Selection and Application of H-type Indices in Book Circulation Data Analysis: A Case Study of Northwest A&F University Library (Postprint)

Authors: Jing Lingdi, Wang Lihong, Wang Qin, Li Ya

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] To optimize collection acquisition and circulation management, and achieve standardized management for timely monitoring, evaluation, and regulation of collection acquisition, total collection volume, and structure. [Method/Process] Through correlation, variability, and adaptability analysis of H-index, A-index, and G-index among H-type indices, the A-index suitable for book circulation data analysis was selected, and according to Pareto's law, the A-index was employed to conduct statistical analysis on the top 20% of book circulation data from Northwest A&F University Library from 2014-2018. [Results/Conclusions] Research results indicate that the average A-index demonstrates satisfactory performance in statistical analysis of first-level categories of the Chinese Library Classification, reader gender, reader major, reader grade, as well as book publisher, publication year, and unit, enabling more comprehensive and accurate identification of core reader groups, core publishers, and core publication years; simultaneously, it is found that Bradford's Law can be applied not only to journal management but also to collection management. This study can provide a reference basis for exploring H-type indices in analyzing university library collection circulation data.

Full Text

Selection and Application of H-family Indices for Book Circulation Data Analysis: A Case Study of Northwest A&F University Library

Jing Lingdi, Wang Lihong, Wang Qin, Li Ya Library of Northwest A&F University, Yangling 712100

Abstract: *[Purpose/Significance]* To optimize collection development and circulation management, and to realize standardized management for timely mon-

itoring, evaluation, and regulation of collection acquisition, total volume, and structure. *[Method/Process]* Through correlation, variability, and adaptability analysis of H-family indices (H-index, A-index, and G-index), the A-index was selected as suitable for book circulation data analysis. Based on Pareto's Law, the A-index was applied to analyze the top 20% of circulation data from Northwest A&F University Library from 2014-2018. *[Result/Conclusion]* Results demonstrate that the average A-index performs well in statistical analysis of first-level Chinese Library Classification categories with reader gender, major, grade, publisher, publication year, and department, enabling more comprehensive and accurate identification of core reader groups, core publishers, and core publication years. Additionally, Bradford's Law applies not only to periodical management but also to collection management. This study provides a reference for exploring H-family indices in analyzing university library circulation data.

Keywords: H-family indices; average A-index; coefficient of variation; correlation; book circulation data
Classification Number: G252
DOI: 10.13266/j.issn.0252-3116.2020.09.008

1. Introduction

The H-index, proposed by American scholar J.E. Hirsch in 2005, aims to evaluate individual academic achievement. It is defined as: a scientist has index h if h of his/her N papers have at least h citations each, and the other (N - h) papers have no more than h citations each [**?**]. Its main advantages include simple calculation, simultaneous measurement of quantity and quality, and relatively slow growth that prevents sudden artificial inflation, ensuring stable evaluation results over time [**?**, **?**]. Jin Bihui and R. Rousseau suggested using the average citation count of papers in the H-core set as a metric, termed the A-index [**?**]. This allows the A-index to reflect changes within the H-core even when the H-index itself remains unchanged, making it an immediately popular metric.

Liu Yuxian and R. Rousseau [**?**] applied the H-index to library borrowing analysis. Liang Xiaohong [**?**] argued that print collections possess irreplaceable advantages over digital resources, and maximizing print collection potential and funding efficiency represents a major challenge for university libraries. Zhou Chunlei [**?**] analyzed reader demand from three perspectives: cumulative utilization rate, cumulative borrowing volume, and H-index, providing analytical foundations for print resource utilization. Kong Chao and Ding Xuan [**?**] conducted feasibility analyses of H-index and G-index in borrowing data, concluding that the H-index more accurately evaluates collection quality. Qian Lingfei and Wang Rong [**?**] noted that reader borrowing data and collection circulation data sequences resemble the paper-citation sequence in the original H-index definition, making H-index calculation feasible. Zhang Xiaoyang, Fang Youliang, and

Song Xinping [?] theoretically proposed many H-index derivatives that address certain limitations of the H-index.

However, the H-index only increases over time, and its reflection of extreme value influence is insufficient, making it insensitive to data changes [?] and unsuitable for cross-domain comparisons [?]. Therefore, L. Egghe proposed the G-index based on the H-index [?, ?]. Previous studies discussed H-family index applications from different perspectives but relied on logical reasoning or direct analysis of A-index and G-index derived from H-index, lacking statistical analysis, selection among H-family indices, and empirical analysis of collection circulation data. This study uses R software to conduct chi-square tests on correlations between first-level Chinese Library Classification categories and Excel column variables, employs LEFT function extraction, performs pairwise correlation analysis of H, A, and G indices, uses coefficient of variation methods to select the A-index, and empirically analyzes circulation data from Northwest A&F University Library (2014-2018) to provide new methods for effectively analyzing collection circulation data and identifying core reader groups, core collections, core publishers, and core publication years.

---

2. Data Sources and Acquisition

2.1 Research Object Selection Based on Pareto's Law [?], we selected the top 20% of circulation borrowing data from Northwest A&F University Library (2014-2018) as our research object. We conducted statistical analysis on seven influencing factors of book circulation and calculated H-index, A-index, and G-index for data showing significant associations through chi-square tests. Data extraction was performed on March 16, 2019.

2.2 Data Sources and Methods First, borrowing data exported from the library management system were processed using Excel's LEFT function to extract the first 1-3 characters of call numbers. Using first- and second-level Chinese Library Classification categories as row variables and factors including copy quantity, title, author, publisher, publication year, ISBN, reader ID, name, gender, department, grade, and reader type as column variables, we formed data matrices for the top 20% of circulation data from 2014-2018 (149,526; 119,801; 105,631; 98,413; and 112,795 records respectively) and performed data normalization and cleaning.

Second, R software conducted chi-square tests between first-level classification categories and influencing factors to screen for correlated indicators. Processed data were saved in Excel files, and Visual C# programming assigned data to two-dimensional arrays (x_1, y_1, x_1, y_2, …, x , y , …, x , y) to generate matrix sequences for calculating H, A, and G indices for row and column variables according to their definitions and reference [?].

Third, we calculated pairwise correlations among average H-index, A-index, and

G-index contingency tables for each influencing factor from 2014-2018, selecting indices with strong correlations that were relatively accessible and highly discriminating for data analysis.

Fourth, the coefficient of variation reflects the degree of dispersion among population values. In multi-index evaluation, if an index shows large variation across observations, it has strong discriminatory power and should be weighted heavily [**?**]. Therefore, we used coefficient of variation to further screen evaluation indices.

Fifth, through correlation analysis and coefficient of variation methods, we selected the A-index to identify core reader groups, core collections, core publishers, and core publication years.

---

3. Data Processing and Analysis

3.1 H-family Index Calculation and Selection Using R software, we analyzed correlations between seven circulation indicators and first-level classification categories from 2014-2018. Simple correlation tests showed p-value < 0.05 for all indicators, demonstrating high correlation between circulation indicators and classification categories.

3.1.1 Correlation Analysis of H-family Indices Standard criteria for linear correlation assessment are: $|r| < 0.3$ (weak), $0.3 < |r| < 0.5$ (low), $0.5 < |r| < 0.8$ (significant), and $0.8 < |r| < 1$ (high). When $r > 0$, x and y are positively correlated; when $r < 0$, they are negatively correlated [**?**]. Results are shown in Table 1 .

Table 1: Correlation Analysis of H-family Indices (2014-2018)

Correlation Analysis	First Category_	First Category_	First Category_	First Category_	First Category_	First Category_	First Category_
H-A Index	0.8287	0.8157	0.4454	0.9621	0.9972	0.7969	0.8388
p-value	5.6871e-07	0.0253	< 2.2220e-16	0.0021	1.1901e-05	1.0173e-08	0.0758
H-G Index	0.7227	0.9119	0.8508	0.9978	0.9717	0.9415	0.3933
p-value	6.6491e-05	0.00422	< 2.2032e-16	7.0603e-06	0.0012	< 2.2e-16	0.5125
G-A Index	0.5733	0.5378	0.6412	0.9699	0.9570	0.5773	-0.1353

Correlation Analysis	First Category_	First Category_	First Category_	First Category_	First Category_	First Category_	First Category_
p-value	0.0034	0.0213	< 2.2350e-16	0.00135	0.0027	0.0003	0.8283

(1) H-index vs. A-index correlation: First-level categories with reader department, major, gender, education level, publication year, and publisher showed correlation coefficients > 0.8 ($p < 0.05$), indicating high correlation. First-level category vs. reader type showed correlation > 0.8 but $p = 0.076 > 0.05$, accepting the null hypothesis of no correlation.

(2) H-index vs. G-index correlation: First-level categories with reader major, gender, education level, department, publisher, and publication year showed coefficients > 0.72 ($p < 0.05$), demonstrating high or significant correlation. First-level category vs. reader type showed $r = 0.39$ ($p = 0.5125$), accepting the null hypothesis of no significant correlation.

(3) G-index vs. A-index correlation: First-level categories with publication year and reader gender showed $r > 0.95$ ($p < 0.05$), indicating high correlation. Categories with department, education level, major, and publisher showed $0.5 < r < 0.8$ ($p < 0.05$), showing significant correlation. First-level category vs. reader type showed $p = 0.8283 > 0.05$, accepting the null hypothesis of no correlation.

In summary, except for first-level category vs. reader type where H, A, and G indices are uncorrelated, all other six factors show high or significant correlations among H-index, A-index, and G-index ($p < 0.05$), indicating information consistency and enabling selection of the most discriminating index through coefficient of variation.

3.1.2 Graphical Correlation Analysis of H-family Indices Although first-level category vs. grade A-G indices showed $r = 0.5773$ ($p < 0.0003$), further analysis of Figure 1 [Figure 1: see original paper]-1 reveals significant correlation only when A-index approaches 400. Similarly, while first-level category vs. publisher A-H indices showed high correlation, Figure 1 [Figure 1: see original paper]-2 shows the relationship only holds when H-index > 10.

Figure 1: Characteristic Correlations of H-family Indices in Book Circulation Data

Comprehensive analysis of Table 1 and Figure 1 shows consistent correlations for first-level categories with publication year, publisher, department, gender, and major. First-level category vs. publication year H-G indices and first-level category vs. grade A-G indices show conditional correlations, while first-level category vs. type shows no correlation among any indices.

3.1.3 Selection of H-family Indices for Circulation Data (2014-2018) We calculated H-index, A-index, and G-index averages for first-level categories with department, grade, reader type, gender, major, publisher, and publication year. Among correlated indices, we selected the most discriminating index using coefficient of variation:

$$v_i = \sigma_i / \bar{x}_i \quad (i = 1, 2, \ldots, n)$$

where σ_i is the standard deviation of indicator i, $\bar{x}_i$ is its mean, and n is the number of indicators [**?**].

Larger coefficient of variation indicates better discrimination [**?**]. As shown in Table 2 , A-index coefficients were largest across all seven factors (0.4966, 0.7661, 2.0218, 1.7727, 0.0700, 1.4380, 1.2263), followed by H-index, with G-index smallest. Coefficient of variation correlates positively with mean value span—larger spans yield better discrimination [**?**].

Table 2: Comparison of H-index, A-index, and G-index Distributions

Category	H-index Range	H-index Mean	H-index CV	A-index Range	A-index Mean	A-index CV	G-index Range	G-index Mean	G-index CV
Department	14-15.5	14.20	0.2230	311-382.7	324.6	0.4966	20.3-21.0	20.6	0.1300
Major	8.2-19.4	13.8	0.2359	84.6-1037.6	797.4	0.7661	17-20.6	18.8	0.6877
Publisher	2.5-2.9	2.7	1.0031	33.3-55.3	44.3	2.0218	6.5-8.6	7.6	0.9856
Publication Year	5.8-6.4	6.1	0.0309	113.7-163.3	131.9	1.7727	4.7-10.5	7.6	0.1746
Gender	16-20.5	18.3	0.0549	2490-3600.8	2881.2	0.0700	20.6-20.6	20.6	0.0042
Education Level	0.2-18.8	9.5	0.8269	0.4-1013.9	507.2	1.4380	16.2-235.1	125.7	0.8909
Reader Type	13.6-20.8	17.2	0.6241	21.2-21.4	21.3	0.0042	0.2-21.4	10.8	0.8269

The A-index considers contributions from extreme values while overcoming the H-index's limitation of only increasing. With the largest coefficient of variation and best discrimination, A-index was selected to evaluate matrices of first-level categories with department, major, gender, education level, publisher, and publication year. For first-level category vs. reader type, where indices are uncorrelated, all three perspectives must be analyzed separately.

Table 3: Selection of H-family Indices for Circulation Influencing Factors

First-level Category	A-G Correlation	H-A Correlation	H-G Correlation	Selected Index
Department				A-index
Major				A-index
Gender				A-index
Education Level				A-index
Publisher				A-index
Publication Year				A-index
Reader Type				H, A, G indices

3.2 A-index Analysis of Core Reader Groups Based on H-family index selection, we used A-index to analyze core reader groups and their borrowing patterns across departments, education levels, grades, gender, publishers, publication years, and majors.

3.2.1 First-level Category vs. Department A-index Analysis Table 4 shows that from 2014-2018, 23 departments were analyzed with an overall A-index average of 324.6. Fifteen departments, including the School of Economics and Management, scored above the mean and constitute core reader groups. The remaining eight departments, including the School of Enology, are non-core reader groups.

Table 4: A-index Ranking by Department (2014-2018)

Rank	Department	Average A-index
1	School of Economics and Management	530.5
2	School of Water Resources and Architectural Engineering	688.8
3	School of Landscape Architecture and Arts	461.0
4	School of Humanities	442.2
5	School of Horticulture	421.9
6	School of Forestry	417.8
7	School of Mechanical and Electronic Engineering	415.4
8	School of Information Engineering	411.4
9	School of Resources and Environment	404.5
10	School of Life Sciences	400.0
11	School of Food Science and Engineering	363.3
12	School of Agriculture	363.1
13	School of Animal Science and Technology	363.0
14	School of Plant Protection	360.0
15	School of Animal Medicine	331.8
16	School of Enology	235.4
17	Department of Foreign Languages	198.9

Rank	Department	Average A-index
18	Innovation Experimental College	176.9
19	School of Chemistry and Pharmacy	151.0
20	School of Science	147.2
21	School of Adult Education	108.1
22	Institute of Soil and Water Conservation	92.7
23	School of Marxism	66.0

Further analysis of Figure 2 [Figure 2: see original paper] reveals that while departmental A-indices fluctuated slightly from 2014-2018, they generally declined. The School of Economics and Management had the highest A-index, forming the first core reader group. The second group includes the School of Water Resources and Architectural Engineering and School of Landscape Architecture and Arts. Twelve schools constitute the third core group, while eight schools are non-core reader groups.

Annual analysis shows that undergraduates in their second and third years constitute core reader groups, with A-indices significantly higher than in their enrollment and graduation years. Master's students in 2013-2017 form core groups, with second-year master's students as the primary core group. Doctoral and postdoctoral reader groups show balanced A-indices across grades with minimal differences.

3.2.2 First-level Category vs. Education Level and Grade A-index Analysis Table 5 shows that undergraduates from 2014-2016 with A-indices above 539.8 constitute core reader groups. Second- and third-year undergraduates are the most core reader groups for print collections. Master's students in 2013-2017 form core groups, with second-year master's students as the primary core group. Doctoral and postdoctoral reader groups show relatively balanced core status across grades.

Table 5: A-index Analysis by Education Level and Grade (2014-2018)

Education Level	Average A-index
Undergraduate (2014-2016)	539.8
Master's (2013-2017)	478.1
Doctoral/Postdoctoral	269.7

3.2.3 First-level Category vs. Reader Type A-index, H-index, and G-index Analysis Since H-index, A-index, and G-index show low pairwise correlation for reader type, we analyzed from all three perspectives. Table 6 shows that undergraduates and master's students have H-indices above average. The overall A-index mean is 1557.5, with undergraduates at 4468.1 (significantly above average) and master's students at 1449.9 (near average). H-index

and A-index results are consistent, while G-index has limited discrimination. Therefore, undergraduates and master's students are core readers.

Table 6: H, A, and G Index Analysis by Reader Type (2014-2018)

Reader Type	H-index	A-index	G-index
Undergraduate	17.1	4468.1	21.3
Master's	17.1	1449.9	21.3
Doctoral/Postdoctoral	17.1	188.5	21.3
Overall Mean	17.1	1557.5	21.3

3.2.4 First-level Category vs. Gender A-index Analysis From 2014-2015, male readers had higher A-indices (5449, 3340.2) than females (1752.6, 2482.6). Over time, female A-indices increased, surpassing males from 2016-2018 (2975.2, 3098.6, 4277.0 vs. 2305.7, 1881.8, 2249.3), with the gap widening.

3.3 First-level Category vs. Publisher A-index Analysis Table 7 shows that from 2014-2018, the library's collection involved 533 publishers. One hundred twenty-eight publishers with A-indices above the average of 43.2 are key publishers. Writer's Publishing House ranked first (A-index = 891.6), nearly 19.6 times higher than Contemporary World Press (rank 128). Following Bradford's Law [**?**], we divided these 128 publishers into three zones where the sum of A-indices in each zone is equal, yielding a ratio of 12:32:84 (core:related:non-related), confirming Bradford's Law applies to publisher classification.

Table 7: Publisher A-index Zoning (2014-2018)

Zone	Publishers (A-index Range)	Count
Core	Writer's Publishing House, Guangzhou Publishing House, Beijing October Literature & Art Publishing House, etc. (891.6-269.2)	12
Related	Changjiang Literature & Art Press, China Renmin University Press, etc. (268.4-124.3)	32
Non-related	Jiangsu Phoenix Literature & Art Press, Beijing United Publishing Company, etc. (43.3-121.2)	84

3.4 First-level Category vs. Publication Year A-index Analysis The library purchased books published from 1952-2018 (including early reprints). Table 8 shows that books published 2004-2017 have A-indices above the average (core publication years). Annual analysis indicates that books borrowed 2-5 years before the borrowing year have the highest A-indices, representing the most important years for collection and acquisition.

Table 8: Publication Year A-index Analysis

Publication Year	Average A-index
2012	1015.2
2013	849.9
2011	794.8
2014	769.9
2010	539.8
2015	478.1
2009	474.6
2016	296.4
2008	269.7
2007	240.7
2017	210.6
2006	175.7
2005	156.6
2004	138.7
Overall Mean	131.9

4. Discussion

4.1 Data Processing Steps Correlation analysis between first-level classification categories and seven circulation factors ensured data reliability. Empirical analysis of correlations, variability, and adaptability among H, A, and G indices validated the A-index for university library circulation data. Correlation diagrams illustrated the applicability scope of H-index and A-index, providing references for libraries to adopt A-index analysis.

4.2 Method Extension and Application This study demonstrates the rationality of average A-index analysis through empirical validation. The method can be extended by: (1) transposing matrices to map core reader groups to core collections; (2) directly analyzing core collections via A-index for more reasonable identification; (3) comparing A-index mean, mode, and median to select the most suitable metric; (4) conducting deeper studies on each factor; and (5) performing annual, quarterly, and monthly analyses to observe trends.

4.3 Analysis of Collection Influencing Factors

(1) The 15 core reader departments are well-established with strong comprehensive strength, correlating with long history, academic culture, and reading habit cultivation. Some non-core departments were established more recently or focus on university-wide teaching and basic research.

(2) Second- and third-year undergraduates show significantly higher A-indices than enrollment and graduation years, forming the core print collection

user group. This relates to curriculum structure and academic focus development. Second-year master's students show the highest A-indices, indicating stable, intensive research-phase reading. Doctoral and postdoctoral groups show balanced A-indices across grades.

(3) With limited collection space in many libraries, core publisher selection using Bradford's Law provides reference for adjusting acquisition patterns and collection allocation.

(4) Despite balanced gender enrollment ratios, female readers' A-indices have gradually surpassed males, with the gap widening annually, indicating the need to consider gender reading preferences in collection development.

4.4 Reader Classification Granularity A-index analyses of first-level categories with department, education level, grade, gender, publisher, publication year, and reader type mutually validate each other, proving A-index applicability. However, first-level category vs. department and vs. degree-granting major analyses are not mutually supportive due to significant differences in college numbers and enrollment sizes within majors, suggesting that appropriate reader classification granularity should be determined for circulation data analysis.

---

5. Conclusions

5.1 Scientific and Reliable Method A-index analysis based on H-family index selection addresses comparability issues in descriptive statistics and the H-index's extreme value 钝化 phenomenon. Average A-index identification of core readers, publishers, and publication years yields mutually verifiable results, confirming the method's rationality and scientific validity.

5.2 Information Consistency Among Indices A-index is suitable for big data analysis. Except for reader type, the other six factors show high or significant correlations among H-index, A-index, and G-index, confirming information consistency. Conditional correlations exist for grade A-G indices (A-index $\geq$ 400) and publisher H-A indices (H-index $\geq$ 10), making A-index appropriate for large-scale data analysis.

5.3 Multi-perspective Scientific Observation Using A-index to analyze the top 20% of circulation data (2014-2018) from multiple perspectives (department, major, gender, education level) successfully identified core reader groups, publication years, and publishers. When H-family indices show weak correlation (reader type), joint analysis from multiple perspectives enables scientific circulation data management.

5.4 Bradford's Law Applicability to Publishers Applying Bradford's Law to divide 128 publishers into three zones (12:32:84 ratio) with equal A-

index sums empirically demonstrates that Bradford's Law applies not only to periodical management but also to identifying core, related, and non-related publishers for optimized collection management.

References

[1] RONALD R, LIU Junwan, MA Jianhua. New developments in Hirsch index research[J]. Science Focus, 2006(4): 23-25.

[2] ZHOU Zhifeng. Exploration of H-index application in library borrowing data analysis[J]. Library Development, 2009(11): 83-89.

[3] JIN Bihui, ROUSSEAU R. R-index, AR-index: Supplementary indicators for H-index function expansion[J]. Science Focus, 2007(3): 1-8.

[4] HIRSCH J E. An index to quantify an individual's scientific research output[J]. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102(15): 16569-16572.

[5] EGGHE L. How to improve the H-index[J]. The Scientist, 2006, 2(1): 8-9.

[6] EGGHE L. An improvement of the H-index: the G-index[J]. ISSI Newsletter, 2007, 2(1): 8-9.

[7] LIU Yuxian, ROUSSEAU R. Hirsch-type indices and library management: A case study of Tongji University Library[C]//ISSI 2007 Proceedings. Madrid: CINDOC-CSIC, 2007: 514-522.

[8] LIANG Xiaohong. Research on utilization of university library print collections based on circulation data[D]. Jinan: Shandong University, 2018.

[9] ZHOU Chunlei. Research on academic credit evaluation based on H-index[D]. Wuhan: Wuhan University, 2010.

[10] KONG Chao, DING Xuan. Exploration of g-index in university library borrowing data analysis[J]. Modern Information, 2014, 34(12): 126-129, 133.

[11] QIAN Lingfei, WANG Rong. H-index based OPAC data analysis and application: A case study of Nanjing University of Aeronautics and Astronautics Library[J]. Journal of Academic Libraries, 2012, 30(2): 61-64.

[12] ZHANG Xiaoyang, FANG Youliang, SONG Xinping. Impact of book citations on academic performance h-index: A case study of library and information science[J]. Studies in Science of Science, 2014, 32(2): 189-194.

[13] LI Ya, YAO Lihui, GUO Hong. Analysis of journal guarantee rate from citation characteristics of plant protection doctoral dissertations[J]. Journal of Library and Information Science in Agriculture, 2006(2): 156-157, 160.

[14] CHEN Junke, LI Hong, LI Jie, et al. Comprehensive benefit evaluation of Xinjiang animal husbandry development based on coefficient of variation

method[J]. Heilongjiang Animal Science and Veterinary Medicine, 2015(18): 17-20.

[15] HUANG Liangwen, CHEN Renen. Principles of Statistics[M]. Beijing: Central Radio & TV University Press, 1996.

[16] PU Xiaoge, LIU Liming. Variation coefficient and grey relational analysis of database performance[J]. Library and Information Service, 2014, 58(14): 71-78.

[17] QI Yanli, LIU Xiqin. Research on core patent identification methods[J]. Information Studies: Theory & Application, 2016, 39(11): 5-9.

[18] LIU Lianhua. Evaluation of mathematical core journals based on distance comprehensive evaluation method[J]. Chinese Journal of Scientific and Technical Periodicals, 2015, 26(8): 871-874.

Author Contributions

Jing Lingdi: Responsible for statistical methodology implementation, paper structure organization, and writing.
Wang Lihong: Responsible for raw data extraction, normalization, and batch A-index calculation.
Wang Qin: Responsible for R software batch correlation calculations and paper revision.
Li Ya: Responsible for overall concept, structure planning, figure/table production, and final manuscript revision.

English Abstract

Selection of H-family Indices for Book Circulation Data Analysis: Taking the Library of Northwest A&F University as an Example

Jing Lingdi, Wang Lihong, Wang Qin, Li Ya Library of Northwest A&F University, Yangling 712100

Abstract: *[Purpose/significance]* To improve book procurement and circulation management, and to realize standardized management for timely monitoring, evaluation, and regulation of collection acquisition, total volume, and structure. *[Method/process]* We compared the correlation, variability, and adaptability of three H-family indices (H-index, A-index, and G-index), selecting the A-index as suitable for book circulation data evaluation. We then used the A-index to analyze the top 20% of NWAFU Library's circulation data from 2014-2018. *[Result/conclusion]* Results show that the A-index is a more comprehensive and precise indicator for identifying core readers, core publishers, and core publication years across multiple dimensions (first-level classification categories, reader gender, major, grade, department, publisher, and publication

year). Additionally, Bradford's Law applies not only to periodical management but also to book collection management. This paper provides a reference basis for exploring H-family indices in analyzing university library circulation data.

Keywords: H-index; average A-index; coefficient of variation; correlation; book circulation data

Note: Figure translations are in progress. See original paper for figures.