DSI DATA SCIENCE AND INFORMETRICS
VOL.1, NO.3, AUG. 2021

# How to classify bibliometrics indicators? A thorough investigation of objective classification and its application

Liping Yu[a], Houqiang Yu[b], Lingmin Chen[c*], Yeyang Zhou[d]

a. School of Management & E-business, Zhejiang Gongshang University, Hangzhou, China
b. School of Economics & Management, Nanjing University of Science & Technology, Nanjing, China
c. School of International Business, Zhejiang Financial college, Hangzhou, China
d. Division of Business & Management, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai, China

**ABSTRACT**

Classification of bibliometric indicators is a fundamental issue in information science. Traditionally, the classification is based on subjective classification. This article presents an empirical study on the mathematics journals listed in JCR 2019 by using objective classification methods including cluster analysis, factor analysis, and principal component analysis to classify bibliometric indicators. Different classification results are compared and further interpreted, major finding are: the classification results of objective classification methods share similarities; objective classification helps better comprehend bibliometric indicators; objective classification should be used in combination with subjective classification; cluster analysis performs better in classifying bibliometric indicators than factor analysis and principal component analysis; not all the results of objective classification are meaningful; cluster of indicators has sufficient influence on subsequent evaluation and regression analysis. This study provides a new paradigm for journal classification and indicator analysis.

**KEYWORDS**

Bibliometrics; Objective classification; Bibliometrics indicators; Principal component analysis; Factor analysis; Clustering analysis

## 1 Introduction

Classification of bibliometric indicators is a fundamental problem in information science and library science and also the basis for further analysis like journal evaluation, scholarship evaluation, institution evaluation, and scientist evaluation. The current classification of bibliometric indicators is mainly based on subjective classification, for example, in one classification system, the indicators are based on source and citation information, while in another classification system, the indicators may be simple or composite. Due to the increasing number of bibliometric indicators and their increasingly complex connotations, relying solely on

* Correspondence should be addressed to Chen Lingmin: clm986@126.com

subjective classification is not enough and must be combined with objective classification. In addition, attention should be given to the diversity of classification systems. So far, many classification systems for classifying bibliometric indicators have been proposed. Majority of the existing classification systems are based on subjective classification, but the reliability of such subjective classification remains to be explored. For example, Immediacy Index is indicator of journal impact as well as indicator of timeliness, which category should it be classified into? Results that are based on subjective classification are quite different from that based on objective classification. Different algorithms adopted in objective classification like cluster analysis will also probably lead to different classification result. What is the ultimate reason behind the different classification results? Furthermore, what is the implication of such different results based on objective classification? The different results of classification of bibliometrics indicators will definitely have influence on the relationship between different types of bibliometrics indicators. What will the influence be? This study, based on a thorough summarization of the mainstream objective classification systems, conducts an empirical study on the mathematics journals listed in the Journal Citation Report (JCR) 2019 by using objective classification methods including cluster analysis, factor analysis, and principal component analysis to classify the relevant bibliometric indicators. The classification results are compared and further analyzed. Some special classification results are discussed to reveal the functions and characteristics of objective classification methods. The relationship between objective classification and subjective classification is also discussed in depth. Moreover, comparison is conducted on the influence brought by different classifications on indicator of journal impact and indicator of timeliness.

Journals of different disciplines also attach great importance to the classification and evaluation of indicators. Jason (2018) was to investigate the correlations among 6 of the commonly used bibliometric indices (Impact Factor, SCImago Journal indicator, SCOPUS h-index, Google h-index, Eigenfactor, Article Influence Score) in neurosurgical and spinal surgical journals. Miot et al. (2021) aim to evaluate the trends of the main bibliometric indicators of? Anais Brasileiros de Dermatologia, in the decade of 2010– 2019,including indices: impact factor, immediacy index, Eigenfactor?, SJR. Wohlrabe and Gralka used 7 indicators to classify economists. These indicators are number of distinct works, weighted number of works, number of citations, weitghted number of citations, number of pages, weighted number of pages, number of downloads.

In classifying bibliometric indicators, principal component analysis has been widely used. Costas et al. (2007) applied principal component analysis (PCA) to bibliometric indicators at the level of individual scholars. Bornmann et al. (2008) adopted principal component analysis to classify nine h-type indices. Hendrix (2008) used principal component analysis to compare indicators at the institutional level, and classified bibliometric indicators into three categories, namely gross research productivity, impact of research and research productivity, and impact at the individual level. Leydesdorff (2009) used principal component analysis to classify bibliometric indicators into two categories, namely component 1 that represents the size factor, including total citing, total documents, self-citations, total cited, and page rank; and component 2 that represents the impact, including impact factor, immediacy index, SJR and h index. In addition, based on medical science data provided by the Institute of Science and Technology Information of China, Yu & Pan (2009) used factor analysis to classify journal bibliometric indicators into impact indicators, source indicators, and immediacy indicators.

Cluster analysis is another effective method for classifying bibliometric indicators, and usually used in combination with PCA. Anderberg (1973) considered cluster analysis an identification method that requires no supervision, i.e., without any guidance from prior information, capable of identifying potential similar pattern in the dataset and grouping the dataset in order to make similarity in the group as large as possible and the similarity between groups as small as possible. Franceschet (2009) proposed different clustering techniques to group bibliometric indicators that are used in quantitative assessment of research quality and compared the results with PCA. The proposed methods group bibliometric indicators into clusters of highly mutual correlated indicators. The set of all the representatives correspond to a base of orthogonal metrics capturing independent aspects of research performance and can be exploited to design a composite performance indicator. Bollen et al. (2009) studied the aggregation of various journal indicators including bibliometric and social network indices computed on both citation and usage networks. Principal component analysis, as well as hierarchical and repeated partition (K-means) clustering techniques were used, leading to similar conclusions.

For contemporary research, both PCA and cluster analysis are widely applied in information science and library science and are very important tools for objective classification. Several methods are prevailingly used in cluster analysis, for example, Euclidean distance, Pearson correlation etc, while other methods for cluster analysis are seldom deployed to classify bibliometric indicators. Because PCA shares similar principles with factor analysis, many studies use PCA to classify bibliometric indicators but neglect the use of factor analysis, which should be explored more. What is more, there are few studies comparing the results of different objective classification methods. This study starts with summarizing subjective classifications of bibliometric indicators, then reviews the objective classification methods. Next, the mathematics journals listed in JCR 2019 are used in an empirical study by adopting different objective classification methods to classify bibliometric indicators. The classification results are further analyzed. At the end, conclusions are drawn and discussed.

The contributions of the study are mainly as follows: (1) Unlike traditional studies that use a single classification method, for example, Bollen et al. (2009) evaluated the similarity between two metrics by computing the Euclidean distance of the measure correlation vectors, Franceschet (2009) used the Pearson correlation coefficient as the similarity metric, this study uses various classification methods and algorithms, such as cluster analysis, principal component analysis, and factor analysis, for classifying bibliometric indicators and obtains various classification results. The new classification system is observed, which extends previous studies. (2) Previous studies mainly focus on how to classify bibliometric indicators in order to do journal evaluation, this study not only focuses on the classification results, but also compares the differences between different classification methods, so as to reveal some characteristics of bibliometric indicators. The results are further interpreted. (3) This study proposes a new research paradigm, which can not only be used to aid in classifying bibliometric indicators, but also be used to study the characteristics of bibliometric indicators.

# 2 Review on contemporary classification systems of bibliometric indicators

## 2.1 Subjective classification systems of bibliometrics indicators

Subjective classification of bibliometric indicators is mainly based on certain predefined conditions. If the indicator possesses the defined characteristics, it is classified as one type, otherwise, it is classified as the other type. There are currently four major such subjective classification systems listed as follows:

(1) Source indicator and citation indicator. This is currently the mainstream classification system for bibliometric indicators. The so-called source indicators refer to indicators that are determined by the characteristics of the journals themselves, such as number of publications, selected publication ratio, number of authors, district distribution, funded publication ratio, citation half-life, overseas publication ratio etc. The citation indicators refer to indicators that are derived from citations, such as total cites, journal impact factor, immediacy index, 5-year impact factor, h index (Hirsch, 2005), Eigenfactor (Bergstrom, West & Wiseman, 2008), article count impact factor (ACIF) (Markpin, 2008) etc.

(2) Simple indicator and composite indicator. The simple indicator refers to an indicator that is relatively concise and encompasses limited information. The traditional bibliometric indicators are mostly simple indicators, like number of publications, funded publication ratio, cited half-life, journal impact factor etc. The composite indicator refers to an indicator that is relatively complicated to compute and encompasses a large volume of information, much research focus on the development and properties of the indicators(Kosten,2016). for example, Source Normalized Impact per Paper(SNIP) proposed by Moed (2010), Successful Paper (SP) proposed by Kosmulski (2011), and SCImago Journal Rank (SJR) index proposed by Gonzá lez-Pereira (2007). Of course, Eigenfactor and h index are also composite indicators.

(3) Productivity metrics, impact metrics, and hybrid metrics (Franceschet, 2009). Productivity metrics include indicators like citable items, number of paper per individual author, number of papers per academic year etc. Impact metrics include total cites, impact factor, immediacy index etc. Hybrid indicators have characteristics of productivity indicators and impact indicators simultaneously, such as h index (Hirsch, 2005), the m-quotient (Hirsch, 2005), g index (Egghe, 2006), the individual h index (Batista et al., 2006), the contemporary h index (Katsaros et al., 2007), the Altmetrics (Melero, 2015), etc.

(4) Flow indicators and stock indicators. This classification is first used in statistics and not yet widely used in bibliometric studies. The majority of bibliometric indicators are flow indicators that report annual statistical data, such as journal impact factor, 5-year impact factor, funded publication ratio, citation half-life etc. A minority of the indicators are stock indicators that report accumulated data, for example, series of h-type indices. Total cites has the attribute of both flow and stock indicators because citation source is a stock indicator.

## 2.2 Objective classification methods for classifying bibliometrics indicators

Since only subjective classification is artificially biased, objective classification should be supplemented (Pronin, 2007). The mainstream objective classification methods of bibliometric indicators include cluster analysis, principal component analysis, and factor analysis. As statistics and information science develop, new classification methods will be proposed in the future.

(1) Cluster analysis. Clustering is the process of organizing objects into groups of which members share some certain similarity (Jain & Dubes, 1988). It classifies sample sets that have no categorical labels into several categories according to certain standards, making similarity in the same sample set as large as possible and similarity between sample sets as small as possible. Meanwhile, cluster analysis makes the distances between different categories as large as possible and distances in the same category as small as possible, and the classification results should be interpreted in a convincing manner. For the same group of

data, cluster analysis can be performed on the records (called Q-type cluster), or the variables (called R-type cluster). These two types of cluster analysis have no mathematical difference. Cluster analysis is a common technique and certainly can be used for classifying bibliometric indicators.

Different algorithms, even different settings of the same algorithm, will generate different categories for the same dataset. Research has shown that there is no single cluster analysis method that can generate the optimal classification of all the datasets (Liang et al., 2010; Pal & Biswas, 1997). This has in fact provided a clue for further analyzing the cluster analysis results, for example, for the classification of bibliometric indicators, what are the differences between fewer classifications and more classifications? How does the difference come into being? Answers to these questions will advance understanding of the meanings of bibliometric indicators and the irrelationships with other indicators.

(2) Factor analysis and principal component analysis. Factor analysis is a dimension reducing technique of which the aim is to select a few comprehensive variables as indicators from multiple variables. The basic principle is to, by studying the covariance matrix of the variables and the structure of the coefficient matrix, find out a few random variables to describe the relationships between multiple variables, and then classify the variables according to the values of the correlation coefficients. Each group is called a common factor. The correlations between variables in the same group are relatively high and the correlations between variables from different groups are relatively low. The contribution of each common factor is different. Ranked by the contribution rate of variance, common factors that have variance contribution over 1 are selected to evaluate so as to reduce the dimension of the dataset. From the principle of factor analysis, it can be used to evaluate the indicators via reducing dimensions, as well as classifying the indicators via classifying common factors.

Principal component analysis is a special form of factor analysis. They share the similar principle. The difference lies in the process of data matrix creation during data processing. Factor analysis spins the factor while principal component analysis does not. Factor analysis requires the extracted common factors to possess practical meaning, while principal component analysis requires only the principal components to encompass the principal information but their meanings are not necessarily interpreted in an accurate way. From the perspective of classifying indicators, factor analysis performs better in interpreting factors than PCA. However, this does not mean that PCA is not applicable for indicator classification.

In summary, cluster analysis classifies the indicators so that the similarity between objects of the same class is stronger than that of other classes. Factor analysis and principal component analysis transform multiple indexes into several comprehensive indexes, which make it have better performance than the original indexes. But there is a difference between the two. Compared with principal component analysis, factor molecules tend to describe the correlation between the original variables. As a multivariate analysis method, the three analysis methods are often used to analyze multiple index problems. Since the indexes do not exist independently, only studying a certain index or separating these indexes can not explain the research problem. To grasp the problem as a whole, it is necessary to group and classify the data, and analyze the interrelationships and internal connections.

## 3 Research design and data

The research design takes the main indicators published by the JCR database as the research object. The indicators are classified by three objective classification methods: cluster analysis, factor analysis, and principal component analysis, and the results are compared and analyzed in depth, and some special analysis. The results are discussed, with a view to in-depth examination of the role and characteristics of objective classification methods, and an in-depth discussion on the relationship between objective classification and subjective classification. On this basis, further compare the influence of different index classifications on the relationship between the influence index and timeliness index of academic journals.

Data are from the JCR 2019 database. In order to improve the robustness of the study, the mathematics discipline is selected because there is a relatively higher number of journals in the discipline. In total, 11 major indicators are published in JCR 2019, namely Total Cites, Journal Impact Factor (IF), Journal Impact Factor without Journal Self Cites, 5-Year Impact Factor, Immediacy Index, Eigenfactor Score, Article Influence Score, Cited Half-life, Citing Half-life, Average Journal Impact Factor Percentile, and Normalized Eigenfactor.

In total, there are 325 mathematics journals, 17 of which are discarded due to missing data. The descriptive statistics of the remnant 308 journals are shown in Table 1.

**Table 1** Descriptive statistics of 11 indicators

| Indicator | Abbr. | Mean | Max. | Min. | Std. |
|---|---|---|---|---|---|
| Total Cites | TC | 1818 | 24068 | 117 | 2764.423 |
| Journal Impact Factor | IF | 0.996 | 8.455 | 0.216 | 0.770 |
| Impact Factor without Journal Self Cites | IFWSC | 0.929 | 8.182 | 0.143 | 0.747 |
| 5–Year Impact Factor | IF5 | 1.040 | 12.862 | 0.204 | 0.956 |
| Immediacy Index | II | 0.312 | 2.000 | 0.000 | 0.248 |
| Eigenfactor Score | EFS | 13.438 | 48.900 | 1.700 | 9.494 |
| Article Influence Score | AIS | 16.178 | 38.000 | 6.800 | 3.773 |
| Cited Half–life | CITEDHL | 0.004 | 0.046 | 0.000 | 0.006 |
| Citing Half–life | CITINGHL | 0.929 | 8.970 | 0.088 | 1.094 |
| Average Journal Impact Factor Percentile | IFP | 46.895 | 99.846 | 1.077 | 27.432 |
| Normalized Eigenfactor | NEF | 0.524 | 5.571 | 0.024 | 0.734 |
| N | | 308 | | | |

# 4 Results of objective classification methods and comparative analysis

## 4.1 Correlation analysis

Correlation coefficient is the foundation of classifying bibliometric indicators. The Spearman correlation coefficients of journal bibliometric indicators are shown in Table 2. Most of the correlation coefficients between indicators pass statistical tests. Bibliometric indicators with similar attributes have high correlation coefficients, for example, the correlation coefficient between IF and 5-Year Impact Factor is 0.984, and the correlation coefficient between IF and impact factor without journal self cites is 0.991. On the other hand, bibliometric indicators with different attributes have low correlation coefficients. For example, the correlation coefficient between IF and citing half-life is -0.012, and the correlation coefficient between Eigenfactor and Cited Half-life is 0.137.

**Table 2** Spearman correlation coefficients between journal bibliometric indicators

| Probability | TC | IF | IFWSC | IF5 | II | CITEDHL | CITINGHL | EFS | AIS | IFP | NEF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| TC | 1 | | | | | | | | | | |
| IF | 0.056 | 1 | | | | | | | | | |
| IFWSC | 0.051 | 0.991*** | 1 | | | | | | | | |
| IF5 | 0.118** | 0.924*** | 0.923*** | 1 | | | | | | | |
| II | 0.009 | 0.134** | 0.144** | 0.090 | 1 | | | | | | |
| CITEDHL | 0.351*** | 0.098* | 0.127** | 0.153** | −0.033 | 1 | | | | | |
| CITINGHL | 0.052 | −0.012 | 0.020*** | −0.024 | −0.075 | 0.366*** | 1 | | | | |
| EFS | 0.861*** | 0.011 | 0.013 | 0.081 | 0.005 | 0.137** | 0.041 | 1 | | | |
| AIS | 0.202*** | 0.731*** | 0.765*** | 0.801*** | 0.111** | 0.364*** | 0.184*** | 0.207*** | 1 | | |
| IFP | 0.099* | 0.207*** | 0.225*** | 0.167** | 0.061 | 0.296*** | 0.303*** | 0.103** | 0.308*** | 1 | |
| NEF | 0.861*** | 0.011 | 0.013 | 0.081 | 0.005 | 0.137 | 0.041 | 1.000*** | 0.207*** | 0.103** | 1 |

*、**、*** means two–tail significance of 0.1、0.05、0.001.

## 4.2 Classification results of clustering analysis

In classifying bibliometric indicators, clustering methods are very important. This study mainly uses two clustering methods. The first method selects the largest distance between groups. The second method selects the smallest distance between groups. For calculation methods of the distance, majority of the frequently used methods including Euclidean distance, cosine similarity, correlation coefficient, Manhattan distance etc are used. Three classification results are obtained.

(1) Classification result one

The first classification result has two categories of indicators. Total cites belongs to the first category and all the other indicators belong to the second category (Figure 1). In total eight clustering methods and algorithms are used of which the classification results remain the same. The first group of analysis considers inter-group connection and the distance algorithms used include Euclidean distance, squared Euclidean distance, Minkowski distance, and Chebychev distance; the second group of analysis considers inner-group connection and the distance algorithms used include Euclidean distance, squared Euclidean distance, Minkowski distance, and Chebychev distance.

In all the existing classification studies of journal evaluation, few scholars separate total cites and the other bibliometric indicators into two categories. Yu (2014) suggested that evaluation of scholarly journals can be divided into stock evaluation and flow evaluation from the perspective of time. The current journal evaluation based on the indicator system mainly uses flow evaluation while seldom uses stock evaluation. Total cites is an indicator with the attributes of both flow evaluation and stock evaluation but is widely used in flow evaluation. This is logically unreasonable as it makes journals with longer history attain higher scores.

Among the 11 indicators in this study, only total cites has the attribute of a stock indicator while the other indicators are flow indicators. This is in accordance with the cluster analysis results, which provide a new perspective on classifying bibliometric indicators and journal evaluation.

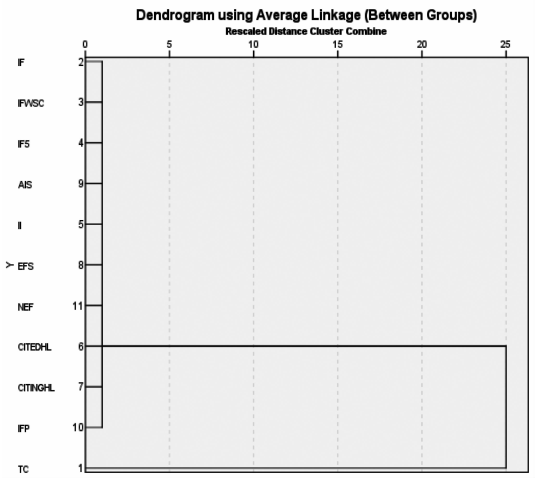


**Figure 1** Classification result one

(2) Classification result two

Classification result two has three categories. The first category includes Eigenfactor, Normalized Eigenfactor, and Total cites; the second category includes IF, IF without Self-cites, 5-year Impact Factor, Article Influence Score, Average Journal Impact Factor Percentile, Immediacy Index; and the third category includes Cited Half-life, and Citing Half-life (Figure 2).

Inter-group connection is used for clustering, with distance calculation algorithms like Pearson correlation and intersection angle cosine. Inner-group connection is also used with distance calculation algorithms like Pearson correlation and intersection angle cosine. The results of the above four algorithm combinations are the same.

The first category is relatively objective and scientific. Normalized Eigenfactor and Eigenfactor should above all be in the same category because Normalized Eigenfactor is calculated based on Eigenfactor. Total cites is also in the category. After removing several journals with missing data or short history, the remnant journals generally have mid to long history. Generally speaking, journals with longer history and higher impact have more stable and continuing impact. Total cites is a good indicator for describing the long term impact of a journal. Eigenfactor and Normalized Eigenfactor when compared with other bibliometric indicators better reflect the impact of a journal. So it is reasonable that the cluster analysis classifies these three indicators into the same category.

The second category presents more logical progressive relationships. IF is first grouped with IF without Self-cites. These two indicators are the closest. Then they are grouped with 5-year Impact Factor, which extends the time period of IF. Next, they are grouped with Article Influence Score because it is a completely new indicator of which the calculation method and principle are both different. After that, they are grouped with Average IF Percentile. Yu (2016) considered Average IF Percentile a non-parametric transformation that transforms continual data into ranking. The transformation makes it different from IF, so it is grouped later. At the end, they are grouped with Immediacy Index because immediäcy index counts the average citations in the year and has special statistical characteristics.

The third category is Cited Half-life and Citing Half-life, both of which are immediacy indicators and should be in the same cätegory.

What is more, the first category can be grouped with the second category to be parallel with the third category, thus forming one big category and a small category. The big category includes cited indicators measuring journal impact, while the small category includes immediacy indicators.

As shown in the cluster analysis results, the progressive classifications and layers are very rich, which is good for further analyzing the meanings and features of bibliometric indicators from anoter perspective.

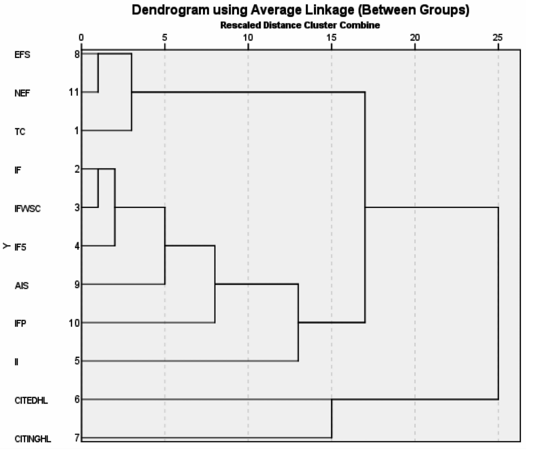


**Figure 2** Classification result two

(3) Classification result three

Classification result three (Figure 3) also comprises three categories. The first category includes IF, IF without Self-cites, 5-year Impact Factor, Article Influence Score, Normalized Eigenfactor, Immediacy Index, Eigenfactor Score, Total cites; the second category includes Average IF Percentile; and the third category includes Cited Half-life, Citing Half-life. The clustering method is based on inner-group connection. The distance calculation algorithm used is the Euclidean distance.

The first category has the widest scope but all the indicators inside share the common feature that they are all continual indicators. The second category includes only one indicator, i. e., Average IF Percentile, which is a non-parametric transformation of IF from continual data to ranking data, and also the only indicator of non-continual data. The third type is citation half-life and citation half-life, which are indicators of the timeliness of journals.

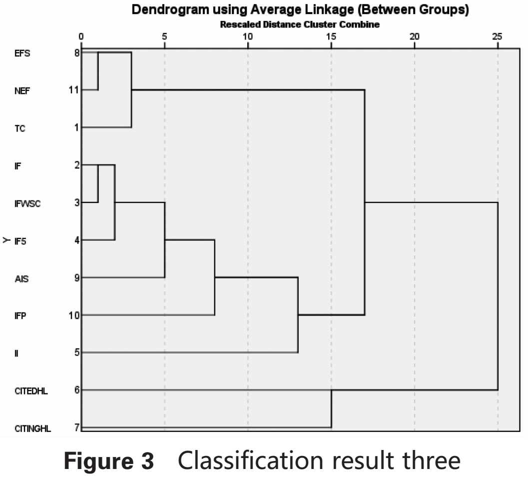


**Figure 3** Classification result three

### 4.3 Classification results of factor analysis and principal component analysis

(1) Classification results of factor analysis

Before conducting factor analysis, KMO test and Bartlett test should be performed to decide whether the preconditions for factor analysis are satisfied. The KMO test value is 0.0.785, higher than threshold value 0.5. The Bartlett test value is 8464.784 with a corresponding p value 0.000. So factor analysis is applicable. In total, there are three factors with variance contribution rates higher than 1. The contribution rates are 40.90%, 26.73%, and 14.10%, respectively, and the aggregate contribution rate is 81.73%, indicating that these three factors explain 81.73% of the information in the original indicators. The rotation matrix and classification are shown in Table 3.

The classification results from factor analysis are the same as the second classification results from cluster analysis, which is a coincident. Differing from cluster analysis, factor analysis lacks the progressive relationships and layers.

**Table 3** Rotation matrix of factor analysis and the classification

| Indicators | 1 | 2 | 3 | Classification |
|---|---|---|---|---|
| IF | 0.967 | 0.115 | –0.095 | Category one |
| IFWSC | 0.974 | 0.108 | –0.064 | |
| IF5 | 0.942 | 0.116 | 0.010 | |
| II | 0.534 | 0.147 | –0.333 | |
| AIS | 0.871 | 0.201 | 0.240 | |
| IFP | 0.741 | 0.284 | –0.252 | |
| TC | 0.164 | 0.927 | 0.128 | Category two |
| EFS | 0.184 | 0.967 | –0.035 | |
| NEF | 0.184 | 0.967 | –0.035 | |
| CITEDHL | 0.080 | 0.154 | 0.828 | Category three |
| CITINGHL | –0.183 | –0.077 | 0.775 | |

(2) Classification results of principal component analysis

The KMO test result and Bartlett test result of PCA are same as those of factor analysis. In total there are threefactors with variance contribution rates higher than 1. The contribution rates are 48.01%, 20.40%, and 13.33%, respectively, and the aggregate contribution rate is 81.74%, indicating that these three factors explain 81.74% of the information in the original indicators. The principal component matrix and classification are shown in Table 4.

**Table 4** Principal component matrix and classification

| Indicators | 1 | 2 | 3 | Classification |
|---|---|---|---|---|
| IF | 0.904 | –0.366 | 0.084 | Category one |
| IFWSC | 0.903 | –0.366 | 0.116 | |
| IF5 | 0.873 | –0.325 | 0.180 | |
| II | 0.563 | –0.209 | –0.239 | |
| AIS | 0.835 | –0.157 | 0.368 | |
| IFP | 0.803 | –0.164 | –0.147 | |
| TC | 0.586 | 0.747 | –0.033 | Category two |
| EFS | 0.636 | 0.727 | –0.193 | |
| NEF | 0.636 | 0.727 | –0.193 | |
| CITEDHL | 0.077 | 0.325 | 0.778 | Category three |
| CITINGHL | –0.260 | 0.233 | 0.720 | |

### 4.4 The influence of classification on the relationship between evaluative indicators of scholarly journal

As an example, the relationship between journal impact indicators and immediacy indicators is investigated. According to cluster analysis result two, factor analysis and principal component analysis, CITEDHL and CITINGHL are classified into one category, the other indicators are classified to another category. Using partial least squares (PLS) to calculate, the relationship between these two categories are shown in Figure 4.

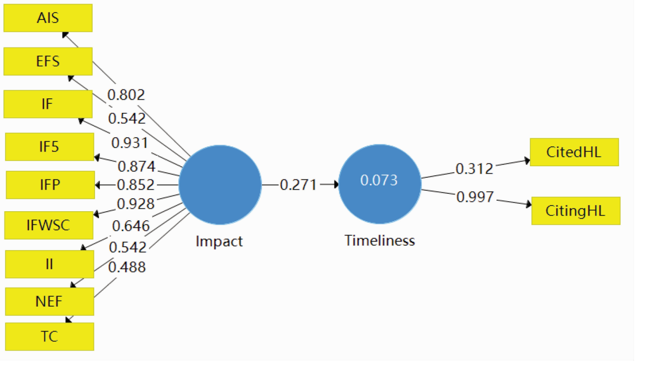


**Figure 4** Relationship between impact indicator and timeliness indicator (1)

From the PLS result, the AVE value of latent variable of impact is 0.567 which is above the threshold value of 0.5, the composite reliability is 0.918 which is above the threshold value of 0.6. The AVE value of latent variable of timeliness is 0.0.546 which is below the threshold of 0.5, the composite reliability if 0.654 which is above the threshold value of 0.6. The results show that the impact indicator can show the influence of academic journals, and CITEDHL and CITINGHL well measure the immediacy of journals. The coefficients of If, IF5, IFP and IFWSC on the influence passed the statistical test, while other indicators did not pass the statistical test; coefficient of CITINGHL is timeliness statistically significant, but CITEDHL has not.

The coefficient between impact and timeliness is 0.271 which passes not the statistical test. It indicates that impact of journal is not related to immediacy.

According to cluster analysis result three, IFP is a special indicator, it is non-parametric indicator and kind of ranking indicator, and thus should not be analyzed together with the other parametric indicators. The regression result after eliminating the IFP is shown in Figure 5.

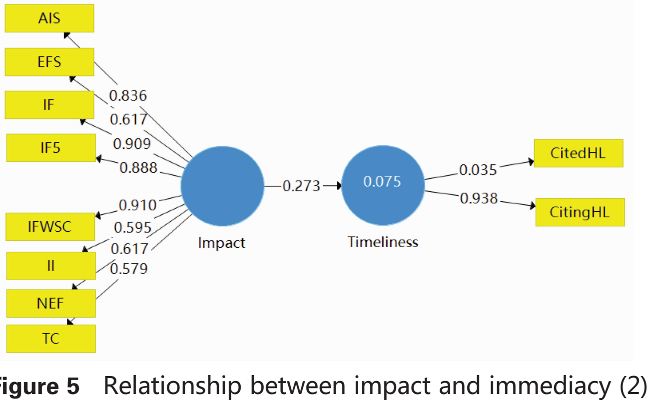


**Figure 5** Relationship between impact and immediacy (2)

From the PLS result, the AVE value of latent variable of impact is 0.572, above the threshold of 0.5 but higher than the previous 0.567, the composite reliability is 0.912, also much higher than the previous result. The AVE value of latent variable of timeliness is 0.440, below the threshold of 0.5, the composite reliability is 0.458, below than the threshold of 0.6. The coefficients of AIS, IF, IF5 and IFWSC for impact have passed statistical test, the coefficients of other indicators have not, but the t test value has increased. The coefficients of CITINGHHL have not passed the statistical test, coefficients of CITEDHL is not. The coefficient between impact and timeliness is 0.273 which is sufficiently lower and has not passed the statistical test.

By comparing Figure 4 and Figure 5, it is shown that IFP problem is revealed by conducting cluster analysis so that modification can be made on the model of relationship between indicator of impact and indicator of timeliness.

# 5 Conclusions and discussion

(1) Total Cites and Average IF Percentile are two special indicators

Based on different classification methods, the study finds that Total Cites and Average IF Percentile can each form a single category by themselves. However, subjective classification never classifies indicators in this way. This prompts us to find the ultimate reason for the phenomenon. This study discovers that Total Cites can form a single category because it possesses the attributes of both flow and stock indicators. Average IF Percentile forms a single category because it is non-parametric transformation of IF from continual data to ranking data. So the classification is entirely determined by the attributes of the indicator and is a logical result.

(2) Objective classification results share similarities

Whatever objective classification method is used, indicators with similar attributes are always classified into the same group, for example, indicators such as IF, 5-year Impact Factor, IF without Self-Cites etc are always in the same group, while Cited Half-life and Citing Half-life are in the same group. Indicators that are classified into different groups in accordance with different classification methods should be particularly analyzed and considered for how to classify these indicators.

(3) Objective classification methods help improve the comprehension of bibliometric indicators

In classifying bibliometric indicators, if we merely rely on subjective classifications, we inevitably go into in a blind alley. Objective classifications help improve the comprehension of bibliometric indicators. Among objective classification methods, progressive cluster analysis has a relatively better sense of layers and can help people understand the nature and emphasis of some indicators, for example, Total Cites as an indicator having the stock attribute is seldom classified into a single category in subjective classification; it helps understand whether Average IF Percentile has the closest relationship with IF or journal impact indicators that include IF, IF without Self-cites, and 5-year Impact Factor. New discoveries by using objective classification methods can guide research like journal evaluation and relationships between bibliometric indicators etc, thus deepening basic and application research in bibliometrics.

(4) Objective classification methods should be used in combination with subjective methods

Subjective classification methods that rely on an expert's expertise, from Dewey's decimal classification to Linnaeus's taxonomy of species, are all top-down systems. In an expert's mind, there is a system tree. The objective classification of bibliometric indicators, however, is bottom-top and is in accordance with the characteristics of data and the principle of distribution, but the logical relationships within the category and between categories need to be further addressed. The problem of subjective classification is that there are always some indicators that cannot be properly classified into any category and are eventually classified into a category with difficulty, or classified into two or more categories. Objective classification has no problem of improper classification by assigning any indicator to the closest category. But the problem arises when the categories must be given a label because the elements of the category are complicated, unlike categories in subjective classification, which are well defined. So it works best to combine objective classification methods with subjective classification methods.

Although subjective classification methods are important for classifying bibliometric indicators, this study shows pure reliance on subjective classification is not enough. In classifying bibliometric indicators, subjective classification must be considered together with objective classification because objective classification is a critical complement to subjective classification. In real practice, classification should be performed based on a comprehensive analysis that combines the researc20215h purpose and specific features of the research object. The procedure for classification is first to clarify the research purpose and object to have a general subjective classification, then conduct cluster analysis, factor analysis, and PCA to make objective classification, next use as many objective classification methods as possible to compare the objective classification results with subjective classification results, and in the end determine the final classification methods of bibliometric indicators.

(5) Cluster analysis has several advantages over factor analysis and principal component analysis in classification process

In cluster analysis, the progressive relationship of classifying can be demonstrated with a rich sense of layers. This is significantly meaningful in improving the comprehension of bibliometric indicators and advance the combined use of subjective classification and objective classification. In addition, algorithms for cluster analysis are various and so are the classification results. On the other hand, PCA and factor analysis give only one definite classification result. So cluster analysis has some advantages, but this does not mean PCA and factor analysis should be abandoned, given that not many objective methods for classifying indicators are usable at present. As a consequence, these methods should be used comprehensively.

It should be noticed that according to the characteristics of the classifying indicators, cluster analysis should adopting inter-group connection or inner-group connection algorithms, and not other algorithms.

(6) Not all objective classification results are meaningful

Objective classification methods rely totally on data, but due to the complexity of bibliometric indicators and heterogeneity of the research objects, classification results that are totally based on objective classification methods are not necessarily meaningful. This study is an exploratory analysis of mathematics journals, so the results of objective classifications can all be interpreted from a common perspective. However, the characteristics of objective classification of bibliometric indicators of other journals remain to be explored.

(7) The study provides a paradigm for classifying bibliometric indicators

The study mainly provides a research paradigm for classifying bibliometric indicators and interprets some classification results, so it does not emphasize the completeness of indica tors. In the future research, more indicators can be brought into analysis.

(8) Clustering of indicators has critical influence on the subsequent evaluation and regression analysis

There are two major types of empirical studies of evaluative indicators of journal. The first type is to do evaluation and classification of indicators undoubtedly has influence on the evaluation. The second type is to analyze based on regression analysis and the different classifications will also have influence on the regression. The aim of adopting objective evaluation method is to, combined with subjective evaluation method, make the classification result more scientific, and simultaneously discover the latent attribute of indicators to make the selection of indicators more scientific.

## References


Anderberg, M. R. (1973) . *Cluster analysis for application*. New York: Academic Press.

Batista, P. D., Campiteli, M. G., & Konouchi, O. (2006) . Is it possible to compare researchers with different scientific interests?. *Scientometrics, 68* (1) , 179–189.

Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008) . The Eigenfactor metrics. *The Journal of Neuroscience, 28* (45) , 11433–11434.

Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009) . A principal component analysis of 39 scientific impact measures. *PLoS ONE, 4* (6) , e6022. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0006022

Bornmann, L., Mutz, R., & Daniel, H. (2008) . Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from Biomedicine. *Journal of the American Society for Information Science and Technology, 59* (5) , 830–837.

Costas, R., & Bordons, M. (2007) . The h–index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics, 1* (3) , 193–203.

Egghe, L. (2006) . Theory and practice of the g–index. *Scientometrics, 69* (1) , 131–152.

Franceschet, M. (2009) . A cluster analysis of scholar and journal bibliometric indicators. *Journal of the American Society for Information Science and Technology, 60* (10) , 1950–1964.

Gonzólez–Pereira, B., Guerrero–Bote, V.P., & Moya–Anegón, F. (2010) . A new approach to the metric of journals´ scientific prestige: The SJR indicator. *Journal of Informetrics, 4* (3) , 379–391.

Hendrix, D. (2008) . An analysis of bibliometric indicators, National Institutes of Health funding, and faculty size at Association of American Medical Colleges medical schools, 1997–2007. *Journal of the Medical Library Association, 96* (4) , 324–334.

Hirsch, J. (2005) . An index to quantify an individual´s scientific research output. In *Proceedings of the National Academy of Sciences, 102* (46) , 16569–16572.

Jain, A. K., & Dubes, R. C. (1988) . *Algorithms for clustering data*. Upper Saddle River, NJ: Prentice–Hall.

Katsaros, C., Manolopoulos, Y., & Sidiropoulos, A. (2007) . Generalized h–index for disclosing latent facts in citation networks. *Scientometrics, 72* (2) , 253–280.

Kosmulski, M. (2011) . Successful papers: A new idea in evaluation of scientific output. *Journal of Informetrics, 5* (3) , 481–485.

Kosten, J. (2016) . A classification of the use of research indicators. *Scientometrics, 108* (1) : 457–464.

Leydesdorff, L. (2009) . How are new citation- based journal indicators adding to the bibliometric toolbox? *Journal of the American Society for Information Science and Technology, 60* (7) , 1327–1336.

Liang, J. Y., Zhao, X. W., Li, D. Y. et al. (2012) . Determining the number of clusters using information entropy for mixed data. *Pattern Recognition, 45* (6) , 2251–2265

Markpin, T., Boonradsamee, B., Ruksinsut, K., Yochai, W. et al. (2008) . Article–count impact factor of materials science journals in SCI database. *Scientometrics, 75* (2) , 251–261.

Melero, R. (2015) . Altmetrics–a complement to conventional metrics. *Biochemia medica, 25* (2) ,152–160.

Miot, H. A., Ianhez, M., & Ramos, P. M. (2021) . Trends of the main bibliometric indicators of Anais Brasileiros de Dermatologia (2010–2019) . *Anais brasileiros de dermatologia, 96* (3) , 309–314.

Moed, H. F. (2010) . Measuring contextual citation impact of scientific journals. *Journal of Informetrics, 4* (3) , 265–277.

Pal, N. R., & Biswas, J. (1997) . Clusters validation using graph theoretic concepts. *Pattern Recognition, 30* (6) , 847–857.

Pronin, E. (2007) . Perception and misperception of bias in human judgment. *Trends in cognitive sciences, 11* (1) , 37–43.

Wohlrabe, K., & Gralka, S. M. (2020) . *Classifying top economists using archetypoid analysis.* CEPIE Working Paper.

Yu, L. P. (2014) . Investigate the misunderstanding of total cites and immediacy index from perspective of time. *China Publishing Journal, 2014* (12) , 8–11. (In Chinese) .

Yu, L. P., Pan, Y. T., & Wu, Y. S. (2009) . Research on classification of academic journal evaluation indicators based on factor analysis. *Library and Information Service, 2009* (8) , 146–149. (In Chinese) .

Yu, L. P., Yu, H. Q. (2016) . Does the average IF percentile make a difference?. *Scientometrics, 109* (3) , 1979–1987.