



# Use of scientific journals in Spanish universities: analysis of the relationship between citations and downloads in two university library consortia

Andrés Fernández-Ramos<sup>1</sup> · Blanca Rodríguez-Bravo<sup>1</sup> · Ángela Díez-Díez<sup>2</sup>

Received: 20 July 2022 / Accepted: 16 February 2023 / Published online: 7 March 2023  
© The Author(s) 2023

## Abstract

We analyse the relationship between downloads of electronic journals included in four big deal bundles subscribed to by public university libraries affiliated to two library consortia in Spain (Castile and León and Galicia) and citations of the same journals by researchers at these universities. Download data on the big deals analysed (Emerald, ScienceDirect, Springer and Wiley) were obtained from COUNTER Journal Reports 1, and citation data were obtained from the bibliographic references given in articles indexed in Scopus between 2010 and 2017. The results show that only a low percentage of the subscribed journals was used in the scientific output of the universities' researchers, with values ranging from 15 to 50%, and that there was a strong correlation between the universities' volume of scientific production and the percentage of cited journals. We also found a strong correlation between downloads and citations, which was higher in the case of universities with a higher scientific output.

**Keywords** Big deals · Citations · Downloads · Scientific production · Electronic journals · University libraries

## Introduction

The growing demand for accountability in the use of public and private funds renders it increasingly important to measure the use and impact of scientific publications. Academic libraries must present quantitative data demonstrating the value of journal subscriptions in order to justify budgetary allocation to these, but journal use is difficult to observe directly

---

✉ Andrés Fernández-Ramos  
afernr@unileon.es

Blanca Rodríguez-Bravo  
blanca.rodriguez@unileon.es

Ángela Díez-Díez  
angela.diez.diez@unileon.es

<sup>1</sup> Department of Library and Information Science, University of León, León, Spain

<sup>2</sup> Department of Systems Engineering and Automation, University of León, León, Spain

in its full magnitude and, therefore, to quantify objectively (Chew et al., 2016). Various methods have been proposed for this, ranging from measuring downloads of electronic subscriptions (Duan & Xiong, 2017; Fernández-Ramos et al., 2019) to analysing bibliographic references in researchers' scientific output (Peñaflor & Aliwalas, 2022; Vaaler, 2018; Wilson & Tenopir, 2008). It is also common to combine these methods with cost indicators (Gumpenberger et al., 2012; Kurtz & Bollen, 2010). Despite the utility of each of these methods, they present limitations when employed in isolation, and only provide a partial view of the use and usefulness of collections.

As regards the analysis of downloads and citations to evaluate collections, Ivanov et al. (2020) view these as complementary indicators of a journal's intellectual value (identifying the frequency with which its articles are cited) and of a publication's usefulness (identifying the frequency with which a journal's articles are consulted and downloaded). Martin et al. (2016) stress, however, that the two metrics are not the same and therefore not comparable, because downloading an article requires less effort than citing an article. Thus, the number of downloads of a widely-used title is likely to be much higher than the number of citations of a frequently cited article (Chu & Krichel, 2007; Gorraiz et al., 2014; Moed, 2005; Wan et al., 2010; Watson, 2009).

It should also be borne in mind that not all content is downloaded for research purposes; some may be downloaded just for information by professionals, or for learning purposes by teachers or students (Gorraiz et al., 2014; Martin et al., 2016). As a logical consequence, articles are often downloaded many times but remain uncited. Thus, on the one hand, downloads may not be a perfect proxy to estimate the overall usage, but they measure at least the intention to use the downloaded material. On the other hand, many citations are included in the lists of references without previous reading of the cited document and citations can just measure the impact in the "publish or perish" community (Gorraiz et al., 2014). Accordingly, usage metrics can be regarded as complementary to citation metrics (Bollen et al., 2005; Chi & Glänzel, 2017; Hitchcock et al., 2003; O'Leary, 2008).

Furthermore, although both metrics change over time, their evolution is not necessarily parallel because citation of an article usually occurs some time after its download due to the interval between consulting it and citing it in a subsequent publication (Wan et al., 2010; Watson, 2009). Meanwhile, Vogl et al. (2018) have suggested that because citations increase over time, they may be the best indicator of an article's quality. In contrast, downloads and other alternative metrics have a shorter half-life, tending to stagnate after publication, and therefore measure immediate influence. However, both measures influenced each other. Early downloads are a predictor of citations and downloads give an idea of the potential of a paper (Bollen et al., 2005; McDonald (2007); and citations influenced latter downloads (Moed, 2005; Watson, 2009). These circumstances condition the correlations between the two variables, which will not always be high, as pointed out by Coats (2008), who highlighted the lack of consensus about the value of an article.

Martin et al. (2016) have observed that although the literature abounds in studies analysing the use of journal subscriptions on the basis of either download data or citations in the scientific output of researchers, far fewer studies have combined both types of data. However, this type of analysis is highly important because it provides a more complete picture of the usefulness of collections in institutions and minimises the limitations and partial view offered by the isolated use of citation or download data.

Thus, in the context of an institution or a group of institutions, a joint analysis of download and citation data would help determine whether there is a relationship between these two variables in such a way that one predicts the other. Examples of this combined use of data include Wical and Vandembark's (2014) study at the University of Wisconsin-Eau

Claire and Faulkner's (2021) study at the Psychology Department of California State University. In both cases, the authors indicated that the results would be used to make decisions regarding journal subscriptions.

Several studies in the specialised literature have analysed this relationship, but the results obtained have been mixed. Before the existence of standardised usage statistics (COUNTER), Tsay (1998) compared the use of journals in a medical library with citations by researchers at the institution over the same period, and found a statistically significant relationship between frequency of use and the number of medical science journal citations. Another early study suggesting a correlation between citations and other measures of journal usage was the one conducted by Blečić (1999) in the health sciences, at the University of Illinois (Chicago). Similarly, after reviewing COUNTER statistics for the California Institute of Technology, McDonald (2007) reported that the use of online journals was a significant variable in predicting citation patterns. Other studies in which positive correlations have been found were those conducted by Feyereisen and Spoiden (2009) in the Department of Psychology and Education Science at the University of Louvain, and by Gumpenberger et al. (2012) at the University of Vienna.

More recently, Wood-Doughty et al. (2019) analysed this association at the ten universities belonging to the University of California System, studying the scientific output of their researchers between 2010 and 2016. They found a positive correlation between the two variables, but with small differences depending on subject area. Other studies that have reported positive correlations include Rodríguez-Bravo et al. (2021), who analysed scientific production on Library and Information Science at universities in Castile and León (Spain), and De Groote et al. (2013), who analysed scientific production in medicine at the University of Illinois (Chicago). In contrast, studies by Gao (2016) at the University of Houston School of Communication, Ke and Bronicki (2015), also at the University of Houston but in the field of psychology, and Fernández-Ramos et al. (2022) in the same field and limited to the university library consortium of Castile and León (Spain), found no significant correlation between citations and downloads.

Besides the reasons given by Vogl et al. (2018), other explanations for this disparity in the results of studies analysing the relationship between citations and downloads include the characteristics of each institution and its users, with very different citation patterns depending on the discipline, and the method employed in each study. The correlation between citation and usage data depends on the discipline's publication output as documented in previous studies focused on particular disciplines, journals (Coats, 2008; Moed, 2005; O'Leary, 2008; Watson, 2009) or platforms (Bollen et al., 2005; Brody et al., 2006; Chu & Krichel, 2007; Wan et al., 2010).

Besides, although relationships have generally been measured using the same method, not all studies have examined the totality of downloads and citations or used the same sampling technique. Thus, for example, in a study conducted at the Galter Health Sciences Library in Chicago analysing the correlation between downloads and citations of dermatology publications issued between 2007 and 2016, Pastva et al. (2018) found that the results obtained when including all the most frequently cited journals differed from those obtained when journals from other disciplines were excluded. Similarly, in the studies by Rodríguez-Bravo et al. (2021) and Fernández-Ramos et al. (2022), the correlation coefficient increased significantly when only discipline-specific journals were included.

Meanwhile, a recent analysis of downloads of journals included in the main big deals (ScienceDirect, Emerald, Springer and Wiley) subscribed to by the Castile and León consortium found that downloads of bundle titles had risen in recent years (2012–2018) (Fernández-Ramos et al., 2019), despite a parallel decline in the number of teaching staff

and students over the same period and despite the proliferation of open access journals, repositories, academic social networks and platforms such as Sci-Hub, which are opening new and increasingly important avenues of access to scientific information for the academic community. However, the same study also found that only a limited number of bundle titles was being used in the consortium universities, with a small number of titles accounting for the majority of downloads.

We believe that the rise in downloads reported in this study is related to the convenient—transparent and direct—access that researchers have to subscribed resources and that the still significant use of subscribed journals is likely to lead to an increase in consultation and citation of articles from these journals, which strongly suggests a need to determine whether there is a relationship between downloads and their citation in scientific output. Thus, the aim of the present study was to ascertain the degree of relationship between downloads of journals subscribed to by the seven universities that make up two consortia of university libraries in Spain (Castile and León and Galicia) and the citation of these journals in the bibliographic references provided in scientific production by these universities' researchers, limiting the study to articles indexed in Scopus over the period 2010–2017.

## Methods

We used an observational and quantitative method to achieve the proposed objectives. Thus, we obtained data on downloads of scientific journals subscribed to by the university libraries included in the study; we searched Scopus for the scientific output from these seven universities, downloading and normalising all relevant bibliographic records from Scopus; we extracted and analysed the bibliographic references included in these records; and we compared downloads of the subscribed journals against their citation in the bibliographic references given in scientific production by researchers at the universities included in the study. Figure 1 depicts the stages included in the research.

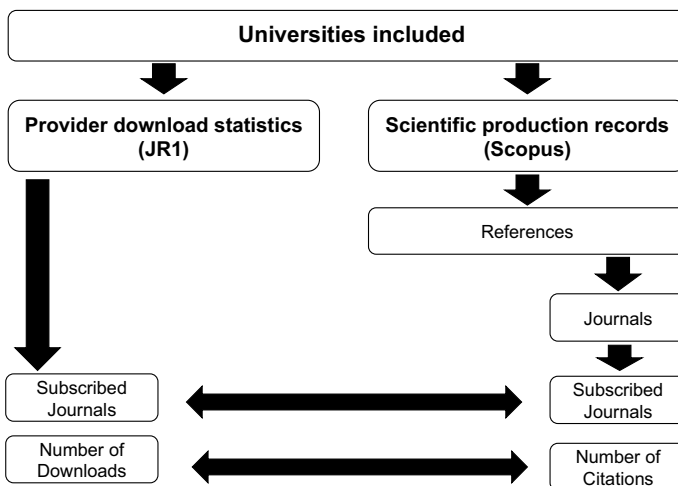


Fig. 1 Research stages

## Download Data Collection

Download data were obtained and standardised for journals included in the Emerald, ScienceDirect, Springer and Wiley bundles subscribed to by the two consortia included in this study, for the study period 2010–2016. This information was provided to us by the participating libraries based on the COUNTER Journal Report 1 (JR1), disaggregated by year, university and provider.

## Search and Download of Scientific Output

We searched Scopus and downloaded indexed scientific production published between 2010 and 2017 and written by researchers from the seven public universities that make up the two consortia of *Castile and León* and *Galicia* (see Table 1). Given the time lag between downloading an article and subsequently citing it, an additional year was considered in the case of scientific production. The search was conducted in July 2018 for each of the seven universities using the university name in the “Affiliation” field. Records were downloaded in.csv format and then imported into Excel and analysed as described below.

## Analysis of Bibliographic References

Bibliographic references were extracted from the “References” field of each of the downloaded scientific production records, and a database was created in which the references corresponding to journal articles were standardised and purified. This process was semi-automated, using an algorithm designed to identify references to journal articles by analysing the structure of the bibliographic references and locating the journal title. However, lack of standardisation rendered it necessary to conduct manual checking of errors in the references (e.g. modifying references written in Chicago style) and ambiguities in some journal names (journals with abbreviated or expanded titles, subtitles or words preceding the journal name). Subsequently, bibliographic references were counted for each journal.

**Table 1** Universities included in the study

University Name	Acronym	Webpage	Consortium
University of Burgos	UBU	<a href="https://www.ubu.es/">https://www.ubu.es/</a>	Castile and León
University of León	ULE	<a href="https://www.unileon.es/">https://www.unileon.es/</a>	Castile and León
University of Valladolid	UVA	<a href="https://www.uva.es/">https://www.uva.es/</a>	Castile and León
University of Salamanca	USAL	<a href="https://www.usal.es/">https://www.usal.es/</a>	Castile and León
University of A Coruña	UDC	<a href="https://www.udc.es/">https://www.udc.es/</a>	Galicia
University of Vigo	UVI	<a href="https://www.uvigo.gal/">https://www.uvigo.gal/</a>	Galicia
University of Santiago de Compostela	USC	<a href="https://www.usc.gal/es">https://www.usc.gal/es</a>	Galicia

## Analysis of the Relationship Between Citations and Downloads

Once the bibliographic references corresponding to journals had been identified, these were matched against the list of journals included in the four big deals mentioned in stage one of the research, for each of the universities. This enabled us to select those references that corresponded to subscribed and cited journals. We used this information to create a table containing the citation and download data for the subscribed journals that had been cited in the scientific output from these seven universities between 2010 and 2017.

These data were then used to calculate the percentage of subscribed journals that had been cited in the study period and the volume of citations corresponding to subscribed journals. To analyse the relationships between citations and downloads (of cited and subscribed journals), we generated scatter plots of the correlations between the two variables. These plots enabled us to identify a series of outliers that might distort the results, and we eliminated all those assigned an anomaly index over 100 by SPSS (v 26). Once these values (which were less than 0.001 of the total) had been removed, Pearson's correlation coefficients were calculated to test the correlation between citations and downloads for each of the seven universities. These coefficients were obtained separately for the following conditions: using data for all subscribed journals at each university, and using only data for subscribed journals that had been cited at least once.

## Discussion of Results

### Scientific Production and Bibliographical References

Table 2 shows the universities' scientific production indexed in Scopus. As can be seen, there was a sustained increase in the volume of publications over the study period, albeit with some differences in the universities' scientific output, which were mainly due to disparities in university size in terms of the number of students and—above all—researchers at each university (Table 3). Thus, the University of Santiago de Compostela was the most productive, while the universities of Burgos and León, which had the fewest students and teaching staff, presented the lowest level of scientific output.

As expected, the bibliographical references cited in publications mainly corresponded to scientific journal articles. Although some differences were detected between universities, ranging from 73.97 to 80.97% of references (Table 4), they were not

**Table 2** Scientific production at the universities included in the study

University	2010	2011	2012	2013	2014	2015	2016	2017	Total
UBU	200	208	219	249	281	304	319	336	2116
ULE	321	360	464	468	464	499	482	543	3601
UVA	896	978	1071	1072	1074	1005	1032	1201	8329
USAL	1175	1270	1365	1453	1623	1612	1678	1783	11,959
UDC	773	791	879	956	893	948	1055	1072	7367
UVI	1204	1294	1301	1374	1401	1363	1407	1457	10,801
USC	1651	1818	1955	1923	1957	1963	2015	2141	15,423
Total	6220	6719	7254	7495	7693	7694	7988	8533	59,596

**Table 3** Number of researchers

University	2010	2011	2012	2013	2014	2015	2016	2017
UBU	692	747	762	733	751	777	776	803
ULE	952	920	942	908	854	883	900	925
UVA	2645	2590	2473	2029	2199	2260	2319	2314
USAL	2481	2469	2425	2334	2306	2312	2301	2272
UDC	1478	1551	1541	1526	1500	1488	1501	1493
UVI	1588	1756	1749	1650	1600	1631	1572	1573
USC	2307	2335	2301	2188	2158	2164	2109	2127

**Table 4** Distribution of bibliographic references

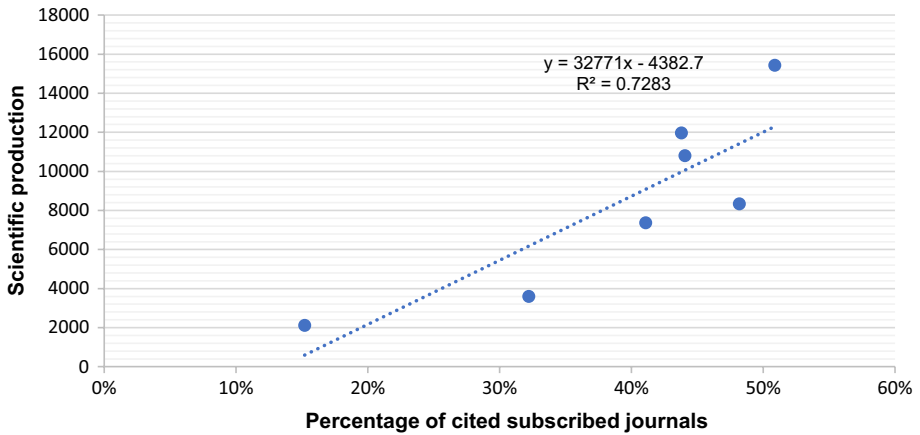
University	Castile and León				Galicia		
	UBU	ULE	UVA	USAL	UDC	UVIGO	USC
Scientific production	2116	3601	8329	11,959	7367	10,801	15,423
References	98,469	153,167	318,476	472,375	273,585	430,968	673,207
Cites to journals	77,799	122,407	239,216	382,492	203,667	318,775	533,113
Percentage of cites to journals	79.01%	79.92%	75.11%	80.97%	74.44%	73.97%	79.19%

**Table 5** Use of subscribed journals

University	Castile and León				Galicia		
	UBU	ULE	UVA	USAL	UDC	UVIGO	USC
Subscribed journals	12,390	11,092	10,502	11,261	9916	10,249	10,011
Cited subscribed journals	1884	3575	5061	4932	4074	4517	5095
Percentage of cited sub- scribed journals	15.21%	32.23%	48.19%	43.80%	41.09%	44.07%	50.89%

particularly significant. These small variations might be due to greater specialisation in one or another subject area at each university. It is well known that not all disciplines present the same citation patterns and that some disciplines primarily use scientific journals, as in the case of the health sciences (Larivière et al., 2006; Tucker, 2013), whereas others rely more heavily on books and book chapters, as in the case of the humanities (Arakaki, 2018; Ezema & Asogwa, 2014), or on conference proceedings, as in the case of engineering (Zhang, 2018).

One of the most striking results of the study was the limited percentage of the subscribed journals included in this study that was cited in the researchers’ scientific production, as can be seen in Table 5. Although there were differences between universities, ranging from 15% at the University of Burgos to more than 50% at the University of Santiago de Compostela, in general we found that a high percentage of the scientific journals subscribed to were not cited by researchers in their publications for a period of time as long as eight years. These results are in line with those of other studies, such as Fernández-Ramos et al. (2022) and Shu et al. (2018), and also agree with studies reporting that many of the



**Fig. 2** Relationship between volume of scientific production and percentage of subscribed journals cited

**Table 6** Downloads corresponding to cited journals

University	Castile and León				Galicia		
	UBU	ULE	UVA	USAL	UDC	UVIGO	USC
Downloads	873,629	1,158,358	1,892,939	2,398,031	1,251,356	2,481,787	3,758,106
Downloads from cited journals	627,096	1,024,901	1,788,365	2,168,919	1,096,435	2,283,772	3,481,811
Percentage of downloads from cited journals	71.78%	88.48%	94.48%	90.45%	87.62%	92.02%	92.65%

journals subscribed to through big deals are rarely if ever downloaded (Fernández-Ramos et al., 2019; Srivastava & Kumar, 2018; Zhu & Xiang, 2016).

Predictably, citation data for subscribed journals are closely related to volume of scientific output. As can be seen in Fig. 2, there was a strong correlation between the two variables: the higher the scientific output, the higher the percentage of subscribed journals that were cited, since the more articles published, the greater the chances of citing any of the subscribed journals (and other non-subscribed journals).

### Relationship Between Citations and Downloads

Most downloads of subscribed journals corresponded to journals that had been cited (at least once in the study period), as can be seen in Table 6, which shows percentages of around 90% for most universities. The exception was the University of Burgos, with a percentage of 71.78%, which, as can be seen in Table 1, was the university with the lowest scientific production.

Our results showed a strong correlation between citations and downloads in the universities analysed; however, as in the case of the number of journals cited, this correlation was not the same for all universities, being greater in the case of universities with a higher scientific output. Table 7 shows the Pearson’s correlation coefficients for each



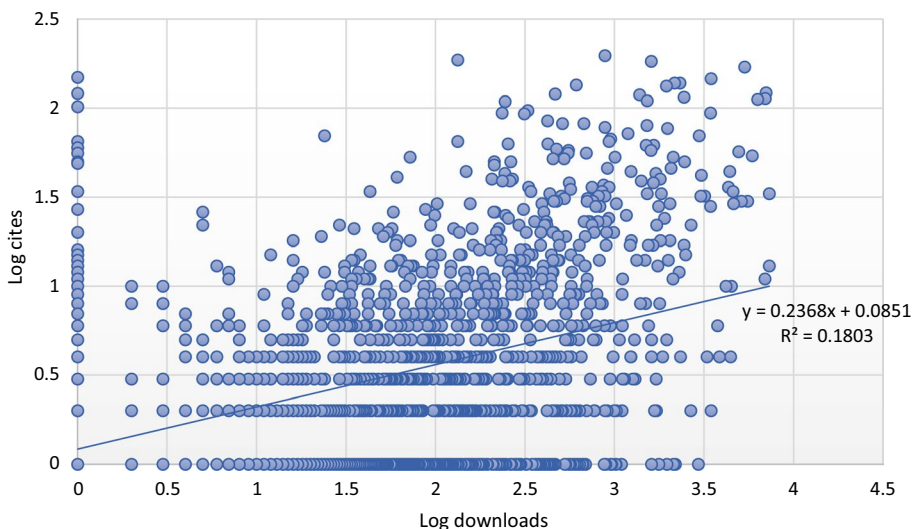
**Table 7** Correlations between citations and downloads according to Pearson’s correlation coefficient

University	Castile and León				Galicia		
	UBU	ULE	UVA	USAL	UDC	UVIGO	USC
All cited journals	0.443	0.716	0.637	0.572	0.671	0.691	0.773
Cited subscribed journals	0.432	0.695	0.616	0.546	0.659	0.679	0.778

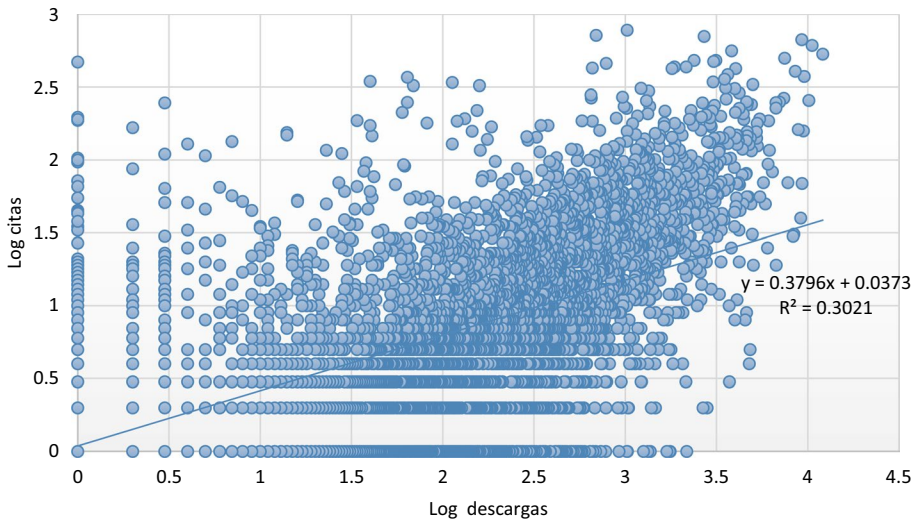
of the universities analysed, giving the correlations between citations and downloads separately for analyses that included (1) all subscribed journals, and (2) only journals that had been cited at least once. We found a slightly higher correlation when all subscribed journals were included than when only cited subscribed journals were included. The probable explanation for this finding is that many journals are neither cited nor downloaded.

The figures below show the dispersion of citation and download values for the journals with at least one citation subscribed to by the universities included in the study, ranked from lowest to highest correlation between citations and downloads. In these figures, a logarithmic transformation has been applied to both variables in order to better illustrate and highlight the correlations between them (Figs. 3, 4, 5, 6, 7, 8 and 9).

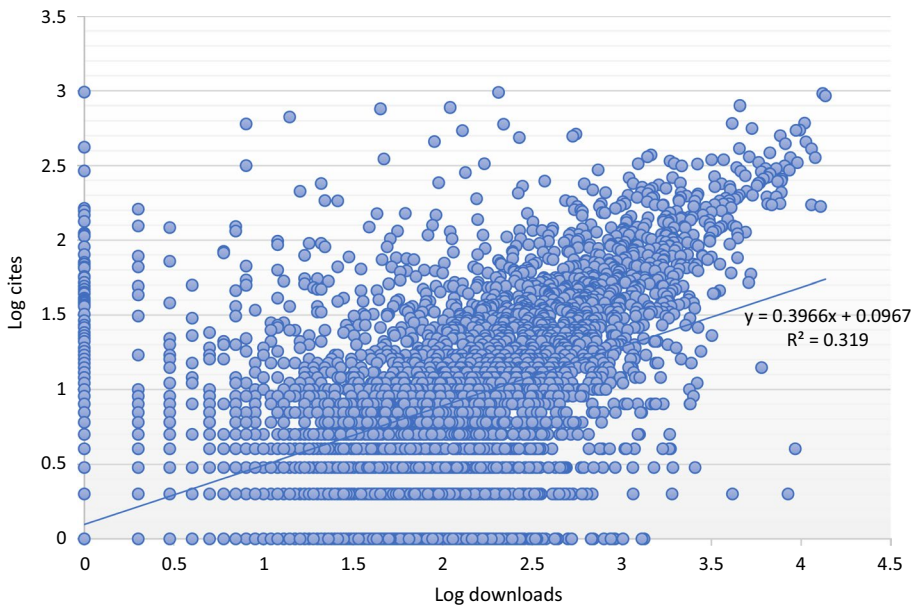
These results are consistent with previous studies showing a similar positive correlation between downloads and citations, the former being a variable that can predict the values of the latter (Feyereisen & Spoiden, 2009; Gumpenberger et al., 2012; McDonalds, 2007; Rodríguez-Bravo et al., 2021; Wood-Doughty et al., 2019;). However, other studies have failed to find significant correlations between the two variables, as in the case of Gao (2016) and Ke and Bronicki (2015).



**Fig. 3** Relationship between citations and downloads at UBU



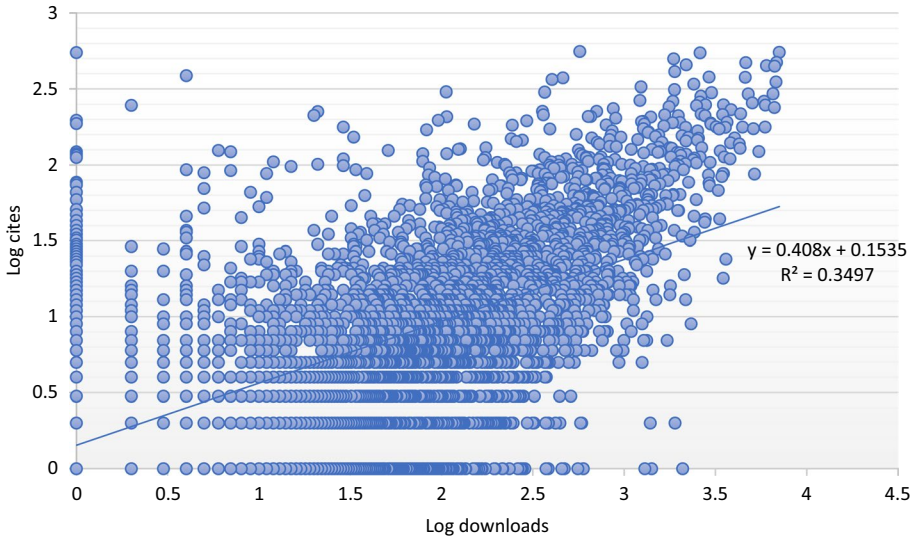
**Fig. 4** Relationship between citations and downloads at USAL



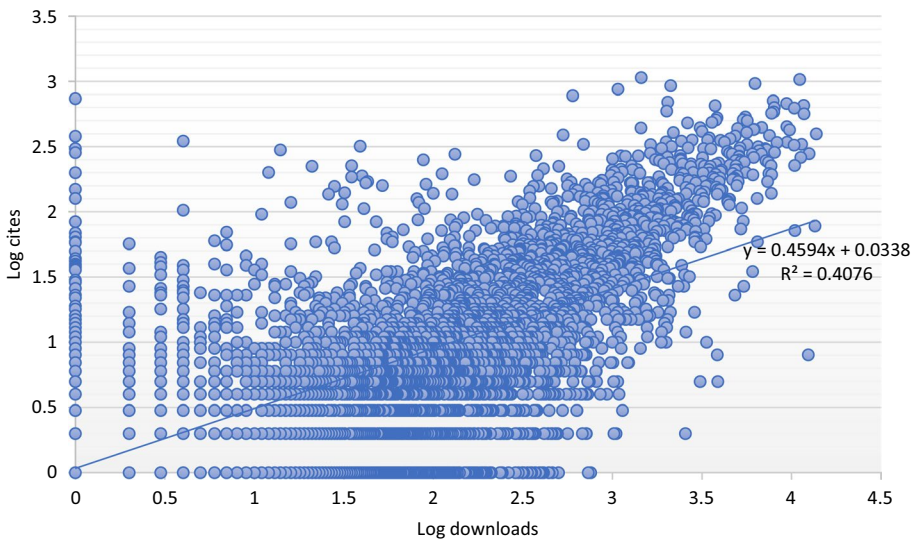
**Fig. 5** Relationship between citations and downloads at UVA

## Conclusions

The results of this study confirm a relationship between the size of the universities analysed and the volume of their scientific production, which increased over the study period. Likewise, they confirm the importance of scientific journals as a fundamental vehicle for

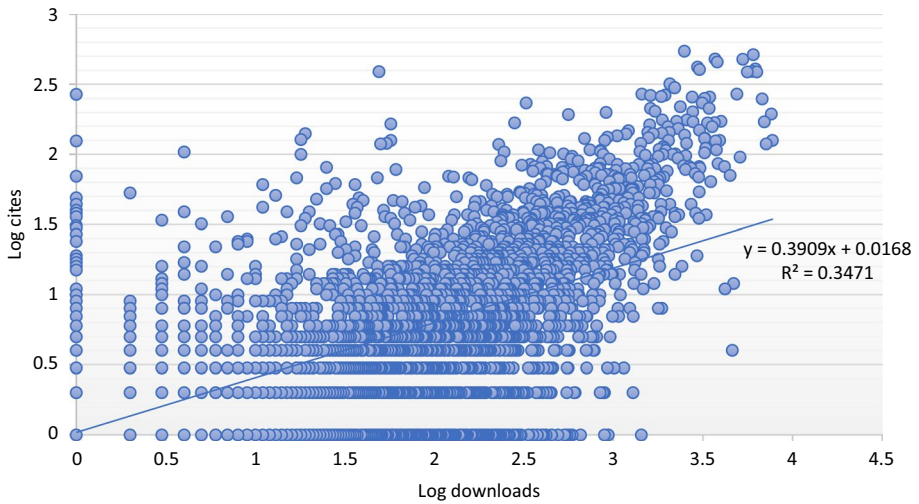


**Fig. 6** Relationship between citations and downloads at UDC

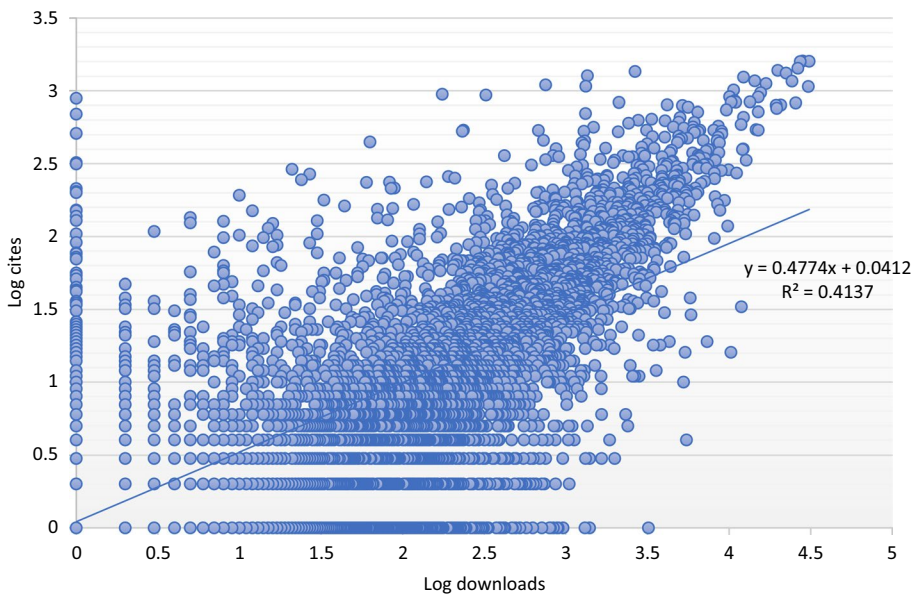


**Fig. 7** Relationship between citations and downloads at UVIGO

the transmission of knowledge, as evidenced by the finding that more than 73% of the references analysed in this study corresponded to this type of document, in line with the results found in other studies (Fernández-Ramos et al., 2022). This importance of scientific journals has recently been highlighted by Kim et al. (2020) and Herman et al. (2020). The latter indicate that journals are the only product that still consistently fulfil all the functions traditionally attributed to them—recording, curation, evaluation, distribution and archiving—and that they remain necessary to institutionalise and confidently add a scholarly



**Fig. 8** Relationship between citations and downloads at ULE



**Fig. 9** Relationship between citations and downloads at USC

contribution to the body of knowledge. It should also be noted that, in the case of Spanish researchers, the current evaluation system influences document type, marginalising monographs or book chapters in favour of journal articles (Osca-Lluch et al., 2019).

The citation of subscribed scientific journals reached a moderate percentage in most universities, the highest being 50% at the University of Santiago. It is important to highlight the strong correlation found between citation of subscribed journals and the volume

of scientific output. According to Shu et al. (2018), researchers only cite a fraction of the journals subscribed to by their libraries, and that fraction is decreasing, reducing the value of subscribed journal bundles, especially when the size of the university is small, as it is the case of some universities in this study. However, citations of journals included in subscription bundles confirm that the publishers distribute and facilitate access to quality—useful—content. Thus, they give visibility to the journals they distribute and promote their reading and subsequent citation, although the use of subscribed journals in scientific production varies considerably depending on discipline, as previous studies have found (Fernández-Ramos et al., 2022; Mongeon et al., 2021; Rodríguez-Bravo et al., 2021). Moreover, it should be kept in mind that articles are often downloaded many times but remain uncited because not all content is downloaded for research purposes (Gorraiz et al., 2014; Martin et al., 2016).

It should be borne in mind that the present analysis was limited to four electronic subscription bundles, not to all the subscriptions maintained by the universities studied, albeit these bundles included three of the main big deals—ScienceDirect, Springer and Wiley—and one of them—ScienceDirect—which contains the most widely-used content, as reported in various studies, including some conducted in the consortium of Castile and León (Fernández-Ramos et al., 2019). Despite the increase in downloads noted in previous studies, this volume of downloads does not strictly parallel the volume of citations. Previous studies (Fernández-Ramos et al., 2022; Rodríguez-Bravo et al., 2021) have found that besides the journals distributed as part of a big deal, widely-used journals also include those from other commercial publishers such as Taylor & Francis, prestigious institutional publishers and publishers that offer open access content. However, we found a significant presence of the most frequently downloaded journals among the cited journals, with high percentages (around 90%) in almost all universities. This result agrees with other studies that saw a bigger correlation when compared the articles more downloaded to the more cited (Chu & Krichel, 2007; Gumpenberger et al., 2012; O’Leary, 2008).

One of the main findings of this study is the high correlation between citations and downloads generally observed in the universities included in the analysis. All of these universities have a Pearson Correlation Coefficient under 0,5 except the one with less scientific production, the University of Burgos. In general terms, a higher correlation has been observed in universities with a higher output. This finding supports the idea, reported elsewhere in the literature (Tenopir & King, 2000), that researchers are the main users of scientific journal articles and that they use them primarily for research purposes. We conclude, therefore, that our results indicate that download values can predict future citation values. This highlights the usefulness of downloads data when making decisions about collection management in academic libraries (Gumpenberger et al., 2012).

These results should be viewed in light of the particularities of the data analysed (Scopus as the source of analysis of scientific production and four big deals as the source of scientific journal download data) and the following limitations: on the one hand, citation and download data for a given period of time were considered in conjunction, which only allows an approximation to reality since the date of download of a cited article is uncertain, although it is generally close to the date of citation. Time delays between downloads and citation show a large variability among users, due to differences in the amount of time they need to prepare a manuscript, and to differences in publication delays among journals selected for publication (Moed, 2005). As pointed out by Brody et al. (2006), the time delay may range anywhere from 3 months to 1–2 years or even longer. Besides, downloads and citations show different obsolescence functions (Ding et al., 2021; Moed & Halevi, 2016). Furthermore, there are disciplinary differences in obsolescence characteristics

between citations and downloads using synchronic and diachronic counts (Gorraiz et al., 2014). Correlation between citations and downloads is dependent on the discipline as well (McGillivray & Astell, 2019; Moed, 2005; Moed & Halevi, 2016; Wan et al., 2010).

On the other hand, regarding downloads, COUNTER JR does not cover downloads made to other versions of papers published (such as preprints or postprints in repositories or academic social networks) or downloads of open access articles made from outside the university domain (Gorraiz & Gumpenberger, 2010; Mongeon et al., 2021). Furthermore, errors may have occurred in the standardisation of journal titles, which may have resulted in duplicate journals. In this respect, it is worth highlighting the intrinsic difficulty of analyses such as the present one because of the time required for manual data cleaning and standardisation (Belter & Kaske, 2016; Mongeon et al., 2021; Rodríguez-Bravo et al., 2021). It is also worth noting the existence of outliers, which corresponded to extreme cases of journals that were frequently cited but rarely downloaded. In some cases, this may have been because the subscription had been discontinued at some point or the journals had changed their names but continued to be cited. Such cases would require an in-depth analysis of each of these journals.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This research has been funded by the State Program for Research, Development and Innovation Oriented to the Challenges of Society 2017, convened by the Spanish Ministry of Economy, Industry and Competitiveness and the Spanish State Research Agency (CSO2017-87956-R), and by the program of grants designed to support the recognized research groups of public universities in Castile and Leon that began in 2018, convened by the Ministry of Education of the Government of Castile and Leon (LE028G18).

## Declarations

**Conflict of interest** All authors declare that they have no conflicts of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arakaki, M. (2018). Uso de información en docentes universitarios peruanos: un análisis de citas en trabajos de investigación (2010–2014). *Anales De Documentación*. <https://doi.org/10.6018/analesdoc.21.2.302651>
- Belter, Ch. W., & Kaske, N. K. (2016). Using bibliometrics to demonstrate the value of library journal collections. *College & Research Libraries*, 77(4), 410–422. <https://doi.org/10.5860/crl.77.4.410>
- Blecic, D. D. (1999). Measurements of Journal Use: An analysis of the correlations between three methods. *Bulletin of the Medical Library Association*, 87(1), 20–25.
- Bollen, J., Van de Sompel, H., Smith, J. A., & Luce, R. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. *Information Processing and Management*, 41(6), 1419–1440. <https://doi.org/10.1016/j.ipm.2005.03.024>
- Brody, T., Harnad, S., & Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. *Journal of the American Society for Information Science and Technology*, 57(8), 1060–1072. <https://doi.org/10.1002/asi.20373>

- Chi, P. S., & Glänzel, W. (2017). An empirical investigation of the associations among usage, scientific collaboration and citation impact. *Scientometrics*, *112*(1), 403–412. <https://doi.org/10.1007/s11192-017-2356-4>
- Chew, K., Schoenborn, M., Stemper, J., & Lilyard, C. (2016). E-journal metrics for collection management: Exploring disciplinary usage differences in Scopus and Web of Science. *Evidence Based Library and Information Practice*, *11*(2), 97–120. <https://doi.org/10.18438/B85P87>
- Chu, H., & Krichel, T. (2007). Downloads vs. Citations in economics: relationships, contributing factors and beyond. In D. Torres-Salinas & H. F. Moed (Eds.), *Proceedings of the 11th International Society for Scientometrics and Informetrics Conference* (pp. 207–215). Madrid, Spain. Retrieved May 10, 2022, from <http://eprints.rclis.org/11085/1/DownloadsVsCitations.pdf>
- Coats, A. J. S. (2008). The top papers by download and citations from the International Journal of Cardiology in 2007. *International Journal of Cardiology*, *131*(1), e1–e3. <https://doi.org/10.1016/j.ijcard.2008.11.001>
- De Groot, S. L., Blecic, D. D., & Martin, K. (2013). Measures of health sciences journal use: A comparison of vendor, link-resolver, and local citation statistics. *Journal of the Medical Library Association*, *101*(2), 110–119. <https://doi.org/10.3163/1536-5050.101.2.006>
- Ding, Y., Dong, X., Bu, Y., Zhang, B., Lin, K., & Hu, B. (2021). Revisiting the relationship between downloads and citations: A perspective from papers with different citation patterns in the case of the Lancet. *Scientometrics*, *126*(9), 7609–7621. <https://doi.org/10.1007/s11192-021-04099-3>
- Duan, Y., & Xiong, Z. (2017). Download patterns of journal papers and their influencing factors. *Scientometrics*, *112*(3), 1761–1775. <https://doi.org/10.1007/s11192-017-2456-1>
- Ezema, I. J., & Asogwa, B. E. (2014). Citation analysis and authorship patterns of two linguistics journals. *Portal: Libraries and the Academy*, *14*(1), 67–85. <https://doi.org/10.1353/pla.2013.0050>
- Faulkner, K. (2021). Faculty use of open Access journals: A case study of faculty publications and citing references at a California University. *Publications*. <https://doi.org/10.3390/publications9030039>
- Fernández-Ramos, A., Rodríguez-Bravo, B., Alvite-Díez, M. L., Santos-De-paz, L., Morán-Suárez, M. A., Gallego-Lorenzo, J., & Olea, I. (2019). Evolution of the big deals use in the public universities of the Castile and Leon region, Spain. *Profesional De La Informacion*, *28*(6), e280519. <https://doi.org/10.3145/epi.2019.nov.19>
- Fernández-Ramos, A., Travieso-Rodríguez, C., & Rodríguez-Bravo, B. (2022). Faculty use of subscribed journals in a Spanish library consortium: Downloads and citations in the field of psychology. *Serials Review*, *48*(1–2), 121–136. <https://doi.org/10.1080/00987913.2022.2066966>
- Feyereisen, P., & Spoiden, A. (2009). Can local citation analysis of master’s and doctoral theses help decision-making about the management of the collection of periodicals? A case study in psychology and education sciences. *Journal of Academic Librarianship*, *35*(6), 514–522. <https://doi.org/10.1016/j.acalib.2009.08.018>
- Gao, W. (2016). Beyond journal impact and usage statistics: Using citation analysis for collection development. *Serials Librarian*, *70*(1–4), 121–141. <https://doi.org/10.1080/0361526X.2016.1144161>
- Gorraiz, J., & Gumpenberger, C. (2010). Going beyond citations: SERUM—A new tool provided by a network of libraries. *Liber Quarterly*, *20*(1), 81–93.
- Gorraiz, J., Gumpenberger, C., & Schlögl, Ch. (2014). Usage versus citation behaviours in four subject areas. *Scientometrics*, *101*(2), 1077–1095. <https://doi.org/10.1007/s11192-014-1271-1>
- Gumpenberger, C., Wernisch, A., & Gorraiz, J. (2012). Reality-check: Cost-related journal assessment from a practical point of view. *Journal of Scientometric Research*, *1*(1), 35–43. <https://doi.org/10.5530/jscires.2012.1.8>
- Herman, E., Akeroyd, J., Bequet, G., Nicholas, D., & Watkinson, A. (2020). The changed—and changing—landscape of serials publishing: Review of the literature on emerging models. *Learned Publishing*, *33*(3), 213–229. <https://doi.org/10.1002/LEAP.1288>
- Hitchcock, S., Brody, T., Gutteridge, C., Carr, L., & Harnad, S. (2003). The impact of OAI-based search on access to research journal papers. *Serials*, *16*(3), 255–260. <https://doi.org/10.1629/16255>
- Ivanov, A. O., Johnson, C. A., & Cassady, S. (2020). Unbundling practice: The unbundling of big deal journal packages as an information practice. *Journal of Documentation*, *76*(5), 1051–1067. <https://doi.org/10.1108/JD-09-2019-0187>
- Ke, I., & Bronicki, J. (2015). Using scopus to study researchers’ citing behavior for local collection decisions: A focus on psychology. *Journal of Library Administration*, *55*(3), 165–178. <https://doi.org/10.1080/01930826.2015.1034035>
- Kim, L., Portenoy, J. H., West, J. D., & Stovel, K. W. (2020). Scientific journals still matter in the era of academic search engines and preprint archives. *Journal of the Association for Information Science and Technology*, *71*(10), 1218–1226. <https://doi.org/10.1002/ASI.24326>

- Kurtz, M. J., & Bollen, J. (2010). Usage bibliometrics. *Annual Review of Information Science and Technology*, 44, 3–64. <https://doi.org/10.1002/aris.2010.1440440108>
- Larivière, V., Archambault, É., Gingras, Y., & Vignola-Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, 57(8), 997–1004. <https://doi.org/10.1002/asi.20349>
- Martin, V., Gray, T., Kilb, M., & Minchew, T. (2016). Analyzing consortial “Big Deals” via a cost-per-cited-reference (CPCR) metric. *Serials Review*, 42(4), 293–305. <https://doi.org/10.1080/00987913.2016.1248218>
- McDonald, J. D. (2007). Understanding journal usage: A statistical analysis of citation and use. *Journal of the American Society for Information Science and Technology*, 58(1), 39–50. <https://doi.org/10.1002/asi.20420>
- McGillivray, B., & Astell, M. (2019). The relationship between usage and citations in an open access mega-journal. *Scientometrics*, 121(2), 817–838. <https://doi.org/10.1007/s11192-019-03228-3>
- Moed, H. 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*, 56(10), 1088–1097. <https://doi.org/10.1002/asi.20200>
- Moed, H. F., & Halevi, G. (2016). On full text download and citation distributions in scientific-scholarly journals. *Journal of the Association for Information Science & Technology*, 67(2), 412–431. <https://doi.org/10.1002/asi.23405>
- Mongeon, P., Siler, K., Archambault, A., Sugimoto, C., & Larivière, V. (2021). Collection development in the era of big deals. *College & Research Libraries*, 82(2), 219–236. <https://doi.org/10.5860/crl.82.2.219>
- O’Leary, D. E. (2008). The relationships between citations and number of downloads in *Decision Support Systems*. *Decision Support Systems*, 45(4), 972–980. <https://doi.org/10.1016/j.dss.2008.03.008>
- Oscá-Lluch, J., González-Sala, F., Haba-Oscá, J., Tortosa, F., & Peñaranda-Ortega, M. (2019). Comunicación científica o cualificación para una carrera académica: ¿Qué uso tienen los artículos en las revistas de psicología? *Anales De Psicología*, 35(1), 166–174. <https://doi.org/10.6018/analesps.35.1.329211>
- Pastva, J., Shank, J., Gutzman, K. E., Kaul, M., & Kubilius, R. K. (2018). Capturing and analyzing publication, citation, and usage data for contextual collection development. *Serials Librarian*, 74(1–4), 102–110. <https://doi.org/10.1080/0361526X.2018.1427996>
- Peñaflor, J., & Aliwalas, A. (2022). Research output and information use: A citation analysis of faculty publications in engineering. *Collection Management*, 47(4), 300–315. <https://doi.org/10.1080/01462679.2022.2081830>
- Rodríguez-Bravo, B., Fernández-Ramos, A., & Travieso-Rodríguez, C. (2021). Relación entre descargas y citas de revistas científicas en el ámbito de la Documentación: el caso de las universidades públicas de Castilla y León. *Revista Española De Documentación Científica*, 44(4), e307. <https://doi.org/10.3989/redc.2021.3.1806>
- Shu, F., Mongeon, P., Haustein, S., Siler, K., Alperin, J. P., & Larivière, V. (2018). Is it such a big deal? On the cost of journal use in the digital era. *College & Research Libraries*, 79(6), 785–798. <https://doi.org/10.5860/crl.79.6.785>
- Srivastava, B., & Kumar, S. (2018). Usage and impact of Science Direct material science package in a material science library. *Desidoc: Journal of Library & Information Technology*, 38(1), 21–26. <https://doi.org/10.14429/djlit.38.1.12124>
- Tenopir, C., & King, D. W. (2000). *Towards electronic journals: Realities for scientists*. Librarians and Publishers.
- Tsay, M. Y. (1998). The relationship between journal use in a medical library and citation use. *Bulletin of the Medical Library Association*, 86(1), 31–39.
- Tucker, C. (2013). Analyzing faculty citations for effective collection management decisions. *Library Collections, Acquisitions, and Technical Services*, 37(1–2), 19–33. <https://doi.org/10.1016/j.lcats.2013.06.001>
- Vaaler, A. (2018). Sources of resources: A business school citation analysis study. *Journal of Business & Finance Librarianship*, 23(2), 154–166. <https://doi.org/10.1080/08963568.2018.1510252>
- Vogl, S., Scherndl, T., & Kühberger, A. (2018). Psychology: A bibliometric analysis of psychological literature in the online media. *Scientometrics*, 115(3), 1253–1269. <https://doi.org/10.1007/s11192-018-2727-5>
- Wan, J. K., Hua, P. H., Rousseau, R., & Sun, X. K. (2010). The download immediacy index (DII): Experiences using the CNKI full-text database. *Scientometrics*, 82(3), 555–566. <https://doi.org/10.1007/s11192-010-0171-2>



- Watson, A. B. (2009). Comparing citations and downloads for individual articles. *Journal of Vision*, 9(4), 1–4. <https://doi.org/10.1167/9.4.i>
- Wical, S. H., & Vandenbark, R. T. (2014). Notes on operations: Combining citation studies and usage statistics to build a stronger collection. *Library Resources and Technical Services*, 59(1), 33–42. <https://doi.org/10.5860/lrts.59n1.33>
- Wilson, C. S., & Tenopir, C. (2008). Local citation analysis, publishing, and reading patterns: Using multiple methods to evaluate faculty use of an academic library's research collection. *Journal of the American Society for Information Science and Technology*, 59(9), 1393–1408. <https://doi.org/10.1002/asi.20812>
- Wood-Doughty, A., Bergstrom, T., & Steigerwald, D. G. (2019). Do download reports reliably measure journal usage? Trusting the fox to count your hens? *College and Research Libraries*, 80(5), 694–719. <https://doi.org/10.5860/crl.80.5.694>
- Zhang, L. (2018). Analyzing citation and research collaboration characteristics of faculty in Aerospace, Civil and Environmental, Electrical and Computer, and Mechanical Engineering. *College & Research Libraries*, 79(2), 158–178. <https://doi.org/10.5860/crl.79.2.158>
- Zhu, Q., & Xiang, H. (2016). Differences of Pareto principle performance in e-resource download distribution. *The Electronic Library*, 34(5), 846–855. <https://doi.org/10.1108/EL-05-2015-0068>