CrossMark

Evaluation of *h*-index, its variants and extensions based on publication age & citation intensity in civil engineering

Muhammad Raheel[1] · **Samreen Ayaz**[1] · **Muhammad Tanvir Afzal**[1]

Abstract The scientific community has proposed diversified set of parameters to rank researchers, including publications, citations, *h*-index, different variants and extensions of *h*-index. However, there is a debate in the scientific Community which index ranks authors in a better way. Current state-of-the-art depicts that these indices are evaluated on imaginary case scenarios and small datasets. Furthermore, these indices are evaluated on different datasets, making it difficult to grasp the contribution and importance of each index over the others. To analyze the individual behavior of each index, these indices should comprehensively be evaluated on some extensive data set. This study emphasizes on the scrutiny of *h*-index, some of its variants and extensions to rank authors. These indices are evaluated using a comprehensive data set of Civil Engineering field. For the evaluation of results obtained from these indices, first correlation was calculated among indices. There exists weak correlation between various indices, which demonstrates that the author's rankings acquired from these indices are not identical. Secondly, occurrences of awardees are checked in all ranked lists. The prestigious award winners of four Civil Engineering societies are considered as a benchmark. In top 10% of ranked list, maximum 47% of the awardees were brought by Wu-index. Overall, none of the index succeeded in bringing 100% awardees to the top rankings. Highest number of awardees on top of all ranked lists are found to be from ACI (American Concrete Institute), which shows ACI might be dependent on these indices for its criterion to honor awards.

✉ Muhammad Tanvir Afzal
mafzal@cust.edu.pk

Muhammad Raheel
raheelmuhammad66@yahoo.com

Samreen Ayaz
samreen.ayaz@gmail.com

[1] Department of Computer Science, Capital University of Science and Technology, Islamabad, Pakistan

🖄 Springer

Keywords ASCE · CEDB · *h*-Index · Variants of *h*-index · Civil engineering field · Author ranking

Introduction

Ranking of researchers/authors in their respective domain can assist in multiple aspects. For instance, ranking of authors can help organizers of journal or conferences to locate a reviewer for a paper, it can help Ph.D. students to find a supervisor and also it can help any organization to recruit a candidate for a job. Researchers/Authors in any field are ranked on the basis of the impact of their research work in the scientific community. Moreover, ranking of authors helps to answer the questions pertaining to the issues like who should get the scholarship; who has more impactful research work; who should get tenure and who should be given promotions?

To date, various techniques have been introduced in the literature for ranking of authors. Each ranking technique has different criteria to rank authors. Some techniques rank quantitatively and some techniques rank authors by considering qualitative expertise. To measure the performance of an author, different author ranking techniques employ the parameters including publication count (Balog et al. 2006; Fang and Zhai 2007), citation count (Cameron et al. 2007; Borgers et al. 2008), co-authors (Deng et al. 2012; Sun et al. 2013) and hybrid approaches (Afzal and Maurer 2011; Hirsch 2005). Furthermore, many variants and extensions of *h*-index have also been proposed (Bornmann et al 2011) which include: *a*-index (Jin 2006), ar-index (Jin 2007), contemporary *h*-index (Sidiropoulos et al. 2007), raw *h*-rate (Burrell 2007), *m*-quotient (Hirsch 2005), *f*-index (Tol 2009), *t*-index (Tol 2009), Maxprod index (Kosmulski 2007), q^2-index (Cabrerizo et al. 2010) and wu-index (Wu 2010) etc.

According to the nature of problem there is no benchmark available to evaluate these techniques and to discover the optimal parameter to rank the authors. In general, whenever a new technique is proposed in the literature, its evaluation is made on either imaginary case scenarios (Alonso et al. 2009; Kosmulski 2007) or on just small and different datasets (Cabrerizo et al. 2010; Bornmann et al. 2011). Furthermore, these techniques are evaluated on different datasets, making it difficult to grasp the contribution and importance of each technique over the others. To identify which technique performs better for ranking of authors, there should be a comprehensive evaluation of these techniques on the wide-ranging dataset from the same domain.

In this paper, we intend to investigate the role of indices based on citation intensity (*h*-index, *a*-index, *f*-index, *t*-index, Maxprod index, q^2-index, Wu-index, tapered *h*-index) and publication age (ar-index, contemporary *h*-index, raw *h*-rate) on a comprehensive data set belonging to the domain of Civil Engineering. The purpose of this research is the evaluation of these indices and identification of the best performing indices for ranking of authors. For evaluation purposes, awards given by four prestigious societies in Civil Engineering field are considered as a benchmark, which include ASCE (American Society of Civil Engineering), CSCE (Canadian Society of Civil Engineering), ACI (American Concrete Institute) and ICE (Institute of Civil Engineering).

Literature review

The analysis of scientific production of authors plays pivotal role in different aspects of research, judgmental decisions, accepting research projects, nominating scientific awards, promoting an individual, allocating tenures to skilled persons and to grant the contracts to the experts. There exist various evaluation parameters upon which the entire research work of an individual author may be ranked. For this purpose, many techniques have been proposed in the literature. One of the earliest proposed parameter for scientist's assessment is publication count. The number of publications is harnessed for evaluation of authors (Balog et al. 2006) but it does not cover the impact and quality analysis of one's work (Cameron et al. 2007). To overcome this deficiency, the citation count was introduced (Bogers and Van 2008). But there are some limitations with citation count too e.g. newly published papers require a lot of time to acquire citations. Therefore, novice authors do not get a chance to be ranked in the top position because of having low number of citations even if they have large number of publication count. Another issue is that sometimes scientists cite the research papers of other authors to criticize them, the count of such citations can't be deemed as a measure to analyze the quality of a publication.

To overcome the issues pertaining to these parameters, h-index was introduced by Hirsch (2005). H-index gained the popularity in a few years and nowadays the h-index has become the dominant parameter over all other measurements for evaluating research output of authors. H-index has some weaknesses as well, for example h-index is mainly based on long term observations, so the old authors have large h-index than novice authors and it is field specific. In order to overcome the deficiencies of h-index, various modifications of h-index are introduced which are known as variants and extensions (Bornmann et al. 2011).Many studies are presented in the literature in which different indices are applied on different data sets, and a few of those studies are presented in this section.

In 2006, Kosmulski tested the relationship between the h-index and h(2)-index. The data on which the relationship was examined includes 19 professors affiliated with the department of Chemistry in a Poland university (Kosmulski 2006). Results of Kosmulsiki's study revealed the existence of a strong correlation between the two indices which indicates the results acquired from two indices are same.

In a recent study, Ayaz and Afzal (2016) have evaluated the h-index (Hirsch 2005), g-index (Egghe 2006) and complete-h (Dienes 2015) on a real data set in the field of Mathematics. Their proposed scheme has considered awarding winners from the Mathematics field as a benchmark. According to this study, complete-h, which is defined to complete the definition of h-index having added community factor in h-index definition, performs better than h-index and g-index in bringing award winners to the top rankings (Ayaz and Afzal 2016).

Another study is conducted by Van Raan (2006) to compare the h-index with other indicators. In this study, to examine the correlation between these indicators and h-index, the author has utilized the results of larger evaluation study of 147 research groups of chemistry in Netherland. The author focused on research groups rather than the individual performances and the citations are restricted to only three year window instead of whole life citation count (Van Raan 2006).

Yan et al. (2016) have considered 29 different indicators including 26 h-index variants, calculated correlation among the indices and with h-index and Wu-index, considering h-index and Wu-index as benchmarks. They found out that with some exceptions, mostly indices which are highly correlated with h-index are less correlated with Wu-index,

similarly which are highly correlated with Wu-index have low correlation with *h*-index. It is also highlighted that those indices which have high correlation show little improvement over *h*-index or Wu-index and should be abandoned or integrated into these indices (Yan et al. 2016).

In 2007, Jin and colleagues proposed a scheme to utilize a combination of indices for the purpose of authors' evaluation (Jin et al. 2007; Liu and Rousseau 2007). The scheme utilizes a combination of the *h*-index and *r*-index, or combination of the *h*-index and the ar-index. Both the pairs were used to find an indicator for scientist's evaluation. In a pair, one index is used as quantitative measure while the other index is used as qualitative measure. According to their findings, pair (*h*-index, ar-index) is a good indicator for research evaluation.

In 2008, Schreiber performed an empirical study of the *g*-index in comparison with the *h*-index, *a*-index, and *r*-index for data sets of 26 physicists (Schreiber 2008). In this study, Schreiber analyzed the citation records of 26 physicists of the Institute of Physics at Chemnitz University of Technology. The data sets for analysis were compiled in January 2007 from Web of Science. According to Schreiber, the *g*-index is more suitable than the *h*-index to characterize the overall impact of the publications of a scientist.

In another interesting study, Wu and Zhang (2017) emphasized that validity of an index should be checked prior to its usage as an evaluation criteria for different purposes, like promotion, funding allocation or other such decisions. It is asserted that for an index to be valid, it should satisfy basic domination relation. Basic domination relation states, if scholar A's papers have citations at least as much as citations of the scholar B's corresponding papers, when sorted in descending order of citation count, then scholar A should not be considered inferior to scholar B. In other words, different indicators or indices values' for scholar A should never be less than scholar B. It is identified that *a*-index, *m*-index, *e*-index and q^2-index violate the basic domination relation so, these indices should not be used independently.

In the above paragraphs, we have discussed different studies in which the indices are used to evaluate/rank the authors. Following are the observations derived from the above discussions.

The indices which are discussed above in literature are mostly evaluated for hypothetical case scenarios to prove their accuracy. Secondly, these indices are applied on different and very small data sets. There is need to evaluate these indices on a comprehensive and large data set of a specific domain. Thirdly, there exist no standard benchmark to evaluate these indices. A benchmark data set is required to evaluate these indices.

Methodology

The research questions focused in this study are:

Q1   Is there any correlation between the ranked lists achieved from the previously known author ranking indices?

Q2   Whether the international prestigious awardees lie in the top ranking obtained from these indices?

Q3   Which index contributes the most to bring the international prestigious awardees in the list of top authors?

Q4   Which award honouring society (ASCE, CSCE, ICE, and ACI) depend upon these indices?

The methodology of proposed solution to answer these questions is discussed in this section and the results are discussed in the next section. The architecture diagram of our methodology is given in Fig. 1.
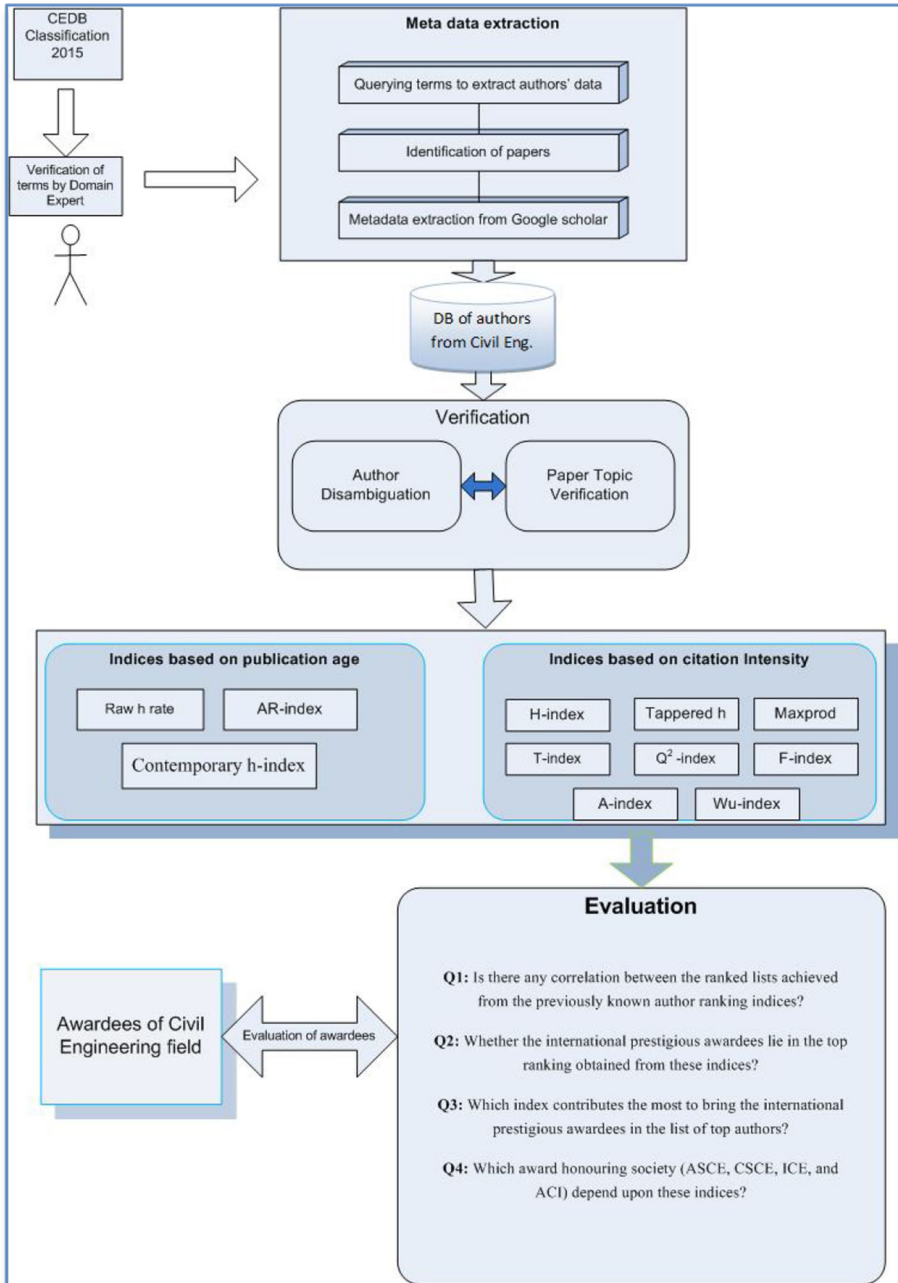

Fig. 1 Architecture diagram of methodology

Domain selection

For the evaluation of indices, we have selected the Civil Engineering field as it is one of the most important fields of science consuming over 40% amount of world's matter (mass and energy) and still not well explored for Scientometrics studies (Czarnecki et al. 2013).

Taxonomy building

Civil Engineering is a diverse and vast field of knowledge. To collect comprehensive data set for evaluation of indices, topic classification for Civil Engineering was acquired. The certified classification of Civil Engineering is CEDB[1] (Civil Engineering Data Base) which is the effort of ASCE (American Society of Civil Engineering). ACSE is a well-known society of Civil Engineering and publishes 37 journals in this field. The CEDB is used by all these journals of Civil Engineering, these journals request the authors to use the subject heading of CEDB. It helps in finding the items of potential interest to users. Every year the database of this classification is updated. We have started to collect data in September 2015, so we have used CEDB 2015 classification.

Search engine

To collect the data of authors, publications and citations against these categories of CEDB, we have used Google Scholar. Other resources like Web of Science and Scopus were also available, but the reason to choose Google Scholar is that it has large coverage of publications than other recourses (Falagas et al. 2008; Moed et al. 2016), Google Scholar allows open access to data, whereas other recourses are not openly available and Web of Science has limited coverage than Google Scholar. Another reason to choose Google Scholar is that it also does citation indexing and data relevant to every field of science is comprehensively available. A recent study reveals that the growth of Google Scholar is 13% more than that of Web of Science. The number of citations in Google Scholar increases by approximately 1.5% per month (Harzing 2014). In another study it has been demonstrated that the citation noise (non-Scholar citations, count citations double) in Google Scholar is less than that of Web of Science (Below 2005) and Scopus (Moed et al. 2016). According to Google scholar, they are adding new papers, several times in a week, so we can say that Google Scholar is also updated often. With respect to updating existing Google Scholar record and insertion of new records Harzing reports in a study that Google Scholar is updated every 2–3 months (Harzing 2014). Many researchers work on the comparison of Google Scholar with other sources like Web of Science, they finally conclude that the Google Scholar is the best source for publication, citation and other metadata (Moed et al. 2016; Meho and Yang 2007; De Winter et al. 2014; Noruzi 2005; Handerson 2005). Hence we can say that the Google Scholar is a preferable source to collect the data of publication, citation and authors of these publications. Therefore, to collect the data set of Civil Engineering the terms of classification categories are given to Google Scholar by using a crawler. It collects title of Publication, names of authors, Venue of publication, Citation Count, URL of Paper and year of publication.

Beel and Gipp (2009) have discussed how Google Scholar extracts the records against the given query. They concluded that the top results are more relevant to query term and this relevance is measured by the occurrence of the query in the paper's title. The results

---

[1] http://cedb.asce.org/CEDBsearch/.

shown by Google Scholar are first ranked according to relevance with query term and then the citations received by searched results. To collect the data we used Google Scholar because it gives top ranked highly cited papers and relevant to the query.

The terms from classification were given to the crawler to collect data of publications and authors. There were many generic terms in the classification which give irrelevant results when given to Google Scholar. The example of this problem is illustrated in the Fig. 2. The top result "construction grammar" is not relevant to construction in Civil Engineering. All such terms were further tuned by domain experts and only those terms were selected which only related to domain of Civil Engineering.

With the help of domain expert fine-tuned taxonomy is built that contains 152 terms. The new fine-tuned taxonomy is available as ''Appendix 3'' at http://cdsc-cust.org/ Appendices/.

Crawling metadata of authors

For data collection, we used a dedicated crawler which retrieves the metadata of authors from Google Scholar. The metadata contains the title of paper, journal or conference in which it was published, author's name, the URL of papers and citations of the paper. In order to collect data, the selected subcategories were given to crawler and crawler first identified the papers against given subcategory on Google Scholar and then extracted the records. The crawler extracted top 600 records from Google Scholar, the reason to select only top 600 records is that the Google Scholar gives irrelevant records beyond 600 records. The records are further saved in a database and the database was maintained in SQL server. It must be noted here that for some queries, the Google Scholar returns results below 600. In such cases, whatever number is returned by Google Scholar is saved into the database.

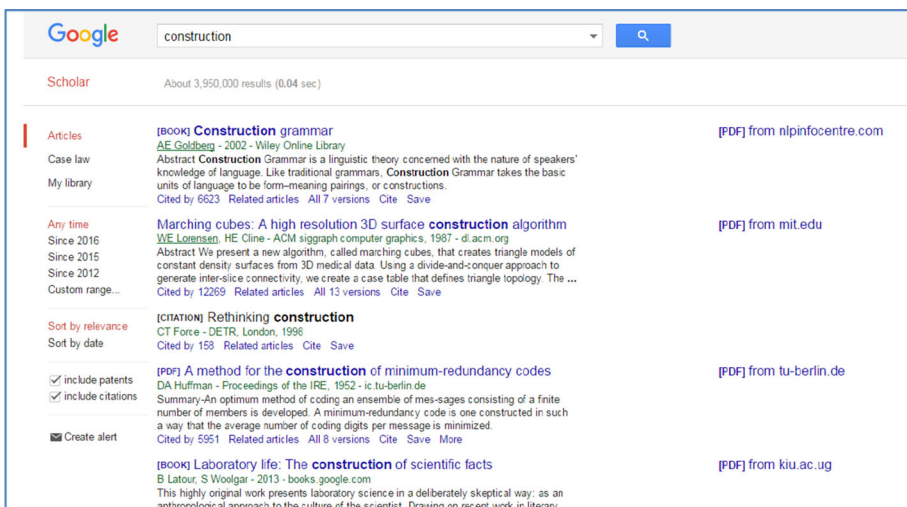Table 1 shows total publications and authors collected through dedicated crawler.

Fig. 2 Irrelevant results against "Construction"

Table 1 Data set before cleaning

Total publications	20,307
Total authors	36,921
Total citations	2,184,638

Data cleaning

Many authors (Deng et al. 2012; Harzing 2014; Jacsó 2008; Ayaz and Afzal 2016) who collected data from Google Scholar pointed that the collected data should be cleaned as it has noise. Therefore, the data was further cleaned and verified, this step of cleaning was performed in two cases. In first case it is ensured whether the data is relevant to Civil Engineering field. In second case it is ensured that the authors are disambiguated (Rosenstreich and Wooliscroft 2009; Aguillo 2011).

To verify data in first case further three steps are followed:

1. Removal of crawled records which have invalid characters in the title ($, £, %, *, &, ?, /).
2. Verification of papers whether they belong to journals or conferences of Civil Engineering.
3. Manual verification of the remaining results.

After performing these steps, some publications were removed and final publications have been shown in Table 2.

In second step, the authors are disambiguated, for this purpose; we checked the duplication of author names. To perform this task first we find the second name shared with other authors. After this we have analysed the first names and checked whether the first names were distinct, if both first name and second name of more than one author were same, it was considered as duplication of authors. Further, we manually checked those authors by browsing their profile on home pages.

In second step, author disambiguation is performed. In data set the total number of authors were 36,921 and among those authors, 17,589 authors shared the same last name, those 17,589 authors further needed to be disambiguated. By verification, we found that 4130 distinct names were shared by 17,589 authors. This means that there were 4130

Table 2 Publications data set after verification

Descriptions	Number of instances effected
Publications which include invalid keywords	34 Removed
Publications which were published in other than Civil Engineering venues	3250
Publications which were published in other venues and were not relevant to Civil Engineering field	404 Removed
Total publications removed from data set	438
Total publications after verification	19,869

authors having a total of 17,589 variations. In 4130 names, some names were shared by 100 plus authors and some names were shared only by 2 authors.

After the data collection the numbers of authors in the dataset were 36,921. There were two cases in name variations, these two cases are highlighted in the Fig. 3. In case 1 the first and last names of authors are same, in case 2 the last names of authors are same but first names are different. We verified such type of record to ensure the author disambiguation and we handled both cases. We accomplished this task manually.

In case 1 the publication of each author among 4130 authors were checked manually to ensure that either the authors belong to different name or same name. Interestingly there was not any author having same first and last name. For case 2, we again manually visited the home profiles of authors and checked their publications to ensure either they are different authors or same. In this process 88 variations were detected as duplications of 45 authors, it means that 45 authors shared same last name with different first names and the duplications were removed from database. After all these verifications we obtained the dataset of 36,876 authors and 2,184,638 total citations against 152 subcategories or terms of CEDB. Table 3 shows number of authors before and after verification.

Author ranking by different indices

After data collection and cleaning, we calculated 11 indices on data, ''Appendix 1'' shows calculation of these 11 indices. After calculation of indices we have separately ranked the authors according to each index. After this step, we get 11 distinct rankings of authors which are further evaluated based on four research questions postulated in this study.

Award benchmark

To evaluate the indices, we have considered international prestigious awardees given in the field of Civil Engineering as our benchmark. The data of awardees is collected from different societies which include ICE, ASCE, CSCE and ACI, these societies bestow awards to competent authors. List of awards is shown in Table 4 and data of awardees is available as ''Appendix 2'' at http://cdsc-cust.org/Appendices/.

Evaluation

After the ranked lists have been calculated, the evaluation of four research questions has been addressed in this section.

	Authors first name	Authors last name
Case 1	K	Adams
	K	Adams
Case 2	M	Adams
	D	Adams

Fig. 3 Show the cases of name variations

Table 3 Authors data after verification

Descriptions	Number of instances effected
Initial no. of authors	36,921
Authors sharing same last name	17,589
No. of names shared	4130
No. of records having duplicated authors	88
No. of duplicated authors	45
Total authors after removing duplications	36,876

Table 4 List of awards

Societies and their awards	Total awardees
ACSE	669
Collingwood prize past award	59
J. James R. Croes medal	62
Samuel Arnold Greeley award	60
Rudolph Hering medal	58
Wesley W. Horner award	64
Karl Emil Hilgard Hydraulic prize	63
Moisseiff award	62
Norman medal	114
Alfred noble prize	62
Raymond C. Reese research prize	65
CSCE	174
Cazimir Gzowski medal	29
Thomas C. Keefer medal	35
P.L. Pratley award	37
Donald Stanley award	34
Horst Leipholz medal	39
ACI	165
Construction award	35
Design award	30
Chester Paul Siess award	33
Wason medal for materials research	36
Wason medal for most	31
ICE	53
Telford medal	28
James Alfred Ewing medal	24
Grand total	1060

Correlation between the indices

The first research question is to find whether there is correlation between ranked lists obtained by these indices. The purpose of finding correlation is just to check as to what

extent these indices are similar to each other. Spearman's rank correlation coefficient is used keeping in view the rank nature of our results (Corder and Foreman 2009).

Awardees' trend in ranked lists

In this step, we evaluated top ranked authors and focused on second question and third question. The second question is whether the international prestigious awardees lie in top ranked authors. Third question is which index contributes more to bring awardees in rank list. For both questions, we checked the percentage of occurrences of awardees in top 10% of ranked lists then 1–10%, 10–20%, 20–30% up to 90–100%. Moreover, the occurrences of awardees are checked in top 100, 500 and 1000 ranked authors.

To find out the percentage of occurrence of awardees in ranked lists, first we have calculated all the indices for all the authors. After that, we have ranked authors separately, according to each index. We have identified/marked the position of award winners in these ranked lists. Then it was calculated, that how many award winners lie in top 10% of each ranked list. Finally we have calculated that out of total award winners, which were found in our data set, what percentage of award winners lie in top 10% of each ranked list. Similarly for 1–10%, 20–30% and so on. Let us explain with an example, let us suppose there are 50 authors in data set, out of which 10 are award winners. We sort the authors according to h-index value, from high to low. We extract top 10% records from data set, it would comprise of 5 authors' record which are on top, on the basis of value of h-index. Suppose in top 10% we have 2 occurrences of award winners. Hence in this case 20% of award winners lie in top 10% of authors ranked on the basis of their h-index value.

Dependency of award honoring societies on these parameters

This section focuses on fourth question which is to check dependency of indices on award giving societies (ASCE, CSCE, ICE and ACI). To accomplish this evaluation task we find the occurrence of awardees in top 10% of ranked lists obtained from these indices and according to our observation the index which brings more awardees is dependent on that society.

In previous research question we have identified/marked the position of award winners in these ranked lists and also calculated how many award winners lie in top 10% of each ranked list. We have separated the award winners found in our data set on the basis of their award giving societies, statistics are shown in table. Finally we have calculated that out of total award winners from one society how many lie in top 10% of ranked lists. We have done this separately for all the authors' lists ranked on the basis of different indices. For example there are 299 award winners from ASCE which are also found in our data set. Let's suppose out of these 299, 120 award winners are found in the top 10% of author's ranked list based upon their h-index value. In this case 40% of award winners from ASCE are found in top 10% of the list ranked on the basis of h-index values.

Results and discussion

In this section the answers of our four research questions are discussed.

Correlation between the ranked lists obtained from indices

The purpose of this evaluation is to find the similarity between the ranked lists and to answer the first research question i.e. to compute the correlation among all the indices. To accomplish this task, the correlation of each ranked list with all other ranked lists is computed. The value of correlation greater than zero indicates the extent to which a specific index is positively correlated with other index, the value of correlation less than zero indicates that the indices are negatively correlated, and the value of zero indicates that there exists no correlation among indices. The obtained values of correlation of each index with every other index are presented in Table 5 and are graphically illustrated in Fig. 4. We followed guidelines regarding strength of correlation from (Evans 1996). The strength of correlation for absolute value of r is very weak (0.0–0.19), weak (0.20–0.39), moderate (0.40–0.59), strong (0.60–0.79) and very strong (0.80–1.0). In Fig. 4, only weak and strong correlation between indices is presented whereas detailed values are shown in the Table 5. The Table 5 shows the correlation between the ranked lists acquired from 11 indices and red color is used to highlight the values which show strong or very strong correlation. The correlation of index with itself is 1, and values in blue color represent weak or very weak correlation.

The correlation values among different indices are graphically shown in Fig. 4. This figure presents only low and high correlation values. The relationship between those indices having values below 0.4 or having values equal to or above 0.6 are exhibited. The rectangle is used to represent indices, edges represent the strength of a correlation between two indices. The edges in red color stipulate the existence of a strong or very strong correlation between the indices. Similarly, the indices connected through blue edges have weak or very weak correlation. The rankings acquired from 11 indices are not similar because most of the edges are of blue color, which depicts the existence of vast cases of a weak correlation.

Table 6 shows the frequency of low, high and negative correlation among indices. For instance, h-index has strong correlation with 5 indices and weak/low correlation with 5 indices. It is not negatively correlated with any of the indices. Overall the number of indices having low correlation is greater than those indices, which have strong correlation.

From these correlation results, it can be observed that overall weak correlation cases prevail the strong correlation cases. Moreover, raw h-rate and a-index have negative correlation with each other. Hence, it can be stated that in most of the cases the indices are not strongly correlated. The findings from first question further motivates to scrutinize next three research questions.

Awardees trend in author ranked lists

This section presents the evaluation of next two research questions, which are to assess the contribution of indices in bringing award winners in top ranks and to find the occurrence of award winners in top ranks.

To evaluate these two questions, first we examined the occurrence of awardees in top 10% of ranked list for each parameter. Secondly, we examined the occurrence of awardees in 1–10%, 10–20%, 20–30% and up to 90–100% of ranked list calculated by each parameter. The total number of considered awardees are 1060, from which 135 awardees are duplicated, after removing duplication, and remaining 925 are unique awardees. Ideally, all the award winners of this field should be present in the data set, but it only contains

Table 5 Correlation of indices

	H-index	Wu-index	A-index	Maxprod	Ht-index	F-index	T-index	AR-index	Q2-index	Row h-rate	Hc-index
H-index	1	0.935**	0.047	0.291**	0.768**	0.992**	0.98**	0.375**	0.307**	0.459**	0.939**
Wu-index	0.935**	1	0.076	0.319**	0.84**	0.911**	0.886**	0.428**	0.343**	0.505**	0.945**
A-index	0.047	0.076	1	0.949**	0.332	0.043	0.04	0.729**	0.931	− 0.02?**	0.089**
Maxprod	0.291**	0.319**	0.949**	1	0.545**	0.284**	0.278**	0.778**	0.981**	0.069**	0.328**
Ht-index	0.768**	0.84**	0.332**	0.545**	1	0.747**	0.724**	0.598**	0.566**	0.334**	0.797**
F-index	0.992**	0.911**	0.043**	0.284**	0.747**	1	0.994**	0.362**	0.295**	0.443**	0.925**
T-index	0.98**	0.886**	0.04***	0.278**	0.724**	0.994**	1	0.35**	0.286**	0.43**	0.909**
AR-index	0.375**	0.428**	0.729**	0.778**	0.598**	0.362**	0.35**	1	0.76B**	0.389**	0.448**
Q2-index	0.307**	0.343**	0.931**	0.981**	0.566**	0.295**	0.286**	0.768**	1	0.077**	0.349**
Raw h-rate	0.459**	0.505**	− 0.02**	0.069**	0.334**	0.443**	0.43**	0.389**	0.077**	1	0.506**
Hc-index	0.939**	0.945**	0.089**	0.328**	0.797**	0.925**	0.909**	0.448**	0.349**	0.506**	1

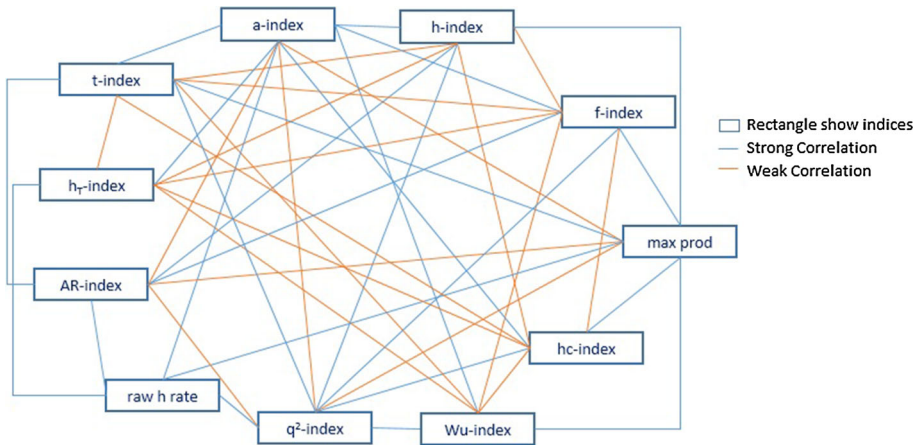**Correlation is significant at the 0.01 level

Fig. 4 Strong and weak correlation

Table 6 Strong and weak correlation

Index	Strong correlation (> 0.60)	Weak correlation	Negative correlation
h-index	5	4	
Wu-index	5	3	
*A-index	3	6	1
Maxprod-index	3	6	
Tapered h-index (h_T index)	5	2	
F-index	5	4	
T-index	5	4	
AR-index	3	4	
Q^2-index	3	6	
raw h-rate	0	4	1
Contemporary h-index (hc-index)	5	3	

*we can read it as: A-index has strong correlation with 3 indices, low correlation with 6 indices and negative correlation with 1 index

501 awardees. Therefore, we checked occurrences of these 501 awardees in the ranked lists.

Usually, it is assumed that award winners hold a strong research background, their number of citations, publications must be high. Based on this assumption, it was expected that all the award winners should lie in top 10% of authors in lists ranked according to these indices. Whereas, our results portrayed a different picture as shown in Fig. 5 which illustrates the occurrence of these awardees in top 10% of ranked lists acquired from indices.

From Fig. 5, it can be observed that maximum occurrences of awardees (around 47%) lies in the top 10% of the authors in the list ranked by their Wu-index value, followed by f-index and t-index. Comparatively the better performance of Wu-index could be due to the
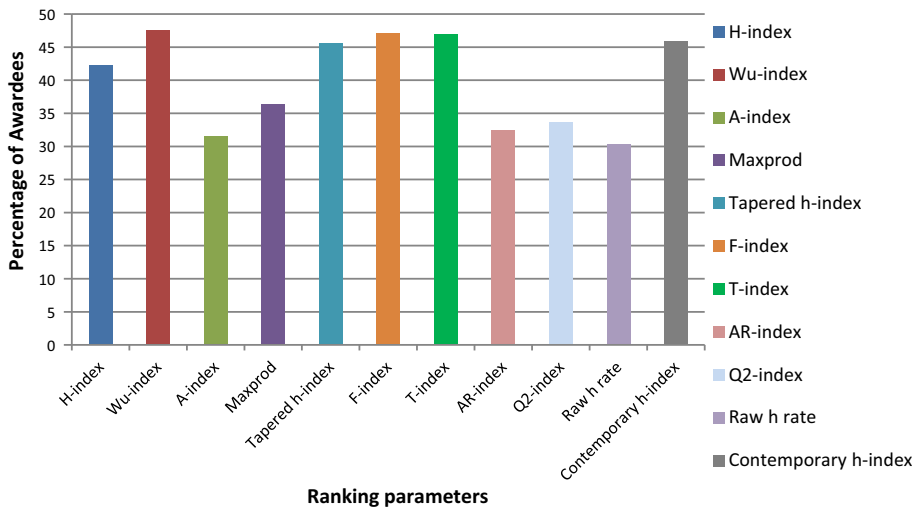
Fig. 5 Percentage of occurrence of awardees in top 10% of ranked lists of each index

fact that it pays more attention to the highly cited papers (Egghe 2011; Schreiber 2010; Kosmulski 2013). The indices that focus on highly cited papers are also found better in terms of evaluating economists (Ellison 2013). Wu-index is also known for more accurately reflecting the influence of a scientist's top papers (Panaretos and Malesios 2009) as well as in assessing the integrated impact of a researcher's work, especially the most excellent papers (Wildgaard et al. 2014). Furthermore, overall performance of tapered *h*-index and contemporary *h*-index is also around 45% and *h*-index brings 42% awardees. In rest of ranked lists, the occurrences of awardees in top 10% are lower, which are astonishing results as raw *h*-rate performs 30%.

None of the indices succeeded in bringing even 50% of award winners in top 10% of the ranked lists. We decided to analyze the trend in 100% results, although it is expected that award winners would merely be present below the top 50% of ranked lists, it is quite surprising that the award winners are dispersed from top 10% to the least 10% throughout the experiments.

The Fig. 6 depicts the overall percentage of occurrence of awardees in 1–10%, 10–20%, 20–30% up to 90–100% from the ranked lists. Wu-index, *f*-index, *t*-index, hc-index and h_T-index perform better in top 10% of the ranked lists, but *a*-index and raw *h*-rate have low performance even in top 10%. Other indices also perform better in 1–10% of the ranked list and bring average number of awardees. The occurrence of awardees gradually decreased as we moved in downward of the ranked list. In 20–30%, 30–40% and other lower portions up to 90–100%, the performance of all indices remained same except a-index and raw h-rate.

Dependency of awarding society on indices' results

This section presents the evaluation of fourth research question i.e. to find out which awarding society depends upon these indices. To answer this question, we examined the occurrences of awardees of each society in top 10% of the ranked list. As we have mentioned that 501 award winners were present in our data set, out of these, 299 were from ASCE, 76 from CSCE, 96 from ACI and 30 award winners from ICE. Figure 7 shows the
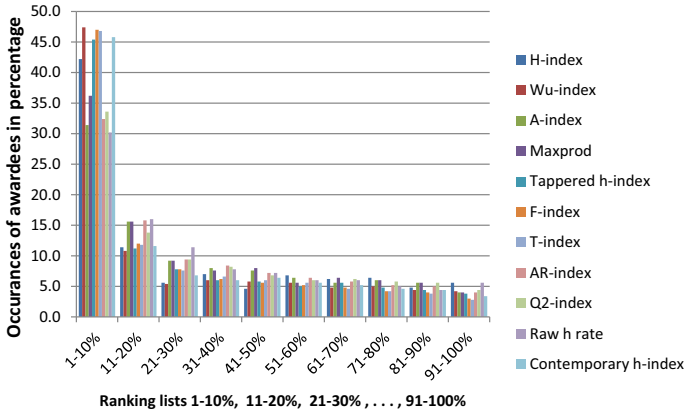
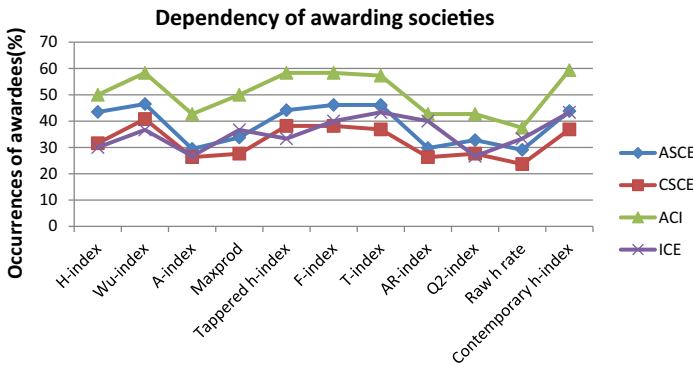**Fig. 6** Occurrence of awardees in partial 10% parts of ranked lists

Fig. 7 Trend of indices on awarding societies

dependency of awarding societies on these indices. Following observations are formed from the obtained results.

1. ASCE
   a. Overall the Wu-index, *t*-index and *f*-index outperformed all other indices and around 46% of awardees from ASCE exist in the ranked lists of Wu-index, *t*-index and *f*-index.
   b. *H*-index, h_T-index and hc-index also performed sufficiently by securing 43% approximately.
   c. The performance of *a*-index, ar-index and raw *h*-rate remained lower which is around 30%.

2. CSCE
   a. The Wu-index brought 41% of the awardees from CSCE. The h_T-index, *f*-index, *t*-index and hc-index almost performed equally, with around 38, 38, 37 and 36% respectively.
   b. The *a*-index, Maxprod index, ar-index, q^2-index and raw *h*-rate didn't perform well.

3. ACI

    a. Hc-index brought highest number of awardees from ACI awarding society and performs 59%.
    b. H_T-index, Wu-index, *f*-index and *t*-index performed equally, brought around 58% awardees from ACI society.
    c. The raw *h*-rate didn't perform well to bring awardees from ACI society.
    d. The performance of other indices, *a*-index, ar-index and q^2-index is not very optimal, but better than raw *h*-rate.

4. ICE

    a. Hc-index and *t*-index performed better than other indices with respect to ICE awarding society, as around 43% of the awardees are brought in by these indices.
    b. AR-index and *f*-index equally performed well and brought 40% awardees from ICE.
    c. A-index and q^2-index performed inferior to all other indices.

The Fig. 7 shows the trend of awardees of each society.

Evaluation of awardees in top ranked authors

We examined the occurrences of awardees in top 100, 500 and 1000 ranked authors. Figure 8 shows the percentage of awardees' existence in top ranked authors.

Figure 8 shows the occurrences of awardees in top 100, 500 and 1000 ranked authors. In the top 100, the maximum number of awardees, which could be brought are just close to 6% which is an astonishing fact. In top 100 ranked authors, the Wu-index and hc-index equally performed well (6%). The performance of *h*-index and h_T-index remained equal, which is 5%. F-index and *t*-index also performed well around 4.6%. The performance of *a*-index and raw *h*-rate remained lower, the ar-index, Maxprod index and q^2-index performed equally but brought only a few of awardees to the top 100. In top 500 awardees, around 17% of the awardees lie in the ranked lists obtained from *f*-index and *t*-index. These indices have brought the highest number of awardees to the top 500. A-index, q^2-index and raw h-rate brought a small number of awardees, which are around 9%. In top 1000 ranked authors, h_T-index, *hc*-index and Wu-index remained on the top in bringing awardees
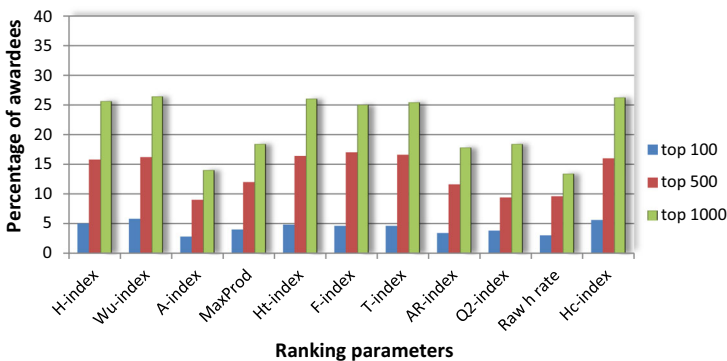

Fig. 8 Occurrence of awardees in top ranked authors

around 26%, *f*-index, *t*-index and *h*-index performed equally by bringing 25% of the awardees to top 1000.

Conclusion

In this study, the *h*-index, its variants and extensions based on citation intensity and publication age are evaluated. We have investigated the role of *h*-index, raw *h*-rate, *a*-index, ar-index, contemporary *h*-index, tapered *h*-index, q^2-index, *t*-index, *f*-index, Max-prod index and Wu-index on comprehensive data set collected from the domain of Civil Engineering. To compare the results obtained from these indices, we have used international prestigious awardees in Civil Engineering as benchmark.

The Spearman Rank correlation is calculated among indices to address the first research question. The results revealed that there exists a weak correlation among most of the indices except *h*-index and *f*-index as these two indices have obtained the strong value of correlation. The negative correlation is found between raw *h*-rate and *a*-index. The positive correlation indicates the extent to which rank lists are similar and negative correlation indicates the extent to which rank list are opposite to each other. Overall, weak correlation among most of the indices implies that 11 rankings are different from each other, these findings motivated us to explore other research questions.

To answer the second and third question i.e. *the evaluation of occurrence of awardees in these ranked lists*, we compared the awardees with the top 10% ranked authors. Unexpectedly, none of the indices succeeded in bringing even 50% award winners in the top 10% authors. Wu-index, *t*-index and *f*-index brought 47% awardees whereas tapered h-index or h_T-index and hc-index equally brought around 45% awardees. 31 and 30% awardees are found in top 10%, when ranked according to *a*-index and raw *h*-rate respectively. Wu-index, *f*-index and *t*-index brought highest number of awardees to the top 10% of the ranked list, it is evident that these indices contribute more to bring awardees at the top.

To answer the fourth question related to the dependency of award giving societies on these indices, we compared the awardees individually from each society with ranked lists. The results depicted that almost 60% of the awardees from ACI awarding society lie in all ranked lists. The existence of more awardees from ACI awarding society shows its dependency on these indices.

Apart from these indices, various other indices have been proposed by researchers' ranking community that should be evaluated on a comprehensive dataset to signify the potential of each index. In future, we intend to evaluate these remaining indices on a comprehensive data set from different domains.

Appendix 1

See Table 7.

Table 7 Indices and their definition/formula

Indices	References	Definitions/calculation formulas
Indices based on age of publications		
Contemporary *h*-index	Sidiropoulos et al. (2007)	$S^c(i) = \gamma * (Y(\text{now}) - Y(i) + 1)^{-\delta} * \|C(i)\|$ where $\gamma = 4$ and $\delta = 1$
AR-index	Jin and Rousseau (2007)	$AR = \sqrt{\sum_{p \text{ in } H} \frac{\text{cit}_p}{a_p}}$
Raw *h*-rate	Burrell (2007)	The *raw h-rate* of an author at time T is given by $h(T)/T$. $H(T)$ denotes the *h*-index at given time T and the T is total time
Indices based on citation intensity		
H-index	Hirsch (2005)	A scientist has index h if h of his or her papers have at least h citations each and the other papers have fewer than $\leq h$ citations each
A-index	Jin et al. (2006)	$A = \frac{1}{h} \sum_{j=1}^{h} \text{cit}_j$
Tappered *h*-index	Anderson et al. (2008)	$h_{T(j)} = \frac{n_j}{2j-1}$, if $(n_j \leq j)$ $h_{T(j)} = \frac{j}{2j-1} + \sum_{i=j+1}^{nj} \frac{1}{2i-1}$, if $(n_j > j)$ where n_j citation of article j
Wu index	Wu (2010)	An author has *w*-index if w of his/her articles receives at least $10\,w$ number of citations and the other articles have less than $10(w + 1)$ citations
q^2-index	Cabrerizo et al. (2010)	$q^2 = \sqrt{h.m}$
Maxprod	Kosmulski (2007)	The highest value in $i \times c_i$ where i is the publication of author and c_i denotes the number of citations of *i-th* most cited paper from all citations
T-index	Tol (2009)	$T_{(p)} = \max_t \, \exp\left[\frac{1}{t} \sum_{i=1}^{t} \ln(c_i)\right]$
F-index	Tol (2009)	$f(p) = \max_f \frac{1}{\frac{1}{f} \sum_{i=1}^{f} \frac{1}{c_i}}$

References

Afzal, M. T., & Maurer, H. A. (2011). Expertise recommender system for scientific community. *Journal of Universal Computer Science, 17*(11), 1529–1549.

Aguillo, I. F. (2011). Is Google Scholar useful for bibliometrics? A webometric analysis. *Scientometrics, 91*(2), 343–351.

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2009). h-Index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics, 3*(4), 273–289.

Anderson, T., Hankin, R., & Killworth, P. (2008). Beyond the Durfee square: Enhancing the h-index to score total publication output. *Scientometrics, 76*(3), 577–588.

Ayaz, S., & Afzal, M. T. (2016). Identification of conversion factor for completing-h index for the field of mathematics. *Scientometrics, 109*(3), 1511–1524.

Balog, K., Azzopardi, L., & De Rijke, M. (2006). Formal models for expert finding in enterprise corpora. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 43–50). ACM.

Beel, J., Gipp, B., & Wilde, E. (2009). Academic search engine optimization (ASEO): Optimizing scholarly literature for Google Scholar & Co. *Journal of Scholarly Publishing, 41*(2), 176–190.

Belew, R. K. (2005). Scientific impact quantity and quality: Analysis of two sources of bibliographic data. arXiv preprint cs/0504036.

Bogers, T., & Van den Bosch, A. (2008). Recommending scientific articles using citeulike. In *Proceedings of the 2008 ACM conference on Recommender systems* (pp. 287–290). ACM.

Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H. D. (2011). A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics, 5*(3), 346–359.

Burrell, Q. (2007). Hirsch index or Hirsch rate? Some thoughts arising from Liang's data. *Scientometrics, 73*(1), 19–28.

Cabrerizo, F. J., Alonso, S., Herrera-Viedma, E., & Herrera, F. (2010). q^2-Index: Quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core. *Journal of Informetrics, 4*(1), 23–28.

Cameron, D. H. L., Aleman-Meza, B., Decker, S., & Arpinar, I. B. (2007). SEMEF: A taxonomy-based discovery of experts, expertise and collaboration networks (Doctoral dissertation, University of Georgia).

Corder, G. W., & Foreman, D. I. (2009). *Comparing variables of ordinal or dichotomous scales: Spearman rank-order, Point-biserial, and biserial correlations* (pp. 122–154). Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach.

Czarnecki, L., Kaźmierkowski, M. P., & Rogalski, A. (2013). Doing Hirsch proud; shaping H-index in engineering sciences. *Bulletin of the Polish Academy of Sciences: Technical Sciences, 61*(1), 5–21.

De Winter, J. C., Zadpoor, A. A., & Dodou, D. (2014). The expansion of Google Scholar versus web of science: A longitudinal study. *Scientometrics, 98*(2), 1547–1565.

Deng, H., Han, J., Lyu, M. R., & King, I. (2012). Modeling and exploiting heterogeneous bibliographic networks for expertise ranking. In *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 71–80). ACM.

Dienes, K. R. (2015). Completing h. *Journal of Informetrics, 9*(2), 385–397.

Egghe, L. (2006). An improvement of the h-index: The g-index. *ISSI Newsletter, 2*(1), 8–9.

Egghe, L. (2011). Characterizations of the generalized Wu- and Kosmulski-indices in Lotkaian systems. *Journal of Informetrics, 5*(3), 439–445.

Ellison, G. (2013). How does the market use citation data? The Hirsch index in economics. *American Economic Journal: Applied Economics, 5*(3), 63–90.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove: Brooks/Cole publishing.

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, web of science, and Google Scholar: strengths and weaknesses. *The FASEB Journal, 22*(2), 338–342.

Fang, H., & Zhai, C. (2007). Probabilistic models for expert finding. In *European Conference on Information Retrieval* (pp. 418–430). Berlin: Springer.

Harzing, A. W. (2014). A longitudinal study of Google Scholar coverage between 2012 and 2013. *Scientometrics, 98*(1), 565–575.

Henderson, J. (2005). Google Scholar: A source for clinicians? *Canadian Medical Association Journal, 172*(12), 1549–1550.

Hirsch, Jorge E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569–16572.

Jacsó, P. (2008). The pros and cons of computing the h-index using Google Scholar. *Online Information Review, 32*(3), 437–452.

Jin, B. (2006). H-index: an evaluation indicator proposed by scientist. *Science Focus, 1*(1), 8–9. **(in Chinese)**.

Jin, B. (2007). The AR-index: complementing the h-index. *ISSI Newsletter, 3*(1), 6.

Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The R- and AR-indices: Complementing the h-index. *Chinese Science Bulletin, 52*(6), 855–863.

Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original h-index. *ISSI Newsletter, 2*(3), 4–6.

Kosmulski, M. (2007). MAXPROD—A new index for assessment of the scientific output of an individual, and a comparison. *Cybermetrics, 11*(1), 1–5.

Kosmulski, M. (2013). Family-tree of bibliometric indices. *Journal of Informetrics, 7*(2), 313–317.

Liu, Y. & Rousseau, R. (2007). Hirsch-type indices and library management: The case of Tongji University Library. In *11th International Conference of the International Society for Scientrometrics and Informetrics*, June 25–27, (pp. 514–522) Madrid, Spain.

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and Google Scholar. *Journal of the American Society for Information Science and Technology, 58*(13), 2105–2125.

Moed, H. F., Bar-Ilan, J., & Halevi, G. (2016). A new methodology for comparing Google Scholar and scopus. *Journal of Informetrics, 10*(2), 533–551.

Noruzi, A. (2005). Google Scholar: The new generation of citation indexes. *Libri, 55*(4), 170–180.

Panaretos, J., & Malesios, C. (2009). Assessing scientific research performance and impact with single indices. *Scientometrics, 81*(3), 635–670.

Rosenstreich, D., & Wooliscroft, B. (2009). Measuring the impact of accounting journals using Google Scholar and the g-index. *The British Accounting Review, 41*(4), 227–239.

Schreiber, M. (2008). An empirical investigation of the g-index for 26 physicists in comparison with the h-index, the A-index, and the R-index. *Journal of the American Society for Information Science and Technology, 59*(9), 1513–1522.

Schreiber, M. (2010). Twenty Hirsch index variants and other indicators giving more or less preference to highly cited papers. *Annalen der Physik, 522*(8), 536–554.

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

Sun, J., Ma, J., Cheng, X., Liu, Z. & Cao, X. (2013). Finding an expert: A model recommendation system. In *Thirty Fourth International Conference on Information Systems*, Milan.

Tol, R. (2009). The h-index and its alternatives: An application to the 100 most prolific economists. *Scientometrics, 80*(2), 317–324.

Van Raan, A. F. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics, 67*(3), 491–502.

Wildgaard, L., Schneider, J. W., & Larsen, B. (2014). A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics, 10191,* 125–158.

Wu, Q. (2010). The w-index: A measure to assess scientific impact by focusing on widely cited papers. *Journal of the American Society for Information Science and Technology, 61*(3), 609–614.

Wu, Q., & Zhang, P. (2017). Some indices violating the basic domination relation. *Scientometrics, 113*(1), 495–500.

Yan, Z., Wu, Q., & Li, X. (2016). Do Hirsch-type indices behave the same in assessing single publications? An empirical study of 29 bibliometric indicators. *Scientometrics, 109*(3), 1815–1833.