Metrics to Measure ESO's Scientific Success

Based on publication statistics, we provide a first assessment of ESO's scientific impact after four years of VLT operation. A brief discussion of the complexity of measuring scientific success and its inherent problems is given. We present publication and citation statistics drawn from the ESO publication database and provide some preliminary interpretation.

B. Leibundgut, U. Grothkopf, A. Treumann (ESO)

ESO's ultimate goal is to advance astronomical knowledge and provide tools for progress in our understanding of the world we live in. There is a long chain of actions, interactions and activities which lead to such results, and it is important for organisations to evaluate the impact they have on their specific research field and on society as a whole.

The scientific procedure is to produce a hypothesis (often based on previous data and well-known and -loved paradigms), design a test for this hypothesis, carry out the test, analyse the outcome and finally publish the results. It is this last bit, the publication and dissemination that closes the loop and enables others to continue the line of thought.

As an organisation providing resources to the European astronomical community, ESO has to evaluate how successful and competitive it is compared to other observatories and astronomical institutions. This can be done in many ways, ranging from assessing user satisfaction, efficiency of observations and quality of delivered data to the achieved scientific breakthroughs. While the first three metrics lie within ESO's purview, the last one is a joint effort of the research community and ESO. The recent questionnaire on user satisfaction with the service mode offered by ESO (Comerón et al. 2003, Messenger, no. 113, 32) and the evaluations given to the Users Committee in the end of run reports for visitor mode observations measure how well ESO is doing compared to the expectations of its users. The statistics on telescope downtime indicate how well the observatories function. However, the final aim of the combined work of astronomers and ESO is scientific progress.

Measuring scientific success

The definition of scientific success is not easy. This becomes obvious when one does a small literature survey of previous studies on scientific success. Comparisons among observatories (e.g. Trimble, 1995, PASP, 107, 977; Bergeron and Grothkopf, 1999, Messenger, no. 96, 28; Benn and Sanchez, 2001, PASP, 113, 385), astronomical institutions (Abt, 1994, PASP, 106, 107) and individual astronomers (Burstein, 2000, BAAS, 32, 917) are all available. Observatories in particular attempt to quantify their impact (Meylan, Madrid and Macchetto, 2003, STScI Newsletter, 20, no.2, 1). Another recent example is the statistical study on the productivity of ESO's La Silla observatory (see Annex I of the La Silla 2006+ report; available at *www.eso.org/gen-fac/commit/ls2006/*). Investigations on publication behaviour and patterns have also been published (Abt and Zhou, 1996, PASP, 108, 375; Abt, 1998, PASP, 110, 210).

Most of the studies apply bibliometrics, i.e., the quantitative analysis of average or total numbers of publications and citations. However, they all use specific methods which are often not reproducible by other organisations, and so far no general scheme has been developed. One reason was that up to a few years ago, no uniform non-commercial database was available to the astronomical community. This has changed with the availability of the ADS system that collects publications and citations of basically all astronomical literature. Astronomy is privileged in that this nearly complete (and free of charge) database exists, as it represents a situation completely different from almost all other sciences.

In the following, we present publication and citation statistics drawn from the ESO publication database with some preliminary interpretations. After describing how the information is assembled (section 2), we will discuss different criteria that could serve for interpretation (section 3). In section 4 we present the ESO statistics.

Assembling the information

Publications resulting from ESO data should be clearly identified as such. The Call for Proposals specifies that papers must list the observing programme(s) within which the data were obtained in a footnote. This serves several purposes; most importantly, it helps to measure the observatory's scientific success. Astronomers are increasingly following this requirement, although not yet as consistently as necessary.

At ESO, the librarians search all major astronomy journals for publications deriving from ESO data. When ESO-related information is not obvious from the publication, a cross-match with the observing schedule is made. Should the records still be incomplete we contact the first author or PI in order to obtain missing details. For completeness, the ADS database is queried at regular intervals.

This procedure leads to the most comprehensive possible record of refereed publications based on observations with ESO facilities. A recent comparison with automated searches in ADS showed that a considerable fraction of papers is not identified as based on ESO data in the ADS database, while others are wrongly attributed to ESO (Grothkopf and Treumann, 2003, LISA IV proc., *www.eso.org/libraries/lisa4/Grothkopf1.pdf*). The main reason is that automated retrieval tools are not capable of interpreting the context in which search terms appear and thus cannot discriminate between relevant and irrelevant papers.

The ESO publication database is publicly available through the libraries webpage at *www.eso.org/libraries/*. Entries contain authors and title of papers, publication year, journal, volume and pages as well as ESO-specific information, such as programme IDs, telescopes and instruments used and the observing mode (service or visitor) in which the observations were carried out. The records are linked to ADS for download of citation information or access to the online version of papers. This database was used to assess how ESO's new and existing observing facilities impact the progress in our science.

Criteria for the interpretation

There are many ways to look at the impact facilities and scientific collaborations have. The two extremes are counting papers and counting Nobel prizes. The latter is not meant as a joke, although of course it is reaching high. It does, however, illustrate that the interpretation strongly depends on the weights assigned to different criteria. Reaching too high will decrease statistics to small numbers and diminish their meaningfulness. On the other hand, simple quantitative statistics, ignoring any quality issues, are just as dangerous. It is essential to find the right balance and to refrain from over-interpreting results.

Papers

Once the data have left the observatory, it is up to the astronomers to convert the bits into knowledge. The most tangible results are scientific publications. The observatory influences the process by providing data products that raise the researchers' interest in writing papers. This requires adherence to strict quality standards, in particular when the user base is growing. Still, some projects may not result in any publication for a variety of reasons and the goal must be to keep this number as low as possible. The publication rate of course depends on the efficiency and size of the observatory. An observatory that serves a large user community will generate more papers.

An increasing number of publications are based on data from more than one telescope, and even more than one observatory. These papers will be regarded by several organizations as originating from 'their' data and hence will appear in several statistics.

Citations

The number of citations indicates how well a publication is accepted and how important it is considered within the community. Similar to observatories that do not produce data suitable for publication, scientific papers that are not cited are useless. But simply counting citations is problematic, and one must be aware of the pitfalls. Some papers create strong reactions because they are wrong or misinterpret data. Catalogues can generate many citations as they are the basis for derivative results. Papers in a popular field will generally attract more attention than those which address very specific topics within a small community. Hence, citation statistics are not necessarily a measurement of quality, but certainly of attention, and they have to be put into context.

Citations critically depend on how much time has elapsed since a paper was published. Publications have a mean life time after which they are either forgotten or superseded by more recent findings. Only very few papers become 'classics' or even enter textbooks; many will be outdated within a rather short time. The mean attention span for papers has shortened from six years in the 1990s (Abt, 1996, PASP, 108, 1059) to about two years today (Crabtree and Bryson, 2001, JRASC, 95, 259). Electronic publishing and a general acceleration in information exchange have contributed substantially to this decrease. Accordingly, there can be strong fluctuations in citation statistics within a relatively short time for recent papers, while older publications that have passed their peak will not be cited much longer.

A frequently discussed issue are self-citations. While self-cites can influence statistics for individual papers, their number typically is too low to cause much concern for whole observatories. Scientists simply can not publish fast enough to increase the number of citations of their own papers considerably.

High impact

Some publications become highly visible and important for the progress of astronomy. They may introduce a paradigm change, produce the fundamental data set for a given subfield or are seminal reviews. Everybody will be able to identify the five most important papers in their field of research; they enter the syllabus of discussions and are the pillars upon which research fields stand. By definition these publications are rare as the selection is so severe.

For an observatory it cannot be the goal to produce high-impact papers itself, but it must aspire to provide the facilities which allow astronomers to explore new territory and make fundamental discoveries.

In bibliometrics, typically citation counts determine which publications are regarded as 'high impact papers'. Such statistics have been used to compare various observatories against each other as well as to argue that certain telescopes are not 'competitive' any longer. One should be careful in using such number statistics blindly as important information may be easily overlooked. An example is the ongoing discussion between 4m- and 8m-telescope science. The extra-solar planets were discovered at small telescopes (1m to 2m diameter) with the important factors of time baseline and progress in analysis software.

The evaluation of papers has to take into account other phenomena as well. If a publication 'finishes off' a field and the trend moves to other problems, it will not be cited very often. Many people will have personal lists of 'most important contributions' that are not necessarily borne out in the statistics. Such contributions are of fundamental importance, yet they do not produce the reaction one would expect.

ESO statistics

The ESO statistics presented in this section were compiled and analysed with the above remarks in mind. They provide a snapshot of the current status (beginning of October 2003) and give a first assessment of the impact of the VLT on astron-

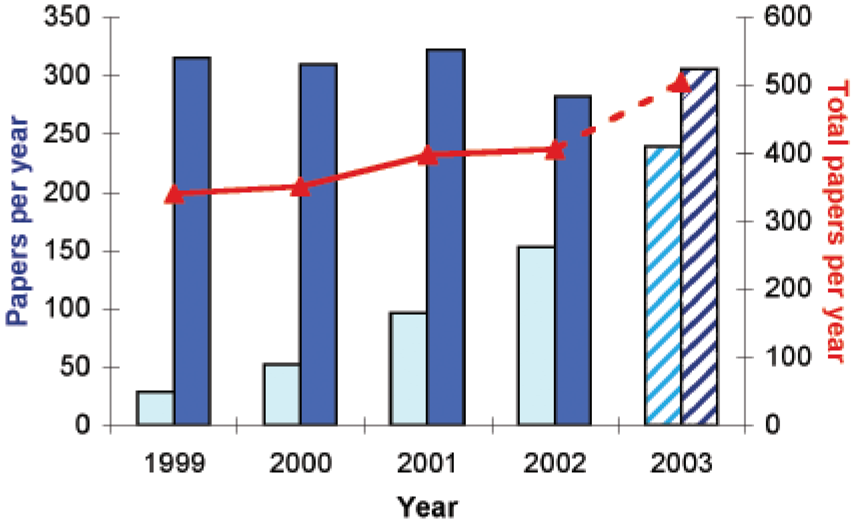

Figure 1: Number of papers based on ESO data published per year. The data for La Silla (dark blue) and Paranal (light blue) are shown individually (left scale). In addition, the total number is shown as the red line (right scale). The 2003 values were extrapolated from the first nine months. Note that papers can be based on data from both observatories, hence the total number is smaller than the sum of papers from individual sites.

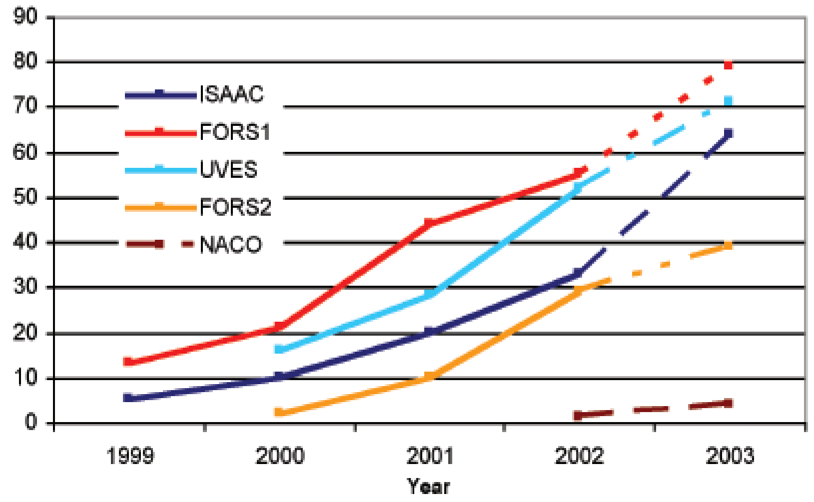

Figure 2: A steady increase in the numbers of papers per year is observed for all VLT instruments. The 2003 values were extrapolated from the first nine months.

omy. Some earlier statistics, in particular those on La Silla, have been presented in the La Silla 2006+ report. We will concentrate here on the early years of the VLT.

Papers

The La Silla 2006+ report showed that around 400 publications based on ESO telescopes were published each year. Since 1999, the share of papers coming from La Silla telescopes has been fairly stable; the VLT has generated a steady increase. The number of publications depends of course on how many facilities are offered. While La Silla has reduced the number of telescopes by closing the smaller ones, Paranal has seen all four VLT unit telescopes come into operation between 1999 and 2002. The number of instruments increased, and so did the observing opportunities and the fraction of the astronomical community that could be attracted. Figure 1 shows the publications in refereed journals per year separately for the two observatories as well as the total number for all ESO-based papers. During the first nine months of 2003 already more VLT papers have been published than in the entire year 2002. Extrapolating from the first three quarters to the end of the year yields a total number near 500 publications for this year with about 240 – or nearly half – of the papers coming from the VLT. This is a higher number of papers based on ESO data than ever before. Overall, it represents an increase in publications of about 20%.

The increase of VLT papers is a good sign. We expect this trend to continue in the near future as more instruments are added to the observatory.

Figure 2 shows the statistics for the individual instruments. The increase of publications per year is comparable for most instruments. NACO is somewhat special as most publications for this instrument currently come from commissioning data. The next years will show whether more complicated instruments produce similar (or higher) publication rates. Other important issues could be the average run length and the larger observing time overheads for IR instruments.

It is interesting to compare these statistics to trends observed at other facilities. Numbers are available for some ground-based observatories as well as space missions, like ISO and HST. The rate of papers continuously increases during the first few years (see Meylan, Madrid and Macchetto 2003). The comparisons are not straightforward, however, as the space missions typically have a smaller instrument suite and different observing patterns. Ground-based telescopes are mostly confined to night-time observing, while this is not necessarily the case for space missions, where other constraints play a role.

We investigated whether VLT service and visitor mode observations lead to different publication behaviour as the way astronomers deal with data may depend on how they were obtained. Also, the two modes offer different astronomical opportunities, for instance monitoring projects in service mode. In Figure 3 a first comparison of the numbers of papers derived from service and visitor mode observations is made. The distribution is fairly even between the two modes. The fluctuations are at this point probably statistical. One has to take into account the delay between observations and their corresponding publications so that the original distribution of modes for the observing programmes per observing semester is blurred by the time of publication. Hence, it is too early to draw firm conclusions. However, both modes appear rather successful in producing results suitable for publication. Among other things this means that the quality of service observations matches that from visitor mode runs, and service mode is accepted as a viable option by the observer community.

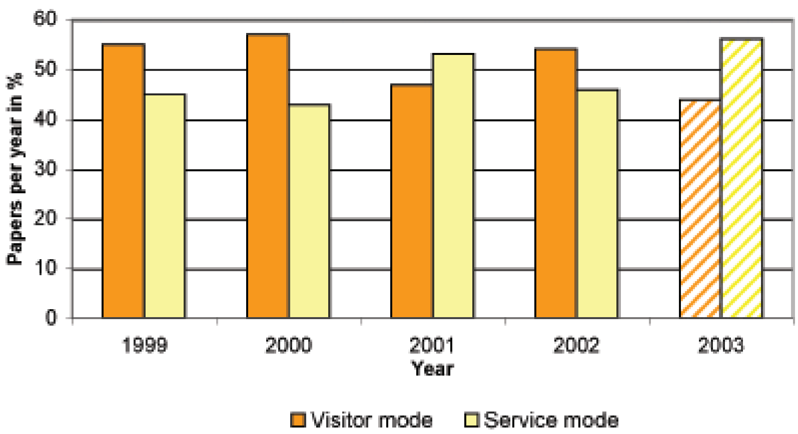

Figure 3: Comparison of publication rates from VLT service and visitor mode programmes. The 2003 values were extrapolated from the first nine months.

Citations

The citation statistics of ESO papers were gathered with information provided by ADS. Although there may be inaccuracies at the individual level, we believe that these are mostly negligible for comparisons at the scale of observatories. There appear to be no known systematics in the ADS system.

It is obvious that citation statistics depend on the time when they are assembled. The dynamics of citation rates are beyond this investigation, but one needs to keep in mind that statistics of recently published papers can change quickly. This is illustrated in Figure 4. The citations for papers published in 2001 nearly doubled during the first nine months of 2003! This trend may continue for papers published in 2002. After a few years, citation rates reach a peak; afterwards they drop. The VLT publications have not yet reached this level, but the early papers are not cited as much as the more recent ones.

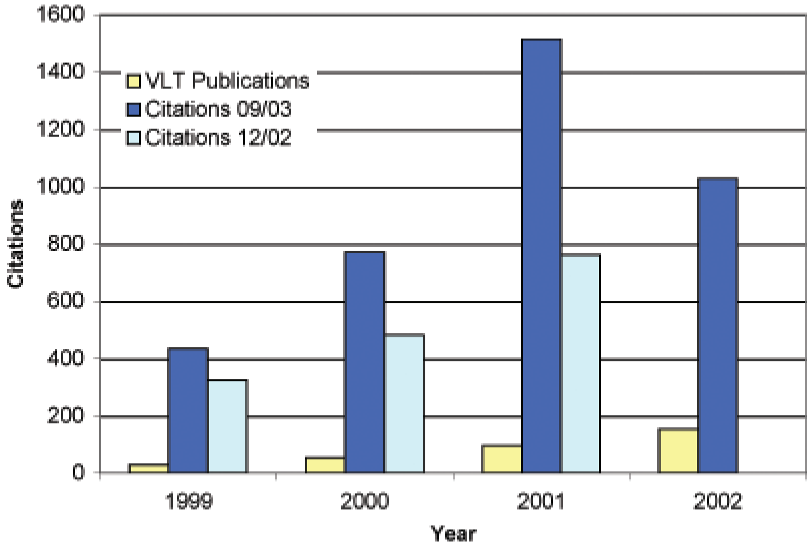

Figure 4: Changes of citations from December 2002 to October 2003.

The number of citations per paper is rather high; on average, VLT papers receive more than 10 cites after a few years. As of October 2003, citation rates for papers published in 2001 and 2002 are significantly higher than those for the very first papers.

Potentially, there are various mechanisms at work. The VLT has received greater attention with time. Also, more extensive projects, e.g. Large Programmes that lead to a larger body of data take longer to complete. Possibly, papers appearing in 2001 and 2002 presented more comprehensive studies, while the early papers may have mostly been letters and short communications which were superseded by the more substantial papers following a bit later. Other possibilities are that with more users and more papers, the rate of self-citations is going up with time. This is, however, unlikely to cause such high citation numbers.

In general, the VLT seems to produce a healthy reaction from the astronomical community and its contributions are recognised.

High impact

As there is no clear definition of what high-impact publications are, we restrict this section to a discussion of some ideas. We are able to identify the most frequently cited papers coming from ESO telescopes. This is the first step in such an analysis. Where to draw the line between average publications and those that change the way we look at the universe is probably somewhat arbitrary. In addition, the perception of the community changes over time and the impact of publications becomes obvious only in retrospect.

We observe that citations for ESO publications typically result in a bifurcated distribution. Some papers receive only few citations even after several years, others increase their total citations over a long time. About 5 to 10% of papers based on VLT data achieve rather high citation rates (more than 10 citations per year). More specifically, over 20% of the papers published in 2001 show citation rates above 10 per year, and 8% received more than 20 citations per year two years after publication. While the first two years of VLT operations have not yielded publications attracting more than 50 citations in total, there are six papers from 2001 that exceeded 50 citations within less than two years of publication. One paper from 2002 has reached this number within one year of publication.

Comparison with other observatories

At present, the different observatories and organisations assemble publication statistics according to their respective policies. As the selection criteria can vary vastly, comparisons have to be done with great care. Meaningful results can only be achieved when statistics are compiled based on the same rules and methodology. This of course requires a close collaboration among observatories and emphasizes their inter-dependence rather than their competition.

Bearing these comments in mind, we note that the first four years of VLT publication statistics are similar to those of HST (see the recently published statistics in Meylan, Madrid and Macchetto, 2003, STScI Newsletter, 20, no. 2, 1). Both VLT and HST experienced comparable start-up phases with an annual increase in published papers of approx. 75%. Meylan et al. also presented the mean number of citations per year for all refereed astrophysics papers and found that publications based on HST data on average are cited twice as often. With a mean citation rate of approx. 15 for papers published in 1999, 2000 and 2001, we recognize a similar trend for VLT papers. A more detailed analysis is only feasible though with a larger baseline.

Conclusion

After four years of VLT operation, we start to see clear signs of the impact the observatory has on the astronomical community. The overall publication rate of papers based on ESO data – approximately one refereed paper published per calendar day – remained constant for the past few years. In 2003, the total number of publications is now increasing, mostly in line with the growing number of facilities offered at Paranal. La Silla still maintains a high publication rate despite a reduction in available telescopes. With regard to observing modes, no clear preference can be stated. Up to now, service and visitor mode programmes result in roughly the same number of publications, which generally corresponds to the time allocation. The large percentage of highly cited papers and the overall citation statistics prove that the scientific results produced by the VLT are highly visible and well recognised within the astronomical community.

Once the number of years over which we look back is sufficiently large to average out misleading short-term effects, we will be able to re-investigate the scientific impact of the VLT in a more comprehensive study. The results will be published in a future *Messenger* article.