Jointly published by Akadémiai Kiadó, Budapest
and Springer, Dordrecht

Scientometrics, Vol. 81, No. 2 (2009) 521–533
DOI: 10.1007/s11192-008-2155-z

A new index for the citation curve of researchers

CLAES WOHLIN

School of Engineering, Blekinge Institute of Technology, Box 520, SE-37225 Ronneby, Sweden

Internet has made it possible to move towards researcher and article impact instead of solely focusing on journal impact. To support citation measurement, several indexes have been proposed, including the h-index. The h-index provides a point estimate. To address this, a new index is proposed that takes the citation curve of a researcher into account. This article introduces the index, illustrates its use and compares it to rankings based on the h-index as well as rankings based on publications. It is concluded that the new index provides an added value, since it balances citations and publications through the citation curve.

1. Introduction

The impact of research is essential. In attempts to quantify impact, different measures have been introduced. This includes the impact factor for journals introduced 45 years ago by GARFIELD & SHER [1963] and GARFIELD [2006]. However, it may be argued that the impact of a journal becomes less important as research articles becomes available on-line and could be easily located even if they are not published in a high impact journal. In general, high impact articles are more important than the journal itself. Historically, highly cited articles have been published in high impact journals, but this pattern may change as for example conference articles are as easily accessible as journal articles. Thus, it becomes more a matter of visibility, accessibility, novelty and quality than where an article is actually published. This also means that time to publication for a submitted paper is likely to become more important as a measure for a journal, i.e. the journal has to ensure the visibility and accessibility.

The advent of the Internet has revolutionized visibility and accessibility, and hence completely changed the rules of the game. For example, many universities have been used to evaluating candidates for positions and promotions through looking at where they have published their research articles. However, this will change. The actual impact of an individual's research will be in focus instead. An article with few citations in a high impact journal is not as good as a highly cited article in a less prestigious journal or conference.

Received November 4, 2008; Published online April 16, 2009

Address for correspondence:
CLAES WOHLIN
E-mail: Claes.Wohlin@bth.se

0138–9130/US $ 20.00

The shift from journal impact to article and researcher impact has just started. The h-index introduced by HIRSCH [2005] has changed the view on impact considerably only in the last couple of years. The index was introduced only three years ago, and several improvement proposals have been published in the last three years. The h-index is defined as: the h-index is equal to X if a researcher has X or more articles with X or more citations. For example, an h-index of 10 implies having 10 articles with 10 citations or more. The h-index has quickly become available in different tools as for example in SCOPUS [2008] and PUBLISH OR PERISH [2008], which is built on top of GOOGLE SCHOLAR [2008]. Thus, the h-index has gained considerable attention since its introduction. However, different researchers have identified different weaknesses in the measures.

One weakness with the h-index not addressed by others is that it takes a point on the citation curve (also called h-graph in SCOPUS [2008]), i.e. the point when the number of articles is equal to the number of citations. However, it is possible to plot a citation curve by sorting articles based on the number of citations, and then the articles are plotted against the number of citations. The point estimate results in ignoring the total production of a researcher, which is evident from the fact that a large number of curves may pass through this specific point. To address this weakness in the h-index, a new index is proposed in this paper.

The main objective of the paper is to define and introduce a new index that captures the citation curve instead of providing a point estimate. It is shown how the new index is capable of differentiating between researchers having the same h-index. Furthermore, it is illustrated how the h-index does not capture some differences between researchers, which the new index is capable of doing.

The remainder of the paper is outlined as follows. The h-index and some of the suggested improvements of the h-index are discussed in Section 2. The new index that takes the citation curve better into account is proposed in Section 3. The new index is illustrated in Section 4, where a set of scholars in software engineering are evaluated based on their publications in 2000 to 2004, and two scholars with similar h-indexes but different citation curves are compared. Finally, some conclusions are presented in Section 5.

2. Measures

Historically, impact has been viewed from a journal perspective, i.e. since Garfield introduced the idea of an impact factor in 1963 [GARFIELD, 1963; 2006]. The journal impact factor is based on the number of citations to the articles published in the journal the last two years. More specifically, the impact factor for 2008 is the number of citations to articles published in the journal in 2006 and 2007 divided by the number of articles. Or more simplistic, the impact factor is the average number of citations per article in the last two calendar years. The factor is normally calculated for journals in the ISI database published by THOMPSON [2008]. However, this is a journal view. Recently, the focus has shifted more towards article impact and hence the impact and influence of the researcher or researchers behind the most-cited articles. This can be illustrated by the introduction of a most-cited special issue in the Information and Software Technology journal [WOHLIN, 2008], and the introduction of measures such as the h-index which is more focused on articles and researchers than the journal or the conference of specific articles or papers.

Before introducing the h-index and some of the proposed improvements, a running example is introduced to be used for illustration purposes. The data from one of the most productive and cited researcher in software engineering in the last 40 years is used. Thus, all values are clearly higher than what can be expected from younger and more "normal" researchers. The data is extracted from PUBLISH OR PERISH [2008] using only the "engineering, computer science and mathematics" category in the tool and the analysis is conducted May 25, 2008. It is chosen to use Publish or Perish, since it covers not only journal articles but also all other types of publications being available on the Internet.

A potential weakness with Publish or Perish is that it sometimes is hard to judge whether an article is from a specific researchers or from a researcher with the same surname and first initial. This weakness has to be handled manually, i.e. it is necessary to judge based on the information provided or have a closer look at the actual paper to ensure that a specific paper is from the researcher actually being analyzed. Having said this, the researcher used in the example has a relatively unique name and after a careful examination it is believed that the papers used in the analysis could be attributed to the researcher studied. It is also worth noting that in Publish or Perish, not only articles are found; occasionally other type of documents are found such as course programs and program committee listings, and hence the actual number of items for a researcher is overestimated. In addition, some items occur more than once since it has been found in different places. The non-publication items are mostly of less importance since they often have a very low number of citations and hence they do not influence the indexes. The fact that some publications can be found in different places is slightly more problematic. However, it probably means underestimating the indexes due to that different articles refer to the same paper, but to different sources. Furthermore, this problem exists for all authors and it is not likely to provide a systematic bias. Anyway, the objective has been to give a fair picture of the researcher's publications and hence the indexes related to the researcher being used in the example. These problems have to be addressed carefully when determining the indexes for different researchers, groups of researchers or universities. A summary of the extracted data is presented in Table 1.

Table 1. Summary of data from *Publish or Perish*

Number of items	834
Items with more than five citations	221
Most cited paper	772
h-index	50
g-index	100
hc-index	28
hi-index	36
hi(norm)-index	20.16

A citation curve can be created based on the citations. This is shown in Figure 1 for the 221 items with more than five citations, where the first item is most highly cited. It can clearly be seen that relatively few items have many citations and that the tail is long. This is a typical citation curve, i.e. it is a skewed curve.

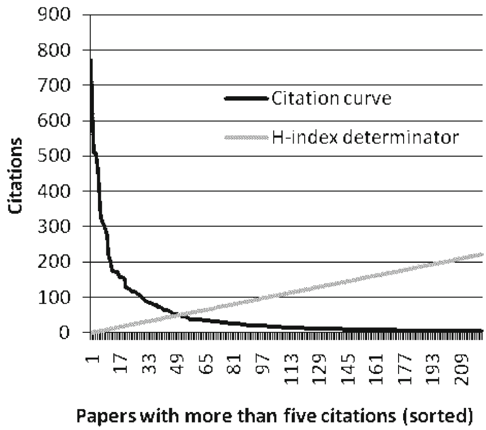

Figure 1. Citation curve and the line to determine the h-index

The h-index is as stated above the point where the number of items is equal to the number of citations. This is illustrated in Figure 1 through the line, which meets the citation curve at 50 items/citations. In other words, Figure 1 shows that the h-index for the researcher is 50. The simplicity and ability to illustrate the h-index easily is one of its main strengths. The drawback is that it may be a too simplistic view, and it does not take a number of other aspects into account. Thus, several extensions and improvements inspired by the h-index have been proposed. For example, the proposal to take the whole citation curve into account as described further in Section 3.

One suggested improvement to the h-index is the g-index. It was introduced by EGGHE [2006]. It is formulated to give higher weight to highly cited articles. This is done by defining the g-index as: the g-index is equal to the number such that the top g items are associated with at least g×g citations. This gives more weight to the highly cited items. In the example, the g-index is equal to 100 since the top 100 items have in total 10128 citations, which is more than 100×100 = 10000, and the top 101 items have in total 10146 which is less than 101×101 = 10201.

Other researchers have proposed that an index should give higher weights to newly published work. This is proposed by SIDIROPOULOS & AL. [2006] in their contemporary h-index denoted hc-index. The hc-index is based on weighting the number of citations. The proposers suggest a weighting where items from the current year receives a weight of four, items published four years ago receives a weight of one, and items published seven years ago receives a weight of 4/7. The weights could be changed, for example instead of having a base of four it is possible to have a base of five or six or any other number helping to define what is recent research. In summary, the index gives more weight to recent and highly cited research than to older research, although maybe being highly cited. The hc-index for the researcher in the example is 28. The actual derivation is not shown here since it would mean describing how the specific publication years are taken into account for the 221 items in the example.

BATISTA & AL. [2006] have proposed an individual normalization of the h-index. The objective of the measure is to take multiple authorships into account. The measure is denoted the hi-index, and it is calculated by dividing the h-index by the average number of authors contributing to the h-index. This is done to reduce the effect of co-authorships. On the one hand, co-authorship may be very positive and there is a risk that if weighting by the number of authors the amount of collaborative papers may go down. This would be unfortunate since some results are based on the combination of expertise. On the other hand, to take co-authorship into account compensates for some papers having a large number of authors. It should also be noted that different research disciplines have different traditions when it comes to co-authorships. A variant to the hi-index has been presented, denoted hi-norm-index. This variant of the hi-index first divides the number of citations with the number of authors, and then calculates the hi-norm-index. The intention is that it should better reflect the original intention of the h-index, i.e. the author impact.

None of the proposed improvements, since the h-index was proposed, has focused on the problem of the index being a point estimate not taking the citation curve into account. Thus, in the next section a new index is introduced that captures the form of the citation curve instead of only focusing on one point of the curve.

3. A new index: w-index

As mentioned in the previous section, several improvements (or additions) have been suggested for the h-index. The improvements have been focused on improving the point estimate based on taking different aspects into account. However, the current indexes do not capture the citation curve and hence they do not take the whole production of a researcher that actually has influenced research into account. To rectify this, an index focused on capturing the form of the citation curve is proposed, which is denoted w-index. The w-index is based on dividing cited items into classes based on the number of citations. GLÄNZEL [2007] proposes a method for defining suitable classes, which could, for example, be different for different disciplines. The latter means that different disciplines may choose to use different classes than suggested here. Here, it has been chosen to divide the citation curve into classes so that the contribution to the index is doubled for each new class.

The classes are defined based on doubling the number of citations in each class. The main reason being that citation curves are often much skewed with many items having few citations and some items having a high number of citations. This is exactly the case in Figure 1. The classes are as follows: the lower class starts with five citations and the next class start with 10 citations, which is followed by a class starting with 20 citations and so forth. The first class starts with five citations. This is based on the approximation that fewer than five citations indicate a very low impact on research.

The 221 items in Table 1, having five citations or more form the basis for the w-index and is used to illustrate it. The 221 items are divided into classes in Table 2, where the length of a class is twice the length of the previous class. In Table 2, the classes, the number of items in each class and the number of items above the lower limit of the class is presented. The latter means for example that all items with more than five citations are listed in the 5–9 citations column in the third row, and hence the number becomes 221. No item has 1280 or more citations in the example, and hence there is no reason to create more classes. The number of items in a class is denoted Nc(i) and the total number of items over the lower limit is denoted t(i), where i is the lower limit of a class, i.e. in the example Nc(5) is equal to 79.

Table 2. Citations divided into classes

Classes	Number in class (Nc)	Above lower limit (t)
5–9 citations	79	221
10–19 citations	46	142
20–39 citations	40	96
40–79 citations	20	56
80–159 citations	20	36
160–319 citations	9	16
320–639 citations	6	7
640–1279 citations	1	1

From the data in Table 2 it is can be seen that despite that the length of the classes doubles the number of items in each class quickly decreases.

A common transformation to normalize data for skewed distributions is to use a logarithmic transformation. This transformation is used here to lessen the impact of the highly skewed distribution. For example, if only counting the number of citations as a whole then a few items will completely dominate that measure and hence not give full credit to the total production and impact of a researcher. In this particular case, the total number of citations for items with five or more citations is 11 180, and the highest number of citations is 772. Thus, the most cited item contributes with 6.9 percent of the total number of citations, and the average contribution is 0.45 percent. Thus, the most cited item contributes with more than 15 times the average. For many researchers, the dominance of the most cited item is higher. This dominance is addressed through the logarithmic transformation.

The following transformation is done. The natural logarithm is taken for the lower limit in each class, for example ln(5) is approximately equal to 1.6094. Doing this means that the doubling of the class length is neutralized, i.e. the differences between classes become approximately 0.6931, since ln(10) is approximately 2.3026 and ln(20) is approximately 2.9957 and so forth. Based on this transformation each item in a class is assigned a value. The value is based on adding the logarithms from the class of the item and all previous classes, for example an item with the number of citations being between 10 and 19 is assigned a value that is equal to 1.6094 plus 2.3026. This is done to capture that it is placed in a class, but it also fulfils the criteria of the lower class(es). Values for an item in a class are shown in Table 3. The logarithm of lower limit of the class is denoted a(i) and the value for an item in a class is denoted v(i). For example, a(10) is equal to 2.3026 (ln(10)), and v(10) is equal to 3.9120 (1.6094+2.3026), which is the weight assigned to an item with between 10-19 citations.

Table 3. Values for items in the different classes

Class	ln(lower limit) (a)	Value of items in this class (v)
5–9	1.6094	1.6094
10–19	2.3026	3.912
20–39	2.9957	6.9078
40–79	3.6889	10.5966
80–159	4.382	14.9787
160–319	5.0752	20.0538
320–639	5.7683	25.8222
640–1279	6.4615	32.2836

The w-index can be calculated in two equivalent ways:

$$\text{w-index} = \Sigma_i t(i)\times a(i) \text{ for all classes, or}$$
$$\text{w-index} = \Sigma_i Nc(i)\times v(i) \text{ for all classes.}$$

In the example, the two calculations become:

$$\text{w-index} = 221\times 1.6094 + 142\times 2.3026 + 96\times 2.9957 \ldots = 1463$$
$$\text{w-index} = 79\times 1.6094 + 46\times 3.9120 + 40\times 6.9078 \ldots = 1463$$

The first calculation illustrates that all items above the lower limit gets credit for that, and in the second case it is shown how much an item in a specific class contributes to the overall w-index. The w-index changes as soon as an item moves from one class to another. This means that the w-index is much more sensitive and fine-grained than the point estimates such as for example the h-index. For example, the w-index would increase with 2.9957 as shown in Table 3, when an article moves from being cited between 10 to 19 times to having between 20 and 39 citations. The total weight for the article changes from 3.9120 to 6.9078, where the difference is 2.9958 (due a rounding error).

4. Illustration: Citation study in software engineering

4.1 Ranking based on w-index

To further illustrate the indexes and compare the outcome from applying the w-index with the h-index as well as with publication-based rankings, an analysis has been conducted of scholars in software engineering. The analysis is based on the following sources:

- A ranking of the top 15 scholars in systems and software engineering is published yearly in the *Journal of Systems and Software* [2008]. For the analysis, a summary of the time period 2000 to 2004 is used. This analysis and ranking are published in [TSE, 2006]. The ranking is based on publications in six journals: *ACM Transactions on Software Engineering and Methodology*, *IEEE Software*, *IEEE Transactions on Software Engineering*, *Information and Software Technology*, *Journal of Systems and Software*, and *Software – Practice and Experience*.
- In [REN, 2007], a ranking of top 50 scholars in software engineering, based on publications in two journals and two conferences, is presented. The journals are: *ACM Transactions on Software Engineering and Methodology*, and *IEEE Transactions on Software Engineering*, and the conferences are: *International Conference on Software Engineering*, and *ACM SIGSOFT International Symposium on the Foundations of Software Engineering*. The latter conference is every second year held in Europe and it is then called the *European Software Engineering Conference and the ACM SIGSOFT International Symposium on the Foundations of Software Engineering (ESEC-FSE)*. The analysis in [REN, 2007] is also done for the time period 2000 to 2004 to make it comparable to the analysis in [TSE, 2006].

Both of the above sources base their ranking on publications. They weigh the contribution based on the number of authors, but none of the studies take actual impact in terms of citations into account. From the analysis in [REN, 2007], it is concluded that the actual ranking is highly dependent on the choice of which journals or conferences that are included in the analysis. One way to address this is to turn the focus to citations rather than focusing on publications, which also can illustrate the differences between rankings. The study here is designed as follows.

First of all, the objective of the study is done to compare the h-index with the newly proposed w-index, as well as compare the citation-based analyses with the publication-oriented rankings published in [TSE, 2006] and [REN, 2007]. Thus, the objective is not to create a ranking of scholars. The latter would require a much broader analysis taking many more researchers into account. The researchers considered in this comparative study are primarily those being ranked among top 15 in [TSE, 2006] or among top 50 in [REN, 2007]. One addition has been made in the analysis. One researcher, whose work is highly cited although he is not on any of the two publication-based rankings, was included to investigate whether a researcher could be highly ranked in a citation study, although not being ranked based on publications in certain journals or conferences.

Thus, the maximum number of researchers in the study is 66. In reality, 61 researchers are included in the study due to that five researchers are both in top 15 in [TSE, 2006] and top 50 in [REN, 2007]. The h-indexes for these 61 researchers were calculated using PUBLISH OR PERISH [2008]. The ranking is based on data available on July 1, 2008. Only publications in the category "engineering, computer science and mathematics" were included. For comparison, only publications listed as published between 2000 and 2004 were taken into account. All publications listed for a researcher were carefully gone through to ensure that the list only included publications of that researcher. It was decided to rank the researchers based on their h-index and then use a cut-off for the h-index of 15. This results in 18 scholars. For these 18 scholars, the w-index is calculated. The scholars are ranked in Table 4 based on the w-index. In addition to the w-index, the h-index and the ranks based on them, the rankings from [TSE, 2006] and [REN, 2007] are provided.

Table 4. Rankings based on indexes and publications.

Scholar	w-index	Rank based on w-index	h-index	Rank based on h-index	Rank from [TSE, 2006]	Rank from [REN, 2007]
W. Emmerich	347	1	23	1		14
B. Boehm	312	2	22	4		
M. Harrold	284	3	23	1		1
G. Rothermel	279	4	23	1		2
L. Briand	266	5	22	4	5	4
A. Mockus	207	6	15	16		9
M. Ernst	204	7	17	6		5
H. Zhuge	197	8	16	11	1	
M. Dwyer	190	9	17	6		29
V. Basili	186	10	16	11		32
P. Tonella	182	11	17	6		31
J. Herbsleb	176	12	15	16		19
G-C. Roman	173	13	17	6		46
K. Sullivan	161	14	16	11		24
P. Inverardi	155	15	16	11		17
B. Kitchenham	152	16	16	11	4	33
K. El Emam	150	17	15	16	2	13
D. Engler	97	18	17	6		49

The Pearson rank correlation between the ranks based on the w-index and the rank based on the h-index becomes 0.65. Unfortunately, it is impossible to calculate any more correlations since not all scholars are ranked in both [TSE, 2006] and [REN, 2007]. However, several observations can be made from Table 4:

- The w-index is capable of providing a unique ranking for all 18 scholars, while many ties exist when using the h-index. In other words, the w-index provides a more fine grained number than the h-index. Furthermore, it balances citations with publications since it is based on the whole citation curve.
- Four out of five of the top scholars from [TSE, 2006] are among the 18 scholars in Table 4. None with a ranking of less than five in [TSE, 2006] makes it into the list in Table 4.
- The scholars ranked highly in [REN, 2007] also appear on the list. Once again four out of five of the top scholars are on the list. However, it is interesting to note that many researchers ranked quite low in [REN, 2007] make it into top 18 when looking at the citations to their work.
- The additional researcher (B. Boehm) added to the analysis, i.e. the researcher neither being ranked in [TSE, 2006] nor in [REN, 2007] came in second on the list when using the w-index. This means that it is possible to publish your work in other venues than the ones used in [TSE, 2006] and [REN, 2007] to get high visibility and hence a high citations.
- It may also be observed that H. Zhuge, who is ranked as number one in [TSE, 2006], does not appear on the list in [REN, 2007] make it into eight place when looking at the citations (using the w-index).
- D. Engler stands out as having a relatively low w-index in comparison to the others in Table 4. This is a consequence of him having few publications in comparison to the others, although his work is well cited when being published. This is visible from him being ranked as 49 in the ranking in [REN, 2007] and still making it into top 18 when being ranked based on the w-index. It illustrates that rankings based on publications or citations can be quite different. However, this is addressed by the w-index, which takes both publications and citations into account by looking at the citation curve for all publications with more than five citations.

Once again it is important to stress that this is not a ranking of the top scholars in software engineering, since the selection of scholars to include in the analysis is based on two other rankings and an additional wildcard. In particular, the latter illustrates that there may be other researchers in software engineering that should be on the list if ranking the top scholars based on the w-index (or h-index for that matter).

Finally, it should be observed that several threats exist to any citation-based ranking. First of all, a number of highly active researchers may be active in a subfield of the discipline and this may bias the result towards that subfield and researchers active in that area. This is particularly problematic for indexes primarily based on citations. The threat is somewhat less for the w-index since it also premiers the number of publications, since the w-index increases as soon as a new publication receives five citations or more. In addition, self-citations have not been removed. To some extent self-citations are healthy since it shows that researchers build on previous work and in many cases it is important to explain how the current work extends the research presented in previous publications, and hence self-citations are needed. Furthermore, it is assumed that highly ranked researchers have a sound attitude towards self-citations and that the proportion of self-citations is about the same for the top scholars in the field.

5. Comparison of two individual researchers

To further compare the different indexes and in particular to highlight what the w-index captures in comparison to the other indexes, a comparison of the citation curves of two researchers in software engineering is provided. In addition, these two individuals are compared with a fictitious researcher having a very uncommon publication record. The curves for two researchers with almost the same h-index are shown in Figure 2.

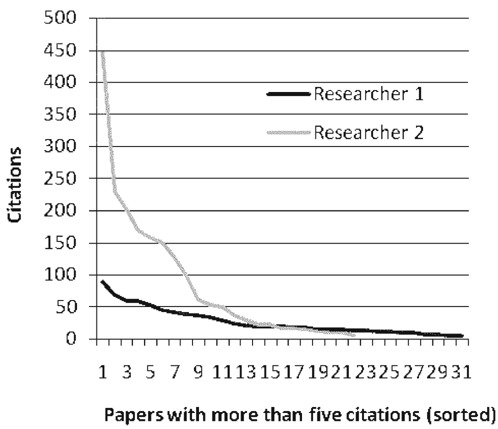

Figure 2. Citation curves for two researchers in software engineering.

From Figure 2, it is obvious that Researcher 2 has more highly cited articles. However, the curves also show that Researcher 1 has more publications with more than five citations than Researcher 2. A comparison between the two researchers with respect to some different measures is presented in Table 5. The table also includes a fictitious researcher to illustrate how indexes such as the h-index and g-index have a difficulty in coping with some forms of the citation curve. The w-index proposed here is better in coping with different shapes of the curve since it takes the whole curve into account. A discussion of the measures for the researchers in Table 5 is provided below.

Table 5. A comparison of measures for Researcher 1, Researcher 2 and a fictitious researcher.

	Researcher 1	Researcher 2	Fictitious researcher
Number of papers	76	80	5
Number of papers with more than five citations	31	22	5
Total number of citations	877	2019	2000
h-index	17	16	5
hi-index	5.56	5.02	N/A
hi-norm-index	10	12	N/A
hc-index	14	14	5
g-index	28	44	5
w-index	182	228	129

Researcher 1 and Researcher 2 were chosen since they had very similar h-indexes and still very different citation curves. The choice is made to illustrate two things. First of all to illustrate the added value of the w-index in relation to the other indexes, and second to emphasize the need to use a set of indexes to describe the track record of individual researchers in terms of publications and citations. Exceptions may exist. For example, if a comparison has a specific goal, which may be to compare how active researchers have been lately in which case the focus would be set on the hc-index.

From Table 5, several observations can be made. Researcher 1 and Researcher 2 are quite equal on most measures, in particular they have almost the same h-index. It is even so that Researcher 1 has a higher h-index despite the fact that Researcher 2 has been cited much more in total. The main differences are that Researcher 2 has higher values on the total number of citations, the g-index and the w-index. At the same time it is interesting to note that Researcher 1 has more papers with more than five citations than Researcher 2. Thus, it may seem that the g-index is capable of capturing the differences. However, this is not completely true, which is illustarted with the fictitious researcher. The fictitious researcher is not likely to exist with five publications having 2000 citations in total, and in the calculations it is also assumed that each publication has 400 citations. However, this example is included to illustrate that the g-index breaks down under certain circumstances, which is not the case for the w-index. The w-index becomes 129. This is lower than for both Researcher 1 and Researcher 2, which is due to the low number of publications by the fictitious researcher. On the other hand, the fictitious researcher gets better credit for the high number of citations for the five publications with the w-index than for any of the other indexes. Thus, the w-index does not break down since it is built on the citation curve and not on an estimation based on a single point on the curve.

6. Conclusions

The w-index has been introduced to address the limitation of just having a point estimate as is the case with the h-index. The w-index takes the whole citation curve into account by dividing the number of citations into classes. The length of one class is twice the length of the previous class, due to that the citation curve is highly skewed. To compensate for the curve being skewed, a logarithmic transformation was introduced. The transformation helped in assigning values to items in each class and through summation the w-index was determined.

The new index better captures the production of a researcher, since it takes more of the citation curve into account than the h-index. The more fine-grained definition means that differences between researchers can be better captured than with the h-index. The w-index provides a more balanced view between citations and actual production of articles/papers.

Future work includes comparing the w-index between different domains and to study how a set of indexes may provide a more comprehensive picture of researchers than a single index.

References

BATISTA, P. D., CAMPITELI, M. G., KINOUCHI, O., MARTINEZ, A. S. (2006), Is it possible to compare researchers with different scientific interests? *Scientometrics*, 68 (1) : 179–189.

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

GARFIELD, E., SHER, I. H. (1963), New Factors in the evaluation of scientific literature through citation indexing, *American Documentation*, 14 (3) : 195–201.

GARFIELD, E. (2006), The history and meaning of the journal impact factor, *Journal of the American Medical Association*, 295 (1) : 90–93.

GLÄNZEL, W. (2007), Characteristic scores and scales: A bibliometric analysis of subject characteristics based on long-term citation observation, *Journal of Informetrics*, 1 : 92–102.

Google Scholar (2008), http://scholar.google.com (accessed October 31, 2008).

HIRSCH, J. E. (2005), An index to quantify an individual's scientific research output, *Proceedings of the National Academy of Science of the United States*, 102 (46) : 16569–16572.

THOMPSON (2008), http://isi3.isiknowledge.com (accessed October 31, 2008).

Journal of Systems and Software (2008), http://www.elsevier.com/locate/jss (accessed October 31, 2008)

Publish or Perish (2008), http://www.harzing.com (accessed October 31, 2008).

REN, J., TAYLOR, R. N. (2007), Automatic and versatile publications ranking for research institutions and scholars, *Communications of the ACM*, 50 (6) : 81–85.

SCOPUS (2008), http://www.scopus.com (accessed October 31, 2008).

SIDIROPOUOS, A., KATSAROS, D., MANOLOPOLOS, Y. (2006), *Generalized h-index for disclosing latent facts in citation networks, arXiv:cs.DL/0607066.* Available: http://arxiv.org/abs/cs.DL/0607066 (accessed October 31, 2008).

TSE, T. H., CHEN, T. Y., GLASS, R. L. (2006), An assessment of systems and software engineering scholars and institutions (2000–2004), *Journal of Systems and Software*, 79 (6) : 816–819.

WOHLIN, C. (2008), An analysis of the most cited articles in software engineering journals – 2001, *Information and Software Technology*, 50 (1-2) : 3–9.