Library and Information Services in Astronomy VII: Open Science at the Frontiers of Librarianship
ASP Conference Series, Vol. 492
A. Holl, S. Lesteven, D. Dietrich, and A. Gasperini, eds.

A Bibliometric Analysis of Observatory Publications 2008–2012

Dennis R. Crabtree

Dominion Astrophysical Observatory, National Research Council of Canada, Victoria, BC V9E 2E7 Canada

Abstract. Refereed publications are the primary output of modern observatories. I examine the productivity and impact of a significant number of observatories, as well as some other interesting aspects of observatory papers.

1. Introduction

Papers in refereed journals represent the key product of modern astrophysical observatories. These publications represent the observatories' contribution to astronomical knowledge. Observatories carefully track the papers based on data from their telescopes as this is of great interest to the various governance bodies and other stakeholders. The number of refereed papers each year is a measure of an observatory's productivity. A telescope's productivity usually ramps up over a period of 5–10 years following the start of operations. While a telescope's productivity is one metric of a telescope's performance, the relevance of the associated publications is another important metric. Citation counts are a standard measure of the impact of a paper: the more citations a paper receives, the more relevant the work is to the broader astronomical community.

In this paper, I examine the productivity and impact of international optical telescopes with an aperture large than 3.5-metres. A small number of other telescopes, such as the Hubble Space Telescope (HST) and James Clerk Maxwell Telescope (JCMT) are also included in this study.

2. Observatory Publications

Observatories maintain bibliographies of papers that have used data from their observatory. To be considered an observatory publication, a paper must base some or all of its conclusions on data generated from that observatory. IAU Commission 5 has developed a document describing the best practices for creating such a bibliography.[1]

For this study, I have retrieved the bibliography for each telescope in this study from the following sources — individual observatory websites; directly from the librarians or other individuals; or, in a few cases from sources such as personal ADS libraries. In all cases, I relied on the observatories to determine whether a paper meets the definition of an observatory paper.

---

[1] http://iau-commission5.wikispaces.com/file/view/Best%20Practices%20Final.pdf

The basic publication information as well as the ADS bibcode for each paper was stored in a relational database (Microsoft Access). The bibcode for each paper was validated via ADS and corrections made to the data in the case of an invalid bibcode. Once a paper's bibcode was verified, further information including the number of citations to the paper was retrieved from ADS and added to the database.

3. Citation Metric

This study includes papers only up to 2012 because it takes at least a year before the citation count is meaningful. Generally, a paper's citation rate — the number of citations per year — peaks 2 years after publication, although a paper's citation history can exhibit quite different behavior. Since citation counts grow with time and I will be comparing papers of different ages, it is necessary to correct the citation counts for this aging effect. My approach is to normalize each paper's citation count by the citation count for the median Astronomical Journal paper[2] of the same year. I refer to this normalized citation count as the impact of a paper.

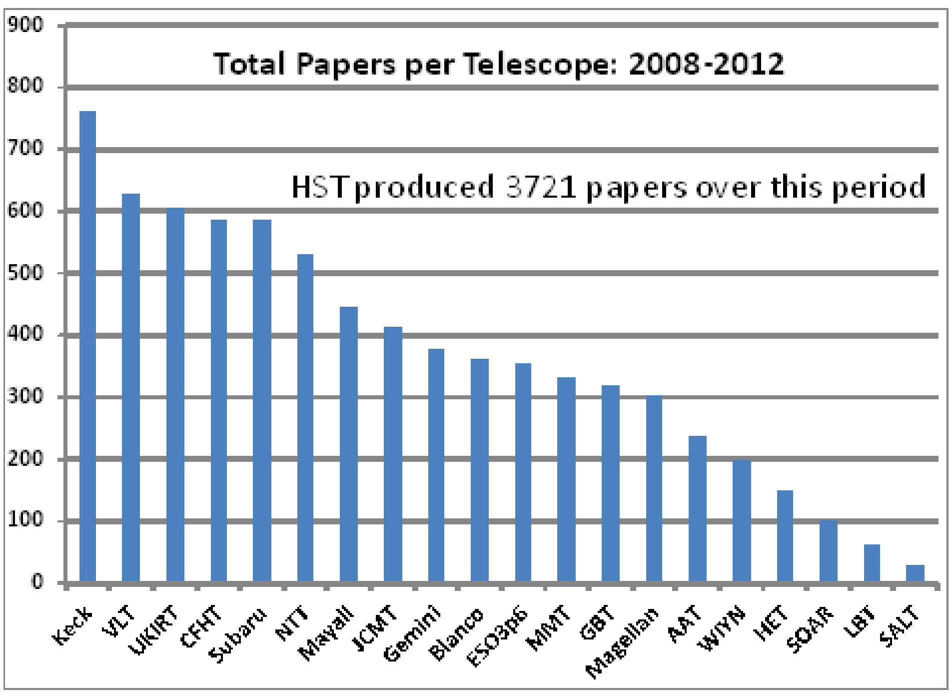

Figure 1. Total papers per telescope: 2008–2012. HST is not included in this figure as it would distort the scale of the plot.

[2]Based on number of citations.

4. Productivity

The productivity of each telescope in this study, measured by the number of refereed papers publisher each year, for the period 2008–2012 is shown in Figure 1. In the case of a telescope actually representing multiple physical telescopes,[3] I have normalized the paper count to produce a productivity figure per telescope.

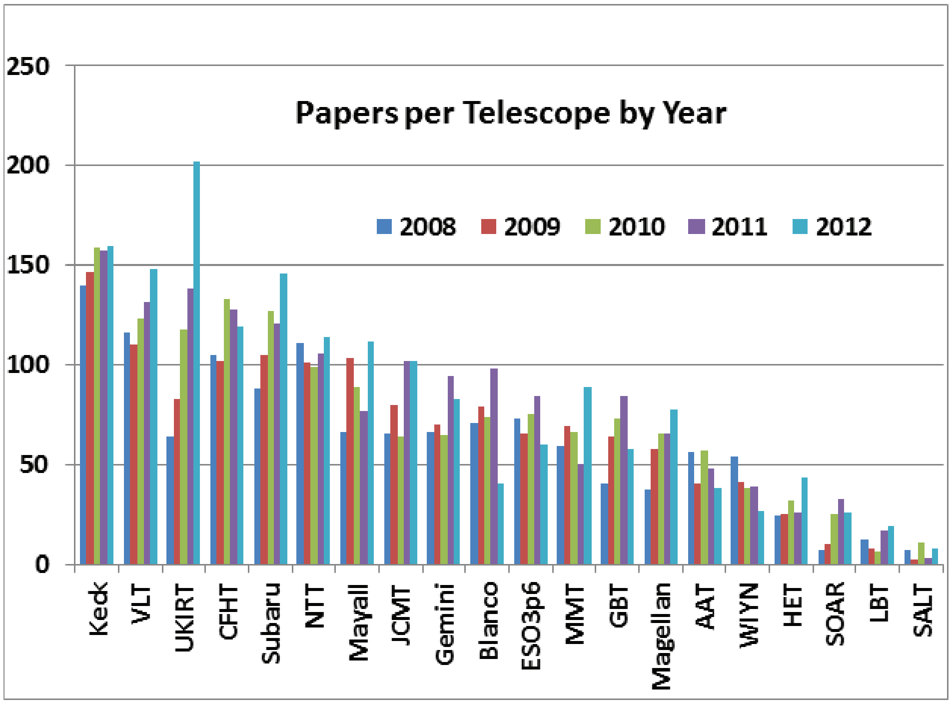

Figure 2. Papers per Telescope by Year

There is a large range of productivity between the most productive (Keck) and the least productive (LBT, SALT) telescopes. The low productivity of the LBT and SALT is most likely due to technical issues that have kept the telescopes offline for periods of time.

When looking at the productivity of each telescope broken down by year, the data shows that some telescopes have increased significantly during the five-year period, such as UKIRT and SOAR. The productivity of most telescopes shows no significant

[3]For example, Keck has two telescopes; VLT has four telescopes. The following short names will be used throughout the article: Keck: Keck Observatory ; VLT: Very Large Telescope ; UKIRT: UK Infrared Telescope ; CFHT: Canada–France–Hawaii Telescope ; Subaru: Subaru Observatory ; NTT: New Technology Telescope ; Mayall: Mayall Telescope ; JCMT: James Clerk Maxwell Telescope ; Gemini: Gemini Observatory ; Blanco: Blanco Telescope ; ESO3p6: ESO 3.6-metre Telescope ; MMT: Multiple Mirror Telescope ; GBT: Green Bank Telescope ; Magellan: Magellan Telescopes ; AAT: Anglo-Australian Telescope ; WIYN: Wisconsin–Indiana–Yale–NOAO Telescope ; HET: Hobby–Eberly Telescope ; SOAR: Southern Astrophysical Research Telescope ; LBT: Large Binocular Telescope ; SALT: South African Large Telescope

trends, only small to moderate changes from year to year. The increase in UKIRTS's productivity is due to the papers from UKIDSS while SOAR's is most likely a result of the normal ramp up in operations.

While not directly related to productivity, the data in this study allows an examination of other aspects of refereed publishing. The growth in the size of research teams and the corresponding increase in the number of authors on papers have often been noted. I examined the long-term trend in the number of authors per paper using the data in my observatory publication database, which extends back to the early 1990s.

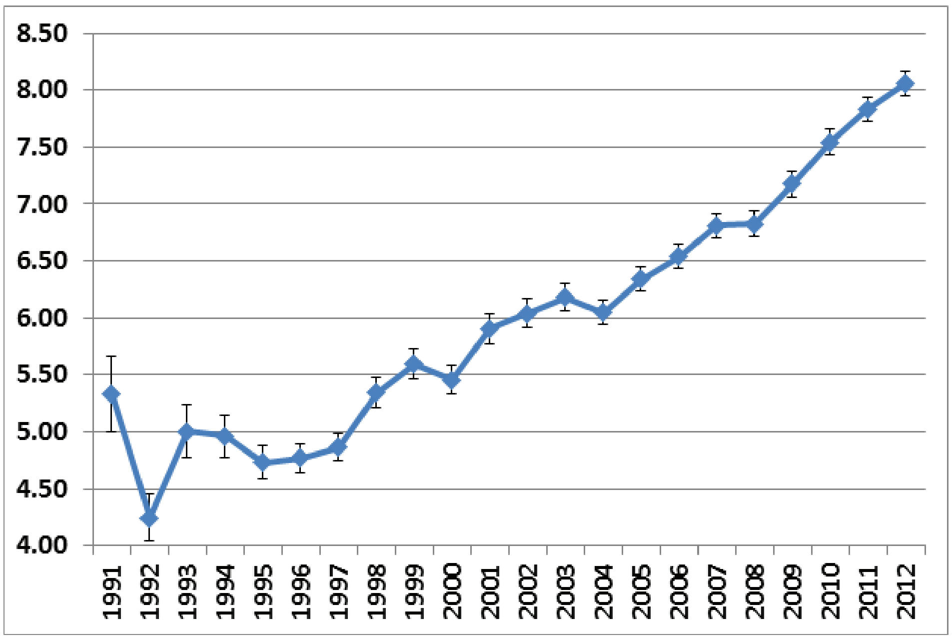

Figure 3. Average number of authors per paper by year

Figure 3 shows the average number of authors on each paper for the period 1991–2012. The error bars represent the standard error of the mean. The data prior to 2000 is based on a smaller number of telescopes and the errors are larger. The data indicate that the increase in the number of authors per paper was relatively flat during most of the 1990's and then began to increase rapidly after 1997. The fairly rapid increase in the size of author lists shows no sign of slowing down.

5. Impact

The impact of the published research is also an important metric of an observatory's performance. The number of times a paper has been cited is usually considered a good quantitative measure of a paper's impact. Impact is not to be confused with the quality of the research. Rather, impact is a measure of the relevance of the paper to other research and researchers in the field. Of course, the number of citations is influenced by factors such as area of research and the citation culture of each sub-field. Since we are studying large aggregates of papers and not comparing individual authors, the effect of these factors should average out.

Before investigating the impact of the observatories in this study, I will look at how the impact of papers correlates with the number of authors. As discussed earlier, research teams are increasing in size, as measured by the increasing number of authors per paper.

Figure 4 shows the average impact per paper (AIPP) and the median impact per paper (MIPP) as a function of the number of authors for all papers in this study. Points in the MIPP curve are labeled with the number of papers included for that number of authors. The AIPP curve has error bars that represent the standard error of the mean. Figure 4 shows the dramatic increase in impact with the number of authors. The average and median impact per paper increases until there are 14 authors per paper and then levels out. This clearly indicates that, on average, papers produced by larger teams have a higher number of citations and thus, a higher impact. The data also indicates a decreasing number of single author papers. For 2008–2012, there more papers with 11 authors (as well as between 2 and 10) than single author papers.

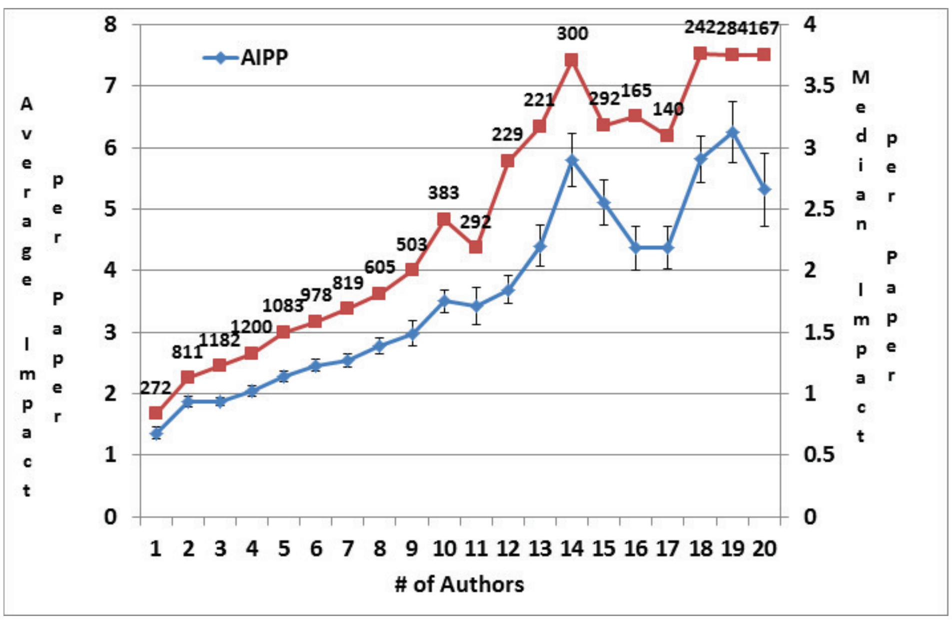

Figure 4. Impact per paper vs. number of authors

Now I will investigate the AIPP and MIPP by observatory. Figure 5 shows the AIPPand Figure 6 the MIPP for each observatory by year for the period 2008–2012. There are considerably smaller differences in impact between observatories than there are in productivity. While Keck is still near the top (#2), the HET has the highest average impact per paper for this period. There is considerable variation in AIPP from year to year for many observatories. UKIRT's impact per paper shows a variation of a factor of 2 over this period, again probably due to UKIDSS publications. SOAR's ramping up of operations is likely the factor in its large increase of AIPP over this period. It is also worth noting that HST's AIPP falls in the middle of the pack. So

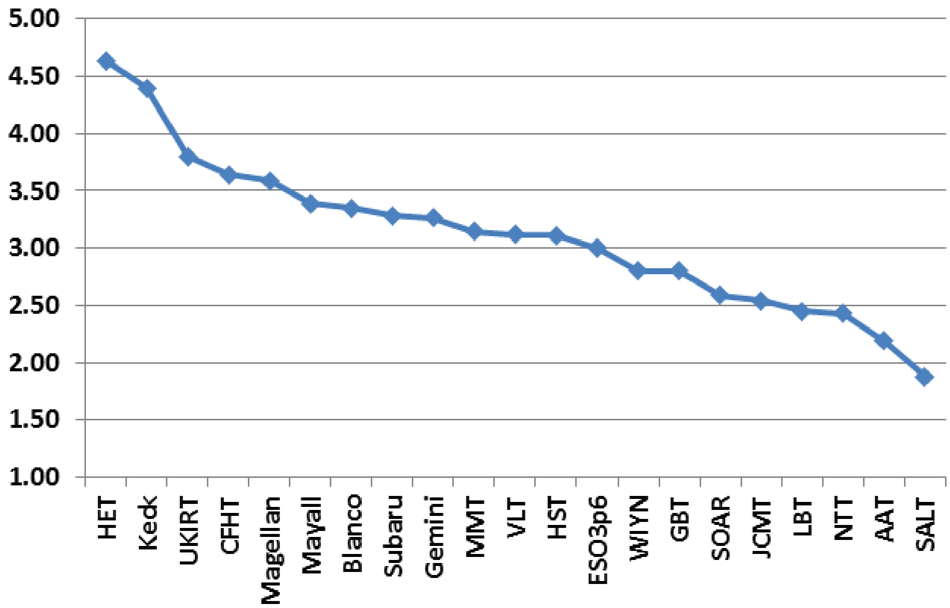

Figure 5. Average impact per paper (AIPP)

while HST produces significantly more publications than ground-based telescopes, the impact of most of its papers is quite average.

While the impact per paper is an important measure, the total impact of an observatory is also a valuable metric. The total impact, which is simply the sum of the impacts of individual papers, is a metric that combines both productivity and the average impact of each paper. The total impact for each telescope for the 2008–2012 period is shown in Figure 7. Even though HET has the highest average impact per paper, its productivity is quite low and so is its total impact.

Several observatories track which instruments are used for each paper. I have included this information in my database when it is either publicly available or has been provided to me by the observatory. Linking the instrument used to each paper allows me to investigate the productivity and impact of each instrument and how much each of them contributes to the overall impact of the observatory.

An example of instrument productivity and impact is shown in Figure 8, which shows the productivity and impact of CFHT instruments. The plot shows the data for the most productive seven instruments during the 2008–2012 period. Although CFHT has only three currently active instruments, papers based on data from earlier instruments are still being published. Figure 8 clearly shows the dominance of Megacam, the wide-field imager, in terms of productivity and impact.

6. Impact Inequality

All papers do not have the same impact with the impact distribution of an observatory's papers is very broad. There are some papers with very little impact — a small number

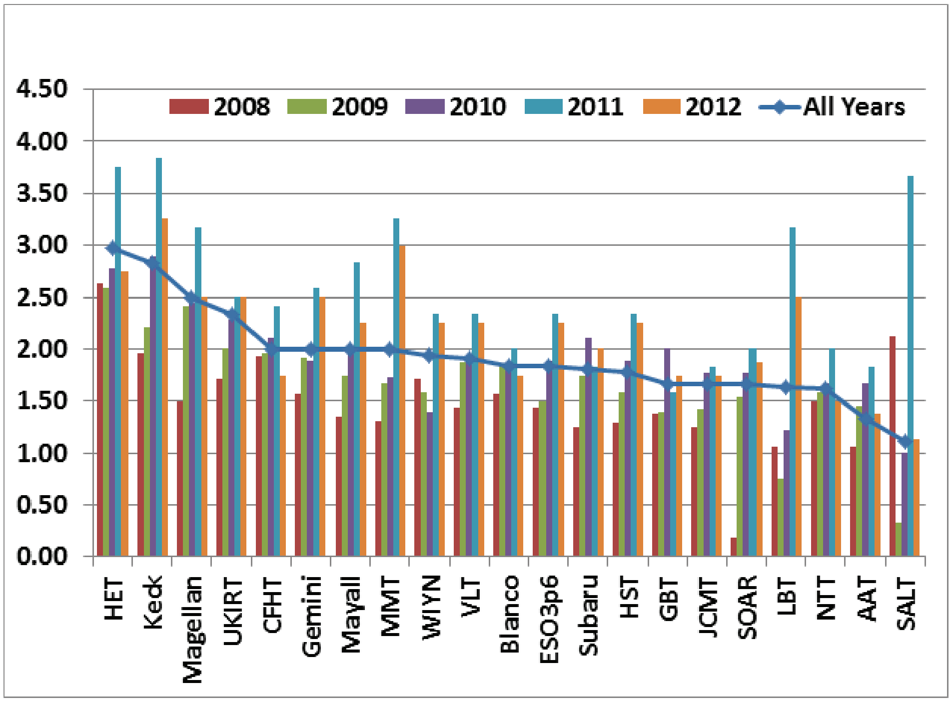

Figure 6. Median Impact per Paper (AIPP)

of citations — and a few papers with impacts of ten times or more the impact of the median Astronomical Journal paper. It is interesting to try and quantify the inequality of this distribution.

In economics the Gini coefficient is a measure used to quantify the income distribution in a country. The coefficient varies between 0, which reflects complete equality and 1, which indicates complete inequality. One can examine the inequality level of the impact of a telescope's publications by calculating the Gini coefficient of the impact distribution. The Gini coefficients vary between 0.5 an 0.6 (with three telescopes slightly below 0.5), which indicates a fairly high level of inequality.

7. Conclusions

Telescope bibliographies provide a record of a telescope's contributions to knowledge and the return on the large investment required to construct and operate large modern telescopes. The productivity and impact of a telescopes as measured by paper counts and citations, provide a quantitative measure of the return on the investment.

Including information in a telescope bibliography is very useful as the productivity and impact of individual instruments can be examined. This type of analysis could be useful when determining which instrument(s) should be decommissioned to make way for a new one.

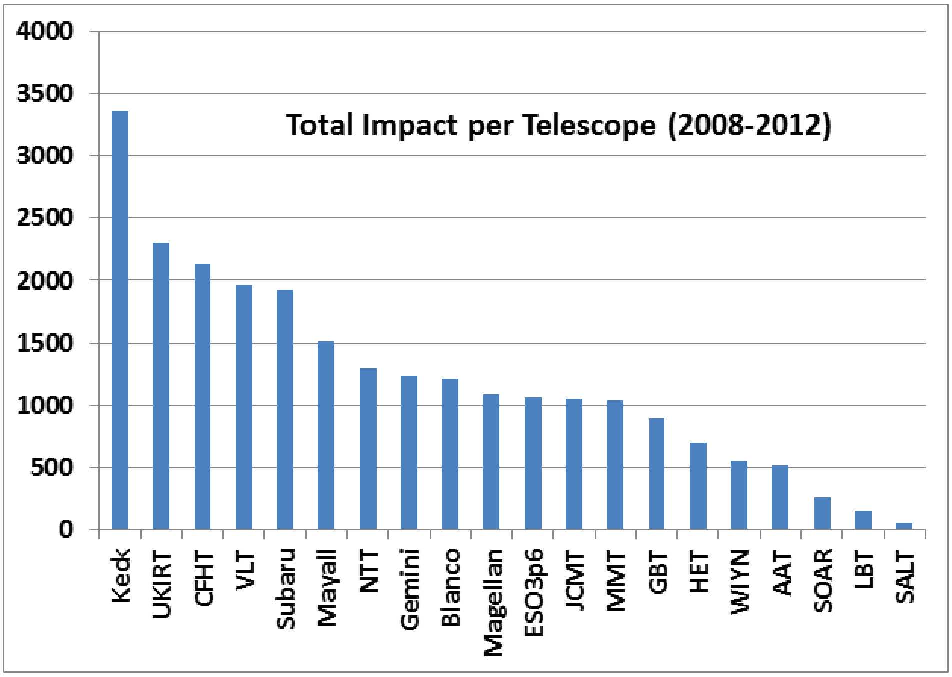

Figure 7. Total impact by telescope: 2008–2012

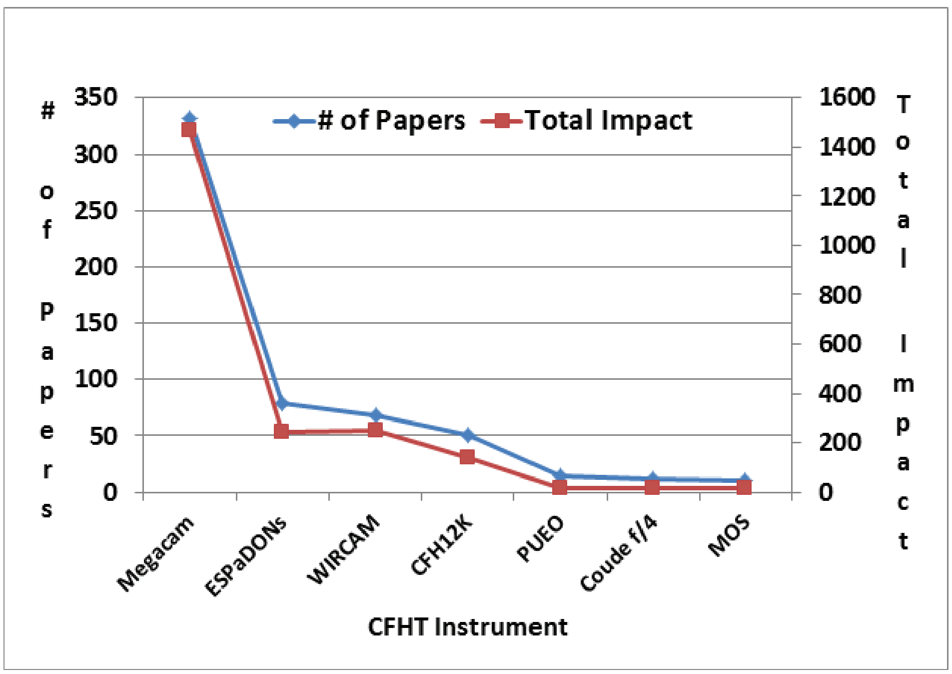

Figure 8. Productivity and impact of CFHT instruments