Publish or Perish?
Impact Factors and the $h$-index

A guide for academics and graduate students

Special Lecture Series: Mona, Cave Hill and St. Augustine - 2014

By Simon F. Mitchell
Professor of Sedimentary Geology
Head, Department of Geography and Geology, UWI Mona
Acting Head, Earthquake Unit, UWI Mona

PDF made available on the 10th June, 2014
Preamble

**Impact Factors** and the *h-index* are examples of **BIBLIOMETRICS** which for our context relates to the study and classifying of scientific publication.

I have, for a number of years, thought about this topic in relation to the UWI, and have discussed the issues at length with various colleagues both at the UWI and elsewhere.

What I am presenting today is some of the outcomes of these discourses that I hope are of value to colleagues and graduate students within the *University of the West Indies*. 
Bibliometrics - why are they important?

Bibliometrics ARE being used

• to assess applicants for tenure and promotion
• when researchers apply for jobs
• when researchers apply for grants
• when researchers submit papers

It is therefore vitally important that you understand these indices and use them to help you.

There are two commonly used bibliometrics

1. Impact Factor (IF) or Impact Factor Score (IFS)
2. $h$-index (and $m$-index)
What are Journal Impact Factor Scores?

The **Impact Factor (IF)** was established in 1955 by **Eugene Garfield**, the founder of the **Institute for Scientific Information (ISI)** which is now a part of **Thompson Scientific**. It was designed as a **tool for libraries** to help them decide which journals they might subscribe to within a discipline with increasing budgetary constraints.

The **Impact Factor (IF)** represents the mean number of times that a paper published in a journal during the previous two years (2010 and 2011) has been cited in a journal during the following year (2012).

But, citations to papers follow a **Bradford Law** (a power law or exponentially diminishing function). For example, for **Nature** (Nature Editorial, 2005), 25% of the papers were responsible for 89% of the citations.

---


So Why Publish in Journals with high Impact Factors?

Most scientists know that the best journals in their respective (sub)disciplines are the journals with the highest impact factors.

If these journals are subscribed to by many libraries (following the reason for Impact Factors to begin with), then publishing in a journal with a high impact factor will get more exposure than publishing in a journal with a lower Impact Factor or no Impact Factor.

BUT, with the Internet, Search Engines and Open Access Journals, this is changing (has changed) and papers that once would not have been found because they were published in an obscure local journal may now be freely available online to everyone.

So lets look at Impact Factors
Impact Factors across different disciplines

Impact Factors within a (sub)discipline are dependent on various factors, for instance, the citation style within that discipline (many papers cited or few papers cited).

For nine broadly defined disciplines (treating mathematics as 1) the average number of citations per paper was (Podlubny, & Kassayova, 2006):

- Mathematics: 1
- Engineering/Technology: 5
- Biology: 8
- Earth/Space Sciences: 9
- Social/Behavioural Sciences: 13
- Chemistry: 15
- Physics: 19
- Biomedical Research: 78
- Clinical Medicine: 78

Impact Factors across different disciplines

While the **Impact Factor** may be a good way to compare journals within a single discipline, there are large variations in **Impact Factors** across different disciplines (cf. citation style above).

So if you want to publish in Ecology you have many journals with impact factors above 4, but not in Mathematics.
Impact Factors across different disciplines

**Impact Factors** within a discipline are also dependent on the number of journals/number of papers published within the discipline. With both citation style varying and differences in size of disciplines; there is therefore little relationship between disciplines, numbers of journals and Impact Factors.

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Number of Journals</th>
<th>Highest IFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics</td>
<td>289</td>
<td>3.841</td>
</tr>
<tr>
<td>Ecology</td>
<td>134</td>
<td>15.389</td>
</tr>
<tr>
<td>Applied Chemistry</td>
<td>71</td>
<td>7.294</td>
</tr>
<tr>
<td>Geology</td>
<td>47</td>
<td>4.306</td>
</tr>
<tr>
<td>Applied Physics</td>
<td>34</td>
<td>32.841</td>
</tr>
<tr>
<td>Parasitology</td>
<td>34</td>
<td>14.067</td>
</tr>
<tr>
<td>Crystallography</td>
<td>25</td>
<td>12.619</td>
</tr>
</tbody>
</table>

With differences across disciplines, **Impact Factors** from one discipline **CANNOT** be compared with those from another discipline.

**Data source:** Thompson ISI Scientific Citation Reports (Data for 2012)
How robust are Impact Factors?

Take the case of the journal *Acta Crystallographica A*. This journal had a maximum Impact Factor of 2.38 before 2009.

In 2008, a paper entitled ‘A short history of SHELX’ by G. M. Sheldrick was published and was cited 5,624 times in 2009 leading to the Impact Factor of *Acta Crystallographica A* rocketing to 49.93, way above other journals at the time, such as *Nature* (IFS = 26.68–34.48) and the *New England Journal of Medicine* (44.01–52.58).

After 2 years, the Impact Factor of *Acta Crystallographica A* returned to its pre-2009 value.

This shows how **even a single paper can significantly affect the Impact Factor of a journal**, and that some care must be taken when looking at Impact Factors.

---


So can we compare different disciplines?

Yes!

While individual **Impact Factor Scores** cannot be compared across disciplines, the relative rankings of journals within each discipline can be.

So, if we have $N$ journals in a discipline (note that **Thompson ISI** provides a list of disciplines) and a journal is ranked (by Impact Factor) $R_i$, we can produce a **Normalized Position of Publication Journal** (NPJ), such that:

$$NPJ = 100 \times \frac{(N+1-R_i)}{N}$$

So a journal with **Rank 1** will have a $NPJ = 100$, etc.

We can then compare journals in different disciplines using the ranking scores for those journals within their respective discipline.
So can we compare different disciplines?

Yes!

So if we want to compare three journals in different disciplines, say, *Geology* from Geology (with 47 journals), *Ecological Monographs* from Ecology (with 134 journals) and *Annals of Mathematics* from Mathematics (with 289 journals), we get:

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Journal</th>
<th>IFS</th>
<th>Rank</th>
<th>NPJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geology</td>
<td><em>Geology</em></td>
<td>3.61</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Ecology</td>
<td><em>Ecological Monographs</em></td>
<td>7.433</td>
<td>5</td>
<td>97</td>
</tr>
<tr>
<td>Mathematics</td>
<td><em>Annals of Mathematics</em></td>
<td>2.928</td>
<td>4</td>
<td>99</td>
</tr>
</tbody>
</table>

So **Impact Factors** are related to differences in the size of a disciplines and differences in methods of citation within those disciplines; whereas the **NPJ** gives a measure of importance of that journal within its discipline. **NPJ Scores** can therefore be compared across disciplines (but remember, we are still talking about journals – **NOT PAPERS OR RESEARCHERS**).
What is the Best Paper?

Each year at UWI Mona for the Principal’s Awards, a committee seeks to identify the Best Papers published in the preceding year within the Faculty of Science and Technology (formerly the Faculty of Pure and Applied Science).

Let’s look at the first three years when the award was made and compare ‘Best Papers’ versus number of citations to date.

[Note: papers for 2002 will have been attracting citations for longer than papers from 2003 or 2004 so the number of citations for papers from 2003 and 2004 has been increased proportionally {this may or may not be appropriate, but is reasonable considering the time lapse since publication}]
Best Papers from Research Days 2002-2004

Plot of number of papers versus number of citations.

Five papers were selected and their citations compared to other papers from the same years are shown.
Five papers were selected and their citations compared to other papers from the same years are shown.
Citations versus Impact Factor of all papers (2002-2004)

Note that the papers with the greatest number of citations are not amongst those published in the journals with the highest Impact Factors. *(Nb. Horizontal lines are one citation per year, based on data corrected to 2002)*
Note that the papers with the greatest number of citations are not amongst those published in the journals with the highest Impact Factors. *(Nb. Horizontal lines are one citation per year, based on data corrected to 2002)*

Note that the papers with the greatest number of citations are in the journals with the highest rank (80 to 100%).

Note that the papers with the greatest number of citations are in the journals with the highest rank (80 to 100%).

Publish or Perish? Ifs and the h-index

Citations from Google Scholar
If the papers published in the journals with the highest rankings were selected, we would have captured 4 of the seven papers which were most cited in the succeeding 10+ years (but missed three)!
But remember we also use other criteria!

Including:

Number of authors

Work done at UWI

Importance for Jamaica and the Caribbean

etc.

But looking at the rank of a journal in a discipline certainly indicates there is increased likelihood that papers published in journals with higher ranks will become better cited than papers published in journals with lower ranks!
The EASE Statement

The European Association of Scientific Editors (EASE) released a statement about Impact Factor Scores in November 2007 (http://www.ease.org.uk/publications/impact-factor-statement) stating:

“that journal impact factors [should be] used only - and cautiously - for measuring and comparing the influence of entire journals, but not for the assessment of single papers, and certainly not for the assessment of researchers or research programmes either directly or as a surrogate.”

This statement has important implications for how we assess researchers and individual papers.

**So do not ask what the Impact Factor is but where a journal ranks in its discipline!**

[The European Association of Science Editors (EASE) is an internationally oriented community of individuals from diverse backgrounds, linguistic traditions and professional experience who share an interest in science communication and editing.]
‘Bogus’ Impact Factors

Be careful – many websites now offer independent Impact Factors for Journals, but many of these are spurious! Do NOT use these!


**Global Impact Factor** (Institute for Information Resources) ([http://globalimpactfactor.com/](http://globalimpactfactor.com/))

**Universal Impact Factor** ([http://www.uifactor.org/](http://www.uifactor.org/))

**Index Copernicus International** (Poland) ([http://en.indexcopernicus.com/](http://en.indexcopernicus.com/))

**General Impact Factor** ([http://generalimpactfactor.com/](http://generalimpactfactor.com/))
So can we assess researchers?

J. E. Hirsch introduced the **h-index** in 2005 as a measure of both **productivity** and **impact** of a published work.

He stated that the index is applicable to **Individual Scientists, Groups of Scientists, Departments** and **Journals**.

The index is based on citations, such that:

*An h-index of h indicates that h papers have at least h citations each, and the other (Np – h) papers have no more than h citations each.*

The **h-index** reflects both **number of publications** and **number of citations per publication** and was designed to improve upon measures such as **total number of citations** (which could be biased by a single paper) or **total number of publications** (some [many?] of which may never have been cited).

Calculating the $h$-index

DATA SOURCE:
online Google Scholar Citation profiles for FST Mona December 2013
Calculating the $h$-index

BUT, over time an $h$-index can only getter larger!

More than $h$ citations per article

Less than $h$ citations per article

DATA SOURCE: online Google Scholar Citation profiles for FST Mona December 2013
So can we access researchers?

J. E. Hirsch also introduced the \textit{m-index} in 2005.

The \textit{m-index} was defined as

\[ m = \frac{h}{\text{academic age}} \]

Where academic age is the number of years since the publication of a researcher’s first paper.

The \textit{m-index} corrects for age, since obviously the \textit{h-index} can only get higher over time.

[However, the \textit{m-index} is sensitive to academics who have a break away from publishing (secondments, birth of children, etc.) – a factor that also needs to be taken into account]

How to calculate your \textit{h-index}

The \textit{h-index} can be calculated from various datasets, including:

**Thompson Scientific ISI**: This is calculated from the \textit{Journal Scientific Reports} which are biased towards English-language journals and journals published in North America and England. It excludes many regional journals (for instance those of relevance to the Caribbean) and books.

**Publish or Perish\textsuperscript{(TM)}**: This programme uses an online search to gather citations, but is difficult to use.

**Google Scholar Citations**: This has a user friendly interface and uses Google’s online Scholar dataset. It has the value that it includes books, foreign language and more regional or local journals that are not included by Thompson ISI. \textit{It can also be updated and saved online.}
What to include in your $h$-index

Care must be taken to ensure that you are including relevant and appropriate publications for the calculation of your $h$-index.

It is your academic research which should be included NOT you editing ability.

**Include:**
- Journal publications
- Book chapters
- Books

**Exclude:**
- Books which you edited (but include book chapters in the book)
- Journals for which you were the editor

If you artificially boost your $h$-index, people will catch you and disregard your GOOD work!
What should you use – ISI or Google?

“UWI’s mission is to unlock West Indian potential for economic and cultural growth by high quality teaching and research aimed at meeting critical regional needs, by providing West Indian society with an active intellectual centre and by linking the West Indian community with distinguished centres of research and teaching in the Caribbean and overseas.”

“UWI recognises that as a regional university supported by the West Indian peoples, and as the sole local organ equipped to meet local requirements and to relate its own developmental programmes to them, it should give priority to regional needs.”

As such, Google Scholar has advantages for UWI because it includes regional journals that are excluded from the other databases (e.g., Thompson ISI). It is therefore more applicable as a measure of scientific output by the University on a regional as well as international basis.

UWI Mission Statement http://www.mona.uwi.edu/about/mission.php
What should you use – ISI or Google?

So by using only ISI we lose 33% of the papers that are picked up by Google Scholar, some of which have been well cited.
What should you use – ISI or Google?

Sanderson (2008) stated that a scientist should use the maximum \textit{h-index} obtainable from the different databases.

To use \textbf{Thompson ISI requires a subscription.}

Whereas \textbf{Google Scholar is free!}

\begin{itemize}
  \item \textbf{Google Scholar Citations} Mona
  \item \textbf{Google Scholar Citations} Cave Hill
  \item \textbf{Google Scholar Citations} St Augustine
\end{itemize}

Applicability to Promotion

Hirsch (2005) suggested (based on his study of physicists) that for a Major University:

- $m = 1$ ($h$-index = 20 after 20 years), characterized a successful scientist.
- $m = 2$ ($h$-index = 40 after 20 years), characterized an outstanding scientist,
- $m = 3$ or higher ($h$-index = 60 after 20 years), characterized truly unique individuals.

Based on typical $h$ and $m$ values, Hirsch (2005) suggested (with large error bars) that for faculty in Physics at major research universities:

- $h$-index = 12 might be a typical value for advancement to tenure.
- $h$-index = 18 might be a typical value for advancement to full professor.

Applicability to Promotion

Kelly and Jennions (2006) studied scientists on the Editorial Boards of several high-ranking (medium Impact Factor) journals in Ecology and Evolutionary Biology and concluded:

That from their dataset, $m$ was lower than in physics and ranged from 0.52 to 1.89 (mode $m = 1.07$).

They concluded that a small $m$ [but note the $m$-index can be much less than 0.52] was poor evidence that a researcher’s work was held in low regard by their peers (as evidenced by the inclusion of these scientists on the editorial boards).

So we have to ask whether the $h$-index and $m$-index can be compared across different disciplines?

Limitations of the $h$-index

So what are the limitations of the $h$-index?

1. The $h$-index will vary across different disciplines, so may not be suitable for such comparisons.

2. The $h$-index for a researcher can only increase, never decrease.

3. The $h$-index depends on the duration of each scientist’s career because the pool of publications and citations increases over time. Scientists with a break in their careers will be at a disadvantage.

4. Highly cited papers are important for the determination of the $h$-index, but once they are selected to belong to the top $h$ papers, the actual number of citations they receive is not considered.

5. The use of the $h$-index could [Has? Will? Should?] provoke changes in the publishing behaviour of scientists.
Other indices

Many other indices have been proposed, for example:

**g-index**: Given a set of articles ranked in decreasing order of the number of citations that they received, the *g-index* is the (unique) largest number such that the top *g* articles received (together) at least *g*\(^2\) citations.

**hc-index**: This adds an age-related weighting to each cited article, giving less weight to older articles.

**h\(_r\)-index**: This divides the standard *h-index* by the average number of authors in the articles that contribute to the *h-index*, in order to reduce the effects of co-authorship.

**AWCR**: The *AWCR* measures the number of citations to an entire body of work, adjusted for the age of each individual paper. It is an age-weighted citation rate, where the number of citations to a given paper is divided by the age of that paper.

**e-index**: The *e-index* is the (square root) of the surplus of citations in the h-set beyond *h*\(^2\), i.e., beyond the theoretical minimum required to obtain a *h-index* of 'h'. The aim of the *e-index* is to differentiate between scientists with similar *h-indices* but different citation patterns.

But the *h-index* has gained popularity and is now widely used.
Using the *Google Scholar Database*, the *h-index* of each member of the *Faculty of Science and Technology* at *UWI Mona* was determined (December, 2012).

All papers contributing to the *h-index* were checked for correctness (that is, that they were produced by the staff member)*

The dataset was then analysed using the (standard) following criteria:

- Publications by Department/Discipline
- Publications by rank/seniority (Assistant Lecturer, Lecturer, Senior Lecturer, Professor)
- Publications by academic age (that is the date of publication of a researcher’s first paper)

*Note – some staff members may have had papers omitted because of name change or absence from the database – hence – you need to do it yourself!*
h-index in FST at UWI Mona by rank

Dataset: UWI Mona, December, 2012
h-index vs. Academic Age by Department

Academic age is number of years since publishing first paper

Dataset: UWI Mona, December, 2012
h-index vs. Academic Age by Department

15-20 years of “lost productivity”

Dataset: UWI Mona, December, 2012
h-index vs. Academic Age by Rank

First Publication year vs. h-index for different ranks:
- **Assistant Lecturer**
- **Lecturer**
- **Senior Lecturer**
- **Professor**

Dataset: UWI Mona, December, 2012
h-index vs. Academic Age by Rank

Dataset: UWI Mona, December, 2012
Hirsch’s \( m \)-index = \( h/\)academic age

Regression line fitted to all data gives \( m = 0.38 \)

Excluding ‘lost productivity’ gives \( m = 0.48 \)

Dataset: UWI Mona, December, 2012
So how does UWI Mona compare with other universities?

I have selected two (Queen’s, Belfast, and Botswana) as these were used by the “Academic Board Taskforce” at Mona on Strategic Challenges Confronting the UWI, Mona, for comparison.

We will look at $h$-index and $m$-index.

I will compare two disciplines that have relatively large numbers of researchers at UWI: Chemistry and Life Sciences.

[And remember, the dataset has been checked for consistency of only what is included not what may have been excluded.]
h-index vs. Academic Age for Chemistry

Data for UWI Mona: December 2013
h-index vs. Academic Age for Life Sciences

Data for UWI Mona: December 2013
So how do you improve your $h$-index?

1. Develop your own publishing strategy

Regular publishing each year will build up your publication list. Remember, one highly cited paper will not make a career!

Consider Peter Higgs, the physicist after whom the Higgs Boson is named. When talking to the Guardian Newspaper he stated that “I wouldn't be productive enough for today's academic system” where academics are expected to “keep churning out papers.”

*BUT – you must look at the ‘rules’ that are in practice today!*

The UWI aims for lecturers (where say half the time is assigned to research) to produce 2 papers per year. It promotion is based on number of papers published (which certainly is an important consideration), then matching, or exceeding, this target is obviously important.
2. Publish in the right places

You must consider many points when working out where to publish. Including:

a) The *Journal’s Impact Factor* – sometimes a publication is better in a discipline-specific journal (with lower impact factor) than in a high impact factor journal, such as, *Nature* or *Science* (which may NOT be regularly read in your field).

b) The rank of the Journal in your discipline. **Definitely publish in Journals with higher ranks.**

c) The availability of the Journal to scientists who may cite your work (Obscure journals, even with high Impact Factors, may not be read by the people you want to read your articles).

d) Colleagues (and supervisors) advice.

e) The cost of publishing! (Make sure you look at issues of page charges or cost of pdfs, etc. – which may be hidden in small print online!).
f) Avoid journals that do not have electronic versions.

g) Make sure the Editor is respected in their (your) field.

h) Make sure the peer review process is adequate (many online ‘open access’ journals advertise unrealistic peer review times).

i) Look at published dates for submission, acceptance and publication of articles (usually available on each article).

j) Check for ‘papers in press’ availability on the Journal’s webpage. Once accepted, you want your paper available for citation as soon as possible.

k) Do not feel bad if your paper gets rejected – use the comments to improve it and make it better!

l) If you do not hear from a journal in ‘due course’, pull your paper and submit it somewhere else.
3. The Open Access Minefield

The term *Open Access* has become widespread in scientific publishing. While meaning that scientific papers are freely available, there are two models that are used.

a) A Journal is published electronically and papers are made freely available on the internet. Often these Journals are published by Learned Societies or Charitable Organisations. They either have no page charges, or page charges start after a minimum number of pages.

b) The ‘*Open Access Model.*’ These Journals are often published by Publishing Houses which aim to *publish for profit.* In this case a paper is accepted for publication dependent on the author paying a fee (anything from US$200 up, although often there are introductory deals). This model has a major *conflict of interest* issue, in that the Publishing House needs to publish sufficient articles to meet its costs. Invariably the peer-review process is compromised, and papers are published because someone can pay, rather than on the merit of the scientific content.
“A spoof paper concocted by Science reveals little or no scrutiny at many open-access journals”

A spoof paper, describing a hopelessly flawed chemistry experiment, with meaningless results was submitted to various ‘Open Access’ journals for publication under an assumed name by the Science editorial office.

Many Journals accepted the paper without picking up the flaws in the science indicating a failure in the peer-review process.

Many of these Journals belong to Publishing Houses in India and China, amongst others, and much of this appears to be a phishing scam to extort money from academics and their institutions.

The complete article is worth a read!

The Caribbean is not immune

Caribbean Journal of Science and Technology inviting article for publication
Editor CJST [caribjscitech@gmail.com]
Sent: 08 December 2013 09:50

Caribbean Journal of Science and Technology is an internationally open access, Annual published peer reviewed online journal.

CJST aimed at publishing recent advances in the form of research and review articles in Science and Technology world.

In this regard, scientists, research scholars and students who have zeal to publish their research findings are invited. Submitted papers must be in technical English, suitable for scientific publication. All articles have to be original articles that have not been published elsewhere or are being considered for publication in other journals. Receipt of the manuscript will be acknowledged by email. Review process of the submitted articles will be done within 7-10 days after the date of submission and the accepted articles will be published in the upcoming issue. Papers should be submitted electronically by visiting the journal website: (http://caribjscitech.com/)

Please refer to the Author guidelines’ on our website http://caribjscitech.com/downloads/ for more detail. Send your resumes to become editorial team member of CJST.

Anybody can submit the soft copy his/her prepared manuscript Anytime in MS Word format by mailing to eco.cjst@gmail.com

With Regards
Editors,
CARIBBEAN JOURNAL OF SCIENCE AND TECHNOLOGY,

Web: www.caribjscitech.com

Received by email: December 2013

Editors: India (10), Iran (4), Egypt (2), Not listed (2), Ethiopia (1), Libya (1), Malaysia (1), Kurdistan (1), Mexico (1).

The only connection it has with Jamaica (that I can find) is that it has a local ISSN 0799-3757 issued by the National Library of Jamaica.

So .... be careful
So how do you improve your h-index?

Publish good-quality papers – a paper with a large dataset applicable across a broad-range of your discipline is more likely to be cited than a narrow short paper.

Publish in the best journals – papers are more likely to be read by more scientists

Make your papers available – a paper freely available on the internet is more likely to be cited than one that is not.

Good Luck and THANK YOU