2019-1023 IJOI
http://www.ijoi-online.org/

THE EXPONENTIAL AUTHOR-WEIGHTED SCHEME USED TO EVALUATE AUTHOR-/COUNTRY-LEVEL RESEARCH ACHIEVEMENTS FOR A SCHOLAR JOURNAL: A BIBLIOMETRIC ANALYSIS

Tung-Hui Jen
Department of Chinese Medicine, Chi-Mei Medical Center, Taiwan
renin9999@gmail.com

Tsair-Wei Chien
Research Department, Chi-Mei Medical Center, Tainan, Taiwan
smile@mail.chimei.org.tw

Willy Chou*
Department of Recreation and Health-Care Management & Institute of recreation Industry Management, Chia Nan University of Pharmacy, Tainan, Taiwan
and Department of Physical Rehabilitation, Chi-Mei Medical Center, Taiwan
*Corresponding Author; ufan0101@ms22.hinet.net

Abstract

Many author-level metrics assume that all co-authors have an equal contribution to an article, which is unfair, unreasonable, and problematic. Numerous author-weighted schemes were proposed in the literature. We proposed the exponential author-weighted scheme for quantifying coauthors' credits. By searching the PubMed database, we used the keywords "Scientometrics" [Journal], and downloaded 203 articles published since 2010. These citable articles were then individually searched for citations in Pubmed Central (PMC), and the citations were analyzed for individual authors and countries/areas. We observed that (1) the top three countries with highest x-index are Netherlands, Denmark, and Germany on the choropleth map,, (2) Netherlands is evident of the citation-oriented using the Kano model to display their research

achievements, and (3) the author Nees Jan van Eck has the highest x-index in Scientometrics in the past. The article with PMID= 20585380 was cited by 124 articles onto the software VOSviewer used for bibliometric mapping. Choropleth maps and the Kano model incorporated with x-index and the AWS are suggested for application and development in the future.

Keywords: Scientometrics, Kano model, x-index, Google Maps, the exponential author-weighted scheme, individual research achievement

Introduction

The works done by scientific research teams have been accompanied by a trend related to the number of authors (Avula & Avula, 2015). The mean numbers of authors in articles in PubMed Central (PMC) are from1. 9 in 1975 to 5. 67 in 2016 (US National Library of Medicine, 2017). Three leading medicine journals (JAMA, The Lancet, and New England Journal of Medicine) have less (from 8–11 in 2005) to 11–18 in 2010, and 2015 (Muth & Golub, 2015). To date, we observed the highest numbers of coauthors at 4, 107 (PMID= 27770180 and 27770183) in Pubmed Central (PMC) Melander, 2999; Veitenhansl, 2004).

Many author-level metrics, such as h-index (Hirsch, 2005) g-index (Egghe, 2006), x-index (Fenner, 2018), and author impact factor (AIF) [13], assume that all co-authors have an equal contribution to an article, which is unfair, unreasonable, and problematic (Sekercioglu, 2008; Vavryčuk, 2018). Numerous author-weighted schemes were proposed in the literature (Hagen, 2010; Vavryčuk, 2018), including the harmonic, the arithmetic, the geometric, the fractional, the exponential (Chien, 2018), and the sequence-determines-credit (Tscharntke, 2007).

If the consensus has been reached, the levels of the participation contribution were highest for first authors, followed by last (namely, corresponding or supervisory) and then second authors. Middle authors had lower levels. Corresponding authors are guarantors for their articles (Duffy, 2017; Bhandari, 2014; Perneger, 2017; Lindsey, 1982; Mimouni, 2016; Baerlocher, 2007). As such, if the percentage of credits more than half allocated to first authors is reasonable, many schemes but the two (i. e., the geometric and the exponential) fail to satisfy the criterion (i. e., first authors own credits more than half). For instance, the harmonic

author credits for first authors are 0. 48, 0. 44, 0. 41, 0. 39, and 0. 37 (i. e., all <0. 5) when the numbers of authors are increased from five to nine. The exponential with the base=2. 72 places more weights than the geometric with the base=2. 0 onto first and last (namely, corresponding or supervisory) authors. Only the exponential has been illustrated in articles before (Chien, 2018). In the current study, we have the motivation to apply the exponential author-weighted scheme (AWS) to evaluate the author individual research achievements (IRA).

Besides the author-level IRA, the countries/areas-level IRAs are also of interest to explore. The calculation of metrics on countries/areas instead of individuals should be illustrated for highlighting the feature of the formula we concern about in this study.

The last we are interested in is about the metric that might be useful and effective to characterize the feature of the actors (author or country) in IRAs. The x-index (Fenner, 2018) ($= \sqrt{\max(i \times c_i)}$), was proposed using the rectangle area to present the individual IRA, different from the h- (Hirsch, 2005) or g- (Egghe, 2006) index using the square box to evaluate the IRA. The drawback of the x-index is hard to recognize whether the IRA is toward the citations (i. e., ci>i) or the publications (i. e., i>ci). The Kano model (Kano, 1984) can classify the actors into three main categories (i.. e, the attract, the one-dimensional, and the must-be). We are thus interested in using the Kano model to describe the feature of the author (or country) in IRAs.

In this study, we attempt to (1) quantify the coauthor contributions with the exponential AWS to evaluate the author- and country-level IRAs, (2) classify the type of author- and country-level IRAs using the Kano model to interpret, and (3) demonstrate a visual dashboard for the high author- and country-level IRAs and show them on Google Maps.

Methods

Data Sources

By searching the PubMed database (Pubmed. org), we used the keywords “Scientometrics” [Journal] on July 27, 2019, and downloaded 203 articles published since 2010. An author-made Microsoft Excel visual basic for application module was used to analyze the data. All the downloaded abstracts were based on the type of journal article (i. e., with the abstract

and the author). These citable articles were then individually searched for citations in Pubmed Central (PMC), and the citations were analyzed for individual authors and countries/areas. All the data used in this study were downloaded from PMC, which means that the study required no ethical approval according to the regulation promulgated by the Taiwan Ministry of Health and Welfare.

Approaches for displaying research results

The exponential author-weighted scheme.

The exponential AWS was proposed according to the Rasch rating scale model (Andrich, 1979) for quantifying the author contributions as Equation (1) :

$$W_m = \frac{\exp(\gamma_m)}{\sum_{m=0}^{m-1}\exp(\gamma_m)} = \frac{(2.27)^{\gamma m}}{\sum_{m=0}^{m-1}(2.27)^{\gamma m}}, \quad (1)$$

Considering a paper of m – 1 author with the last being the corresponding author, Wm in Eq. 1 denotes the weight for an author on the order m in the article byline. The power γ_m is an integer number from m-1 to 0 in descending order.

The sum of author weights in a byline equals 1 for each paper. More importance is given to the first (=exp (m-1), primary) and the last (=exp (m-2), corresponding or supervisory) authors (Chien, 2018), whereas the others (middle authors) are assumed to have made smaller contributions (Lindsey, 1982; Mimouni, 2016). In Eq. 1, the smallest portion (=exp (0) = 1) is assigned to the last second (i. e., penultimate) author with the odds = 1 as the basic reference.

In comparison to the other four author credit schemes below Eqs from 2 to 4 (Hagen, 2010), we plotted their relations with the AWS in Figure 1 and found that the distinct difference is that the exponential owns the highest credit among the five schemes.

$$\text{Harmonic ith author credit} = \frac{\frac{1}{i}}{\sum_{i=1}^{n}\frac{1}{i}}, \quad (2)$$

$$\text{Arithmetic ith author credit} = \frac{n+1-i}{\sum_{i=1}^{n} i}, \quad (3)$$

$$\text{Geometric ith author credit} = \frac{2^{n-i}}{2^n-1}, \quad (4)$$

$$\text{Fractional ith author credit} = \frac{1}{n}. \quad (5)$$

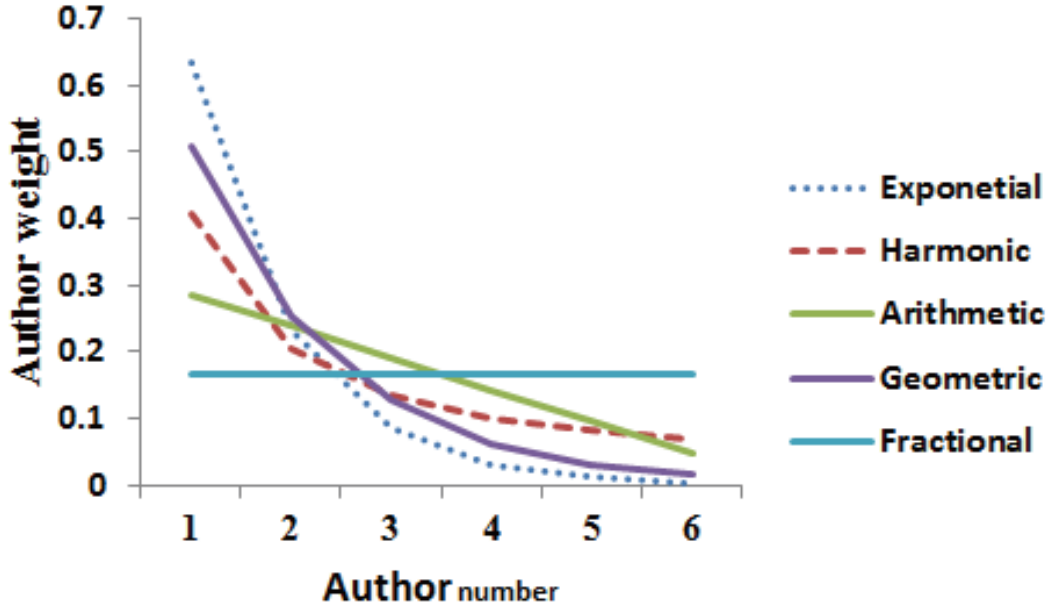

Figure 1. Comparison of author-weighted schemes

The country-level credit scheme.

The county-level credit scheme is defined below:

$$C_{ci} = \sum_{m=0}^{m-1}(W_{cem} \times c_i), \quad (6)$$

Whereas the credit for a country (c) in an article is collected by all weights (ci) on the article citations (ci).

The x-index used in this study

The x-index $(= \sqrt{\max(i \times c_i)})$ (Fenner, 2018) was used to denote the research achievements for a country (or author) by (i) sorting the country-based C_{ci} In Eq. (6), and (ii) determine the number of publications at i and the responding ci. The most cited countries were dispersed with bubbles sized by x-index using the Kano model to display. The most highly-cited ones can be plotted with a dashboard on Google Maps. The citations at x-core are located on the Y-axis and the publication at x-core on the X-axis.

Social network analysis used to explore author clusters

Social network analysis (SNA) was applied to classify the authors in Scientometrics. In alignment with the Pajek guidelines (Batagelj, 2003), using SNA, we defined an author as a node (or an actor) that is connected to another counterpart node through the edge of a line. Usually, another weight

is defined by the number of connections between two nodes. The algorithm of community partition was performed to identify and separate the clusters. The bubbles were sized by the degree centralities in the author collaborations and the citations in articles, see the Eq. (7).

$$AB_i = \sum_{m=0}^{m-1}((W_{Ai} + W_{Bi}) \times c_i), \quad (7)$$

Whereas, the weights for author A and B are denoted by W_{Ai} and W_{Bi}. The number of citations on an article (i) is represented by ci.

Creating dashboards on Google Maps

The metrics and partitioned clusters are yielded by author-made modules in MS-Excel and the SNA algorithms in Pajek. We created pages of Hyper Text Mark-up Language used for Google Maps. The relevant information such as x-index, the citable, and the cited number can be linked to dashboards on Google Maps. Authors' publications in PMC can be easily and quickly retrieved by clicking the author bubble on the dashboard. A choropleth map was used to highlight the countries/areas where authors were affiliated to the high x-index.

Classifying entities into the category under the Kano model

The Kano model is a theory for product development and customer satisfaction developed in the 1980s by Professor Noriaki Kano, which classifies customer preferences into three main categories, namely the "attractive, " the "one-dimensional, " and the "must-be" qualities). These are ploted according to the satisfaction perceived by customers on Axis Y and the effort achieved by providers on Axis X (see the appendix 1). The former three categories have been transformed into various terms such as delighters/exciters, satisfiers, and dissatisfiers, but all refer to Kano's original articles. These categories can be referred to as the three attributes of the citation-oriented, the neutral, and the publication-oriented requirement in this study.

Results

x-indexes for countries/areas on Google Maps

The top three countries with the highest x-index in Scicientometrics are the Netherlands, Denmark, and Germany, see Figure 2. Interested readers are invited to scan the QR-code on Figure 2 and click the country of interest to read the details about the x-index on the Map.

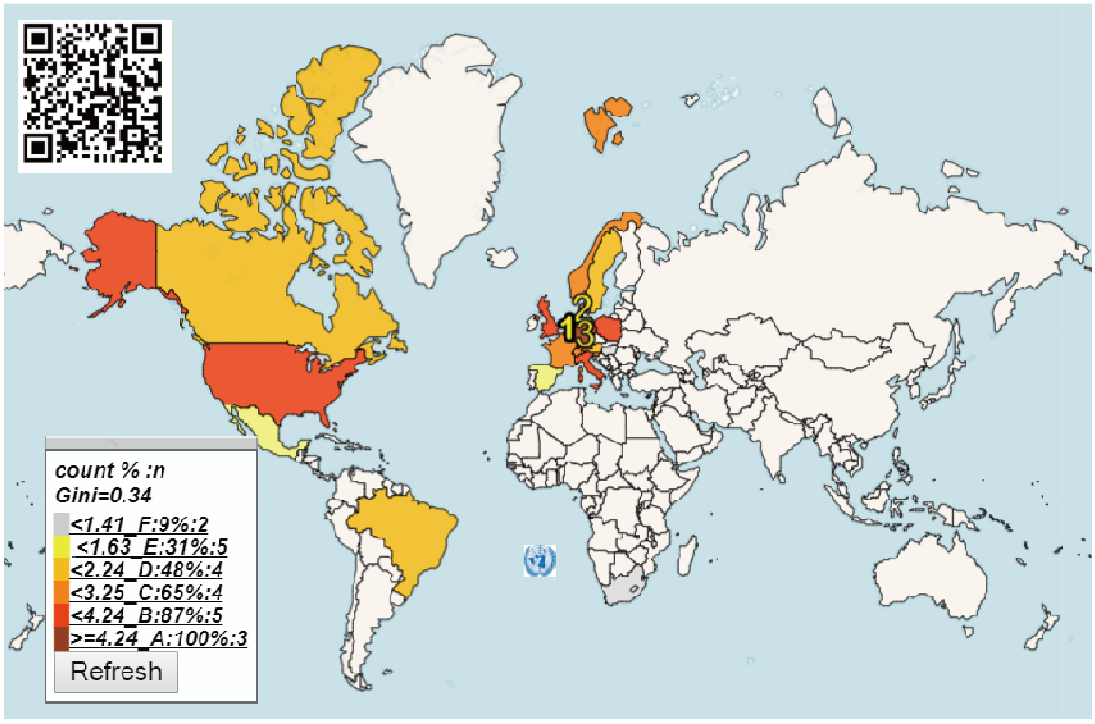

Figure 2. Author distribution by affiliated countries/areas

What types of characteristics for the countries

The x-index fails to tell the characteristics of the IRA in nature. We draw the scatter plots in Figures 3 and 4 to understand the attributes based on the country-/author-level. We can see that the Netherlands is the citation-oriented excitement, and Germany is the production-oriented requirement, see Figure 3.

The author Nees Jan van Eck from the Netherlands has the highest x-index in Scientometrics. Once the bubble in Figure 4 is clicked, the detail information about the x-index= 9. 52, cited=90. 65, and Citable=1 on the x-core are shown on the map.
The article with PMID= 20585380, authored by van Eck (2010) has been cited by 124 articles. The title is "software survey: VOSviewer, a computer program for bibliometric mapping. "

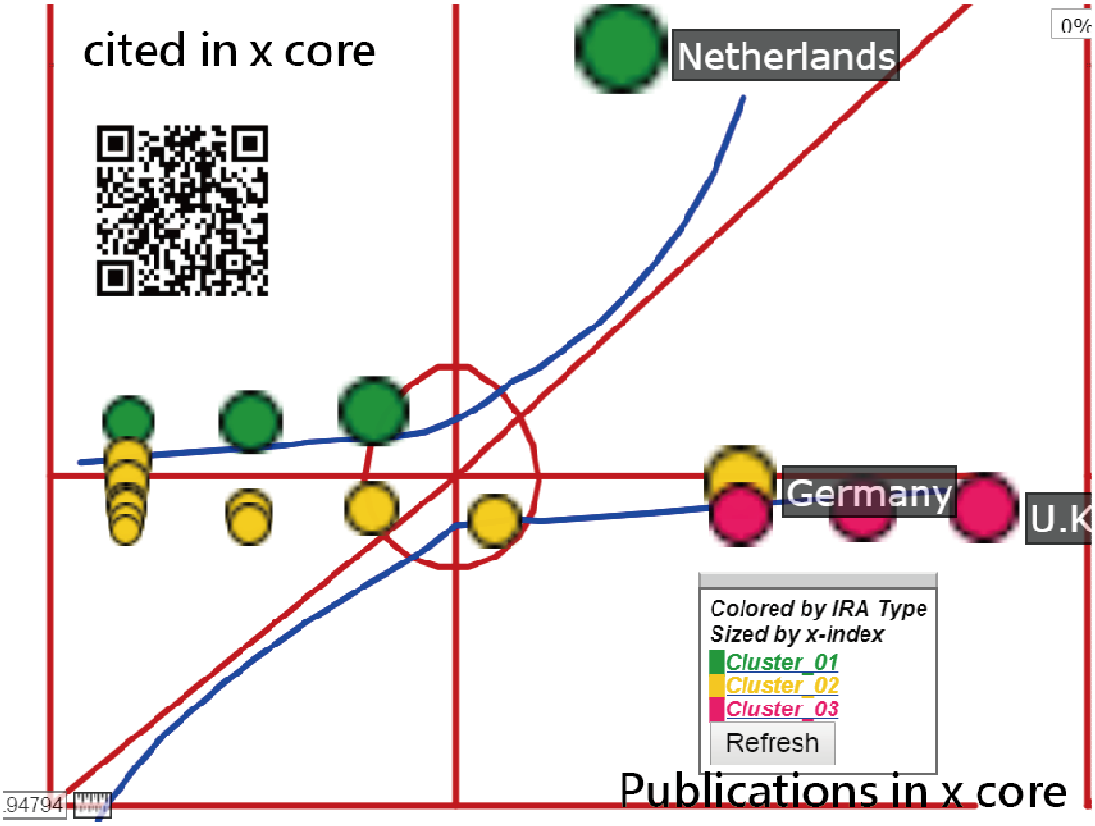

Figure 3. Nation-based research achievements using x-Index to interpretation

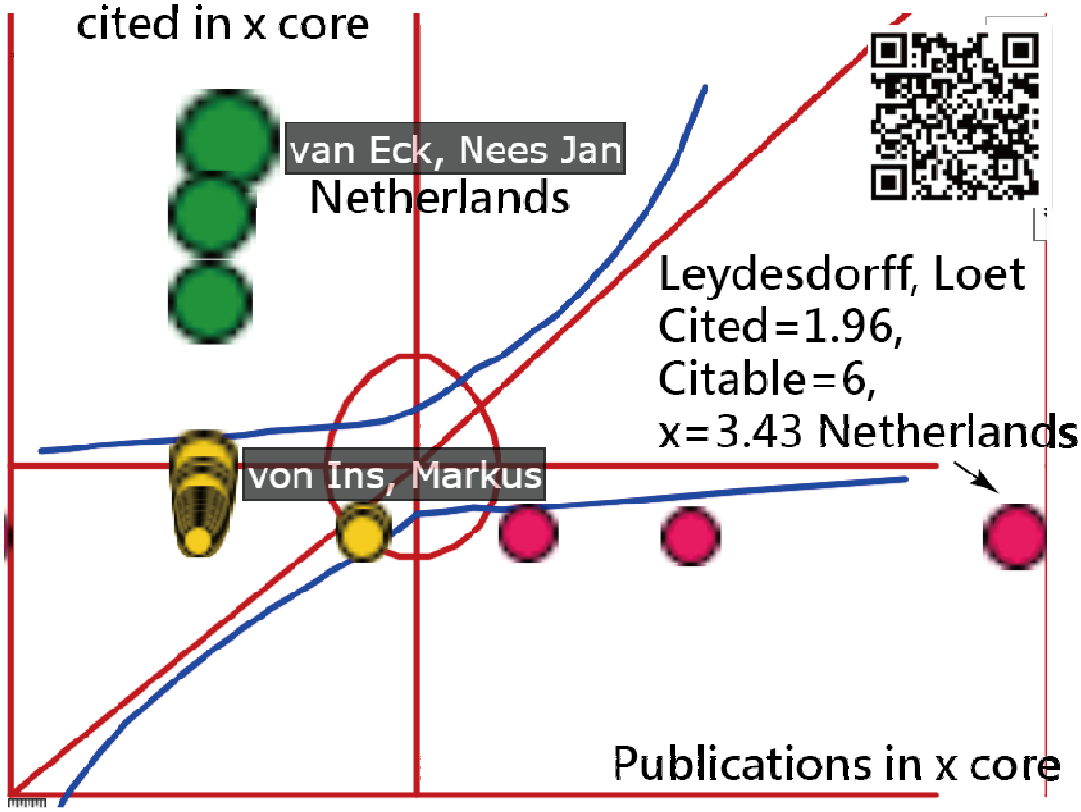

Figure 4. Author-based research achievements using x-Index to interpretation

The author collaborations using SNA to display

In Figure 5, we can see the biggest network on author collaborations is from the Netherlands, indicating the dominant counties in Scientometrics can be easily and quickly disclosed via the visual representations.

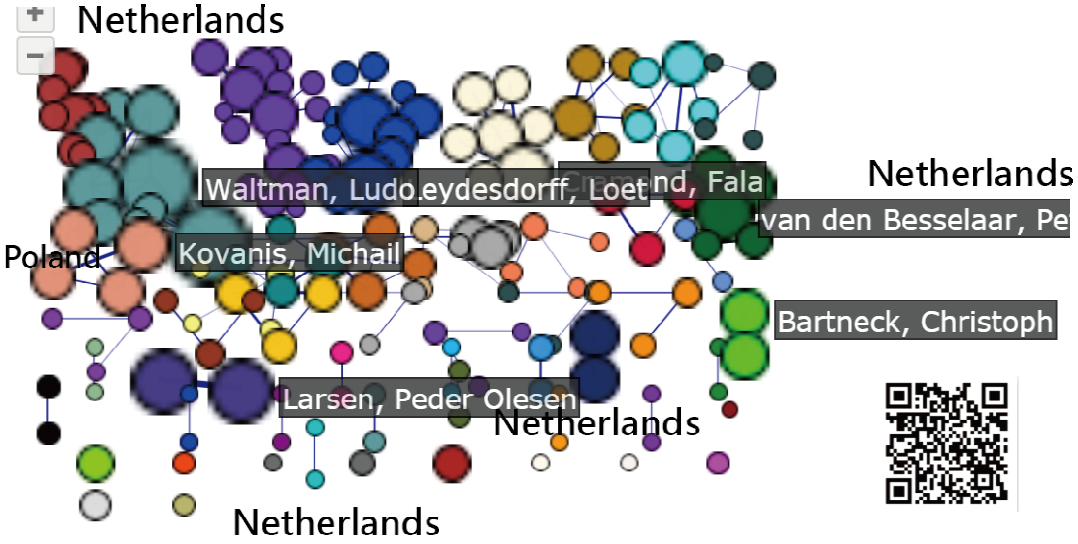

Figure 5. Author network with the citations to interpretation

Discussion

We illustrated several visual representations to present the study results, such as (1) the top three countries with highest x-index are Netherlands, Denmark, and Germany on the choropleth map, (2) the Netherlands is evident of the citation-oriented using the Kano model to display, and (3) the author Nees Jan van Eck from Netherlands has the highest x-index in Scientometrics in the past. The article with PMID= 20585380 was cited by 124 articles onto the software VOSviewer used for bibliometric mapping.

What this study contributes to current knowledge

Several features are worth noting in this study. The exponential AWS was illustrated to show author-/country-level IRAs on the Google platform using Google Maps to display. Not only has the AWS solved the unfair and unreasonable problem of coauthors with equal sizes in an article, but also the visual representations provide readers with a trackable and understandable way using animated dashboard to examine the IRAs for individual authors or countries/areas on Google Maps, which is novel and innovative to present the study results.

Anyone of the other schemes in Equations from (2) to (5) is a special case of the exponential. For instance, geometric ith author credit ($=\frac{2^{n-i}}{2^n-1}$) in Eq. (4) can be expressed by the

$\frac{2^m}{\sum_{m=0}^{m-1} 2^m}$, Where m-1 denotes the number of coauthor in an article. The geometric is thus a special case of the exponential scheme. Similarly, the Eq. (1) can be expressed by the $W_m = \frac{\frac{1}{i}}{\sum_{i=1}^{m} \frac{1}{i}}$, and verified the Harmonic is also special case of the exponential scheme.

Even for the fractional $=\frac{1}{n}$ in Eq. (5) and the arithmetic author credit in Eq. (3), we can apply the $\frac{\exp(0)}{\sum_{i=1}^{m} \exp(0}$ and the $\frac{i}{\sum_{i=1}^{m} i}$ to the weight as another special case of the exponential.

If both h-/g-indexes are applied to the Kano model in Figures 3 or 4, all bubbles will be in the one-dimensional requirements, not like the locations in other two features toward either citations or publications using the x-index to display.

The strength of this study

The first author's credit in the harmonic scheme (Hagen, 2008/2010) will be less 0. 5 when author number is more significant than four. In contrast, the first author's credit will always be greater than 0. 5, or beyond 0. 6, using the geometric or the exponential schemes, respectively.

The clusters in Figure 5 were particularly partitioned by the SNA incorporated with degree centralities and citations to display the most dominant countries/areas, where the most cited clusters are easily highlighted on a map.

Referring to the highly cited article (PMID= 20585380) regarding software VOSviewer used for bibliometric mapping found in this study, we have confidence allowing more authors who are interested in the technique incorporating the SNA with MS Excel to present results on Google Maps.

Another feature is the PMC citations used in this study. Traditionally, over 100 papers were found with a search of "most-cited articles"[Title] in the PubMed library on July 10, 2019. Most of the applied academic databases, such as the Scientific Citation Index (SCI; Thomson Reuters, New York, NY, the United States), Scopus (Elsevier, Amsterdam, the Netherlands), and Google Scholar (Thulesius, 2011), to investigate the most cited ar-

ticles in Scientometrics. None were found using the PubMed library to retrieve the citing articles.

Limitations and suggestions

Despite the findings shown above, several potential limitations require further research efforts in the future. First, the sample of this study only comprised articles in Scientometrics only. It should not be generalized to other journals. As such, the most cited articles and countries or authors are barely determined by the publications in Scientometrics.

Second, there might be some biases when text mining author names from abstracts due to different authors with identical names in the library database, which will affect the result of this study.

Third, we recommend using SNA to partition clusters. They are not limited to Pajek used in this study because many other types of software are used in academics. The style of the visual representations might be somewhat different, but the principle and algorithm of partitions for clusters are similar.

Finally, numerous scientometrics were used for evaluating ARAs. We applied x-index in Figures to illustrate the domain roles in the discipline of the journal of Scientometrics. Authors are familiar with indices, such as h-index, g-/Ab-index, and author factor impact, which can be used to measure ARAs on other topics in the future.

Conclusions

The Netherlands and the author Nees Jan van Eck are evident of the citation-oriented using the Kano model to display their IRAs. Choropleth maps and the Kano model incorporated with x-index and the AWS are suggested for application and development in the future.

List of abbreviations:
AIF: author impact factor
AWS: author-weighted scheme (AWS)
IRA: individual research achievements
AWS:author-weighted scheme
HTML: HyperText Mark-up Language
PMC: Pubmed Central
RSM: rating scale model
SNA: social network analysis

Declarations

Ethics approval and consent to participate.

All the data used in this study were extracted from PMD and published an article, which means the

study does not require ethical approval according to the regulations promulgated by the Taiwan Ministry of Health and Welfare.

Consent to publish.

Not applicable.

Availability of data and materials.

The datasets generated and analyzed during the current study are available on request from the authors.

Competing interests.

The authors declare that they have no competing interests.

Funding.

There are no sources of funding to be declared.

Authors' Contributions.

TH conceived and designed the study, TH, and TW interpreted the data, and WC monitored the process and the manuscript. TW drafted the manuscript. All authors read the manuscript and approved the final manuscript.

References

Alotaibi, N. M., Nassiri, F., Badhiwala, J. H., Witiw, C. D., Ibrahim, G. M., Macdonald, R. L., Lozano, A. M. (2016). The Most Cited Works in Aneurysmal Subarachnoid Hemorrhage: A Bibliometric Analysis of the 100 Most Cited Art. World Neurosurg, 89, 587-592. e6.

Andrich, D. (1978). A rating scale formulation for ordered response categories. Psychometrika, 43, 561-573.

Avula, J., Avula, H. (2015). Authors, authorship order, the moving finger writes. J Indian Soc Periodontol, 19 (3), 258-62.

Baerlocher, M. O., Newton, M., Gautam, T., Tomlinson, G., Detsky, A. S. (2007). The meaning of author order in medical research. J Investig Med., 5 (4), 174-80.

Batagelj, V., Mrvar, A. (2003). Pajek - Analysis, and Visualization of Large Networks. inJünger, M., Mutzel, P., (Eds.) pp. 77-103. Graph Drawing Software, Springer, Berlin..

Bhandari, M., Guyatt, G. H., Kulkarni, A. V., Devereaux, P. J., Leece, P., Bajammal, S., Heels-Ansdell, D., Busse, J. W. (2014). Perceptions of authors' contributions are influenced by both byline order and designation of corresponding author. J Clin Epidemiol., 67 (9),

1049-54. .....

Chien, T. W., Chow, J. C., Chang, Y., Chou, W. (2018). Applying Gini coefficient to evaluate the author's research domains associated with the ordering of author names: A bibliometric study. Medicine (Baltimore)., 97 (39), e12418.

Chien, T. W., Wang, H. Y., Chang, Y., Kan, W. C. (2018). Using Google Maps to display the pattern of co-author collaborations on the topic of schizophrenia: A systematic review between 1937 and 2017. Schizophr Res., 204:206-213...

Duffy, M. A. (2017). Last and corresponding authorship practices in ecology. Ecol Evol. 2, 7 (21), 8876-8887

Egghe, L. (2006). Theory and practice of the g-index. Scientometrics, 69, 131-15.

Fenner, T., Harris, M., Levene, M., Bar-Ilan, J. (2018). A novel bibliometric index with a simple geometric interpretation. PLoS One, 13 (7), e0200098.....

Hagen, N. T. (2008). Harmonic allocation of authorship credit: source-level correction of bibliometric bias assures accurate publication and citation analysis. PLoS One, 3 (12), e4021.

Hagen, N. T. (2010). Harmonic publication and citation counting: sharing authorship credit equitably - not equally, geometrically or arithmetically. Scientometrics, 84 (3), 785-793.

Hirsch, J. E. (2005). An index to quantify an individual's scientificresearch output. Proc. Natl. Acad. Sci. U. S. A., 102, 16569-165728.

Kano, N., Seraku, N., Takahashi, F., Tsuji, S. (1984). Attractive Quality and Must-Be Quality. Journal of the Japanese Society for Quality Control, 41, 39-48.

Lindsey. D. (1982). Further evidence for adjusting for multiple authorship. Scientometrics, 4 (5), 389–395.

Melander, A. (1999), Olsson J, Lindberg G, Salzman A, Howard T, Stang P, etc. 35th Annual Meeting of the European Association for the Study of Diabetes : Brussels, Belgium, 28 September-2 October 1999. Diabetologia, 42 (Suppl 1), A1-A330.

Muth, C. C., Golub, R. M. (2017).

Trends in authorship in team science in major medical journals, 2005-2015. Presented at: Eighth International Congress on Peer Review and Scientific Publication; Chicago, IL

Pan, R. K., Fortunato, S. (2014). Author Impact Factor: tracking the dynamics of individual scientific impact. Sci Rep., 4, 4880.

Perneger, T. V., Poncet, A., Carpentier, M., Agoritsas, T., Combescure, C., Gayet-Ageron, A. (2017). Thinker, Soldier, Scribe: cross-sectional study of researchers' roles and author order in the Annals of Internal Medicine. BMJ Open, 7 (6), e013898.....

Sadoughi, F., Valinejadi, A., Mimouni, M., Zayit-Soudry, S., Segal, O., Barak, Y., Nemet, A. Y., Shulman, S., Geffen, N. (2016). Trends in Authorship of Articles in Major Ophthalmology Journals by Gender, 2002-2014. Ophthalmology, 123 (8), 1824-1828.

Sekercioglu, C. H. (2008). Quantifying coauthor contributions. Science, 322 (5900), 371.

Thulesius, H. (2011). Assessing research impact with Google Scholar: the most cited articles in the journal 2008-2010. Scand J Prim Health Care. 29 (4), 193-5.

Tscharntke, T., Hochberg, M. E., Rand, T. A., Resh, V. H., Krauss, J. (2007). Author sequence and credit for contributions in multiauthored publications. PLoS Biol, 5 (1), e18.

US National Library of Medicine. (2017) Number of authors per MEDLINE/PubMed citation. http://www. nlm. nih. gov/bsd/authors1. html.

van Eck, N. J., Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84 (2), 523-538..

Vavryčuk, V. (2018). Fair ranking of researchers and research teams. PLoS One, 13 (4), e0195509.

Veitenhansl, M., Stegner, K., Hierl, F. X., Dieterle, C., Feldmeier, H., Gutt, B., etc. (2004) 40th EASD Annual Meeting of the European Association for the Study of Diabetes : Munich, Germany, 5-9 September 2004. Diabetologia, 47 (Suppl 1), A1-A464.