Scientific productivity and impact of large telescopes

Dennis R. Crabtree
Gemini Observatory, c/o AURA, Casilla 603, La Serena, Chile

ABSTRACT

The primary scientific output from an astronomical telescope is the collection of papers published in refereed journals. A telescope's productivity is measured by the number of papers published which are based upon data taken with the telescope. The scientific impact of a paper can be measured quantitatively by the number of citations that the paper receives. In this paper I will examine the productivity and impact of the CFHT, Gemini, Keck, Magellan, Subaru, UKIRT and VLT telescopes using paper and citation counts.

Keywords: Productivity, impact, citations

1. INTRODUCTION

The most important output from a modern observatory is the collection of papers based on its data that are published in refereed journals. These papers represent the facilities contribution to knowledge and the return on the capital investment for the construction of the telescope and its instruments. As astronomical observatories have become progressively more expensive the return on investment has come under closer scrutiny. Increasingly, bibliometric measures – the number of publications and the number of citations – are used to measure the quantity and quality of the output of modern observatories.

Productivity, as measured by the number of publications, and impact, as measured by citation counts, are metrics that can be used for many purposes. They can be used to evaluate the performance of a telescope, an individual, a university department or even a country. For example, Blustin(1) used bibliometric measures to compare astronomy groups in the UK. Citations must be used very carefully as they are only one indicator of impact, and an imperfect one. However they are the best quantitative measure that is currently available for studying the impact of papers published in refereed journals.

Abt(2,3) was the first to analyze astronomical publications using bibliometric techniques. One of his goals was to compare public and private American observatories, both in the level of their output and their impact on astronomical research. Trimble(4) studied the papers published between January 1990 and June 1991 that were based upon data obtained with telescopes with apertures of 2-m or greater. She used citation counts to these papers in 1993 as a measure of their impact. Trimble et al.(5) performed a similar study of papers published in 2001 based upon data from ground-based optical/IR telescope (as well as HST and JCMT), and using citations in 2002 and 2003. Benn and Sanchez(6) used the 125 most-cited papers in each year between 1991 and 1998 to compare the impacts of telescopes worldwide.

The studies mentioned in the preceding paragraph are based upon various subsets of the papers produced by optical/IR telescopes. They are either snapshots of papers published during a short period or utilize a subset of papers based upon citation counts. In this paper I will investigate and compare the productivity and impact of several large ground-based optical/IR telescopes as well as HST using complete publication lists that cover a significant time period. The ground-based telescopes included in this study include CFHT, Gemini, Keck, Magellan, Subaru, UKIRT and the VLT.

2. DATA

The raw data for this study are lists of papers in refereed journals compiled by each observatory. Observatories generally maintain a list of papers on the Web that they consider being publications based on data from their telescope(s) and these were generally the source of the data used. The librarians at Keck and CFHT were kind enough to provide me the data for their publication lists in an electronic format which made data ingestion much easier. In general I relied on the observatory to provide an accurate list of their publications. The number of publications in this study is too large to check each paper individually and ensure that it meets a basic eligibility criterion. I have identified a very small number of papers in some of the lists which I felt did not qualify as an observatory publication. However, the number is small enough that it is very unlikely to affect any of the results presented in this paper. The fact remains, that I rely on the integrity of the lists generated by each observatory.

The papers analyzed in this paper include those published though the end of 2006. The citation counts are as of January, 2008. The first publication year for each observatory and the total number of papers included for the period through 2006 are indicated in Table 1.

Table 1. Details on the papers included in this study

Observatory	Year of First Paper	Total Number of Papers
CFHT	1980	1434
Gemini	2000	292
HST	1991	5250
Keck	1994	1683
Subaru	2000	338
UKIRT*	1992	986
VLT	1999	1685

* Note that while UKIRT's first papers were published in 1981, the listings on the UKIRT web pages only include papers from 1992 onward.

The publication information (title, authors, year, journal, volume, page) for all observatory papers are stored in a Microsoft Access database. The publication information for each paper is checked by verifying the resulting bibliographic code (bibcode) against NASA's Astrophysics Data System (ADS) database. The bibcode is a 19 digit identifier which describes the journal article and can be used to retrieve other information such as citations from the ADS. If a publication's bibcode cannot be verified, the publication information is checked. The error is usually caused by an incorrect journal volume, page, or year. Software written in Visual Basic for Applications (VBA) within the MS Access database is used to automate the tasks that utilize the ADS.

Once the bibcode for each publication is verified, the full title, list of authors and range of pages is retrieved from the ADS. The citation count and the number of self citations are also retrieved from the ADS. As I will discuss later, the number of authors on papers is increasing as astronomical research is increasingly undertaken by teams. In identifying a self citation, I use a very broad interpretation. If any author on the citing paper matches any author on the cited paper I count this as a self citation. This has the effect of removing citations by team members of their own papers.

When, as often is the case, a paper is counted by more than one observatory I give each observatory full credit for the paper. Division of the credit (citations) between different telescopes is subjective and with over 10,000 papers in this sample, a careful reading of every paper was infeasible.

A paper accumulates citations as it ages. Previous studies [7] have shown that the peak citation rate (citations/year) peaks approximately two years after publication and declines after that. The accumulating citation counts for papers makes it very difficult to compare papers published in different years. A paper with 40 citations after one year is likely to be having more impact that a paper with 40 citations after 12 years even though they have the same number of citations at this moment.

In order to account for this age effect in the raw citation counts, I determine a paper's impact factor (hereafter called *impact*). A paper's *impact* is determined by dividing the number of citations to the paper by the median number of citations to all Astronomical Journal (AJ) papers published in the *same year*. For example, assume the median AJ paper in 2003 has 15 citations. A 2003 paper with 45 citations has an impact factor of 3.0. This approach treats the median AJ paper as a standard measuring stick (which grows with time as citation counts increase) against which to measure all papers.

3. OBSERVATORY PRODUCTIVITY

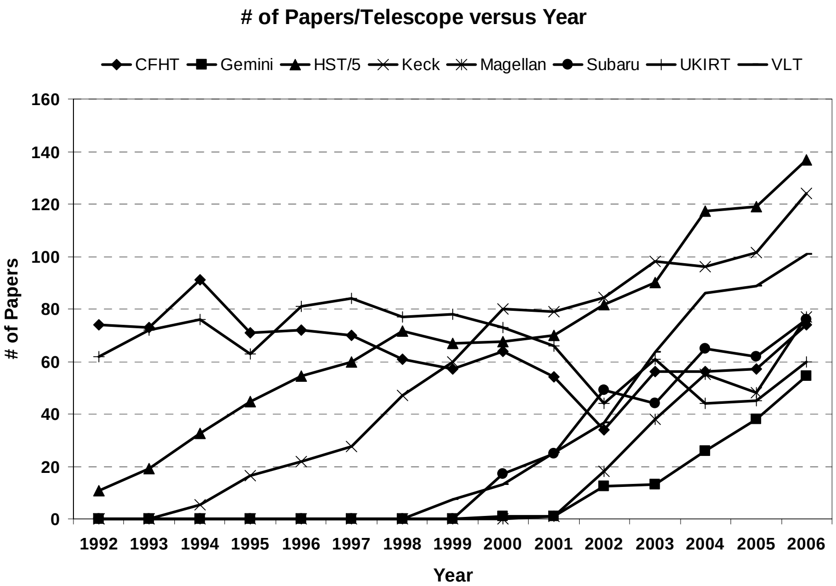

Figure 1 The number of publications per telescope per year for several observatories

Observatories have a tradition of tracking the refereed papers based on data from their telescopes. This list often appears in annual reports and now can almost always be found on observatory websites. The number of papers per telescope for the observatories in this study for the period 1992-2006 is shown in Figure 1. Note that the number for HST is divided by 5 for display purposes. HST is a remarkable paper producing machine with almost 700 papers published in 2006, over half of which were based at least in part on archival data. (This was determined by running simple queries on STScI's HST bibliography page.) One can see how the number of papers for a telescope ramps up after it first begins producing papers. For example, the number of Keck papers continues to increase until it plateaus between 2000 and 2002. The Keck I telescope's first paper was published in 1994 while Keck II's first paper was published in 1997. The oldest two telescopes, CFHT and UKIRT, included in this sample began operations around 1980. Their paper production appears to have reached a plateau at 80 papers per year by the early 1990s. The newer telescopes, including Keck appear to be headed to a plateau that is significantly higher than CFHT or UKIRT.

It is interesting to compare the rate at which new telescopes produce refereed papers as they ramp up their operations. As can be seen in Figure 1 from the newer telescopes, telescope productivity ramps up quickly with age.

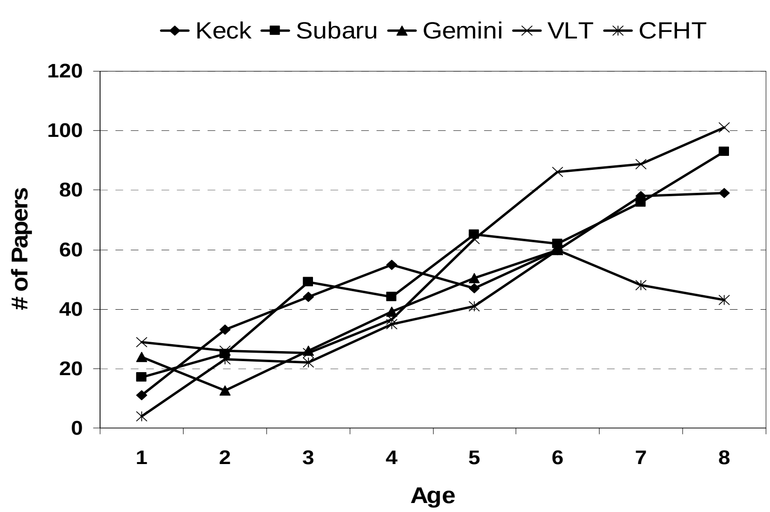

Figure 2 The number of papers per telescope per year as a function of the observatories age as measured by the time from the first significant paper output

Figure 2 shows the number of papers per telescope, for CFHT, Gemini, Keck, Subaru, and the VLT, as a function of the number of years (age) after their first significant number of papers. In the case of multi-telescope observatories I have estimated the age at which the second (or third and fourth) telescopes began producing papers. To first order all these observatories have increased their productivity at the same rate at their telescopes matured. The VLT's early productivity per telescope was initially close to flat as it brought the second, third and fourth telescopes on line. Once all the telescopes were online it appears as though the VLT's productivity ramped up at a slightly faster pace that the other telescopes included here.

CFHT's productivity, after increasing for the first six years, actually fell in year's seven and eight, before increase to the 80 papers per year level as indicated in Figure 1. The reason for this dip may be aging instrumentation, which is not be quite as competitive with instruments available at other telescopes.

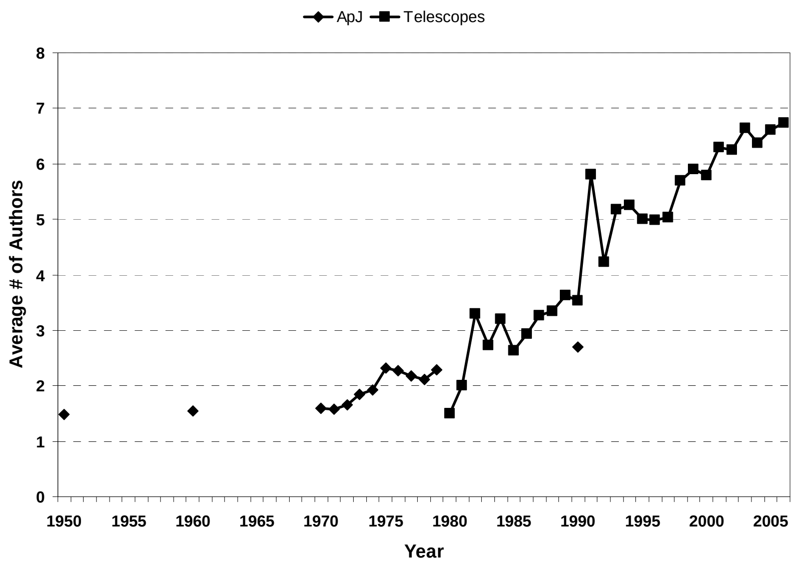

Figure 3 The average number of authors per paper for observatory papers included in this study and for several years of ApJ. Note the rapid linear increase in the average number of authors per paper.

As mentioned earlier, it appears as though the current frontline ground-based observatories' productivity may be reaching a plateau that is higher than that reached by the previous generation of ground-based telescopes. Why might this be the case?

Figure 3 shows the average number of authors per paper for two datasets. The first dataset consists of the observatories included in this study. The data through from 1980 to 1992 is for CFHT only but after that this dataset includes an increasing number of observatories. Note the remarkable linear increase in the number of authors from 1980 to 2006. While in 1980 there were on average 2.5 authors per paper, by 2006 that number had increased to almost seven.

To compare this trend with the general literature and to extend it back in time, I determined the average number of authors per paper for ApJ papers for several years. This is shown in Figure 3 by the dataset labeled ApJ. This dataset shows that the trend of increasing number of authors is indicated in the general literature and that in 1950 the average ApJ article had 1.5 authors. The average number of authors on ApJ papers is less than that on observatory papers (for the same year) as theory work is still done by individuals or small teams.

This rapid increase in the average number of authors per paper indicates a move towards more research being undertaken by scientific teams as opposed to individuals or small groups. This increase in team size is likely related to the larger datasets produced by modern instruments and the fact that many papers are based multi-wavelength data that require a range expertise for reduction and analysis.

The increased productivity of the new generation of telescopes may be higher because they produce larger datasets which result in more papers. Another factor may be the increased size of the teams which are able to produce papers faster than the individuals or small groups that dominated in the past.

This result that teams are playing a more important role in research in not restricted to optical observational astronomy. A recent study [8] of over 19.9 million papers over 50 years has demonstrated that teams are now dominant over single authors in the production of knowledge. Their work shows that teams in general produce more highly cited research that individuals. I will examine the impact of teams in observational astronomy research in the next section.

4. OBSERVATORY IMPACT

The productivity of an observatory is obviously an important metric of an observatory's performance. Increasingly however, the impact of the published research is being recognized as the more important metric. How important or valuable is the contribution of a research paper if it is never or infrequently cited?

The number of citations to a paper is usually considered a good quantitative measure of a paper's impact. While not a perfect measure, it is the best quantitative metric available for measuring 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 other factors such as area of research and the culture of each particular sub-field. However, 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 length of the paper and the number of authors (team size).

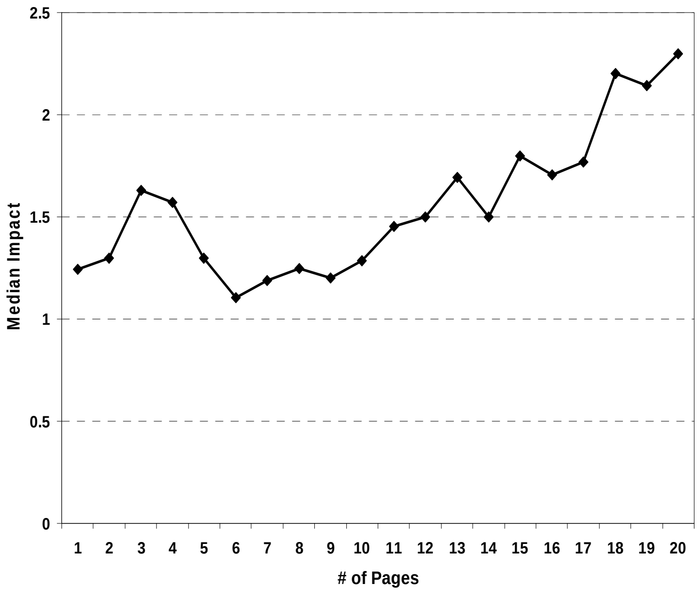

Figure 4 The median impact of a paper as a function of the length of the paper. This includes all of the papers from all of the observatories included in this study.

Figure 4 shows the median impact of papers as a function of the length of the paper. The median is used rather than the mean as the distribution of impact has a long tail towards very high impact papers which significantly affects the mean. The median is not affected by a relatively small number of very high impact papers. The *Letters* effect if clearly visible as the median impact shows a local maximum for articles of 3-4 pages, the usual maximum for a *Letter* in various journals. There are over 1900 papers with a length of four pages, 845 with a length of 5 pages and the number of papers decrease with length until there are 173 papers with a length of 20 pages. Clearly longer papers are more relevant the researchers working in the same field.

Another interesting correlation to investigate is that between impact and the number of authors. As discussed earlier teams are producing proportionately more papers than in the past. How does the impact of papers produced by teams compare to the impact of papers produced by individuals and small groups? Figure 5 shows the median impact of all the papers in this study as a function of the number of authors (team size). First, it is interesting to note that there are more papers in this study that

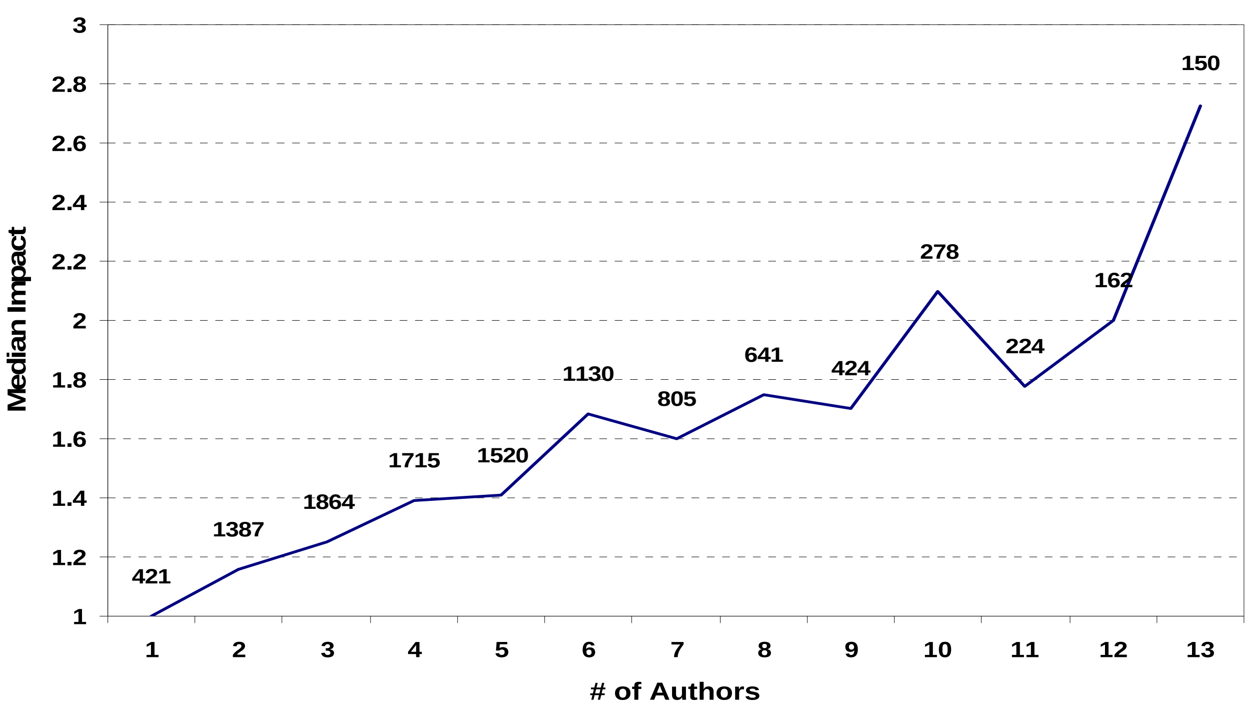

Figure 5 The median impact of all the papers in this study as a function of the number of authors. The number above each point is the number of papers included with that number of authors.

have 9 authors than have a single author. This is another clear indication of the increasing predominance of teams in the area of observational astronomy. It is also clear from Figure 5 that the impact of research papers is a strong function of the size of the team. Larger teams produce papers that are of relevance to a larger number of researchers, and research teams, than papers produced by smaller groups. Papers with a larger number of authors are almost always based on larger datasets and are more likely to include data from more than one facility (including ones not included in this study). Recall that the impact has been adjusted for self-citations so this is not simply the effect of team members citing team papers and hence increasing the impact. This increasing role of larger teams is a trend that will likely continue in astronomy as large *experiment* type instruments are developed to address particularly difficult and significant questions.

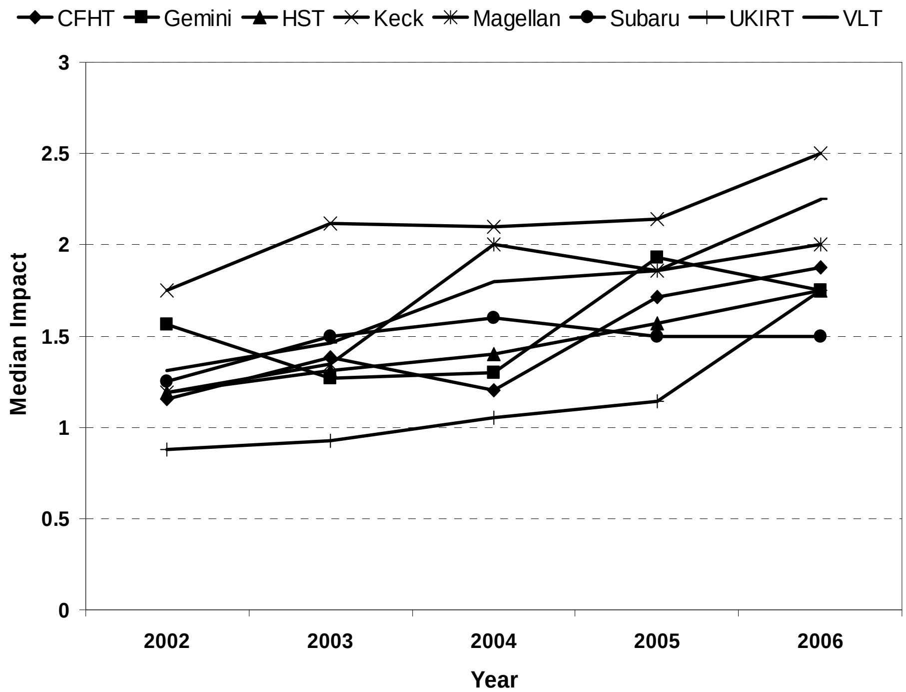

Figure 6 The median impact of observatory papers as a function of year

How do the various observatories compare in the impact of their publications? Figure 6 shows the median impact of observatory papers for the period 2002 – 2006. Again, the median is used rather than the mean to lessen the impact of a small number of very high impact papers.

First note that Keck has the largest median impact for all years. Keck is clearly producing papers that are of the most

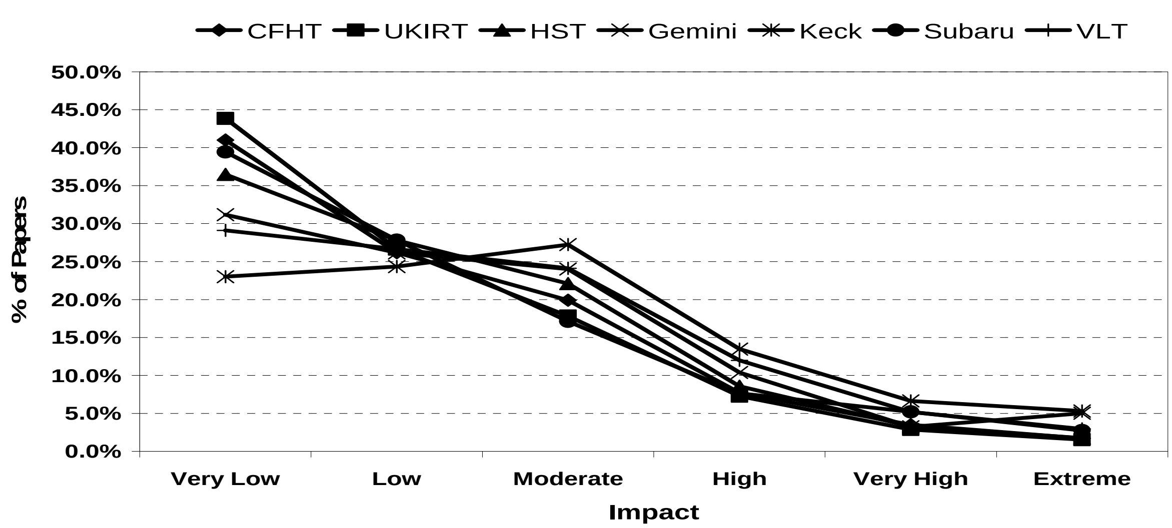

Figure 7 The Impact Distribution Function (IDF) for the indicated observatories. Generally an observatory with a flatter IDF is performing better.

relevance to the broad astronomical community. In 2006 the median Keck paper had twice the number of citations as the median AJ paper of 2006. The VLT appears to becoming relatively stronger as its median impact put it second behind Keck in 2006 while in previous years it third.

Interestingly, while producing about 5 times as many papers as a ground-based telescope, the median impact of an HST paper is lower than that of Keck or the VLT. Of course HST's total impact, i.e., the sum of the individual impacts of each paper, is significantly higher that the other telescopes because of its large productivity. UKIRT's median impact increased significantly in 2006, probably as a result of the initial papers from the UKIDS project being published.

A version of Figure 6 that uses mean impact rather than median looks quite different. In 2006, CFHT has the largest mean impact per paper while when using the median it is the lowest. This is because the effect of one extremely highly cited paper is removed when using the median while it influences the mean significantly.

It appears that it is tricky (and risky) to quantify an observatories impact by a single number such as mean or median impact per paper. An approach that captures the range of impact of observatory publications, including the vey high impact papers, would give a more complete picture of observatory performance.

I calculated the fraction of the total impact of each observatory's papers in six bins of impact which I labeled from *Very Low* to *Extreme*. Papers with an impact factor less than one are considered to be of very low impact while those with impact factors above eleven are considered to be of extreme impact, while the other bins include papers with impact factors in between. The Impact Distribution Function (IDF) is a plot of the fraction of papers in each of the six impact bins.

In general, an observatory is performing better if it has a smaller percentage of lower impact papers and a larger percentage of higher impact papers. This would be exhibited as a flatter IDF. As can be seen in Figure 7, Keck's IDF is characteristically different from the IDFs of the other observatories. All the observatories have an IDF that peaks in the *Very Low* bin and decreases monotonically towards a low in the *Extreme* bin. Keck has the lowest percentage of *Very Low* impact papers and the highest fraction of papers with *Moderate* to *Extreme* impact. HST's IDF is very similar to the other ground-based telescopes included. HST produces a large fraction of Very Low and Low impact papers as do the other telescopes. HST has a larger number of *Very High* and *Extreme* impact papers because of its high production rate, not because a larger fraction of its papers fall in these two bins.

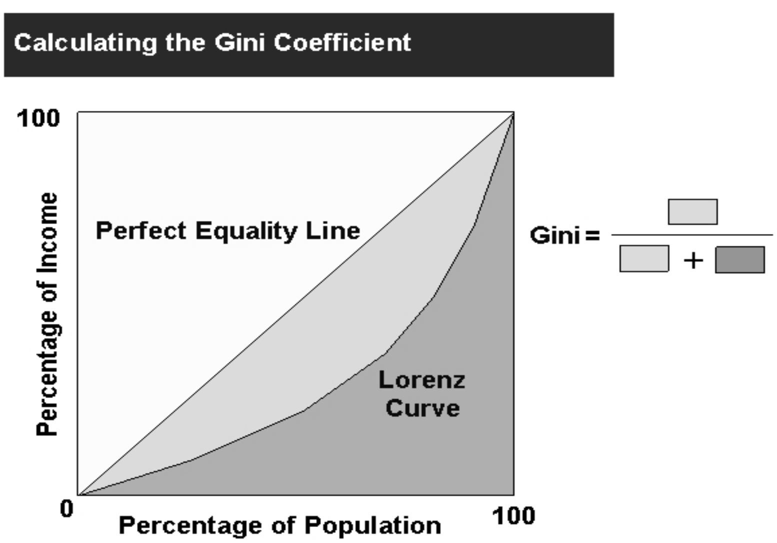

Figure 8 A graphical display of how the Gini coefficient is calculated. In applying this to impact, 'Population' is replaced by 'Papers' and 'Income' is replaced by 'Impact'

A novel approach to quantify the statistical dispersion in the impact distribution of papers from an observatory is to use the Gini coefficient [9]. The Gini coefficient is usually applied in economics to quantify the inequality of income or wealth distribution. It is defined as a ratio with values between 0 and 1 --- a low Gini coefficient indicates more equal income or wealth distribution, while a high Gini coefficient indicates more unequal distribution. A zero corresponds to perfect equality (everyone having exactly the same income) and 1 corresponds to perfect inequality (where one person has all the income, while everyone else has zero income). The Gini coefficient can also be quoted as a number between 0 and 100 and is then referred to as the Gini index.

Figure 8 indicates graphically how the Gini coefficient is calculated for a income distribution. The Lorenz curve shows the actual distribution of income (impact) while the line with a slope of 1.0 shows the distribution for the case where income (impact) is equally distributed.

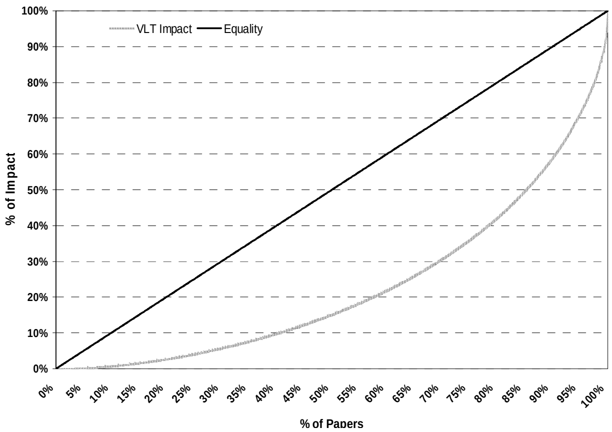

Figure 9 The distribution of impact of individual papers for the VLT. The dotted line indicates the actual distribution while the solid line indicates what the distribution would look like if each paper had exactly the same impact

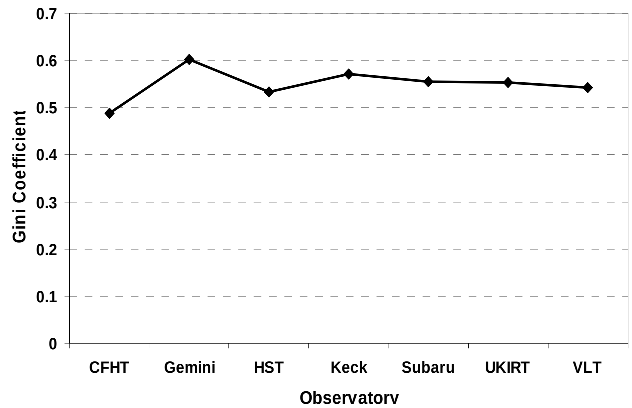

Figure 10 The Gini coefficient for impact distribution of papers for each observatory

In using the Gini coefficient, a zero would indicate that all papers have the same impact factor. It says nothing about the absolute level of impact only how equal the distribution of impact is amongst all the papers. The Gini coefficient for countries ranges from around 0.25 to above 0.6 for the countries with the most unequal income distribution. Move developed countries generally have lower Gini coefficients. The Gini coefficient for the US has risen from around 0.39 in the late 1960's to 0.47 in 2006.

The Lorenz curve for the impact of VLT publications is shown in Figure 9 along with the line indicating perfectly equal distribution of impact. The fact that the distribution of impact is far from equal is not a surprise given the IDF shown in Figure 7. The Lorenz curve for VLT papers shows that the lower 50% of papers, in terms of impact, produce less than 15 percent of the total impact of VLT papers. The top 20% of papers produce approximately 80% of the total impact. The Gini coefficient for the VLT impact distribution is 0.54, which indicates a very unequal distribution of impact. Note that for income distributions in countries, a Gini coefficient of 0.54 is that of a developing country.

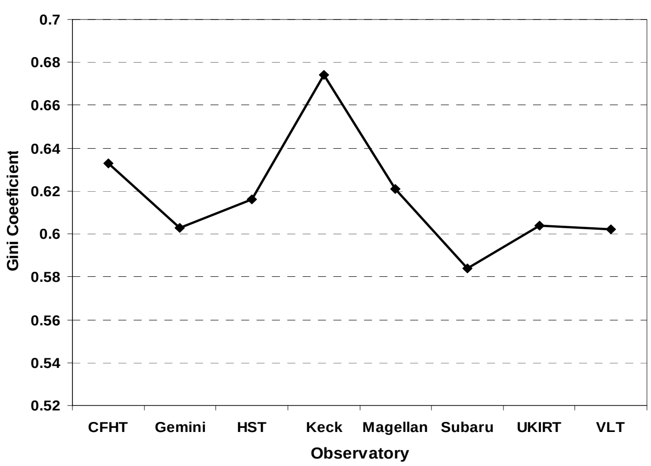

Figure 11 The Gini Coefficient for the distribution of impact by first author for various observatories.

The Gini coefficients for most of the telescopes in this study are indicated in Figure 10. All observatories exhibit Gini coefficients of between 0.5 and 0.6. The most equal distribution of impact belongs to CFHT while the most unequal distribution of impact is that of Gemini. The distribution of impact of observatory publications is far from equal with approximately 80% of the impact being produced by the top 20% of published papers.

Another approach to studying the distribution of the impact of observatory publications is to aggregate impact by first author, i.e., sum the impact of all papers for a given author. One can then investigate how the impact of an observatory's papers is distributed amongst individuals as first authors. Figure 11 shows the Gini coefficients for the distribution of impact of observatory papers by first author. If the total impact of each first author of an observatory's papers was equal, then the Gini coefficient in Figure 11 would be zero. If all of the impact was produced by one author then the Gini coefficient would be one.

Impact across authors is most evenly distributed for Subaru while the distribution is most unequal for Keck authors. All of the observatory Gini coefficients are quite high indicating a very unequal distribution of impact across authors.

5. CONCLUSIONS

This paper has investigated the productivity and impact of a number of optical/IR telescopes. A new telescope's productivity ramps up quickly once publications start appearing based on data from the telescope, with all telescopes demonstrating a very similar rate of increase. A plateau in productivity is reach seven to eight years after the initial publications. A telescope's productivity can be rejuvenated by new instrumentation or, as is the case for HST, having a large number of papers based on archival data. HST is a paper generating-machine producing approximately five times as many papers as a ground-based telescope.

One interesting result of this work that is unrelated to observatory productivity is the increasing size of the teams publishing papers based on observatory data. The average number of authors on a paper is now close to seven which is more than double the number from 25 years ago. This trend of the increasing importance of teams in observational astronomy shows no signs of changing. The immense datasets generated by large panoramic detectors and the increasing use of multi-wavelength datasets require more expertise and a larger number of team members to work effectively with the data.

The Impact Distribution Function (IDF) is a good approach for quantifying the impact of an observatory. The IDF provides a measure of the number of low performance papers as well as the number of high performance papers, unlike a single number metric such as median or mean impact. The IDF for Keck papers shows that Keck produces a significantly smaller fraction of very low impact papers, while producing relatively more papers with higher impact. The IDF for HST shows that it produces a significant fraction of very low impact papers and a relatively small fraction of high impact papers.

The distribution of impact across an observatory's papers, as indicated by the Gini coefficient, is very unequal with approximately 20% of the papers producing 80% of the impact. This same analysis applied to the distribution of impact across authors also shows that impact is distributed very unevenly across authors. A relatively small number of authors produce the majority of the impact from observatory publications.

6. ACKNOWLEDGEMENTS

I thank Liz Bryson and Peggi Kamisato, the librarians at CFHT and Keck respectively, for providing their publication lists in electronic form. This research has made use of NASA's Astrophysics Data System.

7. REFERENCES

[1] Blustin, A., "Publication and citation statistics for UK astronomers", *Astronomy & Geophysics*, **48**, 32-35 (2007).

[2] Abt, H.A., "Long-term citation histories of astronomical papers", *Publications of the Astronomical Society of the Pacific*, **83**, 207-210 (1981).

[3] Abt, H.A., "An Assessment of Research done at the National Optical Observatories", *Publications of the Astronomical Society of the Pacific*, **97**, 1050-1052 (1985).

[4] Trimble, V., "Papers and citations resulting from data collected at large, American optical telescopes", *Publications of the Astronomical Society of the Pacific*, **107**, 977-980 (1995).

[5] Trimble, V., Zaich, P., Bosler, T., "Productivity and impact of optical telescopes", *Publications of the Astronomical Society of the Pacific,* **117**, 111-118 (2004).

[6] Benn, C.R., and Sanchez, S.F., "Scientific impact of large telescopes", *Publications of the Astronomical Society of the Pacific*, **113**, 385-396 (2001).

[7] Crabtree, D.R. and Bryson, E.P., "The effectiveness of the Canada-France-Hawaii telescope", *Journal of the Royal Astronomical Society of Canada*, **95**, 259-266 (2001).

[8] Wuchty, S., Jones, B.F., and Uzzi, B., "The increasing dominance of teams in production of knowledge", *Science*, **316**, 1036-1039 (2007).

[9] See Wikipedia, "Gini Coefficient", http://en.wikipedia.org/wiki/Gini_coefficient, (2008)