Using the *h*-index to Explore the Scientific Impact of the VLT

Uta Grothkopf, Claudio Melo, Christopher Erdmann, Andreas Kaufer, Bruno Leibundgut (all ESO)

The productivity and scientific impact of observatories and individual instruments are one measure of their success. This article presents the results of a study where we have applied the *h*-index, previously proposed for individual researchers, to major ground-based observatories (VLT, Keck, Gemini, Subaru) as well as individual VLT instruments. The concept is expanded by exploring the time-dependence of the *h*-index *h(t)*. Overall, the VLT appears to be among the most successful 8-m-class telescopes. We also show that ESO instruments are making important contributions to progress in astronomy.

Introduction

In order to examine their return on investment, major observatories around the world have developed metrics to trace their scientific output. Such metrics often focus on the observatories' productivity and impact in the scientific community. These two factors are typically measured through the number of scientific publications based on astronomical data and the citations these publications generate, respectively. The methodologies used to compile the ESO Telescope Bibliography, a database that lists all papers based on ESO data, as well as publication and citation statistics derived from this database have been documented in Leibundgut, Grothkopf and Treumann (2003) and Grothkopf et al. (2005). The Telescope Bibliography is publicly available via the web (*http://www.eso.org/libraries/telbib.html*) as well as through the "Select References In: ESO/Telescopes" filter at the ADS (see Delmotte et al. 2005 for more details).

In a recent paper, Hirsch (2005) proposed a new index to measure research output, the so-called *h*-index. While originally meant to analyse the productivity of individual researchers, we recently introduced it into telescope statistics (Grothkopf and Stevens-Rayburn 2007). Based on *h*, we developed *h(t)* which reflects changes of the *h*-index in the course of time. In this paper, we apply *h* and *h(t)* to selected observatories as well as to the VLT instruments.

There are several caveats to bibliometric studies, in particular when used for comparison across various institutes. Despite attempts to synchronise the methods applied to compile science bibliographies, criteria for paper selection are still defined by the individual observatories and are therefore not identical. Similarly, methodologies for building telescope bibliographies vary (for instance, retrieval of relevant publications through database (ADS) searches alone, through screening of paper journals, etc.). Even more importantly, comparing telescope statistics is problematic because of the different features and ways of operation of ground-based and space-based telescopes, different apertures and numbers of instruments, different wavelengths of observations, etc. Any comparison has therefore to be interpreted with utmost care in order to avoid unbalanced or wrong conclusions.

Use of bibliometrics by observatories

In preparation for a presentation given at the IAU General Assembly in Prague, Czech Republic in August 2006, we conducted a survey among large observatories to better understand what kind of telescope statistics are compiled at observatories, at which intervals, and by whom.[1] All respondents regularly gather total and/or average numbers of their refereed publications, the majority also monitor unrefereed publications (e.g. conference proceedings), either regularly or on request. Citation statistics of refereed publications are collected by all respondents, even though only half of them does so on a regular basis, the remaining 50 % only on request.

As can be expected, observatories compile statistics tailored to their individual needs, for instance number of publications and citations per instrument, programme type, observing cycle, observing mode, etc. In addition, almost all respondents investigate their high-impact papers ("most productive instrument, programme, individual authors", etc.) A study of highly cited papers and their distribution among facilities is carried out every year in April at the Space Telescope Science Institute (Madrid and Macchetto 2007).

All methodologies used by observatories have some advantages and some disadvantages: counting publications measures the productivity of facilities, but does not indicate whether or not these publications actually have any influence on the advancement of astronomy. Looking at the numbers of citations does indicate the impact among the astronomical community, but values can easily be inflated by a few extremely highly cited papers. Investigating so-called high-impact papers and their distribution across observatories is less biased, but retrieving such statistics is by far not as straightforward as mere publication and citation counting.

Several in-depth studies have been carried out by Trimble et al. who investigated productivity and impact of optical, space-based, and radio telescopes, respectively (Trimble, Zaich and Bosler, 2005, 2006; Trimble and Zaich 2006). The authors analyse papers, citations and impact factors of articles from 18 journals regarding their distribution among facilities, thus avoiding the bias that typically can be noted in studies from individual observatories.

The *h*-index

In order to overcome some of these disadvantages, J. E. Hirsch of the University of California at San Diego suggested a new and surprisingly simple measure which he called the *h*-index (Hirsch 2005). While the *h*-index originally was meant to be a measure for the research output of individual scientists, we extend it here to observatory publication statistics. In order to determine this value, one needs a list of all relevant papers, ranked by decreasing citation counts; *h* can then be found where the citation count is at least as high as the rank. Thus,

[1] The following observatories responded to our questionnaire (*http://www.eso.org/libraries/telstats-questionnaire.html*): CFHT, Chandra, Gemini, HST, Isaac Newton Group, Keck, NRAO, Subaru, XMM Newton. For more information, contact Uta Grothkopf at *esolib@eso.org*.

the *h*-index combines measure of productivity (number of publications) and impact (average number of citations per paper) and is therefore more balanced than most other measures used for bibliometric studies (Figure 1).

Instead of measuring the research output of individual researchers, the concept can also be applied to entire observatories or specific observing facilities – always bearing in mind the caveats described above. In a recent paper, we introduced the *h*-index into telescope statistics by computing it for selected publication years of some major observatories (Grothkopf and Stevens-Rayburn 2007).

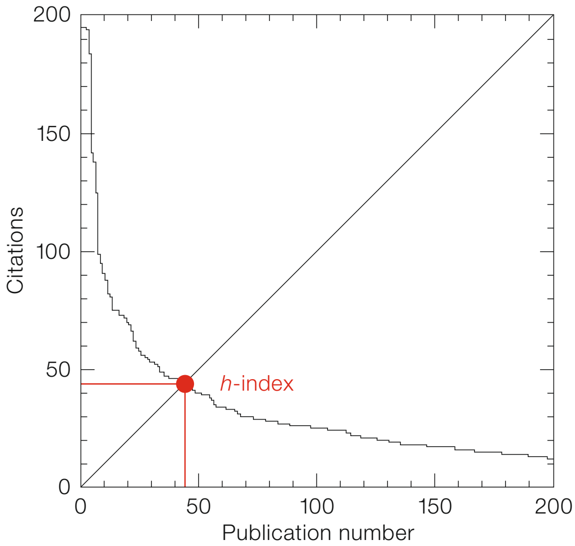

Figure 1: The *h*-index can be found where the counted publication number equals the number of citations. Neither a large number of publications with low citation rates (right edge of the x-axis) nor individual papers with extraordinarily high numbers of citations (upper edge of the y-axis) will alter *h*, therefore this value is more balanced than many other bibliometric measures.

The *m*-parameter

It must be noted that, tempting as it may be, comparing *h* alone among observatories (or researchers) does not lead to a meaningful result. Facilities that have been operative for many years obviously had much more time to produce publications and accumulate citations, hence their *h*-index can be expected to be considerably higher than that of younger facilities. Hirsch therefore introduced the so-called *m*-parameter. For individual researchers, this value is computed by dividing *h* by the numbers of years since publication of the first paper. Correspondingly, when applied to observatories, *h* is divided by the number of years of operation; hence it reflects the various 'lifetimes' of facilities.

Observatory	Range of years of publications	Years since first publication	*h*	*m*
VLT	1999–2006	8	79	9.9
Keck	1996–2005	11	113	10.3
Gemini	2000–2006	7	33	4.7
Subaru	2000–2006	7	41	5.9

Table 1: Range of years of publications, number of years since first publication, as well as *h* and *m* of the observatories included in our study.

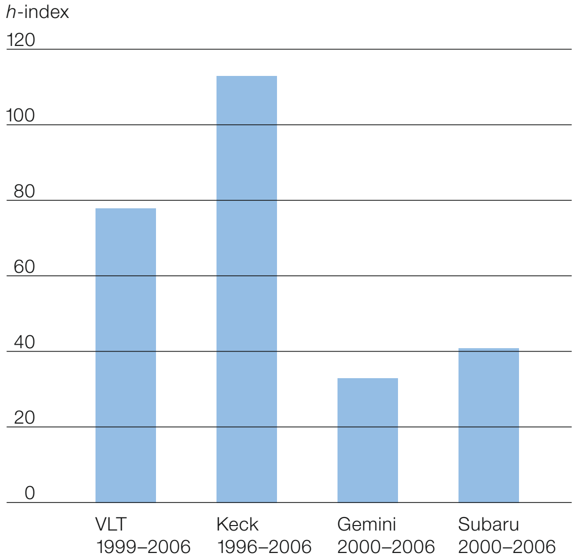

Figure 2: *h*-index for the VLT, Keck, Gemini, and Subaru observatories as of April 2007. Note that *h* correlates with the number of years of operation. Hence, older observatories tend to have higher *h* indices.

Applying *h* and *m* to observatories

We present here *h* and *m* values for the VLT, Keck, Gemini and Subaru observatories. In order to compute *h*, we obtained bibcodes of all papers pertaining to the observatories' publication lists in the following way: for the VLT, bibcodes are stored in the ESO Telescope Bibliography (*http://www.eso.org/libraries/telbib.html*). References of Keck and Subaru papers were retrieved from the web (Keck Science Bibliography at *http://www2.keck.hawaii.edu/library/keck_papers.html* and Publishing Results from Subaru, *http://www.naoj.org/Observing/Proposals/Publish/index.html*) and were translated into bibcodes. Papers using Gemini data were found using the "Select References In" filter on the ADS main search screen. These bibliographies include only refereed papers. The respective ranges of publication years are shown in Table 1. For uniformity, we end the range of publication years for all observatories in 2006. Citation counts for all publications were obtained from the ADS as of April 2007.

For each observatory, citation counts were then ranked in descending order, and *h* was computed. Figure 2 shows the *h*-indices of the respective observatories.

We also computed the *m*-parameter by dividing *h* by the specific number of years since the first publication (see Table 1). Resulting values are given in Figure 3. This brings the VLT and Keck to the top of the list, and also the young facilities Gemini and Subaru perform well.

One should note that both *h* and *m* depend on the number of telescopes per facility. A facility like the VLT with four telescopes will produce more papers more quickly and hence the *h*-index will increase faster, although it is not obvious exactly how much of an effect this is. For a facility like the VLT, with telescopes becoming operational over some time, it will be difficult to quantify this effect in detail.

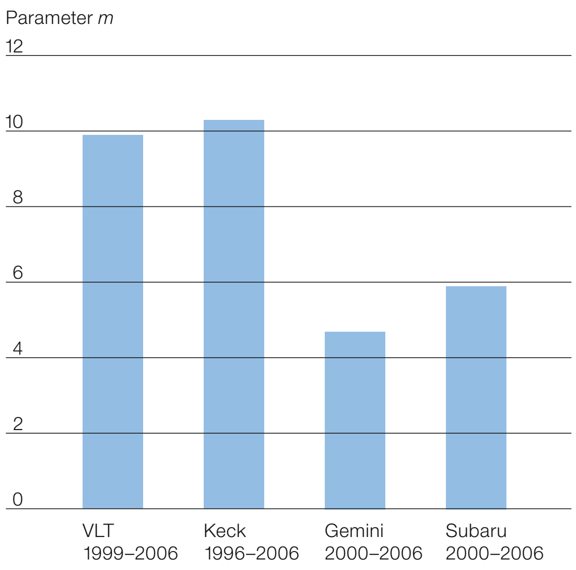

Figure 3: Parameter *m* of the observatories included in this study (as of April 2007).

The *h*-index *versus* time

Although the *m*-parameter reflects the total number of years of operation of a facility, both *h* and *m* are integrated values that don't show how these values were achieved in the course of time. We have therefore further developed the *h*-index to analyse its evolution, *h(t)*.

To compute *h(t)*, we obtained the citation history for each paper included in our study from the ADS. The citation history shows how many citations a paper generated in a given year. For instance, a paper published in 2002 will generate x citations in 2002, y citations in 2003, z citations in 2004, etc. In order to calculate *h* for a year *Y*, we add up all citations up to the year *Y*, using the citation history of each paper. The papers are then listed in decreasing order, and the *h*-index for that year *Y* is computed.

Figure 4 shows *h* over time applied to the observatories in our study. Trends indicated by *h* and *m* are confirmed. Keck has always been performing extremely well, with a steeply increasing *h*-index right from the start. The VLT started well and has even improved during recent years. Both Subaru and Gemini are now on their way up after a slightly slower start.

In this study, all observatories are treated as entities. This does not accurately reflect their early years. Keck I and II, Gemini North and South as well as VLT UT1, 2, 3, and 4 all came online sequentially, increasing the available observing time over the years. A future study may investigate *h(t)* based on the actual observing time of each observatory.

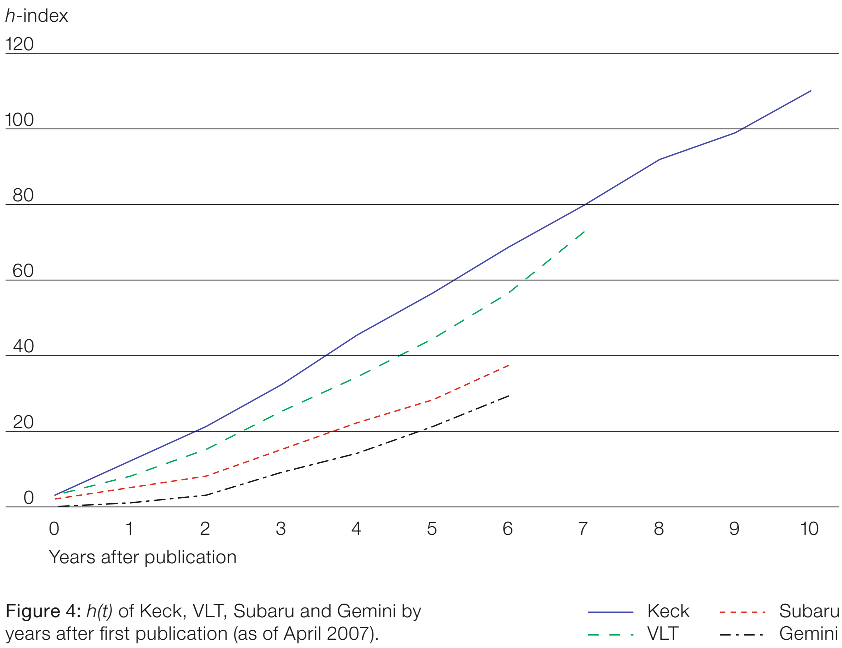

Figure 4: *h(t)* of Keck, VLT, Subaru and Gemini by years after first publication (as of April 2007).

Performance of the VLT instruments

In order to investigate the specific performance of the VLT instruments, we applied *h*, *m*, and *h(t)* to them. We restrict this analysis to the VLT because instrument-level information for La Silla papers has been systematically recorded only starting in 2002. The results are presented in Table 2. Although the absolute number of papers and citations differ, the very first instruments of the VLT (FORS1, ISAAC, UVES, FORS2) have a comparable performance. As mentioned above, the number of years in operation is an important factor. Young instruments necessarily have a lower *h*-index. This can be partially compensated for by normalising the *h*-index by comparing parameter *m*.

As done for the observatories, a better glimpse of how the instruments are performing can be achieved by measuring the *h*-index *versus* time. The results of *h(t)* as of April 2007 for VLT instruments are shown in Figure 5 relative to the year of the first paper of each instrument. In this graph, we included only VLT instruments that have been operative for two or more years and produced at least ten papers.

An obvious result is that the first-generation VLT instruments (FORS1, UVES, ISAAC, FORS2) are alive and kicking with no sign of changing their slopes. They clearly set a very high standard for future instruments.

With respect to the young instruments, we can say that they are … young! Their *h*-indices are growing at a slow pace. Besides this slow increase of the *h*-index for the young instruments, can we predict whether they will look like their older siblings in future?

We see from Figure 5 that most of the young instruments have a much harder time to take off than FORS1 and UVES, but they are not particularly different from the *h(t)* curves of FORS2 and ISAAC. But what is behind the plot shown in Figure 5? Does the *h(t)* have a practical use in judging instrument performance? Can it help observatories to decide whether a given instrument needs to be upgraded, improved or even decommissioned? The answer is not simple and the reader should take Figure 5 with caution.

As defined above, the *h*-index depends on the number of papers and their citation counts. These two quantities are certainly instrument-dependent. For example, detector sensitivity, spectral or spatial resolution, wavelength coverage, quality of the documentation and data-reduction software (pipelines). These are all factors that play a role in the way *h(t)* evolves.

The intrinsic nature of the instrument itself has an influence on its *h(t)*. All instruments have a difficult start with a low *h*. As time goes by the community gets used to the modes, pipeline, etc. At this time the number of papers and citations increases fast, and so does the *h*. Years later, all possible (hot) applications start to run out and the papers gather fewer citations. New (and hopefully better) instruments come. At this point the *h*-index

Instrument	*h*-index	Years since first publication	*m*	Number of papers	Total number of citations	Ratio
VLT						
FORS1	56	8	7.0	525	14267	27.2
ISAAC	49	8	6.1	413	9308	22.5
UVES	48	7	6.9	474	9326	19.7
FORS2	46	7	6.6	342	9006	26.3
NACO	22	5	4.4	111	1942	17.5
FLAMES	14	4	3.5	67	877	13.1
VIMOS	14	3	4.7	44	623	14.2
VLTI						
VINCI[2]	14	5	2.8	38	572	15.1
MIDI	7	3	2.3	19	226	11.9

Table 2: Performance of VLT and VLTI instruments that have been operative for two or more years and produced at least ten papers.

[2] Note that VINCI was offered to the community only from October 2002 to October 2003.

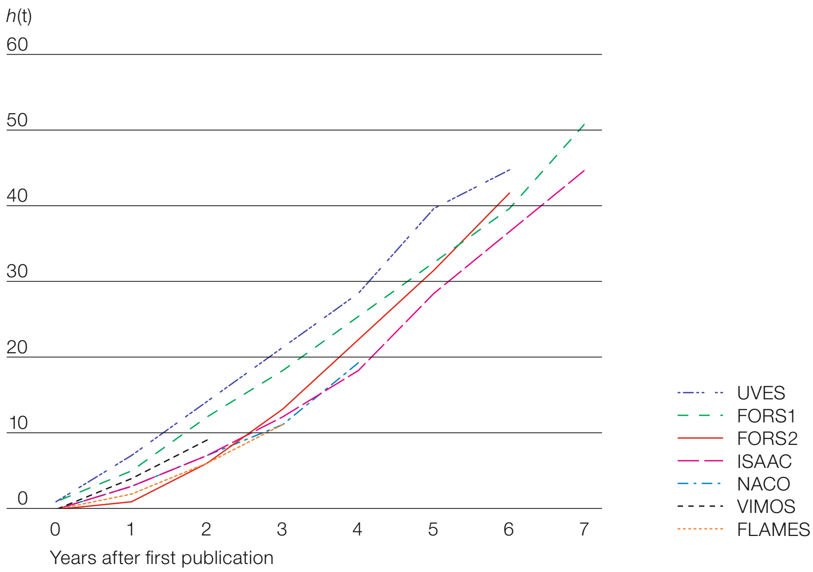

Figure 5: *h(t)* of selected VLT instruments by years after first publication (as of April 2007).

keeps growing, but at a much slower pace. Thus the *h(t)* would have a curve-of-growth shape (similar to those found in chemical abundances studies).

In the case of instruments designed for individual observations, papers can be based on a few observations and the full cycle (from proposal submission to paper acceptance) can be as short as one year. In contrast, papers presenting results based on a large amount of data collected by survey-like instruments (e.g., VIMOS and FLAMES) can take years to be released. But when these papers are finally published, the impact can be immense.

Another bias hidden behind the computation of *h(t)* has to do with the complexity of the new instruments, which in some cases are aimed at tackling problems that require the use of cutting-edge technology. Using a complex (new) technology usually implies large overheads. This is the case of AO instruments and even more pronounced in the VLTI observations whose complexity is of a higher degree than that of any other VLT instrument.

However, the *h(t)* can also be influenced by circumstances not related to the instrument at all. For instance, there is no question that one of the strengths of the first suite of VLT instruments (FORS1, UVES, ISAAC and FORS2) is its versatility, being used from Solar System to cosmological applications. In addition to this versatility, these instruments were favoured by the fact that they were the only ones available to the community during the first years of operation of the VLT. Moreover, they were mounted almost exclusively on each of the UTs. Nowadays, the situation is very different. New instruments have to compete for UT time. For instance, on Kueyen UT2, the time is shared between UVES, FORS1 and FLAMES. Therefore, in order to be fair, the *h(t)* curve shown in Figure 5 has to be corrected by the effective fraction of time used by each instrument.

The bottom line is that we should take the results presented in Figure 5 with caution since none of the potential biases discussed above were taken into account and therefore the results presented here are very preliminary. Having said that, it might be interesting to peruse the effort of using the *h*-index to measure instrument impact and performance in a deeper and more thoughtful study where all biases are considered.

Conclusions

The *h*-index is a simple, yet powerful indicator that has some important advantages compared to other bibliometric methods:
– it combines productivity and impact;
– it can be relatively easily determined using the ADS;
– it is neither affected by a large number of publications with few or no citations (which usually suggests high productivity, but not necessarily high impact) nor by extraordinarily high citations of only a few papers (which inflates citation counts).

The *h*-index is therefore more balanced than other measures if one bears in mind the usual caveats intrinsic to bibliometric comparisons. It is important to note that *h* depends on the number of years of operation and therefore needs to be combined with the *m*-parameter in order to avoid biased interpretations.

Hirsch (2005) points out that the *h*-index is prone to depend on the field of study (Physics, Biology, etc.). This is true even within astronomy. For instance, astronomers in the gamma-ray burst community have higher *h*-indices than their colleagues in other areas, partly because of the importance of the field, but also because of the size of the collaborations and the rate of publications.

If this effect is present with regard to instruments, it may indicate that a given instrument is flexible enough to produce papers in a wide range of astronomical fields. The numbers in Table 2 show that the first-generation VLT instruments have similar *h*-indices. They constitute an example of such versatility, since they are used for observations ranging from the Solar System to cosmological applications. However, other biases influencing *h(t)* need to be carefully corrected before a detailed comparison is made. In this sense, our results concerning the *h*-index for instruments are regarded as preliminary.

Acknowledgements

This research has made extensive use of NASA's Astrophysics Data System Bibliographic Services.

References

Delmotte N. et al. 2005, The Messenger 119, 50
Grothkopf U. et al. 2005, The Messenger 119, 45
Grothkopf U. and Stevens-Rayburn S. 2007, in "Library and Information Services in Astronomy V", eds. S. Ricketts, C. Birdie and E. Isaksson, ASP Conference Series, astro-ph/0610274
Hirsch J. E. 2005, Proc. Nat. Acad. Sci. 102, 16569
Leibundgut B., Grothkopf U. and Treumann A. 2003, The Messenger 114, 46
Madrid J. P. and Macchetto F. D. 2007, in "Library and Information Services in Astronomy V", eds. S. Ricketts, C. Birdie and E. Isaksson, ASP Conference Series
Trimble V., Zaich P. and Bosler T. 2005, PASP 117, 111
Trimble V., Zaich P. and Bosler T. 2006, PASP 118, 651
Trimble V. and Zaich P. 2006, PASP 118, 933