



基于使用数据计量的 CASHL 文献资源保障研究*

□熊泽泉 彭霞 杨莉

摘要 文献资源的价值不仅可以通过学者的引用行为反映,也可以通过读者的下载、阅读、分享等文献使用行为反映。本研究选择 Web of Science 数据库中“经济学与商学”学科中的高使用量论文进行科学计量分析,描述了高使用量文献的一般性特征,在此基础上分析了华东师范大学图书馆及 CASHL 对全球高使用量论文的保障率。研究发现,“经济学与商学”学科的高使用量论文一般为最新发表的论文以及发表在影响因子较高期刊上的论文,且使用量与被引量存在显著正相关;华东师范大学图书馆对“经济学与商学”学科中的全球高使用量论文保障率仅为 83.9%,但 CASHL 对“经济学与商学”学科中的全球高使用量论文保障率达到 99.1%,极大地弥补了一校一馆文献保障率的不足。基于使用数据计量研究文献保障率有望成为基于引文分析研究的有益补充。

关键词 使用数据 使用计量 资源保障 CASHL

分类号 G253

DOI 10.16603/j.issn1002-1027.2022.02.002

1 引言

资源是图书馆的立身之本,文献资源建设是图书馆的核心工作之一。随着数字技术的飞速发展,电子文献数量激增,使文献信息资源结构发生了重要变化^[1]。由于电子文献具有可复制、易传递、易携带等优点,深受读者欢迎,并且对于图书馆物理空间的缓解也大有裨益,因此如何高效地进行电子文献资源建设成为图书馆资源建设的重中之重,亦成为众多学者关注的热点问题。

现有关于文献资源保障的研究主要是基于引文分析^[2-6],但是,基于引文分析的文献资源保障研究存在两方面的问题:(1)文献发表和引用的滞后性造成文献资源保障服务的滞后,特别是人文社会科学领域的研究成果,从引文分析中获得的文献保障情况难以准确反映学者在进行研究时的实际资源保障情况和现实需求;(2)并不是所有的文献需求都能通过引文进行反映,一些文献在课程教学、大众教育等方面发挥着重要作用,但是其用户可能并不发表论,因而这些用户对文献的使用需求也无法通过引

文数量进行反映。

随着互联网技术的发展,在科学计量学领域逐渐兴起的文献使用数据有望成为引文分析的有效补充。文献使用数据是指文献被用户下载、阅读、分享、保存等过程中所产生的可被记录的量化数据。利用文献使用数据来研究文献资源保障情况,可以有效地解决基于引文分析研究中的问题。首先,文献使用行为发生在研究者的研究过程当中,特别是文献下载行为一般发生在科学研究的初始阶段,可实时反映研究者对文献的使用和需求情况;其次,文献使用数据并不关注使用者是否发表论文,而只关注文献是否被使用。本研究认为,文献被使用即发挥了其价值,被使用次数越多,说明其受关注程度越高,而高使用量的论文在文献资源建设中同样应该优先保障。但是,国内外尚未见利用文献使用数据来研究文献资源保障的报道。

此外,现有关于文献保障率的研究多以一校一馆为研究对象,针对中国高校人文社会科学文献中心(CASHL)、中国高等教育文献保障系统(CALIS)

* 上海市哲学社会科学规划青年课题“科学大数据背景下学术文献的使用动机研究”(编号:2019ETQ004)、CASHL“新信息环境下 CASHL 资源与服务拓展设计研究”的研究成果之一。

通讯作者:熊泽泉,ORCID:0000-0002-4349-371X,邮箱:zqxiong@library.ecnu.edu.cn。



等我国高校文献服务保障体系的相关研究则鲜见报道,仅有唐惠燕等运用引文分析法研究了 CASHL 对农业类高校外文资源的保障程度^[7],以及宋姬芳等运用引文和用户调查法研究了 CASHL 对经济学与商学资源的保障程度^[8]。CASHL 作为我国两个规模较大的哲学社会科学文献中心之一^[9],为我国的哲学社会科学文献保障作出了重要贡献,但对其建设成效的评价仍缺乏足够的支撑。本研究从使用数据计量的角度,对 CASHL 在提高高校图书馆资源保障率方面的作用进行分析,以期为文献资源保障体系的评价工作提供有益补充。

2 研究综述

文献资源保障率是指用户对文献需求的满足程度^[10-11],对于文献资源保障率的研究源于文献资源调查的评估方法。1988年,肖自力等归纳了三种文献资源调查的评估方法,即书目核对、引文分析和用户评议^[12],这三种方法经常被用于纸质资源保障情况的综合评价^[13-16]。其中,书目核对是利用各种图书目录来核对图书馆是否藏有某些书,借此衡量馆藏率^[17],反映的其实是文献(图书)覆盖率^[10]或拥有率^[18],这在互联网发展以前,是众多图书馆对馆藏资源建设评价的主要方法^[15, 19];而用户评议方法虽然直接反映用户对文献的需求,但是主观性较强,个体差异较大,不宜单独采用;引文分析法是对某一学科或领域的文献被引用情况进行分析,以此来评价文献被利用的情况,由于这一方法相对简单客观,同时该方法主要反映用户对于学术型文献的满足程度,因此在高校图书馆的资源保障建设中被广泛采用,特别是馆藏资源逐渐向数字化方向发展后,引文分析法逐渐成为评价高校图书馆电子文献资源保障率的主要方法。中山大学^[5]、哈尔滨工程大学^[2, 5]、东北师范大学^[20]、中国政法大学^[21]、复旦大学^[6]等国内知名学府都曾采用该方法对本校的文献资源保障情况进行分析。

引文分析法用于文献资源保障率分析,从本质上来说,也是从文献使用的角度反映用户对文献的需求,认为用户在参考文献中引用了某文献,即反映了用户对该文献(曾经)的需求。然而,用户对文献的使用行为不仅限于引用行为,也包括用户对文献的浏览、下载、阅读、分享、保存等行为。随着互联网的发展,用户的这些文献使用行为能够被系统实时

记录,对于使用数据的挖掘为科学计量研究者提供了崭新的研究视角,如研究论文的使用模式^[22-24],运用使用数据探索科研趋势^[22-24],探索使用数据与引文数据的相关性从而为传统的引文评价提供补充等^[25-27]。

然而,在使用数据计量发展的最初阶段,相关研究发展并不迅速,原因之一主要是缺乏具有国际普适性并相对权威的统一平台^[28],只能依靠不同的出版商提供的使用数据。直到2015年9月,Web of Science(以下简称 WoS)数据库平台在当月更新中,推出了名为“Usage Count”的文献级使用量数据^[29],研究者们有了较为权威而统一的使用数据来源,从而对该使用量数据开展了一系列研究。如王贤文等通过对5本信息科学与图书馆学期刊论文的 WoS 使用量数据分析,发现研究者们更倾向于使用最新的文献,但是高被引论文在发表很长一段时间后仍有可能被引用^[29];梁国强等利用 WoS 使用量数据分析再生医学领域前沿,发现对于同一领域,使用量数据较引文数据能较早地发现研究前沿,并且基于使用量数据所获得的领域前沿多为近两年的研究,表现出更好的即时性和动态性^[30];赵星等发现受基金资助的论文 WoS 被使用次数更多,使用量数据具有明显的学科差异^[31];马库索娃(Markusova V)等比较了被引量 and 长期使用量、短期使用量的相关性,发现在论文层面,引用指标和使用指标之间存在显著的肯德尔等级相关性,并且长期使用量与被引量的相关性更强,但是在期刊层面上这种相关性则非常低^[32];梁国强等通过利用罗杰斯创新扩散理论对 WoS 中的“高使用低引用”以及“低使用高引用”论文进行分析,构建了学术论文的扩散-采纳模型^[33]。由上可见,随着统一平台的出现,对于使用数据计量的研究也从简单地分析使用量指标与传统引文指标的相关性逐渐扩展到前沿探索、知识扩散等领域,但是目前尚未见基于使用数据计量来评估图书馆文献资源保障率的相关研究,本研究将在此方面进行初步探索。

3 数据收集与分析

3.1 数据收集

宋姬芳等曾在2018年利用引文分析法研究了 CASHL 的文献资源保障率^[8],为了与该研究结果进行比较,本文参考该研究的样本选取方式,以 ESI



学科分类作为分类标准,选取 ESI“经济学与商学”(Economics & Business)学科领域中 588 种期刊在 WoS 数据库中近 10 年的发文为数据来源,数据获取时间为 2021 年 3 月 15 日,共获取到 327912 条论文全记录作为原始数据集。

WoS 数据库平台的 Usage Count 记录了用户使用某一文献的次数,其使用行为包括全文链接点击或者各种格式文件的保存^[29],按记录时间分为近期使用量 U1 和长期使用量 U2 两个指标,其中 U1 记录的是近 180 天内的使用量,U2 记录的是自 2013 年 2 月 1 日起至今的使用量^[30]。分别将原始数据集依据近期使用量 U1 和长期使用量 U2 进行排序,选择前 1% 的论文(分别为 3280 条)作为全球高使用量论文数据集,记为 V1 和 V2,并分别对 V1 和 V2 进行描述性统计、相关性分析。然后对 V1 和

V2 进行合并去重,获得全球高使用量论文期刊列表,将该列表与“华东师范大学图书馆全文期刊目录”及“CASHL 全文期刊目录”进行比对以测度文献资源保障率。

3.2 描述性统计

表 1 为高使用量论文数据集 V1、V2 的描述性统计结果,图 1 为近期使用量 U1、长期使用量 U2 的绝对数值分布。长期使用量 U2 的各项检验参数均远大于近期使用量 U1,两者的绝对数值分布均表现为严重的非中心对称的偏态分布,即在高使用量论文数据集 V1、V2 中,80% 高使用量论文的近期使用量 U1 和长期使用量 U2 主要集中在 19—39 和 149—274 之间。经 K-S 检验,结果表明,V1 和 V2 样本的被引量(TC1、TC2)、使用量(U1、U2)的绝对数值分布均不符合正态分布。

表 1 高使用量论文数据集描述性统计

	指标	最小值	中值	最大值	平均值	标准差	K-S Z 值	K-S P 值
V1 数据集	U1	19	27	902	33.18	24.21	0.28	<0.001
	TC1	0	7	3277	74.01	174.65	0.34	<0.001
V2 数据集	U2	149	196	4447	235.47	146.25	0.28	<0.001
	TC2	0	104	3277	151.34	172.08	0.19	<0.001

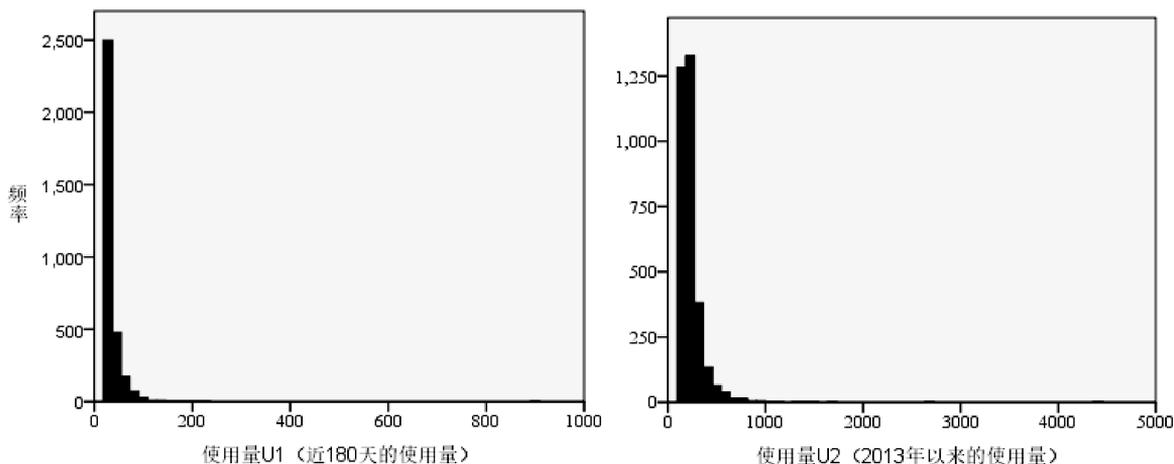


图 1 使用量频次分布

图 2 为高使用量论文 V1 和 V2 发表年份分布及其占年度论文发表总量的比例。基于近期使用量 U1 获取的高使用量论文 V1 多为最新发表的论文,其中近 3 年(2019—2021 年)发文量达到 2272 篇,占总体的 69.27%。而基于长期使用量 U2 获得的高使用量论文 V2 多为较早发表的论文,近 3 年(2019—2021 年)发文量仅 134 篇,占总体的 4.08%。U2 代

表的是 WoS 数据库平台上论文的长期使用量,早期发表论文的长期使用量 U2 会随着时间累积增长,呈现和引文曲线类似的趋势。而具有较高近期使用量 U1 的论文则明显反映出“经济学与商学”学科领域的学者在 WoS 数据库平台中更偏好使用最新发表的论文。

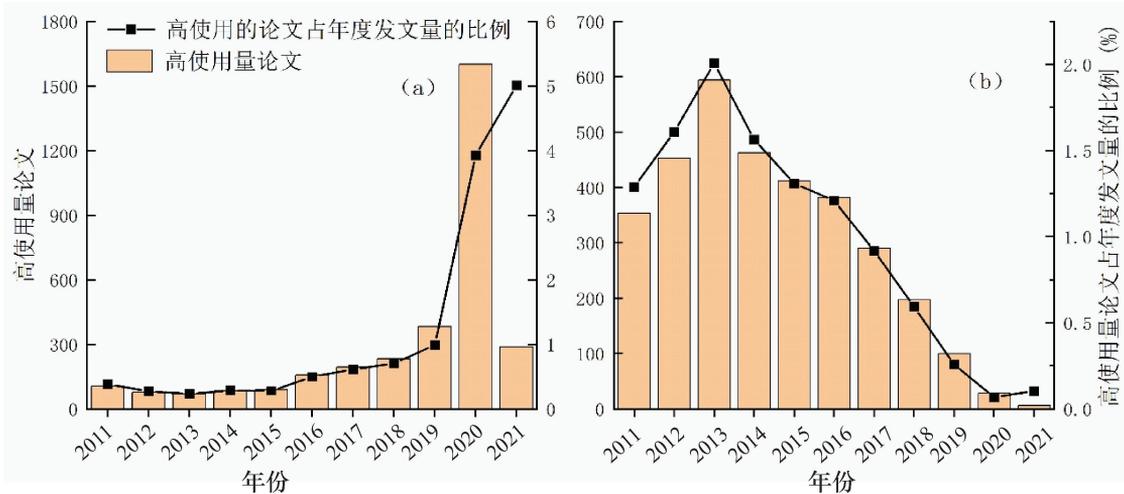


图 2 高使用量论文 V1、V2 发表年份分布及占年度发文总量的比例

(a: 近期使用量 U1

b: 长期使用量 U2)

高使用量论文 V1、V2 共分布在 284 种和 265 种期刊。表 2 中分别列出了 V1 和 V2 中高使用量论文数量排行前五位的期刊,其中四种期刊 (*Journal of Business Research*, *Academy of Management Journal*, *Strategic Management Journal*, *Journal of Management*) 在两个数据集中均位列前五。期刊 *Management Science* 针对近期使用量 U1 上具有较多高使用量论文,而期刊 *Organization Science* 则是针对长期使用量 U2 具有较多高使用量论文。这表明前四种期刊是“经济学与商学”学科领域近 10 年来无论是近期使用量还是长期使用量上均表现较高的经典期刊,而 *Management Science* 则是近几年新晋热门期刊,使用量持续增长,相对地,期刊 *Organization Science* 在 WoS 平台的使用量则逐渐减少,甚至 2020 年和 2021 年均未有论文列入高使用量论文数据集 V2 中。其原因可能有两点:

(1) 期刊载文量的影响。*Management Science* 近 10 年载文量一直处于上升阶段,由 2011 年的 140 篇增长至 2020 年的 301 篇,而 *Organization Science* 的载文量总体表现为下降趋势,由 2011 年的 104 篇减少到 2020 年的 70 篇,载文量的减少可能导致对于该期刊论文的使用量减少。

(2) 期刊影响因子的影响。*Management Science* 的影响因子近 10 年同样表现出较为明显的上升趋势,2011 年的影响因子仅为 1.733,而 2019 年上升至 3.931(见图 3)。*Organization Science* 则由 2011 年历史最高影响因子 4.338,逐渐跌至到

2019 年 2.782(见图 4)。

表 2 高使用量论文排行前五位的期刊及其占期刊发文总量的比例/排序

高使用量论文 V1—近期使用量 U1		
期刊名	高使用量论文	占期刊发文总量比例/排序
<i>Journal of Business Research</i>	143	3.10%/51
<i>Academy of Management Journal</i>	140	17.48%/2
<i>Strategic Management Journal</i>	131	10.95%/6
<i>Management Science</i>	120	5.48%/24
<i>Journal of Management</i>	111	12.31%/3

高使用量论文 V2—长期使用量 U2		
期刊名	高使用量论文	占期刊发文总量比例/排序
<i>Academy of Management Journal</i>	275	34.33%/1
<i>Strategic Management Journal</i>	207	17.31%/5
<i>Journal of Management</i>	199	22.06%/3
<i>Journal of Business Research</i>	124	2.69%/56
<i>Organization Science</i>	97	11.28%/10

期刊影响因子会在一定程度上影响科研人员的使用量,如谢娟等指出 JCR 分区 Q1 论文的平均使用次数约为 Q4 论文的 5 倍^[34],本研究也侧面反映了研究人员偏向使用影响因子较高的期刊论文。但是期刊载文量的变化是否会影响用户对文献的使用量尚未见报道,作者将在后续进一步研究。

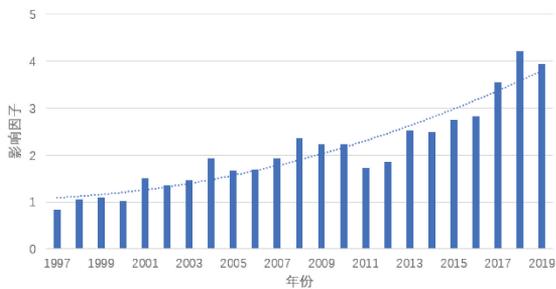


图3 Management Science 期刊影响因子变化趋势

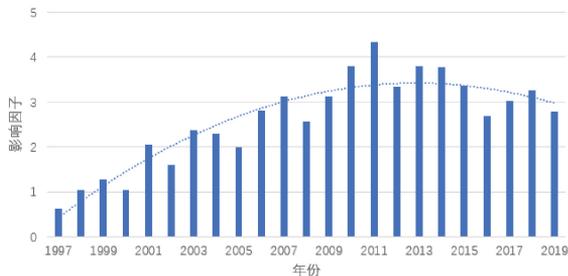


图4 Organization Science 期刊影响因子变化趋势

3.3 相关性分析

前文通过 K-S 检验可知使用量和被引量均不符合正态分布,因此相关性分析采用 Spearman 相关检验。图 5 是从论文角度的高使用量论文的使用量与被引量的散点图,长期使用量 U2 与被引量 TC

之间的 Spearman 相关系数(R 值)要略高于近期使用量 U1,分别为 0.407、0.142,显著性系数均 $P < 0.01$ 。图 6 是从高使用量论文发表的期刊角度的使用量与被引量散点图,长期使用量 U2 与被引量 TC 之间的 Spearman 相关系数同样高于近期使用量 U1,分别为 0.816、0.697,显著性系数均 $P < 0.01$ 。从上述结果可知,相比于近期使用量 U1 而言,长期使用量 U2 与被引量 TC 之间存在更强的相关性。这一相关性的差异也暗示,被引量能够一定程度上反映读者的长期使用倾向,但是在反映读者近期使用倾向上则略显不足。

3.4 全球高使用量论文的保障率分析

(1)华东师范大学图书馆对“经济学与商学”高使用量论文期刊的保障情况

本文考察了华东师范大学图书馆对高使用量论文集 V1 和 V2 所在期刊集合的保障情况。“经济学与商学”学科领域的高使用论文集 V1、V2 所在期刊不含重复项共计 329 种,华东师范大学图书馆可保障 276 种,资源保障率为 83.9%。在可保障的 276 种期刊中,含 11 种 OA 刊。缺藏的 53 种期刊详细信息见表 3,其中属于 Q1、Q2、Q3 和 Q4 分区期刊分别有 13、17、15 和 8 种。

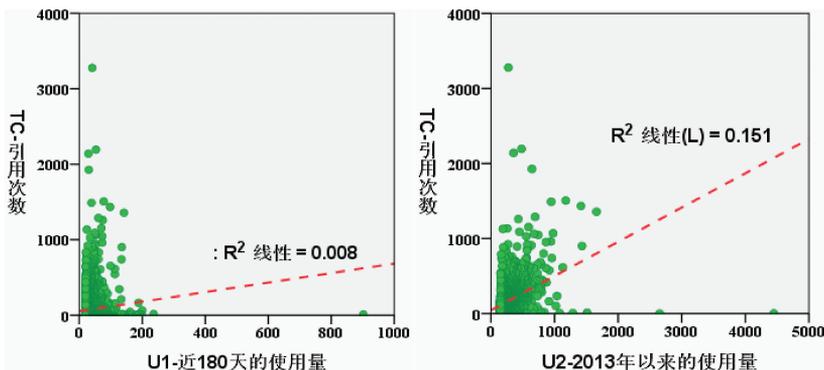


图5 论文角度的近期使用量 U1、长期使用量 U2 与被引量 TC 散点图

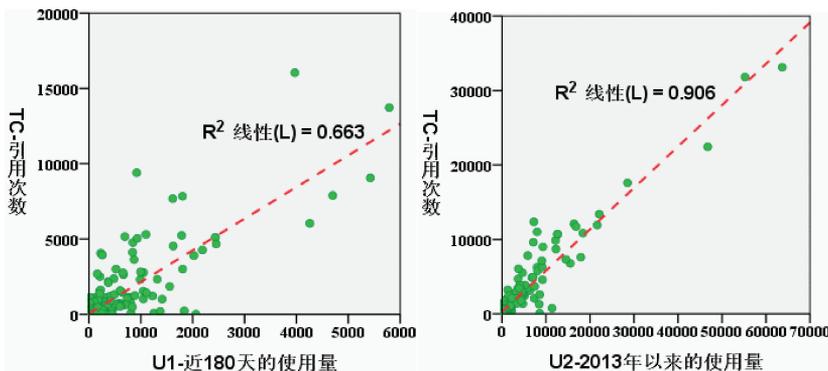


图6 期刊角度的近期使用量 U1、长期使用量 U2 与被引量 TC 散点图



表 3 华东师范大学图书馆缺藏高使用论文期刊(53 种)

期刊名称	ISSN	影响因子	分区	近期高使用量论文数—U1	长期高使用量论文数—U2
<i>Academy of Management Annals</i>	1941—6520	11.75	Q1	49	50
<i>Engineering Construction and Architectural Management</i>	0969—9988	2.16	Q2	42	1
<i>Management Decision</i>	0025—1747	2.723	Q2	31	14
<i>Accounting Review</i>	0001—4826	3.993	Q1	28	17
<i>European Journal of Innovation Management</i>	1460—1060	2.613	Q2	28	2
<i>Journal of Service Management</i>	1757—5818	4.662	Q1	25	14
<i>International Journal of Operations & Production Management</i>	0144—3577	4.619	Q1	18	12
<i>Supply Chain Management—an International Journal</i>	1359—8546	4.725	Q1	16	18
<i>International Journal of Logistics Management</i>	0957—4093	3.325	Q2	14	1
<i>Journal of Intellectual Capital</i>	1469—1930	4.805	Q1	13	0
<i>International Journal of Physical Distribution & Logistics Management</i>	0960—0035	4.744	Q1	12	12
<i>China Agricultural Economic Review</i>	1756—137X	1.775	Q2	12	0
<i>Business Process Management Journal</i>	1463—7154	2.121	Q3	11	2
<i>European Journal of Marketing</i>	0309—0566	2.135	Q3	9	4
<i>Asia Pacific Journal of Marketing and Logistics</i>	1355—5855	2.511	Q2	9	1
<i>Chinese Management Studies</i>	1750—614X	1.036	Q4	9	0
<i>Journal of Product and Brand Management</i>	1061—0421	1.832	Q3	9	1
<i>International Journal of Entrepreneurial Behavior & Research</i>	1355—2554	3.529	Q2	10	0
<i>Leadership & Organization Development Journal</i>	0143—7739	1.977	Q3	8	1
<i>International Marketing Review</i>	0265—1335	2.907	Q2	7	2
<i>Journal of Business & Industrial Marketing</i>	0885—8624	2.497	Q3	6	2
<i>Journal of Services Marketing</i>	0887—6045	3.195	Q2	6	0
<i>Review of International Political Economy</i>	0969—2290	2.312	Q1	6	2
<i>International Journal of Technology Management</i>	0267—5730	1.348	Q3	6	6
<i>International Journal of Conflict Management</i>	1044—4068	1.806	Q2	6	0
<i>International Journal of Manpower</i>	0143—7720	0.953	Q3	5	2
<i>Mis Quarterly Executive</i>	1540—1960	4.088	Q1	5	4
<i>Journal of Organizational Change Management</i>	0953—4814	0.967	Q4	4	1
<i>Marketing Intelligence & Planning</i>	0263—4503	2.164	Q3	4	1
<i>International Journal of Managing Projects in Business</i>	1753—8378	1.989	Q3	4	0
<i>Accounting Auditing & Accountability Journal</i>	0951—3574	3.497	Q1	3	0
<i>Mit Sloan Management Review</i>	1532—9194	2.706	Q2	3	29
<i>Jahrbucher Fur Nationalokonomie Und Statistik</i>	0021—4027	0.658	Q4	3	0
<i>Journal of Labor Economics</i>	0734—306X	3.356	Q1	3	1
<i>Journal of Service Theory and Practice</i>	2055—6225	3.512	Q2	3	0
<i>Review of Environmental Economics and Policy</i>	1750—6816	6.487	Q1	3	25
<i>Employee Relations</i>	0142—5455	1.641	Q3	2	0
<i>Journal of Advertising Research</i>	0021—8499	2.169	Q2	2	5
<i>Post-communist Economies</i>	1463—1377	0.875	Q3	2	0
<i>Academia-revista Latinoamericana De Administracion</i>	1012—8255	0.739	Q4	2	0
<i>Journal of Korea Trade</i>	1229—828X	0.628	Q4	2	0
<i>Pharmacoeconomics</i>	1170—7690	3.563	Q1	1	1



期刊名称	ISSN	影响因子	分区	近期高使用量论文数-U1	长期高使用量论文数-U2
<i>Journal of Medical Economics</i>	1369-6998	1.958	Q2	1	0
<i>Transformations in Business & Economics</i>	1648-4460	1.621	Q2	1	1
<i>Journal of Portfolio Management</i>	0095-4918	0.709	Q4	1	0
<i>International Journal of Bank Marketing</i>	0265-2323	2.8	Q2	1	0
<i>Review of Keynesian Economics</i>	2049-5323	0.689	Q4	1	1
<i>Baltic Journal of Management</i>	1746-5265	0.958	Q3	1	1
<i>Journal of Law & Economics</i>	0022-2186	1.719	Q3	1	1
<i>Journal of Forest Economics</i>	1104-6899	0.958	Q3	1	0
<i>Journal of Electronic Commerce Research</i>	1526-6133	1.875	Q3	1	1
<i>Entrepreneurship Research Journal</i>	2194-6175	1.643	Q4	1	2
<i>American Journal of Health Economics</i>	2332-3493	1.902	Q2	0	4

注:影响因子和分区数据均来源于 2020 年 6 月 Clarivate Analytics 发布的期刊引证报告。

(2)CASHL 对“经济学与商学”高使用量论文期刊的保障情况

“经济学与商学”学科领域 329 种高使用量论文期刊, CASHL 已保障 326 种, 保障率达 99.1%。2018 年宋姬芳等建议优先补藏 24 种 ESI 经济学与商学期刊^[8], 其中 19 种为本研究中的高使用论文期刊, 目前皆已被 CASHL 收录。说明 CASHL 对高使用论文期刊的保障率相较于 2018 年已有较大提高。

CASHL 仅缺藏“经济学与商学”学科领域的 3 种高使用量论文期刊, 分别为 *Review of Keynesian Economics*, *Korean Economic Review*, *American Journal of Health Economics* (见表 4)。其中 *Korean Economic Review* 属于 OA 期刊, 但在 CASHL 馆藏目录中并未揭示。2 种期刊在近期高使用量论文集 V1 中有 2 篇论文, 3 种期刊在长期高使用量论文集 V2 中有 7 篇论文。

表 4 CASHL 缺藏的 3 种“经济学与商学”高使用量论文期刊

期刊名称	ISSN	OA 期刊	影响因子	分区
<i>Review of Keynesian Economics</i>	2049-5323	否	0.689	Q4
<i>Korean Economic Review</i>	0254-3737	是	0.174	Q4
<i>American Journal of Health Economics</i>	2332-3493	否	1.902	Q2

注:影响因子和分区数据均来源于 2020 年 6 月 Clarivate Analytics 发布的期刊引证报告。

4 结论与建议

通过对 WoS 数据库中高使用量论文集 V1、V2 发表年份分布及占年度发文总量的比例进行分析, 发现近期使用量 U1 高的论文多为最新发表的论文, 而长期使用量 U2 高的论文, 多为 7-8 年前的论文, 这种年代分布的不一致提示 U1 和 U2 可能反映了文献使用者不同的使用倾向。其中, U1 表示最近 180 天的使用量, 反映的是对文献短期、即时的使用情况。已有研究证明, 多数研究者为了追踪研究前沿, 倾向于使用最新发表的文献^[29-30], 从而使得最新发表的论文出现较高的近期使用量 U1。U2 代表的是 2013 年 2 月 1 日至今的使用量, 反映的是对文献长期、累积的使用情况, 研究者为了获得更具说服力的文献来支撑自己的研究, 倾向于使用具有较高被引量的文献^[32], 而被引量的累积是一个长期的过程, 从而使得具有较高长期使用量 U2 的论文多为发表时间较长的论文。

通过进一步比较 U1、U2 与被引量的相关性, 发现无论是从论文层面还是期刊层面, U2 与被引量的相关性均高于 U1 与被引量的相关性。马库索娃 (Markusova V) 对俄罗斯出版的 WoS 期刊论文的使用量分析, 也发现除数学学科外, 两个使用量指标与被引量指标具有显著相关性, 且 U2 与被引量的相关性显著高于 U1 与被引量的相关性。作者认为数学学科论文使用量与被引量的相关性较低可能是由于俄罗斯学者特有的文献引用和使用行为^[32], 但作者并未对 U1、U2 与被引量相关性的差异进行解释。本研究认为, 这种相关性的差异一方面是由于



U1 和 U2 的统计意义的差异所致, U2 与被引量一样是一个累积性指标, 指标值只会保持不变或者增加, 而 U1 则是即时性指标, 指标值会随时间上下浮动; 另一方面, 长期使用量 U2 在反映用户文献使用行为上与被引量具有更高的趋同性, 这种趋同性可能来源于使用动机与引用动机在论文发表一段时间后更加契合。熊泽泉等在对论文下载量与被引量动态相关性的研究中, 也发现分年下载量与分年被引量的相关性逐渐增强, 并提出了下载量与被引量随着时间的增长在表征论文的使用和有用程度上逐渐达到统一, 下载动机与引用动机的契合度增强的观点^[35]。笔者将在另一研究中对使用动机展开深入讨论。

虽然对于文献引用动机的相关研究认为引用动机多种多样^[36-38], 但是除了存在少数“被迫引用”“假引”外, 引用行为基本发生在使用者真正阅读、理解了该文^[37], 即“使用并有用”的基础上。长期使用量 U2 与被引量的相关性较高, 说明 U2 较 U1 在反映论文“有用”的方面更为有效。而近期使用量 U1 则更多受到其他因素的影响, 如期刊或平台推送策略等原因使得最新文献的信息更易获取、检索结果默认按发表时间排序使得最新发表的文献更易被发现、以及使用者对最新研究更感兴趣等, 这些因素存在一定的随机性和不确定性, 导致 U1 与被引量的偏离更大, 但 U1 在反映文献使用的即时性方面具有更多优势, 因而能够更好地探测当前研究前沿^[33]。因此, U1 和 U2 在反映文献被使用情况时, 代表了不同的使用动机与使用倾向, 如果利用使用数据来进行文献资源保障分析, 建议利用两种使用量指标进行综合考量以便相互补充。

此外, 本研究通过实证分析, 发现在 329 种“经济学与商学”高使用量论文所在期刊中, 华东师范大学图书馆可保障 276 种, 保障率为 83.9%, 缺藏的 53 种期刊在 4 个 JCR 分区中均有分布, 说明高使用量的论文并不一定分布在高影响因子期刊上, 一些低影响因子期刊上刊载的论文同样受到读者的关注并产生使用量, 这一点与低影响因子期刊上刊载的论文同样有可能成为高被引论文类似^[39-40], 也从另一个角度说明了“以刊评文”的局限性。

而 CASHL 对“经济学与商学”高使用量论文所在期刊的保障率达到 99.1%, 仅 3 种期刊未保障, 极大地弥补了一校一馆文献保障率的不足。但是, 研

究中也发现部分高使用量的 OA 期刊在高校图书馆及“CASHL 全文期刊目录”中并未揭示, 这不利于读者对 OA 期刊的使用, 建议在全文期刊目录中将高质量的 OA 期刊加入其中, 以使 OA 期刊被更广泛地使用。

5 结语

文献资源是图书馆的立馆之本, 是图书馆工作的重中之重。但是对于一校一馆, 受经费所限, 在资源采购时往往只能侧重本校重点学科, 对部分弱势学科的资源覆盖率可能相对较低, 因此需要从用户需求的角度提高资源保障率。基于使用数据计量为文献资源保障研究提供了一个新的视角, 在工作实践中亦可作为引文分析的有益补充。此外, CASHL 等高校图书馆联盟通过成员高校“优势学科重点部署, 联盟全局统筹规划”的保障方式, 促进了学科资源的共建共享, 为成员高校的教学和科研提供了有力支撑, 是高校在有限经费条件下实现学术资源保障的重要途径之一。

参考文献

- 汪涛, 肖希明. 新信息环境下的文献资源保障体系建设[J]. 图书与情报, 1999(1): 33-36.
- 曲岩岩, 童旭. 高校特色重点学科文献资源保障实证研究——以哈尔滨工程大学“核学科”为例[J]. 图书馆研究, 2015, 45(2): 31-35.
- 李海霞. 基于引文分析的图书馆期刊资源建设策略研究——以哈尔滨工程大学图书馆为例[J]. 图书馆建设, 2011(8): 43-45.
- 马建华. 引文分析在图书馆文献采集中的作用——北京大学有机化学专业博士论文的文獻计量研究[J]. 大学图书馆学报, 2003(3): 70-73.
- 徐志玮, 郑建瑜. 高校化学学科用户对纸本/电子期刊需求研究——以中山大学化学学科用户为例[J]. 图书情报知识, 2010(4): 44-50.
- 侯利娟, 郝群, 张立彬. 高校图书馆外文文献资源保障研究——以复旦大学数学学科为例[J]. 图书馆, 2018(1): 92-100.
- 唐惠燕, 包平, 林小娟. CASHL 文献资源保障实证研究——以教育部直属农业院校为例[J]. 图书馆理论与实践, 2012(4): 70-73.
- 宋姬芳, 单向群, 于森, 等. 基于 WoS 和用户调查的 CASHL 文献资源保障研究——以 ESI 经济学与商学学科为例[R]. 2018 年新信息环境下 CASHL 资源与服务拓展设计研究, 2018.
- 廖鹏飞, 韩爽, 李靳元. 基于学科文献需求分析探讨 CASHL 外文资源保障建设——以吉林大学、北京大学哲学学科为例[J]. 情报探索, 2021(9): 68-74.
- 刘凤仪, 叶继元. 我国区域高校图书馆西文期刊保障体系实证研究——以华东五校图书情报学为例[J]. 大学图书馆学报, 2021, 39(2): 28-34.
- 肖自力. 关于文献资源保障率的一点质疑[J]. 情报资料工作, 2000(3): 10.
- 肖自力, 李修宇, 张学华, 等. 文献资源调查评估方法初探[J]. 大学图书馆通讯, 1988(5): 2-8.



- 13 夏远椒. 馆藏文献调查与评估[J]. 图书情报工作, 1990(3): 23-28.
- 14 宋其兰, 朱亚丽. 我馆研究级学科藏书的调查、分析和评估[J]. 晋图学刊, 1989(4): 9-12.
- 15 张洪亮. 运用“书目核对”和“引文查核”的方法来评价我馆“中国文学”类藏书[J]. 晋图学刊, 1989(1): 43-46.
- 16 韩子静, 吴晶晶, 孔令芳. 高校人文社科外文纸质本图书馆藏评估实践与思考——以浙江大学图书馆哲学学科为例[J]. 大学图书馆学报, 2021, 39(4): 28-35.
- 17 林申清. 藏书评价的方法[J]. 广东图书馆学刊, 1988(4): 106-108.
- 18 肖自力. 关于文献信息资源建设的思考(上)[J]. 图书馆杂志, 2000(7): 25-27.
- 19 李曙东. 北京图书馆文献资源的调查与分析[J]. 中国图书馆学报, 1991(2): 29-42.
- 20 刘青华, 钱国富, 梁灿兴. 建立高校图书馆外文期刊保障体系实务研究[J]. 大学图书馆学报, 2012, 30(3): 59-63.
- 21 武莹. 基于引文分析的图书馆文献资源保障率研究——以中国政法大学图书馆为例[J]. 中国管理信息化, 2015, 18(16): 218-219.
- 22 Wang X W, Peng L, Zhang C B, et al. Exploring scientists' working timetable: a global survey[J]. Journal of Informetrics, 2013, 7(3): 665-675.
- 23 Wang X W, Xu S M, Peng L, et al. Exploring scientists' working timetable: Do scientists often work overtime? [J]. Journal of Informetrics, 2012, 6(4): 655-660.
- 24 Duan Y F, Xiong Z Q. Download patterns of journal papers and their influencing factors [J]. Scientometrics, 2017, 112(3): 1761-1775.
- 25 Bollen J, de Sompel H V. Mapping the structure of science through usage[J]. Scientometrics, 2006, 69(2): 227-258.
- 26 Schloegl C, Gorraiz J. Comparison of citation and usage indicators: the case of oncology journals[J]. Scientometrics, 2010, 82(3): 567-580.
- 27 Schloegl C, Gorraiz J. Global usage versus global citation metrics: the case of pharmacology journals[J]. Journal of the Association for Information Science and Technology, 2011, 62(1): 161-170.
- 28 赵星. 学术文献用量级数据 Usage 的测度特性研究[J]. 中国图书馆学报, 2017(3): 44-57.
- 29 Wang X W, Fang Z C, Sun X L. Usage patterns of scholarly articles on Web of Science: a study on Web of Science usage count [J]. Scientometrics, 2016, 109(2): 917-926.
- 30 Liang G Q, Hou H Y, Hu Z G, et al. Usage Count: a new indicator to detect research fronts[J]. Journal of Data and Information Science, 2017, 2(1): 89-104.
- 31 Zhao S X, Lou W, Tan A M, et al. Do funded papers attract more usage? [J]. Scientometrics, 2018, 115(1): 153-168.
- 32 Markusova V, Bogorov V, Libkind A. Usage metrics vs classical metrics: analysis of Russia's research output[J]. Scientometrics, 2018, 114(2SI): 593-603.
- 33 Liang G Q, Hou H Y, Chen Q, et al. Diffusion and adoption: an explanatory model of “question mark” and “rising star” articles[J]. Scientometrics, 2020, 124(1): 219-232.
- 34 谢娟, 龚凯乐, 成颖, 等. 使用数据与引用数据间的补充或替代关系探讨[J]. 情报学报, 2018, 37(5): 486-494.
- 35 熊泽泉, 段宇锋. 论文早期下载量可否预测后期被引量? ——以图书情报领域期刊为例[J]. 图书情报知识, 2018(4): 32-42.
- 36 李卓, 赵梦圆, 柳嘉昊, 等. 基于引文内容的图书被引动机研究[J]. 图书与情报, 2019(3): 96-104.
- 37 马凤, 武夷山. 关于论文引用动机的问卷调查研究——以中国期刊研究界和情报学界为例[J]. 情报杂志, 2009, 28(6): 9-14.
- 38 邱均平, 陈晓宇, 何文静. 科研人员论文引用动机及相互影响关系研究[J]. 图书情报工作, 2015, 59(9): 36-44.
- 39 Lozano G A, Lariviere V, Gingras Y. The weakening relationship between the impact factor and papers' citations in the digital age[J]. Journal of the American Society for Information Science and Technology, 2012, 63(11): 2140-2145.
- 40 Wu L L, Huang M H, Chen C Y. Citation patterns of the pre-web and web-prevalent environments: the moderating effects of domain knowledge[J]. Journal of the American Society for Information Science and Technology, 2012, 63(11): 2182-2194.

作者单位: 华东师范大学图书馆, 上海, 200241

收稿日期: 2021年9月30日

修回日期: 2021年11月26日

(责任编辑: 关志英)

Research on the Guarantee of CASHL Literature Resources Based on Usage Metrics

Xiong Zequan Peng Xia Yang Li

Abstract: The value of literature resources can be reflected not only by the citing behavior of scholars, but also by other using behavior of readers, such as downloading, reading and sharing. In this study, the high-usage papers in the discipline of “economics and business” in Web of Science were selected for scientometric analysis. The general characteristics of high-usage papers were described, followed by the analysis of guarantee rate of East China Normal University and CASHL (China Academic Social Sciences and Humanities Library) for high-usage papers. It is found that the high-usage papers are generally the latest published or with high journal impact factors, and there is a significant positive correlation between the usage counts and the citations. East China Normal University's guarantee rate for the high-usage papers in the discipline of “economics and business” is only 83.9%, but CASHL's guarantee rate in this discipline reaches 99.1%, which greatly makes up for the lack of guarantee rate in a single university. Usage metrics is expected to be a useful supplement to the study of literature guarantee rate based on citation analysis.

Keywords: Usage Data; Usage Metrics; Guarantee of Literature Resources; CASHL