Heliyon 10 (2024) e26235

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

A figure of merit that includes 5 distinct performance indicators to improve research evaluation of academic scholars'

Rayan Mina [a], Farah Homsi [b,*]

[a] Department of Electrical and Mechanical Engineering, Saint-Joseph University of Beirut, Beirut, Lebanon
[b] Department of Civil Engineering and the Environment, Saint-Joseph University of Beirut, Beirut, Lebanon

ARTICLE INFO

Keywords:
Research evaluation
Publications metrics
Research impact
Figure-of-Merit

ABSTRACT

The *h* index has become a widely known indicator to assess the research impact of academic scholars. However, its application has been associated with some criticism regarding its ability to fully capture the quality and significance of an author's research contributions. In this paper, we present a novel approach to improve the evaluation of authors' publications by means of a Figure-of-Merit (FOM) that includes 5 distinct indicators, of which, an enhanced version of the *h* index. Named the Enhanced Research Quality Index (ERQI), it addresses the current limitations of existing solutions and offers a more comprehensive evaluation of research quality. The ERQI builds upon the concept that one metric is never sufficient to capture the performance of an academic scholar, while multiple ones are complex to handle and interpret. The proposed ERQI considers the total number of citations, papers and co-authors and can further differentiate researchers with equal *h* index. By incorporating measurable, and quantitative metrics, ERQI moves away from subjective and indirect factors such as journal reputation, citation context, citation patterns and self-citation righteousness, to offer a more nuanced and accurate representation of research quality. To demonstrate the effectiveness of the proposed metric, we conducted a comparative study using a real dataset of 31 researchers in one of the top 3 engineering faculties in Lebanon, and a randomly generated dataset of 1000 author profiles with >1 million citations. Our findings indicate that ERQI provides a more balanced assessment of research quality by reducing the shortcomings of one indicator. Furthermore, it exhibits a multidimensional effect that captures more efficiently the intrinsic value of scholarly contributions. By adopting ERQI, institutions can make informed decisions that recognize both the quantity and quality of an author's research output and can obtain insightful evaluation enabling fairer recognition of academic scholars' impact and innovation.

1. Introduction

In an era when information is readily accessible, research evaluation is a key mechanism for quality assurance since it provides a means to distinguish between rigorous, well-conducted studies and those lacking in methodological robustness. It is also essential for career development within the academic community. Metrics for research evaluation exist in the literature but have limitations. The main objective of this work is to propose a new method of evaluating research performance based on a novel metric called ERQI, that

* Corresponding author.
E-mail addresses: rayan.mina@usj.edu.lb (R. Mina), farah.homsi@usj.edu.lb (F. Homsi).

https://doi.org/10.1016/j.heliyon.2024.e26235
Received 20 June 2023; Received in revised form 5 February 2024; Accepted 8 February 2024
Available online 10 February 2024
2405-8440/

can differentiate authors more fairly. Furthermore, it helps overcome usual biases in the existing indicators by capturing wider performance aspects of an author's research.

Since its definition in 2005 [1], as the number of significant papers of an author with $\geq h$ citations, the h index was proposed to assess the global interest in a publication. It has been used by institutions worldwide to evaluate academic scholar's performance [2] and in some cases to screen candidates during job interviews [3]. Nevertheless, it has set off debatable reactions as some studies have elaborated on its inconsistency in ranking scientists and proposed alternative metrics such as the highly cited publications indicator [4]. Moreover [5], pointed out drawbacks such as the long-term observation nature of the index and how it fails to assess moderate performance [6]. showed that academic scholars can use the self-citation bias within the h index for rapid career advancement [7–11]. described another limitation and showed how the number of authors contributing to a publication is not considered in the computation of the h value and neither in the total citations.

In the existing literature, many have proposed and demonstrated modifications to compensate for some intrinsic limitations of the h value. For example [7–11], include co-authorship effect in various ways to accurately account for the citations received by a paper with respect to the number of authors [12,13]. studied the self-citation impact on the h index and proposed corrections to mitigate it. In Ref. [14], the number of active years of a researcher has even been used to enhance the evaluation.

In this work we point out and study two additional drawbacks of the existing metrics: (1) authors having equal h values do not necessarily have the same research profile in terms of citations and need to be differentiated. (2) the evaluation should not be based on one single aspect of the author's performance and must include different aspects such as the number of papers produced, and the co-authors patterns.

To cover all the previously cited drawbacks, we propose in this paper a substantially different approach by introducing ERQI, a compact figure composed of five different indicators namely h_e, h_f, CPR, PAR and CAR, each measuring a different aspect of the research quality. These indicators are computed from three directly measurable factors: total number of citations, papers, and cumulative co-authors. Except for h_f, all the proposed indicators are new and address gaps in the current metrics existing in the literature. First, h_e enables ERQI to differentiate researchers having equal h index in a fairer way by computing the number of needed citations to increase h value by 1 point. To capture the co-authorship effect, we incorporated the fractional count of citations per author (h_f, adopted from Ref. [8]). CPR, Citation to Paper Ratio, is included to reduce the bias from producing too many papers with little research interest. PAR, Paper to Author Ratio, is used to measure the productivity of an author disregarding the citations received. Finally, CAR, Citation to Author Ratio, enables differentiating researchers having equivalent h_f values yet different profiles. This work answers the following research question: How to enhance the research evaluation of an author using multiple indicators?

This paper is organized as follows: Section 1 describes the enhancement of the h-index (h_e) for authors with equal h values. Integration of co-authorship effect is detailed in Section 2. New Indicators for research evaluation are proposed and described in Section 3, while Section 4 provides the mathematical foundation behind the proposed compound ERQI and details the advantages as well as motivations of using a multiple-indicator system. Results of the comparative study carried out using a real dataset as well as a much larger randomly generated dataset of profiles are presented in Section 5.

The main contributions of this work are: (1) the definition of a FOM that incorporates multiple indicators condensed into a single value for easier comparison between researchers; (2) the definition of 4 new metrics namely CAR, PAR, CPR and h_e that can capture complementary performance of researchers not measured by the existing metrics in the literature, the latter being an enhanced version of the Hirsch index that differentiates between authors having equal h values.

2. Enhancement of the *h* index for authors with equal *h* values

Researchers having equal h values may conceal different publication profiles, due to the definition of the Hirsch index as an integer number. In fact, it cannot capture how "close" or likely is an author to increasing his/her h value by 1 point thanks to the number of potential coming citations. To demonstrate this limitation in differentiating authors, let us consider a simple example of two researchers whose publications and citations counts are sorted in descending order and listed in Table 1. Although the two authors in Table 1 have 7 published papers, 130 total number of citations and $h = 6$, the last paper of Author-1 needs to be cited 7 times before his/her h index increases.

Table 1
A data showing how two authors having the same h index may differ by how likely they are to reach a higher h value.

# Citations	Author-1	Author-2
Paper-1	53	47
Paper-2	30	30
Paper-3	20	20
Paper-4	10	10
Paper-5	9	9
Paper-6	8	8
Paper-7	0	6
Total citations	130	130
h	6	6
# of citations to increase *h* index by 1	7	1

On the other hand, Author-2 needs only 1 citation and is therefore more likely to reach $h = 7$ faster than Author-1.

This effect can be seen as a noise that conceals the difference in research output between the two. There is indeed useful and impactful information to be taken out within this noise that the h index cannot extract due to its intrinsic definition.

To capture the above effect, we propose a new metric called the distance d, defined as the number of citations N_g needed to reach the next higher h value. In the example of Table 1, we obtain $d_1 = 7$ and $d_2 = 1$. There are several comments that should be stated at this stage to get a better understanding of the distance d. First, theoretically the smallest value corresponds to $d = 1$ and it occurs when a researcher needs just one paper to be cited one time, to increase his/her h value. Second, in the particular case wherein an author has exactly h published papers, the distance d is actually infinite. Third, when several papers have exactly h citations, it means that any of those papers could increase the author's h index if cited just once more.

Table 2 illustrates the above cases for three researchers having $h = 6$. Author-3 has only 6 papers ($d = \infty$), while Author-4 and Author-5 have one and two papers with 6 citations respectively.

Worth noting that it is more likely for Author-5 to reach $h = 7$ before Author-4 does, since any of his/her two papers could be cited once more to increase the h value. This behavior should be reflected in the definition of the metric d, to allow further differentiation between profiles having $N_g = 1$. Therefore, we define the mathematical formulation for the metric d using equation (1):

$$d = \begin{cases} N_g & \mathit{if}\ N_g > 1 \\ \dfrac{1}{N_{ph}} & \mathit{if}\ N_g = 1 \end{cases} \tag{1}$$

where $N_{p,h}$ is the number of papers having h citations exactly. For Author-4 and Author-5 in Table 2, $N_{ph} = 1$ and $N_{ph} = 2$ respectively which leads to $d_4 = 1$ and $d_5 = \frac{1}{2}$.

From equation (1), it is trivial that the lower the distance the better the research profile for academic scholars having equal h values. To include the distance effect, we propose the enhanced h index, h_e, using equation (2):

$$h_e = h + exp^{-d} \tag{2}$$

h_e, is a decimal number lying between the current h value of an author and $h + 1$. For example, if $h = 5$ then $5 \le h_e < 6$. For ease of visibility, the enhanced h index is rounded to 2 significant digits only. Worth noting that the smaller the distance the higher the h_e value, and for very high distances h_e is very close to h (a lot of citations are required before the h value increases) as shown in Fig. 1.

3. The fractional h-index: co-authorship effect

The problem of estimating the individual contributions of published papers having multiple authors has been studied before [8,11,15]. It is obvious that counting the number of citations and disregarding the number of authors is somewhat unfair to researchers who publish single contribution papers. Two main reasons support this claim.

(1) the contribution to the work presented in any publication is certainly divided between all authors and this should be reflected somehow in the research evaluation.
(2) The reach of a multi-authored paper increases thanks to the wider network of connections of the contributors as well as to the higher advertising made on social platforms (e.g., ResearchGate sends invitations to researchers to confirm authorship).

[16] suggested an equal fraction of c/m for an m-authored paper that received c citations, leading to a decimal number. Other methods propose to include effective shares in the fraction calculation when they are disclosed in the text [11], which has the limitation of being sometimes inaccessible.

In this paper, we adopted the fractional h index, h_f, as proposed by Ref. [16] and studied by Ref. [8] to enhance the evaluation of the

Table 2
Data showing how the distance d is computed for 3 authors with similar profiles.

# Citations	Author-3	Author-4	Author-5
Paper-1	53	47	41
Paper-2	30	30	30
Paper-3	20	20	20
Paper-4	10	10	10
Paper-5	9	9	9
Paper-6	8	8	8
Paper-7		6	6
Paper-8			6
Total citations	130	130	130
h	6	6	6
# of citations N_g to increase h index by 1	∞	1	1
# of papers N_{ph} having 6 ($= h$) citations	0	1	2
Distance d	∞	1	½

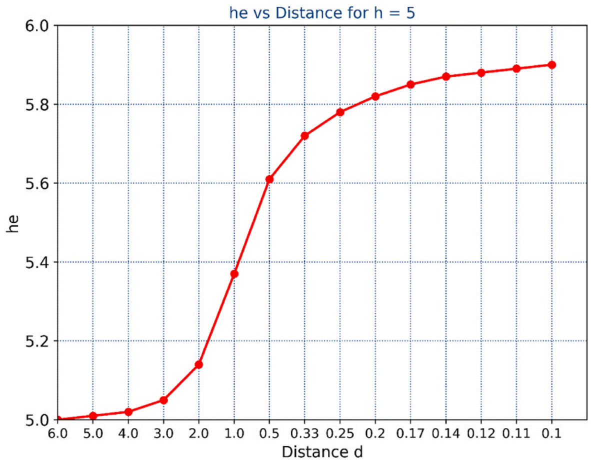

Fig. 1. Plot of the enhanced h index h_e as a function of the distance d, for $h = 5$.

research. The computation of h_f follows the same algorithm used for h except that it is applied on fractions of citations over authors, instead of citations. Table 3 shows an example of how to compute h_f for two authors who published an equal number of papers, have the same h index and received an equal total number of citations. Author-6 has all paper fractions below 2, leading to $h_f = 1$, while Author-7 fractions yield $h_f = 3$. Clearly the fractional h index have captured a non-negligible difference between both profiles, which are equivalent by means of other indicators. Worth noting that h_f considers the co-authorship effect only and cannot be used to draw more conclusions.

Moreover, we show in the next section, that in some scenarios it has blind spots and should be complemented with another indicator to accurately capture the co-authorship impact.

4. New indicators for research evaluation

When studying a complex system or process, a quantitative indicator is defined to evaluate a specific performance. Usually, several indicators are needed to describe the global performance and to cover the complexity of the system under study. Applying this concept to research publications, one can state that the total number of citations is just a measure of the global volume of research interests that exist for the work of an author. On the other hand, the h index is a more elaborate indicator used mainly to assess if an author is well cited across all his/her publications simultaneously. It was originally defined to identify potential scenarios where one researcher's paper is heavily cited while his/her other papers are under cited. For this reason, they are both used in scientific social platforms (e.g., Google Scholar, ResearchGate) to obtain a wider angle of evaluation.

In the previous sections, we have shown that two additional indicators should be considered to overcome two individual oversights of the h index: the co-authorship effect (h_f existing in literature) and the inability to differentiate authors with equal h values (h_e proposed in this work). In the same logic, we will define in the coming paragraphs additional relevant indicators to capture other important performance aspects of a researcher's publications.

We will construct three examples of authors' profiles to represent scenarios in which the h index and the total number of citations lead to somehow narrow conclusions. For each case, a new indicator will be proposed to describe the associated performance being captured.

In Table 4 two authors possess equivalent research profiles with equal total number of citations, number of published papers, as well as equal h, h_f and h_e values. Nevertheless, Researcher-1 co-authored 9 papers with 2 other authors only, while Researcher-2 shares his/her papers with 9 co-authors, thereby making a smaller contribution in those cases. Even the h_f index could not perceive this

Table 3
The fractional h index computed for two authors having identical profiles in terms of h values, total number of citations and number of papers.

#	Author-6		Author-7	
	Citations/Authors	Fraction	Citations/Authors	Fraction
Paper-1	10/8	1.25	10/3	3.33
Paper-2	7/4	1.75	7/2	3.5
Paper-3	6/4	1.5	6/2	3
Paper-4	5/3	1.67	5/3	1.67
h	4		4	
Total Citations	28		28	
h_f	1		3	

Table 4
The *CAR* indicator differentiates two authors having apparent identical profiles.

	Researcher-1			Researcher-2		
	Citations	Authors	Fraction	Citations	Authors	Fraction
Paper-1	5	5	1	5	5	1
Paper-2	5	3	1.67	5	10	0.5
Paper-3	5	3	1.67	5	10	0.5
Paper-4	5	3	1.67	5	10	0.5
Paper-5	5	3	1.67	5	10	0.5
Paper-6	5	3	1.67	5	10	0.5
Paper-7	5	3	1.67	5	10	0.5
Paper-8	5	3	1.67	5	10	0.5
Paper-9	5	3	1.67	5	10	0.5
Paper-10	5	3	1.67	5	10	0.5
h	5			5		
Total Citations	50			50		
Total Papers	10			10		
Cumulative Authors	32			95		
h_f	1			1		
CAR	1.56			0.53		

difference in the profiles making it slightly unfair to Researcher-1. Therefore, we propose a new indicator, the Citation to Author Ratio *CAR* that evaluates such scenarios more fairly using equation (3):

$$CAR = \frac{N_c}{N_a} \tag{3}$$

where N_c is the total number of citations and N_a is the cumulative number of authors across all published papers. As can be seen from Table 4, for Reseacher-1 $CAR = 1.56$ and for Researcher-2 $CAR = 0.53$ due to his/her bigger number of co-authors.

In Table 5 two different research profiles are presented with Researcher-1 having 20 published papers compared to just 10 for Researcher-2. However, they have equal total number of citations, total cumulative authors, h, and h_f values. Moreover, the *CAR* indicator is $118/29 = 4.07$ for both. Nevertheless, Researcher-1 has managed to produce 10 single-authored papers that are very much under-cited (1 each) and hence can be considered as low-impact contributions, which virtually inflated his/her total number of citations. None of the previous metrics were able to detect this pattern, therefore we propose a new indicator in this work, the Citation to Paper Ratio *CPR*, that evaluates such scenarios more fairly using equation (4):

$$CPR = \sqrt{\frac{N_c}{N_p}} \tag{4}$$

where N_c is the total number of citations and N_p is the total number of published papers for an author. As can be seen from Table 5,

Table 5
The CPR indicator accounts for an author having published a lot of under cited papers.

	Researcher-1			Researcher-2		
	Citations	Authors	C/A	Citations	Authors	C/A
Paper-1	18	5	3.0	28	7	2.8
Paper-2	10	5	0.67	10	6	0.4
Paper-3	10	2	2.0	10	2	2.0
Paper-4	10	1	1.0	10	2	2.0
Paper-5	10	1	1.0	10	2	2.0
Paper-6	10	1	1.0	10	2	2.0
Paper-7	10	1	1.0	10	2	2.0
Paper-8	10	1	1.0	10	2	2.0
Paper-9	10	1	1.0	10	2	2.0
Paper-10	10	1	1.0	10	2	2.0
Paper-11	1	1	1.0			
...	…	…	…			
Paper-20	1	1	1.0			
h	10			10		
Total Citations	118			118		
Total Papers	20			10		
Cumulative Authors	29			29		
h_f	3			3		
h_e	10.00			10.00		
CPR	2.43			3.44		

Reseacher-1 has a low *CPR* of $\sqrt{118/20} = 2.43$ while Researcher-2 have a higher *CPR* of $\sqrt{118/10} = 3.44$ thanks to achieving same number of citations in less papers. The previous example shows how an author could easily produce too many papers with little research interest from the scientific community, a certain bias in the evaluation of the quality that should be accounted for and measured somehow using the proposed CPR indicator.

The square root is added in equation (4) to avoid rapid growth of the ratio N_c/N_p due to two main reasons: First, N_c is naturally much higher than N_p; Second, N_c/N_p can get rapidly high in some "lucky" scenarios when just one paper is excessively cited (e.g., 1 paper is cited 200 times and 9 others are cited 10 times leads to a ratio of 29, almost 3 times bigger than 10 corresponding to the case when all papers are equally cited 10 times).

In Table 6 two equivalent research profiles are given with equal h, h_f, *CAR* and *CPR* values. Nevertheless, Researcher-2 has produced 4 more papers, which are moderately cited, than Researcher-1. Data in Table 6 shows that this additional contribution is not reflected in any of the previous indicators, therefore we propose a final indicator in this work, the Paper to Author Ratio *PAR*, that evaluates such scenarios more fairly using equation (5):

$$PAR = \frac{N_p}{\sqrt{N_a}} \tag{5}$$

where N_a is the cumulative number of authors across all papers and N_p is the total number of papers. The square root term $\sqrt{N_a}$ has been added to reward authors having equal ratios N_p/N_a but higher N_p. As can be seen from Table 6, Reseacher-1 has a $PAR = 1.72$ while Researcher-2 has a higher $PAR = 2.41$ since he/she published more relevant papers.

Finally, it is worth noting that the *PAR* is the only indicator that is independent of the citations received or to be received by a paper. It cannot change with time unless the author produces a new paper, and in that logic, it represents a significantly different performance measure for research excellence.

5. ERQI: a figure-of-merit for research evaluation

The previous sections showed how important it is to capture every single performance aspect by means of the five indicators: h_e, h_f, *CAR*, *CPR* and *PAR*. At the same time, taking all previous metrics as a set to evaluate the research quality of an academic scholar is somewhat difficult to handle.

Humans and institutions often look for more concise and unified representation of the overall performance to enable easy comparison and facilitate decision-making. Condensing various indicators into a single value called Figure of Merit (FOM) simplifies the analysis and enhances clarity in the research assessment. It represents an aggregated view and has been adopted widely as a useful tool for summarizing the performance of engineering systems [17–19].

At this stage, it is interesting to evaluate the relations between the proposed 5 indicators. For this purpose, a random data set of 1000 researcher profiles was generated and the corresponding indicators calculated using a computer program that we developed specifically for this purpose. The Pearson correlation coefficients between each pair of indicators of the random data set are presented in a symmetrical 5×5 matrix in Table 7.

The results in Table 7 show that 6 out of 10 coefficients are < 0.5, meaning that the corresponding indicators are slightly correlated. The high correlation (0.96) between h_f and h_e (suggesting that these parameters vary simultaneously), is anticipated since $h_e = h + e^{-d}$ and h_f is computed using the same algorithm as h but applied to fractions of citations to authors for each paper.

Moreover, the high correlation between *PAR* and the two indicators h_e and h_f (0.98 & 0.96) is explained by the fact that *PAR* is

Table 6
The PAR indicator differentiates two authors having apparent identical profiles.

	Researcher-1			Researcher-2		
	Citations	Authors	C/A	Citations	Authors	C/A
Paper-1	4	3	1.0	4	3	1.0
Paper-2	4	3	1.0	4	3	1.0
Paper-3	4	3	1.0	4	3	1.0
Paper-4	4	3	1.0	4	3	1.0
Paper-5				4	3	1.0
Paper-6				4	3	1.0
Paper-7				4	3	1.0
Paper-8				4	3	1.0
h	4			4		
Total Citations	16			32		
Total Papers	4			8		
Cumulative Authors	12			24		
h_f	1			1		
h_e	4.02			4.02		
CAR	1.33			1.33		
CPR	4.00			4.00		
PAR	1.15			1.63		

Table 7
The correlation matrix of the 5 indicators.

	PAR	CAR	CPR	h_f	h_e
PAR		0.26	0.35	0.96	0.98
CAR	0.26		0.83	0.37	0.29
CPR	0.35	0.83		0.46	0.40
h_f	0.96	0.37	0.46		0.96
h_e	0.98	0.29	0.40	0.96	

proportional to the number of papers, which is likely to make the number of citations increase (more papers mean more citation chances).

Having observed the above relations between the 5 indicators, the proposition of a multilinear equation for the FOM is not optimal as some of them are linearly correlated. On the other hand, it was proven in section 3 that each indictor has an important contribution to differentiate between researcher profiles and must not be disregarded.

For this reason, ERQI is proposed in this work as the geometric mean of the previous 5 indicators, to allow capturing diversified aspects of research quality.

Equation (6) shows the mathematical expression of ERQI:

$$ERQI = \sqrt[5]{1 + h_f \times h_e \times CAR \times PAR \times CPR} \tag{6}$$

In equation (6) we have in fact a mathematical product of the proposed indicators (i.e., non-linear function), which enables each one to have a relevant contribution in the final value, even though some are linearly correlated. In addition, since each variable captures a distinct performance, the geometric mean formula prohibits one high value of any indicator from excessively "inflating" or growing the ERQI. Using this definition, the ERQI can capture a much wider range of performance than what each indicator is able to do while being efficient in balancing between them to avoid misleading inflations: for example, if one or two parameters are high compared to the others, ERQI will reduce their effect in the final condensed value. Most importantly, if a researcher is good enough at all levels, then all five indicators will be high and ERQI will increase steadily (as it should). Our proposed single value FOM provides an incentive for academic scholars to increase their performance at all levels equally and thereby helps in reducing the biasing effect of increasing one indicator while neglecting the others. Finally, a factor 1 is added inside the root formula to avoid nulling the result for cases where $h_f = 0$. Worth noting that ERQI formulation in Equation (6) is highly flexible and allows to add any relevant new indicator capturing additional future performance aspect.

6. Study of the proposed ERQI

To demonstrate the efficiency and usefulness of the proposed ERQI as well as to show how ERQI differs from h index quantitatively, we must compute them on data coming from diversified researchers' profiles. The optimal methodology would be to collect all real academic scholars' public profiles available on the internet and compare the corresponding distributions of ERQI and h index values. Unfortunately, this operation either requires a lot of time and effort to complete data collection from the web or it requires a costly paid account to obtain access to existing databases. To overcome this, we decided to conduct a study on two distinct sample datasets: First, on real data corresponding to authentic academic researchers. Second, on a much larger dataset that is randomly generated using a computer script to cover more diversified research profiles in a time-efficient way. The random generation is very fast, and the size of the data can be enlarged as much as we need (for better representation) at very little processing cost.

6.1. The sampling methodology

The real dataset was collected manually from the official Google scholar profiles (public source of information) of the academic authors affiliated in the engineering faculty of one of the top 3 Lebanese universities according to Ref. [20] (from this point onward referred to as Univ-1). Although the real data is just a sample composed of 31 researchers only, it has an acceptable statistical representation covering both females and males, different authors' titles, as well as several engineering domains (see Table 8). The real data comprise 1000 publications and 11000 citations in total and is available publicly at https://github.com/RayanMinaESIB/ERQI_

Table 8
Statistical information pertinent to the real dataset used in this work.

Gender	Female	Male			
	33%	67%			
Professor Title	**Assistant**	**Associate**	**Full Professor**		
	35%	23%	42%		
Research Field	**Computer**	**Civil**	**Mechanical**	**Electrical**	**Chemical**
	39%	26%	9%	20%	6%

codes/blob/main/Real_Dataset.xlsx.

An example of this real dataset is given in Table 9 showing details of the researcher profile: total number of papers, cumulative number of co-authors, total number of citations across all publications, and a list of the number of citations per paper. Also present are the computation results for ERQI, *h* index, the distance *d*, as well as all the 5 indicators composing ERQI.

To enlarge the sample, we generated 1000 random authors profiles by a computer script in Python that we developed specifically for this purpose. Each generated profile contains 3 random elements: (1) an integer N_{pap} representing the number of published papers for an author, (2) a list of integers of size N_{pap} representing the number of citations each paper received, and (3) a second list of integers of size N_{pap} also representing the number of authors for each publication. The random generation was realized using 'numpy.random' library in Python according to simple bounds stated in Table 10 and according to some basic common-sense rules. Uniform sampling is not appropriate for the random generation of number of citations per paper and for the number of papers per author. In fact, only a few researchers possess a significant or a small number of publications, and for any author only a few papers receive a high number of citations. Therefore, we have manually included two probability distributions that take into consideration these trends (Table 11). For example, there is a low probability of 0.125 for generating profiles having <10 or >100 papers, and there is also a low probability of 0.15 for a publication to be cited >60 times. Uniform sampling was only used for the generation of the number of authors within a paper. In total the random dataset contains 53,000 papers and 1,100,000 citations and the Python script can be found at https://github.com/RayanMinaESIB/ERQI_codes/blob/main/random_dataset.py.

To make sure that the random dataset is statistically representative of the researchers' community worldwide, which is estimated in Ref. [21] to approximately 20 million authors, we must compute the minimum required sample size for this population. We have assumed a 95% sampling confidence level and a 97% confidence interval to perform sample size computation. According to Ref. [22] these assumptions lead to a sample of 1068 author profiles, which validates our random dataset size. Table 12 summarizes the statistical assumptions used in this study. The complete dataset is publicly available at https://github.com/RayanMinaESIB/ERQI_codes/blob/main/Random_Dataset.xlsx to provide the readers with all necessary information and tools to reproduce the study.

6.2. Results and analysis

The results computed from the previous two datasets are presented in this paragraph to illustrate how ERQI differs from conventional indicators (especially the *h* index) in evaluating more accurately researchers' profiles. The studied researchers of the Univ-1 are all indexed in Google Scholar and their updated profiles have been retrieved online on May 30th' 2023, as the information is publicly available. All authors were sorted using Python's 'numpy.sort' library in decreasing order of their *h* index as shown in the red curve in the left plot of Fig. 2. We computed, for each one, the 5 indicators (i.e., h_e, h_f, CAR, PAR, CPR) as well as the distance *d* to evaluate the individual ERQI. Finally, the results were sorted again, this time in decreasing order of the ERQI as shown in the blue curve in the left plot of Fig. 2. The same procedure is repeated for the 1000 randomly generated researchers' profiles as shown in the right plot of the same figure. All plots were created using 'matplotlib.pyplot' library in Python within the same script used to generate the random data. Each plot represents a vector of data that is included in this script.

Before comparing ERQI and *h* index results in Fig. 2, it is important to clarify one point related to the study methodology. ERQI is based on the use of 5 indicators that are merged using the geometric mean equation, which is a deterministic and hence a non-probabilistic formula. Statistical hypothesis testing is not genuinely applicable in our case since the outcomes of ERQI equation are in fact mathematically predictable. The geometric mean is widely used in science to handle cases where the impact of extreme values needs to be minimized. In our study, if an author has one indicator that is extremely large then the geometric mean formula will prevent ERQI from being disproportionately high. Each indicator within ERQI captures one aspect of a researcher's performance in contrast with *h* index that grows fast with the number of citations only.

Going back to the results of Fig. 2, the ERQI values are bound to 9 and 12 while the *h* values go as high as 29 and 50 for Univ-1 and random datasets respectively. The magnitude distribution of authors' ERQI values is more compressed than that of the *h* index,

Table 9
Statistical information pertinent to one example of the real dataset used.

	Example
Researcher Gender	Male
Researcher Title	Assistant Professor
Field of research	Electrical Engineering
Total number of papers	14
Cumulative number of co-authors	61
Total number of citations	125
Citation List per paper	[0,0,0,0,2–5,5,9,12,12,25,48]
PAR	1.79
CAR	2.05
CPR	8.93
h	5
h_f	3
Distance d	0.5
h_e	5.61
ERQI	3.54

Table 10
Random generation conditions for the dataset.

.	Min Value	Max Value
# of papers N_{pap}	1	200
# of citations per papers	0	100
# of co-authors per paper	1	10

Table 11
Random generation conditions for the dataset.

.	Range	Probability
# of papers N_{pap}	1–10	0.125
	10–100	0.75
	100–200	0.125
# of citations per papers	0–20	0.5
	20–60	0.35
	60–100	0.15

Table 12
Assumptions made for the sample size of the random dataset.

Population Size	Confidence Level (%)	Confidence Level (%)	Sample Size
20,000,000	95	97	1068

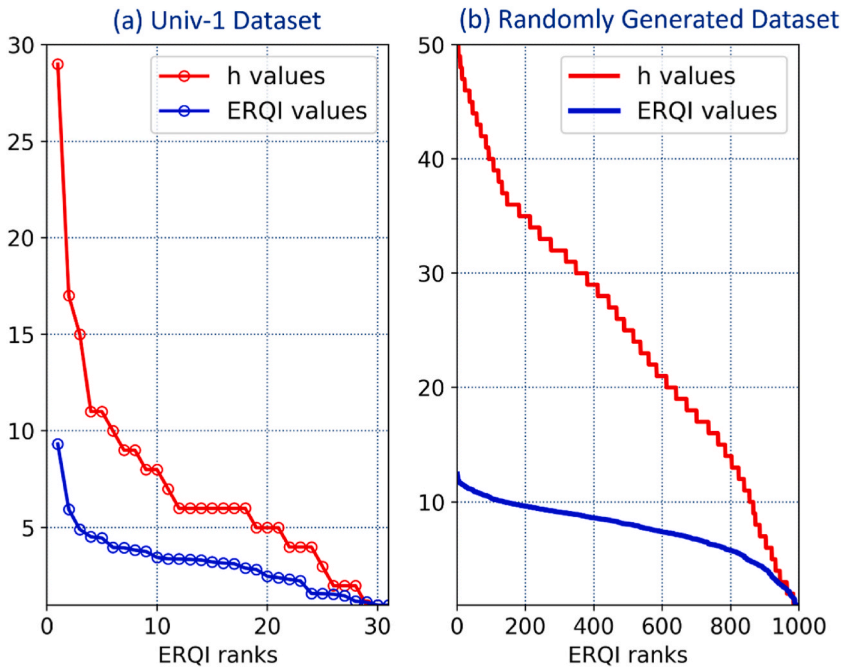

Fig. 2. Plot of ERQI and *h* values for (a) Univ-1 Dataset and (b) Randomly Generated Dataset.

Table 13
Standard deviation and mean of ERQI and *h* indices.

	Mean/Std. Dev.	
	Univ-1 Example	**1000 Random samples**
h	6.9/5.6	24.4/12.0
ERQI	3.3/1.6	7.6/2.4

meaning that the latter significantly boosts the overall performance value while basing the whole evaluation on a single performance only. This leads to shortcomings in evaluating the quality of research appropriately. On the other hand, ERQI is difficult to grow excessively since there are 5 distinct indicators that need to increase simultaneously to boost its value. Furthermore, to better understand how ERQI differs from and surpasses h index in terms of central tendency and variability, we computed their mean and their standard deviation values in both datasets using Equation (7) and Equation (8) respectively. The results are given in Table 13. The much lower mean value of ERQI with respect to h (3 times lower) confirms how it helps reducing excessive performance growth. Moreover, its lower standard deviation further supports this conclusion by exhibiting a close dispersion of the results around the mean value. Finally, to offer a direct visual and descriptive representation of the spread of results, boxplots of h and ERQI values are added in Fig. 3. The plots corroborate the previous findings and show consistency in the spread with very little outliers or extreme values.

$$\mu = \frac{\sum_{k=1}^{M} X_k}{M} \tag{7}$$

$$\sigma = \sqrt{\frac{\sum_{k=1}^{M} (X_k - \mu)^2}{M}} \tag{8}$$

In addition, both plots in Fig. 2 show the existence of too many plateaus in the red curves (stairs-like pattern of the h values) meaning that a lot of authors are likely to have equal h performance while effectively having different publications profiles. In fact, 70% of samples share the same h value with at least another author. This creates a kind of ambiguity when comparing researchers, a serious disadvantage that is eliminated by the gradual progress of ERQI values shown in the blue curves of Fig. 2 plots. In fact, our proposed FOM eliminates ambiguity by exhibiting a smooth evolution of the values thanks to two factors: (1) the presence of h_e which intrinsically differentiates authors having equal h indices; (2) the compound definition of ERQI with 5 distinct indicators.

In the red plots of Fig. 4, h ranks are plotted as a function of ERQI ranks for both datasets and the results are sorted in descending order (along the x-axis the first value is 1 and represents the highest ERQI value or rank = 1). A blue line representing the curve $y = x$ allows to see how both indices differ in ranking the authors. A peak in the red curve above or below the blue one means that the h rank of this researcher is much higher or lower than his/her ERQI rank. Results in Fig. 4 show that for both datasets 75% of samples have different ranks between ERQI and h (reaching sometimes 200) demonstrating thereby how our proposed method captures alternative performance while evaluating research excellence. Worth noting that for the highest and lowest ranks, both indices are consistent in capturing poor or excellent performance and give equivalent ranks as shown on the blue and red curves at the starting and ending points in the x-axis.

Fig. 4 shows both the benefits and relevance of replacing h index by the ERQI for research evaluation. Once again, this benefit comes from the FOM definition, which takes into consideration complementary indicators that are missing in the h index formulation. To the best of our knowledge, a compound FOM to measure research quality has been used only once in literature in Ref. [23]. However, several points distinguish the approach in this work from the latter. First, the authors have used the h index itself while here we proposed an enhanced version, h_e. Second, they have included the number of citations to papers as a single author, as single or first author and as single, first or last author as 3 different indicators. There is some redundant information in those 3 variables and are more prone to subjectivity in ranking author's contribution compared to our approach in which all indicators are measurable and quantitative metrics. Third, the co-authorship effect in Ref. [23] were included using the h_m index from Ref. [7] while we used the formulation studied by Ref. [8].

Finally, to illustrate best the effectiveness of ERQI compared to h index let us take a couple of author profile examples. Profile-1 corresponds to an author who works systematically with a cluster of 10 authors, while Profile-2 has more reasonable co-authorship behavior. Both have published 20 papers, have the same citations list per paper and equal h index of 11. The detailed computations and results are given in Table 14 where one can clearly see how *PAR* and *CAR* indicators are low for Profile-1 thus leading to low ERQI value of 3.23. On the other hand, they are moderate to high for Profile-2 leading to a much higher ERQI score of 6.53. This is a typical example of how ERQI detects patterns of author clusters (large group of researchers who systematically publish their work with their mutual names) and moderates the overall score of such researchers.

6.3. Implications of the study and recommendations

The question of research evaluation is crucial to authors obviously, but also to academic institutions as well as the funding agencies for research projects. Based on the findings of this study, ERQI's capability of evaluating different performance aspects simultaneously offers the previous stakeholders an interesting complement to the h index. There are two major positive implications of the current study that are worth pointing out:

(1) In some cases, significant clusters of authors may agree to publish together systematically by adding each other's names on all their papers, thereby multiplying meaninglessly their number of cumulated citations. The h index is uncapable of detecting such scenarios and will praise those authors individually, however ERQI can easily identify those cases (through *CAR*, *PAR* and h_f indicators) and will heavily penalize the score of such authors. Therefore, authors clusters will become an unappealing option

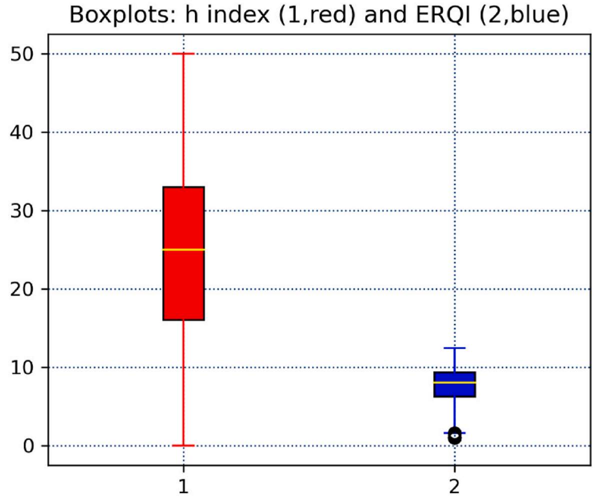

Fig. 3. Boxplots of ERQI and h values for the random Dataset.

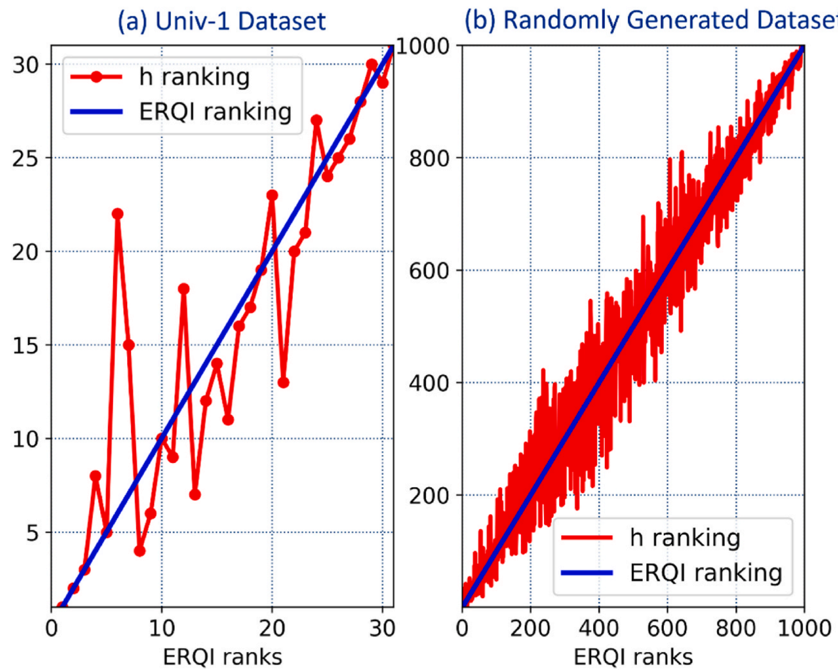

Fig. 4. Plot of authors h ranking vs ERQI ranking for (a) Univ-1 Dataset and (b) Randomly Generated Dataset.

for most academic scholars especially if the funding agencies and universities adopted ERQI as a serious complement to the h index when evaluating research quality.

(2) In some institutions, producing more research papers is by itself a criterion for career advancements disregarding the interest of the community in those publications (number of citations received). While h index remains constant in such cases (since it cannot decrease) ERQI detects such strange patterns through the *CPR* indicator that will decrease rapidly, thereby reducing the overall score of an author. For academic institutions, adding ERQI to their research evaluation can benefit them by making sure that authors who produce quality papers only are being considered for promotion.

6.4. Limitations of the study and possible enhancements

The computation of ERQI is based on several indicators that are related to the published papers and number of citations. Researchers tend to cite publications in indexed journals more frequently than in local or non-indexed journals and conferences. This is a

Table 14
Computation of ERQI and h index on two different author profiles.

	Profile-1
Citations list	[5,5,5,5,5,5,5,5,10,15,15,20,20,20,25,25,30,30,50,50]
Authors list	[10,10,10,10,10,10,10,10,10,10,10,10,10,10,10,10,10,10,10,10]
Fractions list	[5,5,3,3,2.5,2.5,2,2,2,1.5,1.5,1,0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5]
h	11
h_f	3.0
h_e	11.37
CAR	$= 350/200 = 1.75$
PAR	$= 20/\sqrt{200} = 1.41$
CPR	$= \sqrt{350}/\sqrt{200} = 4.18$
ERQI	$= \sqrt[5]{1 + 3.0 \times 11.37 \times 1.75 \times 1.41 \times 4.18} = 3.23$
	Profile-2
Citations list	[5,5,5,5,5,5,5,5,10,15,15,20,20,20,25,25,30,30,50,50]
Authors list	[2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2]
Fractions list	[25,25,15,15,12.5,12.5,10,10,10,7.5,7.5,5,5,2.5,2.5,2.5,2.5,2.5,2.5,2.5]
h	11
h_f	3.0
h_e	11.37
CAR	$= 350/40 = 8.75$
PAR	$= 20/\sqrt{40} = 3.16$
CPR	$= \sqrt{350}/\sqrt{40} = 4.18$
ERQI	$= \sqrt[5]{1 + 3.0 \times 11.37 \times 8.75 \times 3.16 \times 4.18} = 6.53$

limitation of the proposed study since this affects the computations of *CAR*, *CPR*, h_e, h_f and consequently ERQI value for an author. Nevertheless, the geometric mean equation allows to add more indicators and hence this aspect could be included as a future enhancement for ERQI. Pre-publications are not included in any indicator composing ERQI and constitute another minor limitation of our study. Moreover, an additional shortcoming is that the self-citation bias within h index studied in Ref. [6] remains an issue in ERQI and needs to be mitigated in future enhancements. Finally, it could be interesting for future research directions in this field to include in the mathematical definition of the FOM in Equation (6) an indicator that can differentiate the order in which authors are listed within the publication. This could be a valuable enhancement for future work because usually the contribution is not equally split amongst authors.

7. Conclusion

In this paper, we have proposed a novel figure-of-merit as a potential complement of the h index to evaluate research performance for academic scholars. By including 5 distinct indicators, each one formulated to capture a certain aspect of the author's performance, our proposed ERQI offers a more robust evaluation and addresses the main limitations of the h index: the rapid and excessive growth of h values, and the inability of differentiating authors having different research profiles. We have demonstrated the superior performance of our proposed ERQI over the traditional h index in providing a more nuanced and accurate representation of an individual's research achievements. A comparative study conducted over one small real dataset from one of the top 3 Lebanese university and another much larger randomly generated dataset showed good mutual agreement in the conclusions and how ERQI offers a wider angle in capturing the research performance. While our work presents relevant evidence in favor of the proposed FOM, further research and validation across a broader range of disciplines and datasets would be beneficial to validate the approach.

CRediT authorship contribution statement

Rayan Mina: Conceptualization, Formal analysis, Software, Writing – original draft, Writing – review & editing. **Farah Homsi:** Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J.E. Hirsch, An index to quantify an individual's scientific research output, Proc. Natl. Acad. Sci. USA 102 (46) (2005) 16569–16572, https://doi.org/10.1073/pnas.0507655102.
[2] G. Marini, New promotion patterns in Italian universities: less seniority and more productivity? Data from ASN, High Educ. 1–17 (2016), https://doi.org/10.1007/s10734-016-0008-x.
[3] I. Reymert, Bibliometrics in academic recruitment: a screening tool rather than a game changer, Minerva 59 (2021) 53–78, https://doi.org/10.1007/s11024-020-09419-0.

[4] Ludo Waltman, Nees Jan van Eck, The inconsistency of the h-index, J. Am. Soc. Inf. Sci. Technol. 63 (2) (2011) 406–415, https://doi.org/10.1002/asi.21678.
[5] T. Braun, W. Glanzel, A. Schubert, A Hirsch-type index for journals, Scientometrics 69 (2006) 169–173, https://doi.org/10.1007/s11192-006-0147-4.
[6] Marco Seeber, Mattia Cattaneo, Michele Meoli, Paolo Malighetti, Self-citations as strategic response to the use of metrics for career decisions, Res. Pol. 48 (2) (2019) 478–491, https://doi.org/10.1016/j.respol.2017.12.004.
[7] M. Schreiber, A modification of the h-index: the hm-index accounts for multi-authored manuscripts, J. Informetrics 2 (3) (2008) 211–216, https://doi.org/10.1016/j.joi.2008.05.001.
[8] L. Egghe, Mathematical theory of the h- and g-index in case of fractional counting of authorship, J. Am. Soc. Inf. Sci. Technol. 59 (2008), https://doi.org/10.1002/asi.20845.
[9] J.E. Hirsch, An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship, Scientometrics 85 (2010) 741–754, https://doi.org/10.1007/s11192-010-0193-9.
[10] Denis Bouyssou, Thierry Marchant, Ranking authors using fractional counting of citations: an axiomatic approach, J. of Informetrics 10 1 (2016) 183–199, https://doi.org/10.1016/j.joi.2015.12.006.
[11] H. Sauermann, C. Haeussler, Authorship and contribution disclosures, Sci. Adv. 3 (2017) 11, https://doi.org/10.1126/sciadv.1700404.
[12] M. Schreiber, Self-citation corrections for the Hirsch index, Europhys. Lett. Assoc. Europhys. Letters 78 (3) (2007), https://doi.org/10.1209/0295-5075/78/30002.
[13] J.H. Fowler, D.W. Aksnes, Does self-citation pay? Scientometrics 72 (2007) 427–437, https://doi.org/10.1007/s11192-007-1777-2.
[14] Kaptay George, The k-index is introduced to replace the h-index to evaluate better the scientific excellence of individuals, Heliyon 6 7 (2020), https://doi.org/10.1016/j.heliyon.2020.e04415.
[15] J.K. Wan, P.H. Hua, R. Rousseau, The pure h-index: calculating an authors hindex by taking co-authors into account, COLLNET J. Sci. Inf. Manag. 1 (2) (2007) 1–5, https://doi.org/10.1080/09737766.2007.10700824.
[16] Q. Burrell, R. Rousseau, Fractional counts for authorship attribution: a numerical study, J. Am. Soc. Inf. Sci. Technol. 46 (1995) 97–102, https://doi.org/10.1002/(SICI)1097-4571(199503)46:2<97::AID-ASI3>3E3.0.CO;2-L.
[17] Alejandro C. Olivieri, Graciela M. Escandar, Chapter 6 - Analytical Figures of Merit, Practical Three-Way Calibration, Elsevier, 2014, pp. 93–107, https://doi.org/10.1016/B978-0-12-410408-2.00006-5.
[18] R.H. Walden, Analog-to-digital converter survey and analysis, IEEE J. Sel. Area. Commun. 17 (4) (1999) 539–550, https://doi.org/10.1109/49.761034.
[19] Rayan Mina, Youmni Ziadeh, Towards an optimal solution for environmental protection using wireless sensor networks, in: 3rd IEEE International Conf. On Electrical, Electronics, Comput. Eng. and Their Appl, EECEA, 2016, pp. 120–125, https://doi.org/10.1109/EECEA.2016.7470777.
[20] https://www.universityrankings.ch/results/QS/202?ranking=QS&year=202&region=&q=Lebanon.
[21] T. Ryan, Sample Size Determination and Power, John Wiley & Sons, Inc, 2013, pp. 17–55, https://doi.org/10.1002/9781118439241.ch2.
[22] Research gate https://www.researchgate.net/about accessed on September 9th 2023.
[23] J.P.A. Ioannidis, J. Baas, R. Klavans, K.W. Boyack, A standardized citation metrics author database annotated for scientific field, PLoS Biol. 17 (8) (2019) e3000384. https://www.researchgate.net/about?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6ImluZGV4IiwicGFnZSI6ImluZGV4In19.