ARCHIVING

Carlos Jaschek
Observatoire de Strasbourg
11 rue de l'Université
67000 Strasbourg
France

Astronomy is a natural science based upon observations. Observations are made by somebody at a given time and place, obtained with certain instruments using certain techniques and are interpreted with the help of certain models or theories. The latter may change with time, but observations can be used with any new theory, provided the observational data and the reduction techniques are known.

If one wishes to keep astronomy based upon controllable facts and not upon fiction, we must accordingly preserve the observations adequately.

During centuries this was done by all astronomers. It was thus possible to extract the occultation of Jupiter's satellites from the old log-book and the light curve of 3C273 from the plates stored in the Harvard archives. Then came an explosive development of new techniques where observations were registered on magnetic tape - like in radiotelescopes, satellites and in CCD receivers. New instrumentation became available every few years, so that for a certain time newer results were much better in resolution or signal-to-noise ratio, thus progress appeared to be continuous. People then concluded that old observations were useless (since they could be replaced by better ones) and as a result archiving became old-fashioned.

Nowadays in some sectors, like space astronomy, astronomers came to realise that old observations should not be discarded, but carefully preserved. It is curious to see that such a change was not imposed by a reaction of the astronomers themselves, but because space administration organisations like NASA found themselves under criticism from taxpayers because results obtained through considerable expense of both money and energy were stored very inefficiently. The reality is such that, ironically, space observations are now carefully preserved in most cases, whereas for earth-bound observations the situation is still very bad, to the point that one can conclude that most earth-bound observations carried out in the last decade are lost. A result published in the eighties is based upon observations which, after publication of the paper, became inaccessible. Clearly such a situation can only happen if astronomers and institutions are disinterested in the problem of archiving. The most popular reason given by astronomers is that it costs too much for little reward. One should always be doubtful when astronomers, who are usually disinterested in economic reality, come up with economical arguments !

J. Bergeron (ed.), Highlights of Astronomy, Vol. 9, 739–742.

In order to have a better insight into the problem one is led to examine the two essential questions: why do we preserve data and why do we wish to preserve them.

We preserve data essentially because we have learnt that all astronomical phenomena are variable in time. Observations made at a given epoch are thus irreplaceable and have to be stored in memory. The often given answer, that modern observations are much more precise than old observations and can thus replace them advantageously, is based upon the unproven assumption that astronomical processes either do not vary in time or are strictly periodic - assumptions which every student of astronomy knows to be wrong.

What should be preserved? The answer is specific to each problem and depends on a number of considerations. Photoelectric photometry for instance has reduction procedures which are well known and well standardised, so that one no longer needs to store the observational data (photon counts as a function of time). In this case one can simply store the reduced data (magnitudes and/or colour indices) vs. time. Please observe that with few exceptions, astronomers only publish results in the form of either average values or of lightcurves, without publishing the observing dates. This constitutes a serious fault, since one is implicitly using the (wrong) assumption that observations do not vary in time. Astronomers reply usually that journals, for reasons of space, do not want to publish observing dates; this is also only partially true because there do exist data centres where such data can be stored.

Leaving aside this case, the situation is different for CCD observations, where a large number of frames is produced and put on magnetic tape. Some time afterwards the tape is re-used and the data are thus destroyed. Why? The answer usually is **a**) not all observations are finally used (so that archiving them would be superfluous); **b**) that there does not exist a convention of what should be kept; and **c**) that organising an archive would be too costly (again an economic argument!). It is easy to see that all three arguments are only partially true. In effect, one needs only to store those observations which have been used - as in the old days when an underexposed plate was not entered into the archives. But it seems very appropiate that observations which were used in a publication be preserved so that anyone interested might be able to see the observational evidence for the conclusions. With regard to point **b** this is an easy excuse for not doing anything. The best thing is to sit down and try to define at each observatory the policy of what should be kept. With regard finally to argument **c,** it is very doubtful if it is true in the long run. With a cost of telescope time of the order of several hundreds or thousands of dollars per hour, it is very doubtful in my mind if the archiving price, which comes to a fraction of that cost, justifies the use of the argument. The point seems rather to be that since nobody has a working archive right now (with a few exceptions, like La Palma), one thinks that one must start archiving everything from the start of the operations of the telescope on. This is certainly not a good solution; any archive should start with the latest observations, and go forwards and not backwards.

Besides photoelectric photometry and CCDs there are of course other chapters of astronomy, where the situation is even worse.

Good examples are radioastronomy and long base-line interferometry. In the first case radioastronomers have been unable to put their catalogues on a standard format which can be read by everybody and, to the dismay of newcomers in the field, there exists not even an updated catalogue of all radio-catalogues published so far.

A subfield which is in similarly bad shape is that of long base-line interferometry, where archives are cleaned out every few years (or even months) under the excuse that it would be impractical to store all data. Since the interferometric technique is rather new, one might expect that reduction procedures are not yet final, so that very probably old data will need a re-reduction within the next few years. If the original data are missing, this is simply impossible. As a consequence, in time the old objects will have to be observed anew. To the cost of repetition one has to add the disadvantage of being unable to detect possible time variations which occurred meanwhile.

After that much criticism of what is done now, let me write down a few recommendations of how I think the situation could be significantly improved. What we need is

a) to define at each observatory an archiving policy - what is to be kept, how and for how long

b) to apply this policy in a very strict way, for all present and future observations

c) to make sure that observing files (i.e. basic information concerning.the object observed) be stored in a computer-readable way.

d) to make arrangements with organisations like data centres what could be stored by them, or with organisations like the "archive of unpublished photometric observations" for discharging part of the data there

e) to make sure that, together with the observational data, a reasonable amount of engineering data on both telescopes and receivers be stored and that descriptions of the reduction procedures used be available.

As can be seen, none of these measures is neither revolutionary nor difficult to put into practice. What is really required is that the astronomical community take conscience of the importance of the problem.

To the preceding points I would like to add a series of counter-arguments against arguments which turn up regularly in discussions on archiving:

a) one must start the archive with all the old observations carried out at the observatory. Since this implies a large amount of work and money it is usually the best justification for not doing anything on archiving.

b) the establishment of observing files (so runs the argument) would require at least a new post for a technician in charge of the log-book. But it is clear that it is a duty of the observer himself to provide the data required for the observing files - after all pointing the telescope on the object implies knowing most of the information which goes into the observing files.

c) the establishment of an archive needs specific space and the argument then goes that with the present budget squeeze, room is simply not available. This is curious reasoning which (happily !) I have never heard in the discussions over ever-expanding libraries. The fact is that the library has an official status in most observatories, whereas archives do not - despite the fact that both are archives in a broad sense.

d) Another objection to archiving is that the institution has nothing to do with archiving because, supposedly, it is the observer himself who should take care of that (ESO for instance has no plate archive - each observer keeps the material as long as he wishes). This is a curious reasoning because if applied to libraries it would also imply that observers have to care for the storage of books. Individual observers in most cases are unable to document their own observations after some years - either they lost the archive when moving (to a new office or another observatory), or because of a change of interest or any one of ten other good reasons.

In conclusion I think that the most important result to be expected from meetings of this type is to make astronomers aware of the problems involved. Once the astronomers are convinced of the importance of the operation, everything else will follow easily.

Note: Originally it had been planned that I should give the concluding remarks of the meeting. Because of the well-known accident which prevented holding the meeting, the concluding remarks should be taken as a rather personal view of the problem.

https://doi.org/10.1017/S1539299600010248 Published online by Cambridge University Press