



开放科学
(资源服务)
标识码
(OSID)

材料科学高水平论文使用特征与影响力比较研究

匡登辉

南开大学图书馆 天津 300071

摘要: [目的/意义] 探索材料科学高水平论文基于使用数据、被引数据、合作者数据的科学评价,从定量角度揭示高质量学术论文的使用情况及其影响因素。[方法/过程] 选取 Web of Science 提供的数据,以 *Nature Reviews Materials*、*Nature Energy* 刊载的高水平论文 (HCP、HP) 为分析对象,遴选使用数据 (U1、U2), 引文数据 (TC), 合作者数量 (Coauthor) 作为研究指标,分析论文使用和被引的相关关系以及论文质量、文献类型及作者数量对两者关系的调节作用,根据研究发现提出相应的建议。[局限] 用户使用动机与使用请求的关系及特定使用请求类型对文献的兴趣程度需要进一步研究,样本数据目前限于材料学科,需要扩充。[结果/结论] 顶级期刊高水平论文的引用率超过 52%,超过 61% 的论文使用次数在两次以上。高影响力论文的指标数据较为分散、随机,差异较大,但均值均高于样本总体的均值。高影响力论文的引文数据和使用数据存在正的相关关系,高被引论文的合作者数量与使用数据存在正的弱相关,热点论文的合作者数量与使用数据相关关系不明确。两本顶级期刊的热点论文,使用数据无显著差异。综述类论文获得较高的使用次数和引用频次。

关键词: 材料科学; 学术文章; 高水平论文; 使用次数; 影响力

中图分类号: G35; G250.252

Comparative Study on the Characteristics of Usage and the Impact of Articles from Materials Science High-level Articles

KUANG Denghui

Nankai University Library, Tianjin 300071, China

Abstract: [Purpose/Significance] Explore the establishment of scientific evaluation of high-impact papers in Materials Science based on top journals usage count, citation data and coauthors. From a quantitative point of view, the use of high-quality academic papers and their influencing factors are revealed. [Methods/Processes] Select data of Web of Science platform, with *Nature Reviews Materials*, *Nature Energy*'s high-impact papers (HCP, HP) as the analysis object, select usage data (U1, U2),

基金项目 2020 年天津市哲学社会科学规划项目“多源异构数据的学术论文评价体系构建与实践探索研究”(TJTQ20-001)。

作者简介 匡登辉(1981-), 硕士, 副研究馆员, 研究方向为科学评价计量与知识产权信息服务, E-mail: dhkuang@nankai.edu.cn。

引用格式 匡登辉. 材料科学高水平论文使用特征与影响力比较研究[J]. 情报工程, 2023, 9(6): 57-68.

citation data (TC) and Coauthors as the research indicators, the correlation between usage and citations of papers and the moderating effects of paper quality, document type and number of authors on the relationship are analyzed. Finally, corresponding recommendations are made based on the findings of the study. [Limitations] The relationship between user motivation and use requests, and the degree of interest in literature of specific use request types, require further research, and sample data are limited to material disciplines and need to be expanded at present. [Result/Conclusion] The citation rate of top journals was more than 52%, and more than 61% of papers were used more than twice. The index data of high-impact thesis are more dispersed, random and different, but the mean is higher than the mean of the sample population. There is a positive correlation between the citation data and the usage data of high-impact papers, the number of collaborators in high-impact papers is positively correlated with the use of data, and the number of collaborators in hot papers is not clearly related to the use of data. There was no significant difference in Hot Paper data between the two top journals. Review papers are used more often and cited frequently.

Keywords: Materials Sciences; Academic Papers; High-level Articles; Usage Counts; Impact

引言

ESI 高水平论文 (Top Papers) 是 ESI 发布的以论文被引频次为标准的高学术影响力的论文, 通常指高被引论文 (Highly Cited Paper, 简称 HCP) 和热点论文 (Hot Paper, 简称 HP), 分别根据同一年同一 ESI 学科统计的近十年发表论文中被引用次数进入世界前 1% 和按 ESI 学科统计最近两年发表、在最近两个月里被引用次数进入世界前 0.1% 的论文^[1], 代表着相关学科领域动态变化着的研究前沿和热点, 对指引学科发展、展现成果学术影响力具有重要意义。基本科学指标 (Essential Science Indicators, 简称 ESI) 是基于 Web of Science (WoS) 核心合集数据库的深度分析型研究工具, 数据每两个月更新一次。当前, 高水平论文备受重视, 其科研价值是评价学科、人才和机构国际影响力的重要指标之一。ESI 高水平论文的甄选基于引文分析方法, 引文分析作为科学产出质量的测度, 被广泛采用, 研究显示, 被引频次与同行评议之间存在着较强的正相关^[2], 与学术共同体的认可密切相关。但引文的滞后

性和“马太效应”等问题不可避免。此外, 引用通常为撰写学术文章的研究人员的行为, 但是对文献的使用不止研究者的行为, 因而并不能代表论文就具有真正的高影响力, 论文影响力的产生是多方面的, 影响力也是由多重因素共同决定的。

Garfield^[3] 曾提出利用下载量来弥补被引次数的时滞性问题, 受到广泛关注。其实, 论文下载是使用的一个方面, 使用次数反映论文发表后引起同行关注的程度, 其数据范围包括但不限于下载。一般来说, 使用次数越高其传播速度越快。此外, 使用次数避免了文献著录格式不规范、引用时的人为因素造成的缺陷。使用数据提供了科学过程各阶段、各层面的活动相当全面的信息, 远超引文提供的基本信息, 而且对科学交流的影响更广泛, 更普遍。因此, 能否建立基于使用数据并结合引用数据的高水平论文的科学评价体系, 成为科学计量学研究的热点。此外, 合作者数量对于高水平论文的引用有积极影响^[4], 但综合合作者数量指标, 对高水平论文的使用的对比评价研究较少。有鉴于此, 本文拟采用 WoS 提供的使用次数 (usage

count) 作为使用数据源, 结合被引频次、合作者数量等指标, 对高水平论文的不同使用特征及与被引频次的关系以及合作者数量、文献类型对论文使用与引用的调节作用进行研究。

1 文献综述

对于论文的使用研究, 由于技术限制, 以往的使用数据仅限于在图书馆印刷品的使用记录^[5-6]。如 King 等^[7]利用纸质资源重新上架、流通等数据来确定文献使用率, 研究期刊的使用情况和引文排名。纸质文献是静态的, 但网络服务是动态的, 因此, 学术性资源的网络服务被视为复杂知识基础设施网络中的节点。有研究者选择“阅读”“使用”“下载”和“点击”方面讨论文献使用^[8-11]。近年来, 经过严格审查的大型数据库的使用数据可用于文献计量分析, 2015年9月, WoS 推出反映单篇文献受关注程度的新数据: Usage Count, 在用户使用论文后的第二天, 使用数据就能够被 WoS 数据库记录, 时滞短, 可用于直观展示用户感兴趣的研究。

对于高水平论文的使用研究, 国内学者如丁佐奇等^[12]、苟莉等^[13]对 Altmetric.com Top 100 论文、ESI 热点论文进行 Altmetrics 分值、使用次数和被引频次的相关性分析; 匡登辉^[14]、黄晓等^[15]则研究高被引论文中不同学科的 Altmetrics 指标特征, 分析高被引论文的引用频次与在线使用、分享的相关性及表现, 验证 Altmetrics 指标在揭示优质文献方面的有效性; 梁国强等^[16]以再生医学论文为研究对象, 探索高质量论文使用次数和被引次数在论文类型、年代、期刊

和学科水平的分布特征, 发现发表时间越早, 高质量论文使用次数与被引次数的相关系数越高; 刘运梅等^[17-18]研究 PLOS 论文的使用类指标, 发现其具有高及时性、低波动性、高持续性的特点, 可用于挖掘论文评价指标的学科差异性; 外国学者如 Batooli 等^[19]基于 ResearchGate 指标评估伊朗研究人员的高被引论文和热点论文, 结果显示 ResearchGate 可显著提高科学作品的可见度; Wang 等^[20]对一组高被引论文的计量学指标实证比较, 发现论文早期的引用特征、网络使用量和首作者声望可预测论文未来潜力; Martinez 等^[21]描述 IRRODL (《开放和远程学习研究国际评论》) 的主要计量学指标, 并与 IRRODL 出版的所有格式论文的总浏览量进行比较, 显示论文总被引量与论文浏览量存在显著的正相关。

上述研究多是基于 Altmetrics 指标中的使用数据, 其中研究者对使用数据的归类不同, 缺乏统一的数据来源, 另外在统计合并中存在错误风险, 因此研究结果各不相同。本文采用 WoS 的统一数据源, 获取高水平综合性材料期刊论文的使用数据, 综合引用数据、合作者数量, 重点关注以下 3 个问题: (1) 高水平论文的使用模式有怎样的特征? (2) 高水平论文的使用数据与被引数据的关系如何? (3) 合作者数量、文献类型对论文使用与引用的调节作用有哪些特点?

对上述问题的关注, 对研究和推动我国的高水平期刊发展建设有重大意义。我国为实现从“论文大国”到“科研强国”的转变, 积极推进“双一流”建设, 材料科学进入 ESI 全球前 1%, 效果显著, 但高水平研究仍处于“追赶”

态势。本研究一方面有助于了解高水平论文的使用特征,为推动我国高水平研究提供有益参考;此外通过对高水平论文使用数据与被引数据、合作者数量、文献类型的关系探讨,有助于明确高水平论文的影响因素,对促进高水平学术产出有借鉴意义,探讨高水平论文使用次数、被引次数的相关性,克服科研评价中被引指标的滞后性。

2 研究对象与过程

2.1 研究对象

据中国科学院文献情报中心发布的《2016-2017 学科发展报告综合卷》显示:从 ESI 各学科领域论文数量看,我国材料科学居世界第一^[22],但大多数论文发表在国外的水平期刊上。数据显示,2018 年中国对自然指数材料学论文的贡献年增幅为 15.8%^[23]。据 2020 年期刊引证报告, *Nature Reviews Materials*《自然综述:材料》和 *Nature Energy*《自然·能源》(均创刊于 2016 年)一跃成为综合类材料科学期刊的影响因子前两名,和 *Nature Materials*《自然·材料》(以下称其英文刊名)占据材料科学期刊的前 3 名^[24]。这些期刊创刊之初即位于学科领先地位,首个期刊影响因子取得学科领先排名,一般出版第 2 年即可被 SCI 收录。因此,本研究选择 *Nature Reviews Materials* 和 *Nature Energy* 作为研究样本,对比高被引论文和热点论文的使用特征和影响力。

2.2 数据获取

检索数据采集以下步骤:在 Science Cita-

tion Index Expanded (SCI-EXPANDED) 数据库中进行检索,检索策略选择各期刊名称,全记录格式下载数据,数据下载时间为 2022 年 1 月 6 日,获得 *Nature Reviews Materials* 373 篇论文,其中, HCP 87 篇,比例达 23.32%, HP 8 篇,占比 2.14%; *Nature Energy* 916 篇论文, HCP 172 篇,占比为 18.78%, HP 36 篇,占比 3.93%。数据保存成文本格式,导入 Excel 2019。

2.3 研究方法

引文数据需要至少 2-3 年的时间积累才稳定,引文数据选择 WoS 平台提供的 TC 字段,来自 WoS 核心集内的被引频次;使用数据包括两类:U1(数据下载之前 180 天内论文的使用次数,短时使用次数)和 U2(2013 年 2 月 1 日至数据下载时论文的使用次数,长时使用次数)。Coauthor(作者数量)即论文是否为合作论文,在用户论文引用和下载的众多决策影响因素中占据重要地位。因此,将 Coauthor 纳入研究范围,但 Coauthor 数据无法直接获取,因此利用 WoS 提供 AU 字段间接获得 Coauthor 数据。

对 Coauthor 的统计,根据 WoS 导出数据特征,采用 LEN 函数,假设作者的单元格是 A,则公式为:

$$\text{Coauthor} = \text{LEN}(A) - \text{LEN}(\text{SUBSTITUTE}(A, ";", "")) + 1 \quad (1)$$

本研究利用分组比较、描述性统计分析、Spearman 相关性分析、2 个独立样本的非参数检验等方法,探析顶级期刊高水平论文使用指标,探索性分析这些差异,以期完善科技期刊分类评价体系,推动我国高水平期刊的建设提供新的数据支撑和数据参考。对于两种期刊的使

用差别分析，采用 2 个独立样本的非参数检验，SPSS 提供 4 种独立样本的非参数检验方法，在本文中选用 Kolmogorov-Smirnov Z 双样本检验（下文简称“K-S Z 检验”），判别两种期刊两个独立样本所属的总体是否有相同的分布。

数据统计采用 Excel 2019，数据分析采用 SPSS 25.0，可视化工具采用 Tableau 10.5。

3 研究结果

3.1 高水平论文的计量指标统计

使用率衡量文献当前的使用情况，引文数据衡量所有过去使用情况。使用是即时行

为，而科研发文、出版、引用则需要很长的周期。从图 1 可以看出，两本期刊的样本数据，超过 52% 的论文得到了多于 1 次引用，其中 *Nature Reviews Materials* 的引用率为 54.4%，*Nature Energy* 的引用率为 52.8%。对于使用数据，超过 61% 的论文使用次数在两次或以上，其中 *Nature Reviews Materials* 的短时使用率为 67.8%，*Nature Energy* 的短时使用率为 61.5%，长时使用数据在两次及以上的论文比例接近或达到 90%。材料科学的高水平论文在 WoS 中得到了极高的使用和引用，发表当年的论文也取得了高水平的使用次数（短时使用率高）。

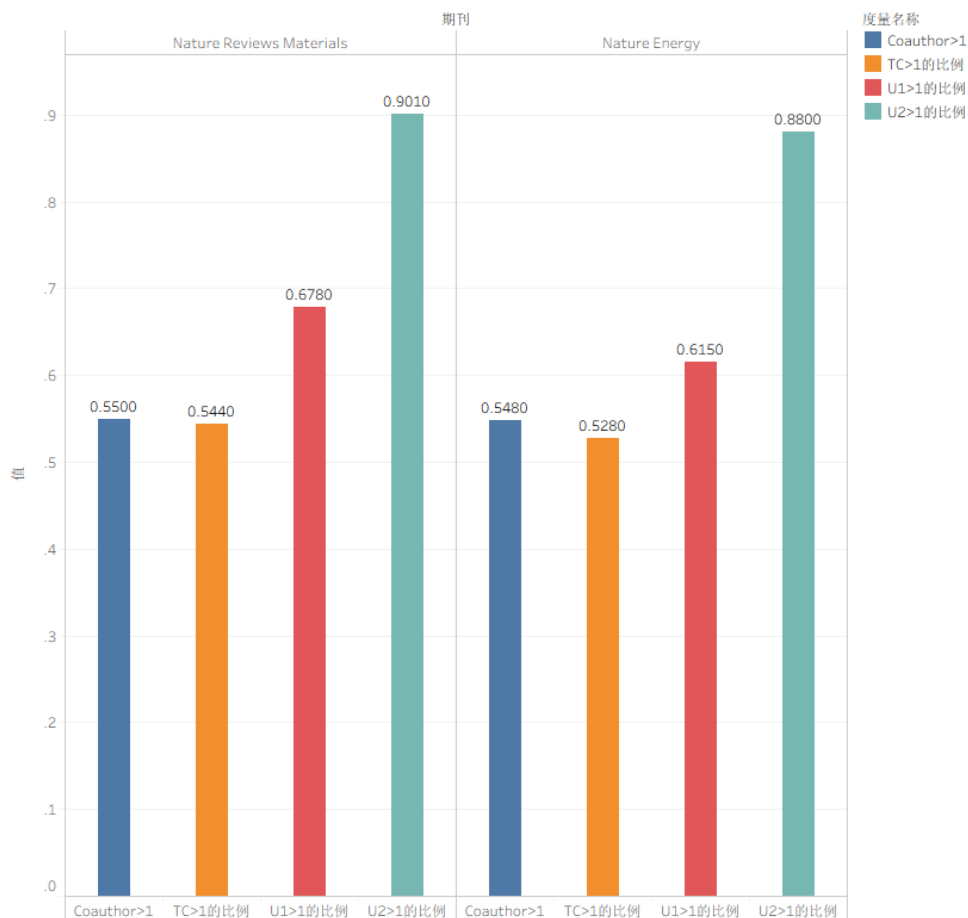


图 1 各指标相对比例的直方图

3.2 高水平论文指标分析

表 1 为高水平论文各指标绝对值分布图。两本期刊高水平论文作者数量最小值仅 HCP 相同, 均为 1 人, 最大值变化大, 如 *Nature Reviews Materials* HCP 为 15 人, 而 *Nature Energy* HCP 为 23 人。使用数据两本期刊的差异较大, 对于短时使用、引文数据的最小值, *Nature Reviews Materials* 的高水平论文均高于 *Nature Energy* 的最小值, 最大值低于 *Nature Energy* 的最大值; 对于长时使用, *Nature Reviews Materials* 的高水平论文, 最小值和最大值均高于 *Nature Energy* 的最小值和最大值。对于 Coauthor、U1、U2 这三种指标的均值, HP 均高于 HCP, 对于引文数据的均值, HP 低

于 HCP, 这是因为数据的统计时段、样本范围不同, 前者选用的是最近两年发表、在最近两个月里被引用次数进入世界前 0.1% 的论文, 后者则是近十年发表论文中被引用次数进入世界前 1% 的论文。对于 *Nature Reviews Materials*, 除长时使用数据外, 其余指标的标准差均为 HCP 高于 HP, 随机变量不稳定, 取值分散。对于 *Nature Energy*, 则完全相反, 除长时使用数据外, 其余指标的标准差均为 HCP 低于 HP。这说明 HCP、HP 的指标数据较为分散、随机, 差异较大。使用数据的积累, 很容易地超过引用总数, 从而显著提高计量学研究的可信度和范围。高水平论文的使用, 比普通论文频繁, 在一定程度上, 使用数据表征论文的受欢迎程度。

表 1 高水平论文各指标绝对值分布

| 指标 | 论文类型 | 最小值 | 最大值 | 均值 | 标准差 | 方差 |
|--------|-------|-----|------|--------|---------|------------|
| 作者数量 | 高被引论文 | 1 | 15 | 4.23 | 2.250 | 5.063 |
| | | 1 | 23 | 8.62 | 4.504 | 20.283 |
| | 热点论文 | 3 | 7 | 4.75 | 1.488 | 2.214 |
| | | 2 | 20 | 9.53 | 4.855 | 23.571 |
| 短时使用次数 | 高被引论文 | 14 | 521 | 113.37 | 92.767 | 8605.654 |
| | | 1 | 623 | 77.22 | 78.532 | 6167.199 |
| | 热点论文 | 63 | 339 | 153.13 | 88.194 | 7778.125 |
| | | 28 | 623 | 129.17 | 105.996 | 11235.114 |
| 长时使用次数 | 高被引论文 | 112 | 2344 | 576.99 | 465.214 | 216424.244 |
| | | 5 | 1976 | 336.05 | 288.388 | 83167.857 |
| | 热点论文 | 200 | 1687 | 634.38 | 485.194 | 235413.125 |
| | | 99 | 953 | 373.50 | 217.142 | 47150.714 |
| 被引频次 | 高被引论文 | 27 | 1059 | 238.39 | 201.695 | 40680.939 |
| | | 9 | 1375 | 174.58 | 185.205 | 34300.876 |
| | 热点论文 | 138 | 522 | 228.75 | 127.510 | 16258.786 |
| | | 25 | 660 | 137.44 | 113.790 | 12948.140 |

注: 表 1 为双层数据, 其中上层为 *Nature Reviews Materials* 高水平论文数据, 下层为 *Nature Energy* 高水平论文数据, 表 2 相同。

3.3 高水平论文使用特征比较分析

将高水平论文的长时使用次数进行对比，从整体看，*Nature Energy* 使用次数集中在 500 次以下，且最高的使用次数也低于 2000 次。从个体看，只有少数论文以较高的使用次数产

生了较高的被引次数，使用次数高于 1100 而被引次数低于 80 的论文有 2 篇，使用次数最高的论文（1976 次）其被引次数为 909 次，处在较高水平，论文的使用次数对被引次数有积极的影响。

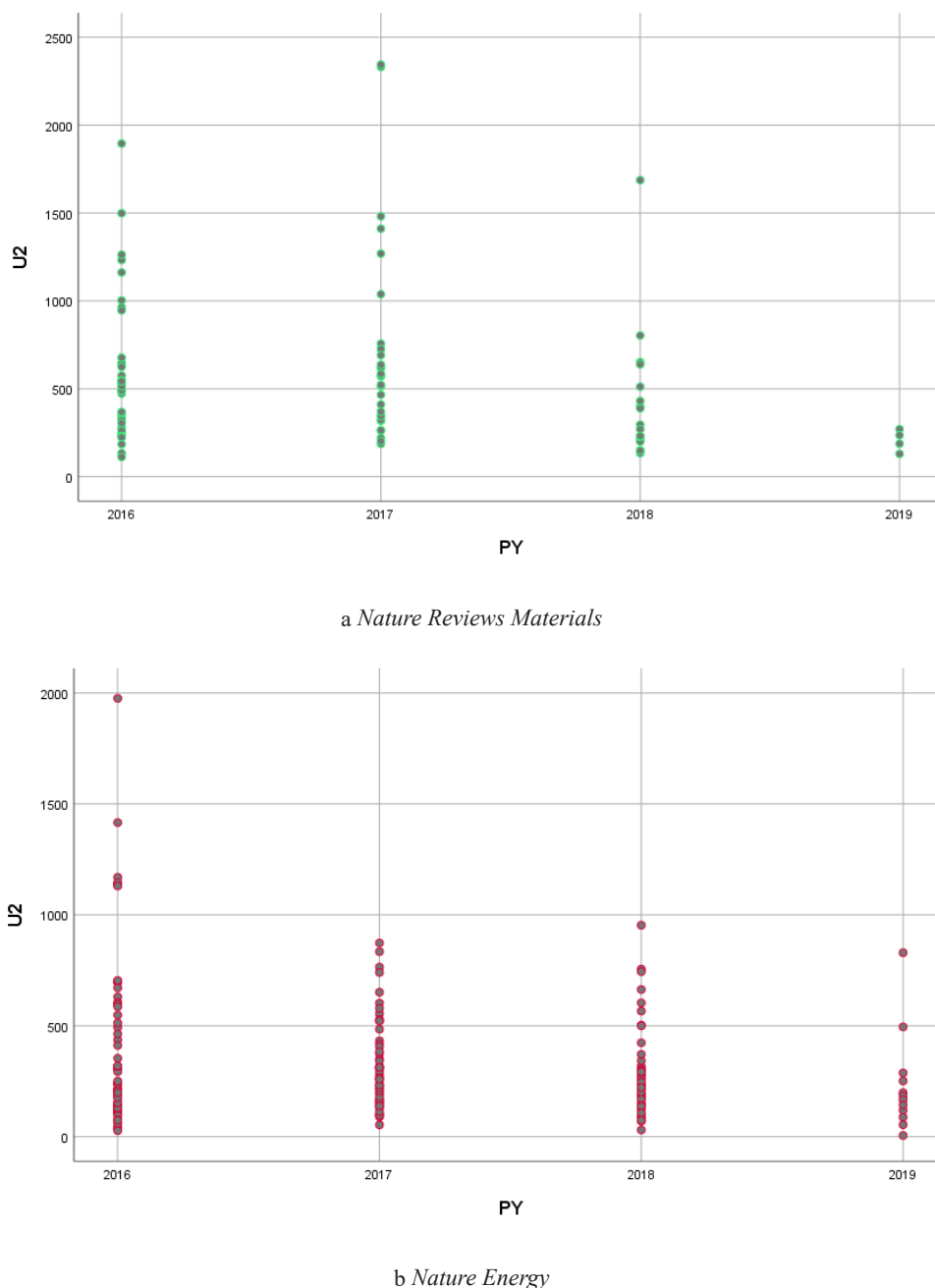


图 2 两种期刊高水平论文使用次数绝对值分布

使用次数与被引次数是相互联系的。在本研究中,高水平论文的引文频次和使用次数存在正的相关关系,使用数据处于较低水平,被引次数也相应较低,对使用次数和被引次数进行相关性检验,最大相关

系数为 0.522,呈中度正相关。在某种程度上,论文被引也属于论文使用的范畴,被引是论文的事实使用行为,但使用次数仅是一种潜在的引用行为,这主要体现为论文下载和阅读。

表 2 高水平论文 Kendall 相关系数

| 论文类型 | 被引次数 VS 短时 使用次数 | 被引次数 VS 长时 使用次数 | 作者数量 VS 短时 使用次数 | 作者数量 VS 长时使 用次数 | 被引频次 VS 作者 数量 |
|-------|--------------------|--------------------|--------------------|--------------------|------------------|
| 高被引论文 | 0.341** | 0.522** | 0.178* | 0.081 | 0.125 |
| | 0.186** | 0.453** | 0.095 | 0.034 | -0.015 |
| 热点论文 | 0.214 | 0.429 | -0.265 | -0.189 | -0.038 |
| | 0.093 | 0.350** | 0.036 | -0.147 | -0.260* |

** 在置信度(双侧)为 0.01 时,相关性是显著的; * 在置信度(双侧)为 0.05 时,相关性是显著的。

科研合作的确会形成较好的研究成果,HCP 的合作者数量与使用数据存在正的弱相关,HP 的合作者数量与使用数据相关关系不明确,其中 *Nature Reviews Materials* 合作者数量与使用数据存在负相关,*Nature Energy* 合作者数量与短时使用存在正的弱相关,与长时使用存在负相关。

科研合作作为重要的学术交流方式,可以共享科研思路和专业知识,但是为了学术质量,需要保持最佳规模,并非合作者越多越好。对于高水平论文的科研合作,本研究显示高水平论文的引文数据和合作者数量的相关关系不明确,其中 HP 的合作者数量与引文数据存在负相关,HCP 的引文数据与合作者数量相关关系不明确,其中 *Nature Reviews Materials* 合作者数量与引文数据存在正的弱相关,*Nature Energy* 合作者数量与引文数据存在负的弱相关。Chuang 等^[25]对“化学工程”论文的研究表明,高被引论文的作者数量较少,可能是单国论文,

总体上还没有达到引文峰值。Toivanen 等^[26]的研究也有类似的结论,他们以 1995—2009 年芬兰的研究成果为数据样本,使用完全归一化方法评估作者数量和被引用次数的相互关系,发现芬兰的小型作者团队比大型的国际作者团队生产力更高,引用次数更多。乔家昌^[27]研究图情领域的 SCI 期刊论文发现,2—3 位作者的高影响力论文相对较多。

在本文中,合作者最多的高被引论文和热点论文均来自 *Nature Energy*,作者数分别为 20 和 23,合作者分别来自美国和英国;引用频次分别为 500、69,长时使用次数分别为 829、764,分别研究“实用的高能量长循环锂离子电池的途径”和“效率提高了 23.6% 的整体钙钛矿 / 硅串联太阳能电池”,为综述文章和研究论文;*Nature Reviews Materials* 高被引论文合作者最少的为 2 位,引用频次为 1069,长时使用次数为 1895,论文类型为综述,热点论文合作者最少的为 3 位,引用频次为

247, 长时使用次数为 1687, 论文类型为综述。*Nature Energy* 高被引论文合作者最少的为 2 位, 引用频次为 190, 长时使用次数为 374, 论文类型为综述, 热点论文合作者最少的为 2 位, 引用频次为 1077, 长时使用次数为 500, 论文类型为综述。前期的研究也证实, 在一段时间内, 论文类型影响文献的使用, 综述类获得较多的使用和引用^[28], 在材料科学高水平论文中依然如此。

3.3 高水平论文使用指标的非参数检验

使用指标数据所属总体不服从正态分布, 可采用非参数检验判断两种高水平论文的总体分布是否存在显著差异性。SPSS 提供 4 种独立样本的非参数检验方法, 在本文中选用 K-S Z 检验, 判别两本期刊的两类高水平论文独立样本所属的总体是否有相同的分布。

首先, 对样本进行分组: 1.0 组为 *Nature Reviews Materials* 的 HCP, 2.0 组为 *Nature Energy* 的 HCP, 以 U1 为例;

其次, 提出本研究的虚无假设 (H_0)。 $H_0: \mu_1 = \mu_2$, 即 *Nature Reviews Materials* 的 HCP 和 *Nature Energy* 的 HCP 的 U1 无显著差异;

最后, 将分组后的数据读入 SPSS, 运行 K-S Z 检验, 结果如表 3、表 4 所示。

表 3 K-S Z 检验频率

| 分组 | 短时使用次数 | 长时使用次数 |
|-----|--------|--------|
| 1.0 | 87 | 87 |
| 2.0 | 172 | 172 |

表 3 显示参与检验的两组的样本量, 1 组即 *Nature Reviews Materials* 的 HCP 的样本量为 87,

2 组即 *Nature Energy* 的 HCP 的样本量为 172。

表 4 K-S Z 检验结果 ($p=0.05$)

| | 绝对 | 统计量 | |
|------------|----|--------|--------|
| | | 短时使用次数 | 长时使用次数 |
| 最极端值 | 正 | .365 | .341 |
| | 负 | -.006 | .000 |
| K-S Z 值 | | 2.777 | 2.593 |
| 渐进显著性 (双尾) | | .000 | .000 |

分组变量: group

表 4 为 K-S Z 双样本检验结果, U1 双尾显著性概率为 0.000, 小于 0.05, 因此在 95% 的置信区间内, 拒绝原假设, 即 *Nature Reviews Materials* 的 HCP 和 *Nature Energy* 的 HCP 的 U1 有显著差异。

同理, 对 HCP 的 U2 指标进行检验, 结果发现: 在 95% 的置信区间内, *Nature Reviews Materials* 的 HCP 和 *Nature Energy* 的 HCP 的 U2 有显著差异 (双尾显著性概率为 $0.000 < 0.05$)。HP 的使用数据的检验过程和 HCP 相同, 发现: 在 95% 的置信区间内, HP 的使用数据无显著差异 (双尾显著性概率为 $0.574 > 0.05$, $0.361 > 0.05$)。

4 讨论

(1) 论文的引用与可见性、使用的辩证分析。论文被引的前提是“可见”, 即论文在索引数据库中是否可以被检索会影响其能否被下载和使用。论文的“可见性”在论文的使用与论文被引用之间起到“中介”作用。使用数据记录在给定的一段时间内各个使用事件的数据

集合：①点击链接，获取全文；②将文献题录信息下载到文献管理软件中；③将数据保存至文件。使用数据规模非常大，请求元数据记录在单篇论文级。用户对特定资源感兴趣时，发出使用请求，使用与特定学术资源相关的特定信息服务。与论文被引频次相比，使用数据通常由于用户对资源感兴趣，体现研究人员对文献的关注度，具有更广泛的影响力^[29]。特别是短时使用数据，反映出用户在论文发布 180 天内，更喜欢使用新文献，读者的兴趣，反映在阅读、学习、传递、实践应用等广泛的学术交流行为，也更多地反映出近来本领域内研究的趋势和热点，尤其是重要的新兴科学主题或突破性论文^[30]。但是，它们只能作为用户动机、兴趣、需求或请求后临时操作导致的使用，论文被使用与否，具有一定的随机性，文献外部特征项如作者、关键词、期刊等影响用户的判断，影响文献进一步的使用，部分使用会转化为被引^[31]。陈悦等^[32]的研究结果显示：使用新颖、易于理解的词汇的标题和摘要的文章，可以获得较高的使用量。俞立平等^[33]基于 JCR 纯数学期刊数据，指出论文数量与论文影响力负相关，说明单纯增加载文量无法提高论文影响力（引用和使用）。此外，WoS 的使用数据，产生的前提是研究人员利用该平台进行检索，产生论文使用数据，由于研究习惯和获取文献习惯的不同，很多研究者通过学术搜索引擎或期刊全文数据库来浏览获取相关的论文，该类使用数据不在统计之列。

（2）使用类型对科学交流的影响探讨。贝尔纳^[34]指出，科学交流不仅包括科学家之间的交流，还包括与大众的交流。因此，提高成果可见性成为学术交流中最重要的一环。此外，

不同的文献请求类型表示不同程度的用户兴趣，例如，查看文章摘要的请求仅表示用户对文章的兴趣，需求级别不高，而要求全文下载可能在于获得相关信息，积累研究资料，相关研究也表明，论文的高被引与其访问类型（开放式或受限式）无关^[35]。Hicks 等^[36]指出科学论文近一半的下载行为是出于学术研究的目的，但更多的下载次数是由教育环境（严格意义上的）以外的人贡献的，表明用户对终身学习的强烈意愿。用户具体的使用动机与使用请求的关系，以及特定使用请求类型对文献的兴趣程度，将是下一步研究的内容。

（3）材料科学论文的使用特征。材料科学的高水平论文的使用数据体现研究人员对文献的关注度，具有更广泛的影响力。被引频次对使用次数有积极的影响；高水平论文的引文数据和合作者数量的相关关系不明确，科研合作需要保持最佳的合作规模；论文类型对高水平论文的使用有积极影响，综述类论文获得较多的使用和引用，使用次数转化为被引次数具有更高的效率；材料科学的高水平论文的影响力与高效率的使用、高质量的引用、规模适度的合作团队是密不可分的。论文发表后，短时间内引起同行关注的程度不同，使用次数越高传播速度越快，但使用数据和被引频次，同为累积性指标，无法全面反映时间因素的影响，应结合论文类型。综述类论文获得较高的使用次数和被引频次。

（4）科学传播对论文使用、引用的影响。成果是否被引用，最终取决于学术成果的质量和研究的的相关性，论文的传播和宣传（即“可见性”）起到促进作用^[37]。因为科学家采用“最

省力法则”，通过各种工具和技术，以最少的精力去筛选相关的研究资源，从同行那里收获最大化注意力。因此，我国未来的科技信息交流平台建设，应满足和适应开放创新条件下的科研用户新需求、新范式、新行为，从而收获学术共同体更多的注意，使我国的科技论文“可见性”增强，提升学术影响力。

5 结语

高影响力论文是被广泛关注的论文，在科学计量学实践中，这类关注一般用论文被引频次来量化和测度，代表着科技发展趋势前瞻和未来创新方向。论文的引用数据体现论文的学术质量，使用数据则表征论文的可见度、关注度和传播速度。但是，论文的发表、出版、引用，需要一定的周期，文献使用之后，评价滞后明显。然而，高影响力论文的遴选和评价，是基于引文分析的方法，但引用频次不是评价学术水平和创新贡献的直接依据。要获得更全面、准确、及时的科学计量，必须进行权衡取舍，简单的引文指标终将被更复杂但更准确可靠的多元评价指标所取代。本文利用使用数据，结合引文数据、论文合作者信息、论文类型等，对材料学科高水平论文进行分析。

在研究过程中，由于受到动态使用数据来源的限制，仅利用材料科学的高水平论文，所选择的样本量相对小，结论的普适性具有一定的局限。同时，除论文的使用数据与引用数据外，社交媒体已经成为重要的学术“导流”工具，即论文在学术社交网站的收藏数据、讨论数据、新闻媒体的提及数据等是否都可以用来衡量高

水平研究成果的影响力？在随后的研究中，将进一步细化研究方法，扩大数据样本，对社会关注度、合作者类型和机构的多样性进行更深入的探索。

参考文献

- [1] Clarivate.com. ESI 数据库 [DB/OL]. (2020-02-03) [2020-02-06]. <https://esi.clarivate.com>.
- [2] GARFIELD E. The 250 Most-Cited primary Authors, 1961-1975. Part II. The Correlation Between Citedness, Noble Prizes and Academy Memberships[J]. *Current Comments*, 1977(50): 5-15.
- [3] GARFIELD E. How can impact factors be improved? [J]. *BMJ Clinical Research*, 1996, 313(7054): 411-413.
- [4] DONG Y, MA H, TANG J, et al. Collaboration diversity and scientific impact[J]. *arXiv preprint arXiv:1806.03694*, 2018.
- [5] SCALES P A. Citation analyses as indicators of use of serials: Comparison of ranked title lists produced by citation counting and from use data[J]. *Journal of Documentation*, 1976, 32(1): 17-25.
- [6] GALVIN T J, KENT A. Use of a university library collection: A progress report on a Pittsburgh study[J]. *Library Journal*, 1977, 102(20): 2317-2320.
- [7] KING D W, TENOPIR C, CLARKE M. Measuring Total Readings of Journal Articles[J]. *D-Lib Magazine*, 2006, 12(10).
- [8] HENNEKEN E A, KURTZ M J, ACCOMAZZI A, et al. Use of the astronomical literature: A report on usage patterns[J]. *Journal of Informetrics*, 2009(3): 1-8.
- [9] DAVIS P M, LEWENSTEIN, B. V., BOOTH, J. G., et al. Open access publishing, article downloads and citations: Randomised trial[J]. *British Medical Journal*, 2008(337): 586.
- [10] BOLLEN J, VAN DE Sompel H, SMITH J, et al. Toward alternative metrics of journal impact: A comparison of download and citation data[J]. *Information Processing & Management*, 2005, 41(6): 1419-1440.
- [11] BRODY T, HARNAD S, CARR L. Earlier Web usage statistics as predictors of later citation impact[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57(8): 1060-1072.

- [12] 丁佐奇, 郝海平. Altmetrics 与传统计量指标的相关性分析及对科技期刊传播的启示 [J]. 科技与出版, 2019(11): 134-139.
- [13] 苟莉, 陈一龙, 王雁, 等. Altmetrics 视角下临床医学科技期刊学术影响力提升策略初探——基于 ESI 热点论文被引频次与 Altmetrics 相关性分析 [J]. 中国科技期刊研究, 2019, 30(11): 1240-1244.
- [14] 匡登辉. 高被引论文的在线使用与分享研究——基于 ALMs 的实证分析 [J]. 现代图书情报技术, 2016(11): 54-63.
- [15] 黄晓, 高嘉慧, 吴江. 不同学科高被引论文的 Altmetrics 指标特征分析 [J]. 情报理论与实践, 2019, 42(9): 56-63.
- [16] 梁国强, 侯海燕, 任佩丽, 等. 高质量论文使用次数与被引次数相关性的特征分析 [J]. 情报杂志, 2018, 37(4): 147-153.
- [17] 李长玲, 刘运梅, 牌艳欣. 基于学科差别的论文影响力评价指标特征分析及体系构建 [J]. 情报资料工作, 2019, 40(6): 23-29.
- [18] 刘运梅, 李长玲, 杜德慧. 时间序列视角下 PLOS ALM 指标特性识别模型构建与应用 [J]. 情报资料工作, 2019, 40(6): 6-15.
- [19] BATOOLI Z. The relationship between Web of Science and ResearchGate Indicators of Iranian researchers' Top Papers [J]. Iranian Journal of Information Processing Management, 2017, 33(1): 163-186.
- [20] WANG M, WANG Z, CHEN G. Which can better predict the future success of articles? Bibliometric indices or alternative metrics [J]. Scientometrics, 2019, 119(3): 1575-1595.
- [21] AVELLO Martinez R, ANDERSON T. Are the most highly cited articles the ones that are the most downloaded? A bibliometric study of IRRODL [J]. International Review of Research in Open and Distributed Learning, 2015, 16(3): 18-40.
- [22] 中国科学技术协会. 2016-2017 中国科协学科发展报告 综合卷 [M]. 北京: 中国科学技术出版社. 2018: 128.
- [23] WU B, MALLETT D. The fastest rising institutions in materials science: visualized [J/OL]. (2019-12-12) [2020-02-06]. <https://www.nature.com/articles/d41586-019-03766-z>
- [24] Clarivate.com. Journals in MATERIALS SCIENCE, MULTIDISCIPLINARY [EB/OL]. (2020-02-03) [2022-02-11]. <https://jcr.clarivate.com/jcr/browse-journals>
- [25] CHUANG K Y, WANG M H, HO Y S. High-impact papers published in journals listed in the field of chemical engineering [J]. Malaysian Journal of Library & Information Science, 2017, 18(2): 47-63.
- [26] TOIVANEN H, SUOMINEN A. Research Collaboration and Production of Excellence: Finland 1995-2009 [C]//14th International Society of Scientometrics and Informetrics Conference, ISSI 2013. 2013: 1506-1527.
- [27] 乔家昌. 作者数量与论文影响力的关系探究 [J]. 情报探索, 2019(2): 36-43.
- [28] 匡登辉. 科学评价视角下文献使用比较研究——基于 Usage Count 的实证分析 [J]. 现代情报, 2019, 39(4): 115-124.
- [29] DUY J, VAUGHAN L. Can electronic journal usage data replace citation data as a measure of journal use? An empirical examination [J]. The Journal of Academic Librarianship, 2006, 32(5): 512-517.
- [30] 王丽婷, 匡登辉. 面向影响力评价的学术文献引用与使用比较研究 [J]. 情报理论与实践, 2019, 42(5): 171-176.
- [31] 谢娟, 龚凯乐, 成颖, 等. 论文下载量与被引量相关关系的元分析 [J]. 情报学报, 2017, 36(12): 1255-1269.
- [32] 陈悦, 宋超, 周京生, 等. 文献计量学视角下的论文被引频次影响因素研究——兼评使用与被引之间关系 [J]. 情报杂志, 2019, 38(4): 96-104.
- [33] 俞立平, 隆新文, 武夷山. 特征因子与其他文献计量指标关系研究——基于面板数据的估计 [J]. 科研管理, 2012, 33(8): 41-47.
- [34] 贝尔纳. 科学的社会功能 [M]. 陈体芳, 译. 北京: 商务印书馆出版, 1982: 417.
- [35] PACHECO A, SOUSA A, YANAI A, et al. Metrics and Altmetrics: an exploratory study of a possible correlation between the most cited papers in open and restricted access in 2016-2018 [C]//Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality. 2018: 258-264.
- [36] HICKS D, ZULLO M, DOSHI A, et al. Widespread use of National Academies consensus reports by the American public [J]. Proceedings of the National Academy of Sciences, 2022, 119(9): e2107760119.
- [37] 匡登辉, 樊振佳. 一文两面——对 PLOS 期刊论文可见性与学术影响的分析 [J]. 图书馆论坛, 2023, 43(6): 29-37.