diseñada al efecto para obtener los resultados que den respuesta a las cuestiones planteadas como objetivo de este estudio.

Se realizan análisis descriptivos univariantes de los datos para caracterizar la muestra, seguidos de técnicas bivariantes de correlación a través del coeficiente de Pearson (en tanto los datos son de carácter numérico) y modelos de regresión lineal simple para determinar la relación entre variables.

Finalmente, se aplica test de normalidad de Kolmogorov-Smirnov junto con el test no paramétrico con el estadístico U de Mann-Whitney para contrastar la igualdad de medias entre grupos.

4. Resultados

4.1. Análisis descriptivo de la muestra

Para cada uno de los 201 artículos se han obtenido las siguientes variables:

- años transcurridos entre su publicación y 2019 (Yearsexp);
- citas recibidas desde su publicación a junio de 2016 (Citjun16);
- citas recibidas desde su publicación a diciembre de 2018 (Citdic18);
- citas recibidas desde su publicación a marzo de 2019 (Citmar19);
- conteo de uso desde 2013 hasta diciembre 2018 (UC2018);
- conteo de uso desde 2013 hasta marzo 2019 (UC2019).

La tabla 1 muestra los estadísticos descriptivos de cada variable. En las citas se observa cómo el transcurso del tiempo incide en la evolución al alza de las medidas, siendo el aspecto más destacable el incremento del valor medio en prácticamente un 72% desde junio de 2016 a marzo de 2019. Por su parte, los conteos de uso presentan un crecimiento más reducido quedando el valor mínimo invariable entre 2018 y 2019 y el máximo con un ligero incremento de 56 usos. Finalmente, los años de vida de cada artículo oscilan entre los 8 años y los 51, con un valor medio de alrededor de 20 años.

Tabla 1. Estadísticos descriptivos de citas, conteos de uso y años de exposición

	Mínimo	Máximo	Media	Desviación estándar	Varianza
Citjun16	201	2.602	387,9	283,6	80.427,1
Citdic18	250	3.781	643,3	473,8	224.531,8
Citmar19	259	3.893	666,9	491,8	241.889,6
UC2018	5	1.548	210,2	195,4	38.182,2
UC2019	5	1.604	218,1	203,5	41.400,5
Yearsexp	8	51	20	6,6	43,6

4.1.1. Influencia de los años de exposición sobre las citas y los conteos de uso

En este primer análisis se contrasta la existencia o no de una correlación significativa entre las citas pasadas y presentes de los artículos que configuran la muestra de estudio con sus años de exposición (tabla 2)

La variable referida a los años de exposición pública del documento¹ únicamente presenta un coeficiente significativo en cuanto a las citas obtenidas en 2016, perdiéndose la correlación en las siguientes mediciones de citas de 2018 y 2019. Este resultado es compatible con la conocida *Ley de obsolescencia de la literatura científica* enunciada por Price, así como con las observaciones de Mafahi y Thelwall (2016) sobre la obsolescencia de documentos altamente citados, ya que, incluso en ellos puede producirse una disminución en su frecuencia de uso (Moed, 2005) o citación a lo largo del tiempo.

Tabla 2. Coeficiente de correlación de Pearson entre citas y años de exposición

		Citjun16	Citdic18	Citmar19
Yearsexp	Correlación de Pearson	0,151**	0,050	0,041
	Sig. (bilateral)	0,032	0,482	0,566

*p < 0,01 **p < 0,05

El gráfico de dispersión entre Yearsexp y la variable Citmar19 (valor más reciente del recuento de citas), revela la inexistencia de una influencia clara de los años de exposición del documento sobre las citas recibidas, encontrando puntos

que representan artículos altamente citados sin encontrarse entre los más antiguos, o por el contrario, un elevado número de documentos con muy baja cifra de citación que tienen un alto número de años de exposición.

Seguidamente se analiza la correlación entre años de exposición y, en este caso, los conteos de uso recibidos por los artículos seleccionados. La tabla 3 muestra los coeficientes de Pearson entre UC2018 y UC2019 con Yearsexp, quedando en evidencia la existencia de una correlación significativa, pero de signo negativo, entre los dos conteos de uso y los años de vida del artículo (los diagramas de dispersión 2a y 2b muestran asimismo dicha correlación negativa). En consecuencia, cuanto mayor es la “edad” de la investigación, menor es el interés que suscita en el mundo académico y los conteos de uso van reduciéndose paulatinamente, resultados en línea con las investigaciones sobre la obsolescencia en el uso de **Moed** (2005) y **Kurtz et al.** (2005).

Este resultado lleva a la decisión de subdividir la muestra inicial en dos grupos de acuerdo con su mayor o menor período de exposición para determinar las posibles diferencias entre ambos casos. Para ello, se ha tenido en cuenta como punto de corte el valor de la mediana² de Yearsexp que se sitúa en 18 años. De este modo queda un primer grupo de 105 artículos (51%) con 18 años o menos de exposición, y un segundo grupo de 96 (el 49% restante) con un período superior a 18 años desde su publicación. En la tabla 4 se reflejan los principales descriptivos de cada uno de estos dos grupos.

Tabla 4. Descriptivos según el grupo por años de exposición

Grupo	Media Yearsexp	Media Citjun16	Media Citdic18	Media Citmar19	Media UC2018	Media UC2019
1 (105 artículos)	15,28	324,43	584,05	609,53	254,71	264,69
2 (95 artículos)	24,92	454,68	705,48	727,10	163,37	169,07

Se aplica el test de Kolmogorov-Smirnov a cada una de las cinco variables analizadas en ambos grupos de pertenencia. Los resultados indican que en ningún caso las frecuencias siguen una distribución asimilada a la curva de la distribución normal. En consecuencia, los tests aplicados son de tipo no paramétrico.

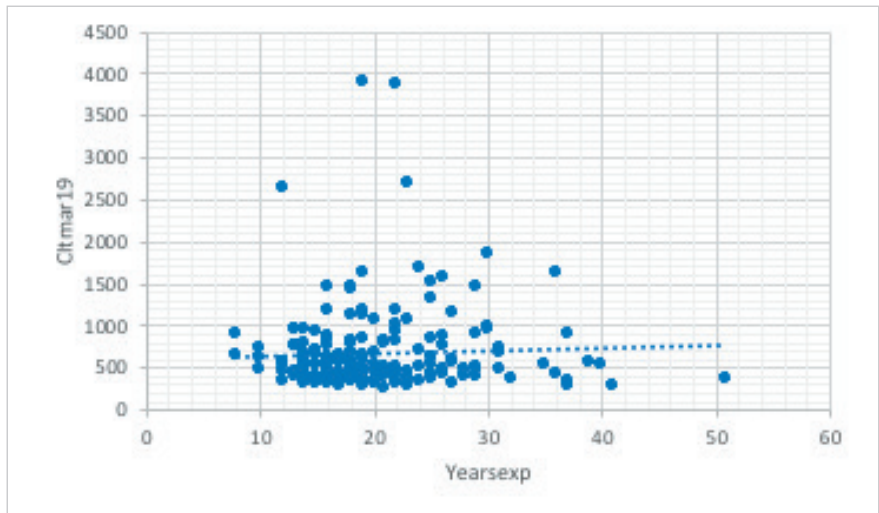


Gráfico 1. Diagrama de dispersión entre años de exposición y citas en marzo 2019

Tabla 3. Coeficiente de correlación de Pearson entre conteos de uso y años de exposición

		Yearsexp
UC2018	Correlación Pearson	-0,326*
	Sig. (bilateral)	0,000
UC2019	Correlación Pearson	-0,327*
	Sig. (bilateral)	0,000

*p < 0,01

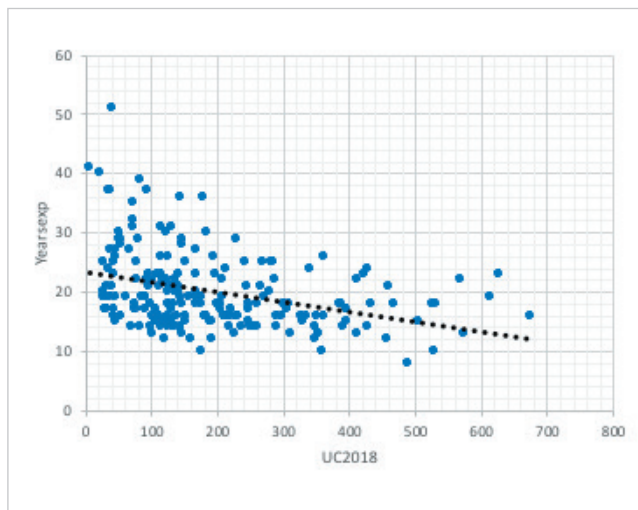


Gráfico 2a. Dispersión Yearsexp-UC2018

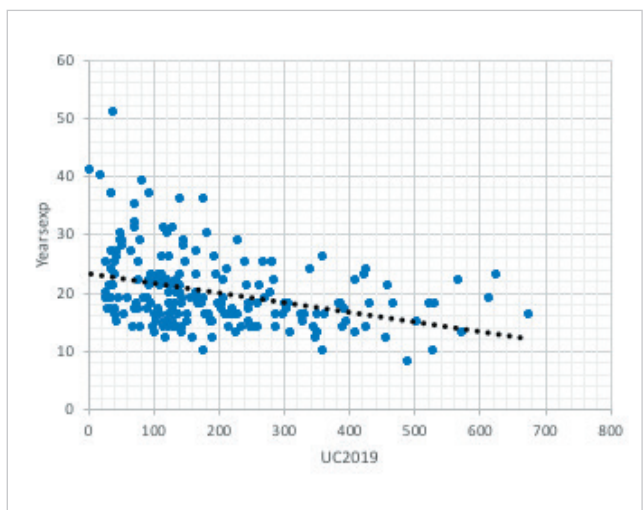


Gráfico 2b. Dispersión Yearsexp-UC2019

Para verificar la existencia de diferencias estadísticamente significativas entre el valor medio de las fechas de citas y conteos de uso, en ambos grupos se ha utilizado la prueba U de Mann Whitney, arrojando los resultados de la tabla 5.

Tras analizar previamente las correlaciones entre años de exposición, citas y conteos de uso, de la tabla 5 se deduce que las diferencias entre las medias de los dos grupos son significativas para las citas en la fecha más temprana y con ambos conteos de uso. Mientras tanto, las citas en diciembre de 2018 y de 2019 son distintas y mayores en el segundo grupo por su mayor tiempo de exposición, pero el contraste de hipótesis realizado confirma la hipótesis nula de igualdad de medias, por lo que la diferencia no se puede considerar relevante.

Los conteos de uso son inferiores en el segundo grupo tanto en 2018 como en 2019, por lo que se confirma nuevamente que, conforme aumenta el período de exposición y pese a que las citas totales han aumentado ligeramente, no es significativo este incremento, pero sí lo es, en cambio, el descenso que han experimentado los conteos de uso.

4.1.2. Relación citas-citas, citas-conteos de uso y conteos de uso-citas

Analizando la correlación existente entre citas-citas y citas-conteos de uso (tabla 6) se observa correlación significativa y casi perfecta entre las citas recibidas en 2019 con respecto a las que se habían obtenido en 2018 y 2016, así como entre estas dos últimas (el gráfico 3a muestra a modo de ejemplo esta correlación entre las citas de 2016 y de 2019).

Tabla 6. Coeficiente de correlación de Pearson entre citas y años de exposición

		Citjun16	Citdic18	Citmar19	UC2018	UC2019
Citjun16	Correlación de Pearson	1,000	0,969*	0,963*	0,526*	0,519*
	Sig. (bilateral)		0,000	0,000	0,000	0,000
Citdic18	Correlación de Pearson		1,000	1,000*	0,664*	0,658*
	Sig. (bilateral)			0,000	0,000	0,000
Citmar19	Correlación de Pearson			1,000	0,674*	0,669*
	Sig. (bilateral)				0,000	0,000

*p < 0,01

Pero lo más destacado es la correlación significativa y positiva entre las citas recibidas en los tres momentos temporales estudiados y los conteos de uso hasta diciembre de 2018 y marzo 2019. El valor del coeficiente siempre supera el 50% y va creciendo conforme la variable de citas se va acercando al presente³ (el gráfico 3b muestra a modo de ejemplo la correlación positiva y significativa entre UC2018 y Citmar 19, más moderada que en el caso de citas con citas). Profundizando en esta cuestión, se ha individualizado la correlación entre Citmar2019 y UC2019 para estudiar su evolución en función del año de publicación de los documentos que configuran la muestra. El gráfico 4 presenta los resultados agrupando los años de publicación de manera que hubiera al menos 10 documentos en cada grupo (resultando intervalos de tiempo más amplios o, en su caso, años individuales en función de esta condición).

Tal como se desprende del gráfico 4, la correlación entre citas y conteos de uso siempre es positiva, significativa y muy elevada, ya que en todo caso el coeficiente de Pearson se sitúa por encima de 0,7. Ello implica que los años de exposición de los artículos no suponen una reducción en la relación existente entre sus usos y las citas recibidas puesto que no

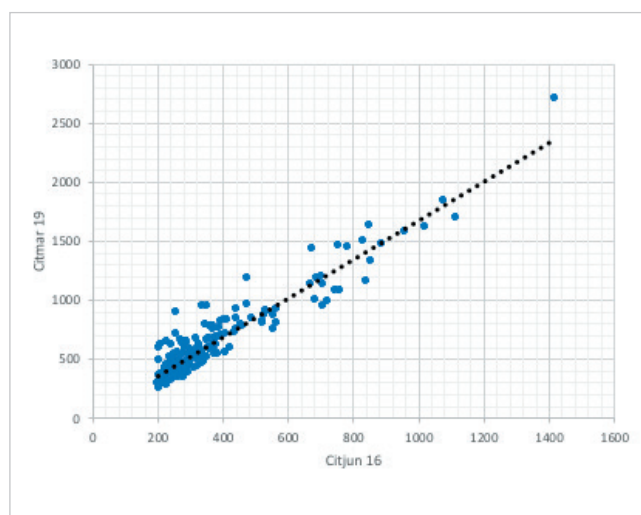


Gráfico 3a. Dispersión Citjun 16-Citmar 19

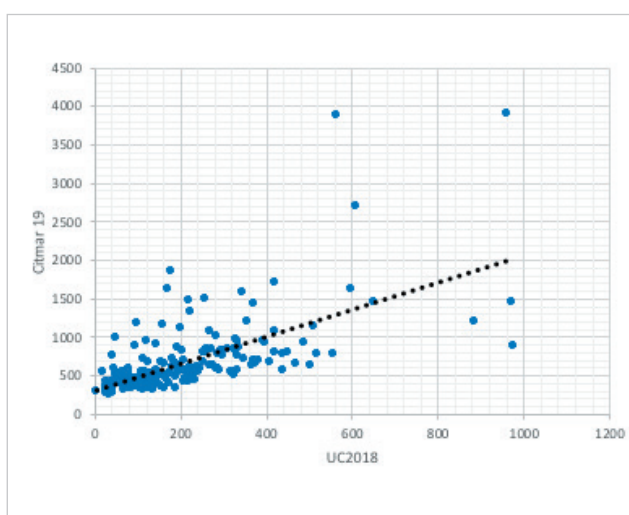


Gráfico 3b. Dispersión UC2018-Citmar 19

Tabla 5. Prueba U de Mann Whitney de diferencia de medias

Variable	U-Mann Whitney Estadístico Z (p-valor)
Citjun16	-2,977 (0,005)*
Citdic18	-0,123 (0,902)
Citmar19	-0,257 (0,797)
UC2018	-4,830 (0,000)*
UC2019	-4,813 (0,000)*

*p < 0,01 (significativo 99%)

se observa una tendencia decreciente en la línea temporal.

Sin embargo, estos resultados no implican que este tipo de correlación exista en otras áreas de conocimiento o incluso para el conjunto total de documentos sobre emprendimiento, ya que este estudio se focaliza en trabajos multidisciplinares, altamente citados, pero de una temática concreta. Para intentar verificar la correlación entre citas y conteos de uso se ha replicado el análisis en otro conjunto de documentos publicados entre 2014

y 2018, recopilado en junio de 2019 de la *Core Collection* de la *Web of Science*. La metodología se basa en la misma raíz de búsqueda que para los trabajos del análisis principal, y fueron seleccionados los 83 documentos más citados en esta fecha ya que el h-core se situaba en este punto⁴. La tabla 7 muestra los resultados de las correlaciones entre las citas de estos documentos en la fecha de su recogida (3 de junio de 2019) y los conteos de uso en los últimos 3 meses y totales desde 2013. Tal como se deduce de los coeficientes de correlación obtenidos, nuevamente se encuentra una relación positiva y significativa entre ambas variables, verificándose un resultado similar al obtenido en la muestra original para este nuevo conjunto de documentos altamente citados sobre emprendimiento.

Tras verificar que los conteos y las citas recibidas se hallan correlacionados, se retorna a la muestra original de 201 documentos y se estima una serie de modelos de regresión, considerando la influencia de los primeros sobre las segundas. Para ello, con el fin de buscar una influencia que tenga sentido, se ha tomado como variable independiente los conteos de uso de los artículos a diciembre de 2018 y se ha analizado su efecto sobre las citas en esta misma fecha y posteriormente en marzo de 2019. La tabla 8 muestra cuatro modelos (dos por cada variable dependiente subdivididos en otros dos según se incluya o no ordenada en el origen). El último modelo es el que mayor sentido bibliométrico tiene, los conteos de uso requieren tiempo para transformarse en citas reales. En cualquier caso, se puede apreciar que el valor de los coeficientes, siempre significativos, es muy similar en las dos variables dependientes (modelo 1 con modelo 3; modelo 2 con modelo 4). Igualmente, dado que los modelos sin ordenada poseen una mejor bondad de ajuste por el mayor valor del coeficiente R² ajustado, será el modelo 4 el que se interprete y utilice en la predicción posterior (tabla 8).

Tabla 8. Modelos de regresión lineal entre citas y conteos de uso

Variable dependiente	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
	Citdic18		Citdic18		Citmar19		Citmar19	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Constante	304,68	8,27*	--	--	310,08	8,21*	--	--
UC2018	1,611	12,54*	2,39	23,66*	1,70	12,89*	2,491	24,09*
R ² ajustado	0,438		0,735		0,452		0,742	

*p < 0,01

De acuerdo con el valor del coeficiente del modelo 4, cada descarga de texto completo o registro en un gestor de referencias de un artículo de la muestra hasta final de 2018 ha supuesto prácticamente 2,5 citas totales recibidas en marzo de 2019.

Si la modelización se realiza en el orden de causalidad inverso, es decir, determinar la influencia de las citas sobre los conteos de uso, que actuarán en este caso como variable dependiente, obtenemos otra serie de modelos (tabla 9). Para mantener el sentido bibliométrico de la influencia se comparan los conteos de uso de 2018 según las citas de 2016 y los conteos de uso 2019, bien según las citas de 2016 o las citas de 2018⁵. Se presentan los resultados de los modelos sin ordenada en el origen pues ofrecen mejor bondad de ajuste.

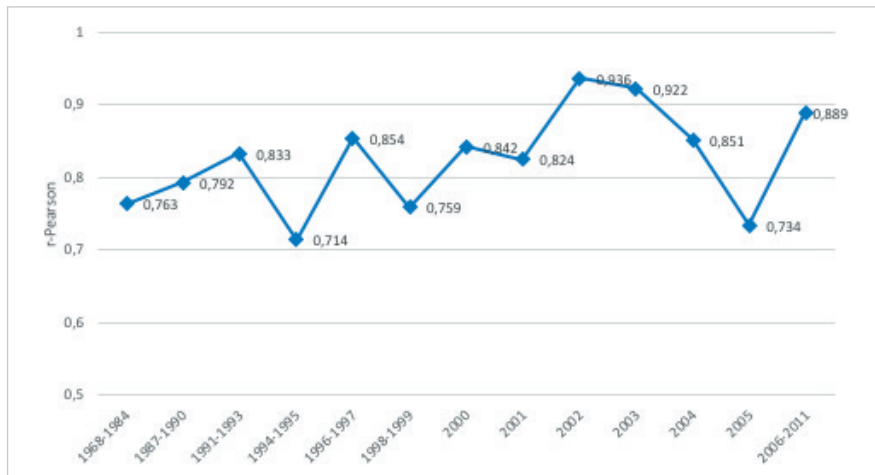


Gráfico 4. Correlación de Pearson entre Citmar2019-UC2019 a lo largo del tiempo

Tabla 7. Coeficiente de correlación de Pearson entre citas y años de exposición para muestra de 83 documentos recopilados en 2019

Citas	Correlación de Pearson	Conteo de uso desde 2013	Conteo de uso últimos 180 días
		0,626*	0,485*
	Sig. (bilateral)	0,000	0,000

*p < 0,01

Tabla 9. Modelos de regresión lineal entre conteos de uso y citas⁶

	Modelo 1	Modelo 2	Modelo 3
Variables dependientes:	UC2018	UC2019	UC2019
Variables independientes:	Coeficientes		
Citjun16	0,480*	0,496*	--
Citdic18	--	--	0,319*
R ² ajustado	0,643	0,638	0,731

*p < 0,01

Los resultados muestran cómo las citas influyen positivamente en los conteos de uso posteriores. Concretamente, las citas recibidas a junio de 2016 generan alrededor de un clic por cada dos citas en diciembre de 2018 y en marzo de 2019. Cuando el período de tiempo transcurrido es menor (citas a diciembre de 2018 sobre conteos de uso de marzo de 2019), la influencia es algo menor (0,319 conteos por cada cita).

Todos los modelos presentan coeficientes significativos y R² ajustado elevado, por lo que no se podría afirmar con rotundidad cuál es el orden de causalidad más correcto entre las citas obtenidas por un artículo y los conteos de uso/ descargas que ha recibido o quién explica a quién realmente.

4.1.3. Predicción de citas a partir de conteos de uso

Una cuestión de interés para los académicos puede centrarse en tratar de predecir si un artículo será más o menos citado en función de las descargas que tenga en un momento del tiempo.

En la muestra obtenida en este estudio, se puede realizar una predicción de las citas estimadas en marzo de 2019 a partir de los conteos de uso alcanzados a final de 2018. Para ello se utiliza el modelo 4 de la tabla 8 en tanto que reflejaba el mejor ajuste. Dado que se poseen los datos de citas reales obtenidas, se pueden comparar éstas con las citas estimadas por el modelo. Se refleja en el gráfico 5 el paralelismo existente a lo largo de los 201 artículos de la muestra.

Incluso las diferencias positivas/negativas observadas ofrece información valiosa, ya que, se podrían identificar y aislar los artículos que presentan mayor desviación con la predicción, agruparse según tipo documental, disciplina, temática y otras características para estudiar los factores que podrían motivarlas.

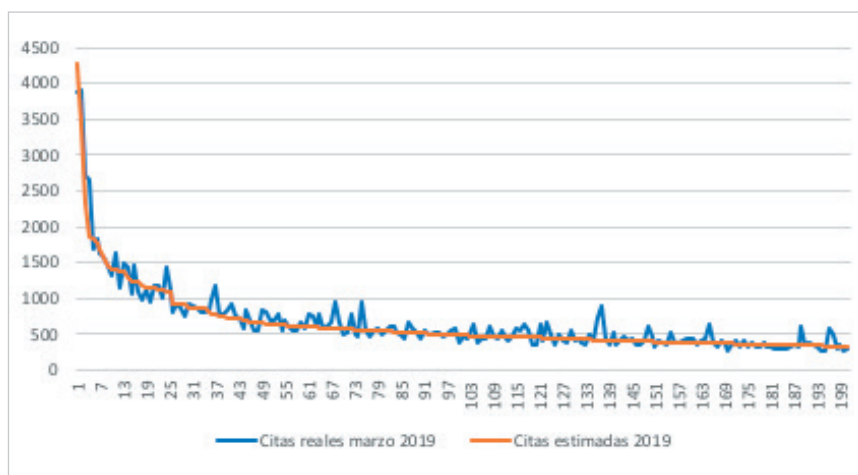


Gráfico 5. Citas recibidas en marzo 2019 vs estimación de citas mediante conteo de uso a diciembre 2018

5. Conclusiones

La utilización de conteos de uso como indicador bibliométrico es una cuestión que debería recibir más atención por parte de los científicos, incluso podría llegar a convertirse en una métrica estratégica complementaria a las tradicionales. El trabajo de Wang, Fang y Sun (2016) señaló que se requiere más experimentación con datos adicionales de diferentes campos, ya que se ha detectado que los documentos altamente citados son los que acumulan más conteos de uso. Los artículos más citados están más consolidados dentro un área de conocimiento, son más visibles, se han convertido en un referente tanto para investigadores noveles como senior, por tanto, tienen un potencial mayor para atraer nuevas citas, vistas y/o descargas. Son especialmente atractivos para detectar frentes de investigación, conocer tendencias y seguir la evolución temática dentro del campo. Por ese motivo, este trabajo se ha centrado en analizar una muestra multidisciplinar de documentos con altas tasas de citación por encima de su h-index.

Los artículos más citados están más consolidados dentro un área de conocimiento, son más visibles, se han convertido en un referente tanto para investigadores noveles como senior, por tanto, tienen un potencial mayor para atraer nuevas citas, vistas y/o descargas

El trabajo realizado revela en primer lugar que:

- La mayor exposición a la cita incide de manera positiva en su evolución en un principio; sin embargo, la correlación con el tiempo va perdiéndose progresivamente arrojando finalmente síntomas de agotamiento u obsolescencia. Los conteos de uso experimentan correlaciones significativas de signo negativo con el transcurso del tiempo a pesar de tratarse de documentos altamente citados.
- Se ha observado que existen correlaciones significativas positivas tanto entre citas pasadas y presentes como entre los conteos de uso y citas. Por tanto, es factible modelizar las citas en función de los conteos de uso, así como la dependencia de estos últimos aplicando las citas recibidas como variable independiente. En ambos casos los modelos poseen un buen ajuste y coeficientes significativos.
- No es posible determinar en esta muestra el orden de causalidad entre citas y conteos de uso a diferencia de otros trabajos como el de **Moed y Halevi (2016)** que concluyen que “las descargas son un buen predictor de las citas y las citas un predictor menos válido de las descargas”.
- Se ha verificado que los conteos de uso en muestras de documentos multidisciplinares muy citados actúan como predictores adecuados del impacto futuro, consiguiéndose un alto paralelismo tras tres meses desde la recogida de los conteos de uso. Incluso las diferencias observadas tienen su utilidad ya que pueden servir para analizar los posibles factores que las provocan.

Este trabajo tiene limitaciones, los hallazgos presentados se relacionan con una muestra particular de únicamente 201 documentos altamente citados, que provienen de diferentes áreas de conocimiento, asociados a una temática concreta: el emprendimiento. Por tanto, podría considerarse más como un estudio de caso, aunque fácilmente replicable y aplicable a otras muestras a través de la metodología de selección de documentos. Por otro lado, los datos de uso se han obtenido del indicador que proporciona la *Web of Science*, que a su vez presenta sus propias limitaciones asociadas con la disponibilidad o no del texto completo alojado en la web de la editorial o vinculadas a la suscripción académica de la institución del investigador. No obstante, el interés que despiertan precisamente los grupos de documentos altamente citados y las posibles implicaciones prácticas en la elaboración de predicciones y/o para el seguimiento de sus desviaciones proporcionan una nueva perspectiva sobre la aplicación práctica de las métricas de uso.

Finalmente, el análisis abre la puerta a futuras investigaciones: aumentar los períodos de exposición de los conteos de uso, observar si se dan correlaciones similares en muestras de documentos de otras temáticas o detectar frentes y tendencias de investigación a partir de predicciones. El uso se puede utilizar como referencia para construir agendas de investigación, así como un nuevo elemento para la evaluación del desempeño temprano de las disciplinas.

6. Notas

1. La fecha de publicación de los documentos de la muestra oscila entre 1968 y 2011 por lo que los años de exposición son diferentes en cada caso.
2. Aunque el valor medio de los años de exposición se sitúa en aproximadamente 20 años de acuerdo con los estadísticos descriptivos vistos en la tabla 1, interesa aquí utilizar el resultado de la mediana en tanto es el valor que divide la muestra en dos grupos con alrededor del 50% de los casos en cada uno.
3. En cualquier caso, el coeficiente no proporciona información acerca de la dirección de la causalidad, dado que, por ejemplo, los conteos de uso en 2018 no pueden influir en las citas que se habían recibido en 2016, sino que el orden causa-efecto bien podría ser el contrario. En otros momentos temporales, en cambio, la influencia se puede producir a la inversa, de manera que los conteos de uso recibidos se van transformando paulatinamente en citas que el artículo obtiene.
4. Muestra completa disponible bajo petición.
5. No se ha modelizado el efecto conjunto de dos variables independientes puesto que se daría un problema claro de multicolinealidad al estar relacionadas entre sí Citjun16, Citdic18 y Citmar19 como ya se comprobó en la tabla 6. Habría que realizar una transformación en una sola variable como producto de las independientes, con lo que la interpretación del coeficiente resultante del modelo de regresión lineal perdería su sentido.
6. No se presentan los valores del estadístico t-student para simplificar la información, dado que todos los coeficientes obtenidos han sido significativos.

7. Referencias

- Bollen, Johan; Van-de-Sompel, Herbert; Smith, Joan A.; Luce, Rick (2005).** “Toward alternative metrics of journal impact: A comparison of download and citation data”. *Information processing & management*, v. 41, n. 6, pp. 1419-1440. <https://doi.org/10.1016/j.ipm.2005.03.024>
- Borgman, Christine L.; Furner, Jonathan (2002).** “Scholarly communication and bibliometrics”. *Annual review of information science and technology*, v. 36, pp. 3-72. <https://doi.org/10.1002/aris.1440360102>

- Brody, Tim; Harnad, Stevan; Carr, Leslie** (2006). "Earlier web usage statistics as predictors of later citation impact". *Journal of the American Society for Information Science and Technology*, v. 57, n. 8, pp. 1060-1072.
<https://doi.org/10.1002/asi.20373>
- Cabezas-Clavijo, Álvaro; Torres-Salinas, Daniel** (2010). "Indicadores de uso y participación en las revistas científicas 2.0: el caso de *PLoS one*". *El profesional de la información*, v. 19, n. 4, pp. 431-434.
<https://doi.org/10.3145/epi.2010.jul.14>
- Chen, Bikun** (2018). "Usage pattern comparison of the same scholarly articles between Web of Science (WoS) and Springer". *Scientometrics*, v. 115, n. 1, pp. 519-537.
<https://doi.org/10.1007/s11192-017-2616-3>
- Chi, Pei-Shan; Glänzel, Wolfgang** (2017). "An empirical investigation of the associations among usage, scientific collaboration and citation impact". *Scientometrics*, v. 112, n. 1, pp. 403-412.
<https://doi.org/10.1007/s11192-017-2356-4>
- Chi, Pei-Shan; Glänzel, Wolfgang** (2018). "Comparison of citation and usage indicators in research assessment in scientific disciplines and journals". *Scientometrics*, v. 116, n. 1, pp. 537-554.
<https://doi.org/10.1007/s11192-018-2708-8>
- Chu, Heting; Krichel, Thomas** (2007). "Downloads vs. citations in economics: Relationships, contributing factors and beyond". In: *Proceedings of Issi 2007: 11th Intl conf of the International Society for Scientometrics and Informetrics*, Vols. I and II. Leuven: Int Soc Scientometrics & Informetrics-Issi.
<http://eprints.rclis.org/11085>
- Costas, Rodrigo; Zahedi, Zohreh; Wouters, Paul** (2015). "Do 'altmetrics' correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective". *Journal of the Association for Information Science and Technology*, v. 66, n. 10, pp. 2003-2019.
<https://doi.org/10.1002/asi.23309>
- Davis, Philip M.; Price, Jason S.** (2006). "eJournal interface can influence usage statistics: Implications for libraries, publishers, and project Counter". *Journal of the American Society for Information Science and Technology*, v. 57, n. 9, pp. 1243-1248.
<https://doi.org/10.1002/asi.20405>
- De-Sordi, José-Osvaldo; Conejero, Marco-Antonio; Meireles, Manuel** (2016). "Bibliometric indicators in the context of regional repositories: Proposing the D-index". *Scientometrics*, v. 107, n. 1, pp. 235-258.
<https://doi.org/10.1007/s11192-016-1873-x>
- Duy, Joanna; Vaughan, Liwen** (2006). "Can electronic journal usage data replace citation data as a measure of journal use? An empirical examination". *Journal of academic librarianship*, v. 32, n. 5, pp. 512-517.
<https://doi.org/10.1016/j.acalib.2006.05.005>
- Garfield, Eugene** (1977). "Introducing citation classics-human side of scientific reports". *Essays of an information scientist*, v. 3, pp. 1-2.
<http://www.garfield.library.upenn.edu/essays/v3p001y1977-78.pdf>
- Glänzel, Wolfgang; Gorraiz, Juan** (2015). "Usage metrics versus altmetrics: confusing terminology?". *Scientometrics*, v. 102, n. 3, pp. 2161-2164.
<https://doi.org/10.1007/s11192-014-1472-7>
- Gregory, Ann T.; Denniss, A. Robert** (2016). "Impact by citations and downloads: What are *Heart, lung and circulation's* top 25 articles of all time?". *Heart lung and circulation*, v. 25, n. 8, pp. 743-749.
<https://doi.org/10.1016/j.hlc.2016.05.108>
- Henneken, Edwin A.; Kurtz, Michael J.** (2017). "Usage bibliometrics as a tool to measure research activity". *arXiv preprint arXiv:1706.02153*.
<https://arxiv.org/abs/1706.02153>
- Hirsch, Jorge E.** (2005). "An index to quantify an individual's scientific research output". *Proceedings of the National Academy of Sciences of the United States of America*, v. 102, n. 46, pp. 16569-16572.
<https://doi.org/10.1073/pnas.0507655102>
- Khan, Muhammad-Salman; Younas, Muhammad** (2017). "Analyzing readers behavior in downloading articles from IEEE digital library: A study of two selected journals in the field of education". *Scientometrics*, v. 110, n. 3, pp. 1523-1537.
<https://doi.org/10.1007/s11192-016-2232-7>
- Kurtz, Michael J.; Bollen, Johan** (2010). "Usage bibliometrics". *Annual review of information science and technology*, n. 44, pp. 3-64.
<https://doi.org/10.1002/aris.2010.1440440108>

- Kurtz, Michael J.; Eichhorn, Guenther; Accomazzi, Alberto; Grant, Carolyn; Demleitner, Markus; Murray, Stephen; Martimbeau, Nathalie; Elwell, Barbara** (2005). "The bibliometric properties of article readership information". *Journal of the American Society for Information Science and Technology*, v. 56, n. 2, pp. 111-128.
<https://doi.org/10.1002/asi.20096>
- Li, Zhi; Wu, Fei-Xiang; Yang, Li-Qun; Sun, Yu-Ming; Lu, Zhi-Jie; Yu, Wei-Feng** (2012). "Citation classics in main pain research journals". *Journal of anesthesia*, v. 26, n. 1, pp. 85-93.
<https://doi.org/10.1007/s00540-011-1254-9>
- Line, Maurice B.; Sandison, Alexander** (1975). "Practical interpretation of citation and library use studies". *College & research libraries*, v. 36, n. 5, pp. 393-396.
https://doi.org/10.5860/crl_36_05_393
- Liu, Xue-Li; Fang, Hong-Ling; Wang, Mei-Ying** (2011). "Correlation between download and citation and download-citation deviation phenomenon for some papers in Chinese medical journals". *Serials review*, v. 37, n. 3, pp. 157-161.
<https://doi.org/10.1016/j.serrev.2011.02.001>
- Mafahi, Nabeil; Thelwall, Mike** (2016). "When are readership counts as useful as citation counts? Scopus versus Mendeley for LIS journals". *Journal of the Association for Information Science and Technology*, v. 67, n. 1, pp. 191-199.
<https://doi.org/10.1002/asi.23369>
- Markusova, Valentina; Bogorov, Valentin; Libkind, Alexander** (2018). "Usage metrics vs classical metrics: Analysis of Russia's research output". *Scientometrics*, v. 114, n. 2, pp. 593-603.
<https://doi.org/10.1007/s11192-017-2597-2>
- Martínez-Sánchez, María-Ángeles; Herrera, Manuel; Jiménez-Contreras, Evaristo; Ruiz-Rodríguez, Antonio-Ángel; Herrera-Viedma, Enrique** (2015). "Characterizing highly cited papers in social work through H-classics". *Scientometrics*, v. 102, n. 2, pp. 1713-1729.
<https://doi.org/10.1007/s11192-014-1460-y>
- Martínez-Sánchez, María-Ángeles; Herrera, Manuel; Lopez-Gijon, Javier; Herrera-Viedma, Enrique** (2014). "H-classics: Characterizing the concept of citation classics through H-index". *Scientometrics*, v. 98, n. 3, pp. 1971-1983.
<https://doi.org/10.1007/s11192-013-1155-9>
- Moed, Henk F.** (2005). "Statistical relationships between downloads and citations at the level of individual documents within a single journal". *Journal of the American Society for Information Science and Technology*, v. 56, n. 10, pp. 1088-1097.
<https://doi.org/10.1002/asi.20200>
- Moed, Henk F.; Halevi, Gali** (2016). "On full text download and citation distributions in scientific-scholarly journals". *Journal of the Association for Information Science and Technology*, v. 67, n. 2, pp. 412-431.
<https://doi.org/10.1002/asi.23405>
- Naude, Filistea** (2017). "Comparing downloads, Mendeley readership and Google Scholar citations as indicators of article performance". *Electronic journal of information systems in developing countries*, v. 78, n. 1, pp. 1-25.
<https://doi.org/10.1002/j.1681-4835.2017.tb00572.x>
- Nieder, Carsten; Dalhaug, Astrid; Aandahl, Gro** (2013). "Correlation between article download and citation figures for highly accessed articles from five open access oncology journals". *Springer plus*, v. 2, art. 261.
<https://doi.org/10.1186/2193-1801-2-261>
- O'Leary, Dan** (2008). "On the relationship between citations and appearances on 'top 25' download lists in the *International journal of accounting information systems*". *International journal of accounting information systems*, v. 9, n. 1, pp. 61-75.
<https://doi.org/10.1016/j.accinf.2008.02.001>
- Ortega, José-Luis** (2015). "Relationship between altmetric and bibliometric indicators across academic social sites: The case of CSIC's members". *Journal of informetrics*, v. 9, n. 1, pp. 39-49.
<https://doi.org/10.1016/j.joi.2014.11.004>
- Perneger, Thomas** (2004). "Relation between online 'hit counts' and subsequent citations: Prospective study of research papers in the BMJ". *British medical journal*, v. 329, n. 7465, pp. 546-547.
<https://doi.org/10.1136/bmj.329.7465.546>
- Priem, Jason; Hemminger, Bradely H.** (2010). "Scientometrics 2.0: New metrics of scholarly impact on the social Web". *First Monday*, v. 15, n. 7.
<https://doi.org/10.5210/fm.v15i7.2874>
- Rousseau, Ronald** (2006). "New developments related to the Hirsch index". Preprint.
<http://eprints.rclis.org/7616>

- Sanz-Casado, Elías; García-Zorita, Carlos; Rousseau, Ronald** (2016). "Using h-cores to study the most-cited articles of the twenty-first century". *Scientometrics*, v. 108, n. 1, pp. 243-261.
<https://doi.org/10.1007/s11192-016-1956-8>
- Schlögl, Christian; Gorraiz, Juan** (2010). "Comparison of citation and usage indicators: The case of oncology journals". *Scientometrics*, v. 82, n. 3, pp. 567-580.
<https://doi.org/10.1007/s11192-010-0172-1>
- Schlögl, Christian; Gorraiz, Juan; Gumpenberger, Christian; Jack, Kris; Kraker, Peter** (2014). "Comparison of downloads, citations and readership data for two information systems journals". *Scientometrics*, v. 101, n. 2, pp. 1113-1128.
<https://doi.org/10.1007/s11192-014-1365-9>
- Subotic, Sinisa; Mukherjee, Bhaskar** (2014). "Short and amusing: The relationship between title characteristics, downloads, and citations in psychology articles". *Journal of information science*, v. 40, n. 1, pp. 115-124.
<https://doi.org/10.1177/0165551513511393>
- Taylor, Mike** (2013). "Exploring the boundaries: How altmetrics can expand our vision of scholarly communication and social impact". *Information standards quarterly*, v. 25, n. 2, pp. 27-32.
<https://scite.ai/reports/10.3789/isqv25no2.2013.05>
- Thelwall, Mike** (2018). "Early Mendeley readers correlate with later citation counts". *Scientometrics*, v. 115, n. 3, pp. 1231-1240.
<https://doi.org/10.1007/s11192-018-2715-9>
- Thelwall, Mike; Kousha, Kayvan** (2015). "Web indicators for research evaluation. Part 1: citations and links to academic articles form the web". *El profesional de la información*, v. 24, n. 5, pp. 587-606.
<https://doi.org/10.3145/epi.2015.sep.08>
- Torres-Salinas, Daniel; Cabezas-Clavijo, Álvaro; Jiménez-Contreras, Evaristo** (2013). "Altmetrics: New indicators for scientific communication in Web 2.0". *Comunicar*, v. 21, n. 41, pp. 53-60.
<https://doi.org/10.3916/C41-2013-05>
- Van-Raan, Anthony F. J.** (2006). "Statistical properties of bibliometric indicators: Research group indicator distributions and correlations". *Journal of the American Society for Information Science and Technology*, v. 57, n. 3, pp. 408-430.
<https://doi.org/10.1002/asi.20284>
- Vaughan, Liwen; Tang, Juan; Yang, Rongbin** (2017). "Investigating disciplinary differences in the relationships between citations and downloads". *Scientometrics*, v. 111, n. 3, pp. 1533-1545.
<https://doi.org/10.1007/s11192-017-2308-z>
- Wan, Jin-Kun; Hua, Ping-Huan; Rousseau, Ronald; Sun, Xiu-Kun** (2010). "The journal download immediacy index (DII): Experiences using a Chinese full-text database". *Scientometrics*, v. 82, n. 3, pp. 555-566.
<https://doi.org/10.1007/s11192-010-0171-2>
- Wang, Xianwen; Fang, Zhichao; Sun, Xiaoling** (2016). "Usage patterns of scholarly articles on Web of Science: A study on Web of Science usage count". *Scientometrics*, v. 109, n. 2, pp. 917-926.
<https://doi.org/10.1007/s11192-016-2093-0>
- Wang, Xianwen; Mao, Wenli; Xu, Shenmeng; Zhang, Chunbo** (2014a). "Usage history of scientific literature: Nature metrics and metrics of Nature publications". *Scientometrics*, v. 98, n. 3, pp. 1923-1933.
<https://doi.org/10.1007/s11192-013-1167-5>
- Wang, Xianwen; Wang, Zhi; Mao, Wenli; Liu, Che** (2014b). "How far does scientific community look back?". *Journal of informetrics*, v. 8, n. 3, pp. 562-568.
<https://doi.org/10.1016/j.joi.2014.04.009>
- Wang, Xianwen; Wang, Zhi; Xu, Shenmeng** (2013). "Tracing scientist's research trends realtimely". *Scientometrics*, v. 95, n. 2, pp. 717-729.
<https://doi.org/10.1007/s11192-012-0884-5>
- Watson, Andrew B.** (2009). "Comparing citations and downloads for individual articles at the *Journal of vision*". *Journal of vision*, v. 9, n. 4, pp. 4.
<https://doi.org/10.1167/9.4.i>
- Xing, Zhao** (2017). "Exploring the features of usage data for academic literatures". *Journal of library science in China*, v. 3, n. 5.
http://en.cnki.com.cn/Article_en/CJFDTOTAL-ZGTS201703005.htm
- Zhang, Lin; Glänzel, Wolfgang; Ye, Fred Y.** (2016). "The dynamic evolution of core documents: An experimental study based on h-related literature (2005-2013)". *Scientometrics*, v. 106, n. 1, pp. 369-381.
<https://doi.org/10.1007/s11192-015-1705-4>