

## ASTROMETRY AND DATA CENTERS

C. JASCHEK  
Centre de Données Stellaires  
Observatoire Astronomique de Strasbourg  
11, Rue de l'Université  
67000 Strasbourg  
France

**SUMMARY.** The paper tries to provide some background to answer the problem if old astrometric data should be stored. The cases of the AC and the CdC are examined, as well as the storage of some more recent catalogues on optical disc. Finally the necessity of producing general catalogues (of the SAO type) for non-astrometrists is underlined, to satisfy the needs of the astronomical community.

One can define data centers (DC) as the memory of the astronomical community. Usually biological analogs have weaknesses when applied to physical sciences, and this also happens here: human memory is full of tricks, and psychology has shown us that what is retained is only the tip of an iceberg, the largest part of which remains invisible. We would certainly not like to have a DC with such defects. So better let us say that the CD's are the archives of the astronomical community.

One knows from historical archives that not everything can be kept in memory, because the amount of data becomes too large, and this happens also in astronomy. Therefore we should start asking ourselves why do we want to archive and what.

Just to get a starting point, I use the compilation of Sevarlic et al.: "Epitome Fundamentorum Astronomiae, vol. I"(1978) which all includes the astrometric catalogues published in the past, from 400 BC on. The authors list 2087 catalogs. Teleki (1985) has provided later on sample counts, which are summarized in Fig. 1. It is amazing to see how regular this growth was, at least from 1700 on. He found that the number of catalogs grew exponentially, according to the formula

$$N = N_0 \exp [0,020(t-t_0)],$$

where  $N_0$  and  $t_0$  are conveniently chosen constants. The formula implies a doubling time of the number of catalogues of 33 years. This exponential growth had two marked breaks around 1915 and 1940 and a general slowdown after 1950. In the history of sciences such slowdowns are connected either with a change of interest of the community or with a stagnation of instrumentation.

Turning next to the total number of star positions contained in these catalogues, we observe that the number of stars in a given catalogue may vary between 10 and several  $10^5$ . We even have the GSC (Guide Star Catalogue) (not included in our statistics, because as yet not available) with  $3 \times 10^7$  stars. If one assumes a gross average of 500 stars per catalogue, we arrive at a minimum of  $2 \times 10^3 \times 5 \times 10^2 \sim 10^6$  star positions. Teleki (1985) quotes  $3,5 \times 10^6$  for all observational catalogues. Observe on the other hand that since many bright stars were observed repeatedly, the total number of different stars observed should be certainly smaller -let us assume  $5 \times 10^5$  stars. Obviously this depends on the definition we use for "catalogues of star positions"- if the Durchmusterungen are counted the number at least goes up to  $1,1 \times 10^6$ .

This leads us quite naturally to the next question: should all these data be kept in memory? Obviously this depends on the use one wants to make of the data, but on the other hand let us not be too ambitious on that. History has shown that often data are used for very different purposes than for the one they were collected for. Ptolemy surely would not have dreamed that his star catalogue would have been used to derive proper motions, whose existence he denied. To see how useful the old positions are nowadays, let us use the precision of the star positions. We reproduce from Teleki (1985) those data which provide a measure of the precision of stellar positions as a function of time.

TABLE 1

Positional precision as a function of time

Author	T	$\epsilon$
Ptolemy	138	900"
Brahe	1601	120
Hevelius	1661	120
Flamsteed	1689	10
Bradley	1755	2
Gill et al.	1890	0.4
Hog	1970	0.2

 $\epsilon$  = Mean error of two observations

T = Epoch of observations

From table 1 it seems clear that only the catalogues obtained in the last century are of a comparable order of precision so that they are usable for modern work. If this be taken at face value, it would imply that one could scrap all older work and keep only the "recent" tail of Sevarlic's list. If one does that, one arrives at the following :

TABLE 2. Number of catalogues

Total number of cat.		2087 catalogues
Keep from	1900 on	1394 "
	1920	1031 "
	1940	598 "

Such an operation alleviates the storage problem somewhat, but does not solve it because the decrease is not that large. This is due to exponential growth of the number of catalogues.

An omission of the old and less precise measurements is what was done essentially for the construction of the FK5 -only catalogs after 1900 were included. So for instance only 100 "recent" catalogs were included (Schwan, 1986) for the construction of the system. For proper motions 190 catalogues with a preference for recent ones were included. The 290 catalogues correspond to a total of  $7 \times 10^5$  star positions.

To use the existing catalogues in the best way, a large effort was undertaken by both the US Naval Observatory and the Astronomisches Rechen-Institut to put a large number of modern catalogues in machine readable form. The complete list of these catalogues will be published by the CDS; colleagues interested in it should let us know.

It is probable that the ARI shall use all the stars available in those catalogues which entered in the FK5, to build up a kind of "super SAO". (This is just a synthetic name.) But I will certainly not take my time to describe this. My remark goes simply in the sense that once you have that many data available, one will certainly be induced to do something with them.

So far so good, but let us turn back to our problems of "memory". Is the oblivion of all old catalogues really justified? One thing one should always keep in mind is that old material, even if "apparently" useless, will be used by somebody for purposes one never thought of. So my reflection would be that we should certainly keep in memory what we have now in machine readable form. Of course the main reason for this is that the old catalogues may be used for deriving proper motions. It is clear that the longer the time base and the smaller the positional error at each one of the epochs is, the better the proper motions are. Because of both facts, we may use observations obtained one century ago, but not much more than that. This is in short the justification for using the Astrographic Catalogue (AC) and the Carte du Ciel (CdC) material -or better said, to start using it. Now this is difficult because in first

place, the majority of the AC material is not machine-readable and in second place measurement, reduction and publication of the different AC zones was carried out in different ways, so that the results are certainly heterogeneous.

One solution of these difficulties, often discussed, is to get hold of the original plates and measure them anew; reduced with modern computing facilities the results should be more precise than what was published. The first condition is to find out what became of the original plates -if they still exist, if they are measurable and so on. To find out about the general state of the two projects, we made at the CDS an enquiry, based upon a circular written by H. Eichhorn and C. Jaschek. The answers were published in CDS Information Bulletins 27, 28 and 29 and can be summarized as follows (see table 3).

TABLE 3

Summary of the answers received to AC and CdC enquiry.  
From C. Jaschek (1985)

There was practically unanimity in the answer to question no.1, implying that one century after the start of the operation, the participating institutions are still interested in both projects.

Question no.2 -to make an effort to get the published measures on magnetic tape- was answered in general negatively. The only large effort made in the past was the one by Lacroute, who passed all French AC zones onto tape. This effort for the moment has found only one imitator (see below).

Question no.3 -remeasurement of AC plates- was also answered negatively, because of various reasons (small improvement of the published measures, impossibility to remeasure all plates, lack of automatic measuring engines). What seems to be more justified is the measurement of the CdC plates. One drawback is that not all zones were really observed.

Question no.4 -plate loan- was answered in general favorably. Appropriate rules will have to be worked out in each case, but this seems to be a fairly obvious requirement.

Question no.5 -inclusion of the CdC plates in the loan- was also answered favorably, whenever the plates do exist.

Question no.6 -state of preservation- produced in general the foreseeable answer that a majority (but not all) plates seemed to be in good condition.

In conclusion it seems that the answers to our questionnaire have

produced an update of the "final" document of the IAU Commission "Carte" du Ciel" (IAU Transactions XIV B, p. 175). It is hoped that these answers may help to provoke some timely discussions on the projects which it is possible to carry out with the help of the AC and CdC material.

Let us observe, however, that so far we have had no extensive experiences on the remeasuring of the old plates. Only one attempt was made (Rousseau and Guibert, 1984) and the results are encouraging, but do not guarantee that the remeasurement would necessarily be a success. From the answers it is also clear that not all plate collections were archived under the best conditions and so the re-measurement is sometimes not even feasible. Therefore for the AC plate material, the best solution seems to be use the published values -but the use passes through a computer readable version. To date this has only been done by P. Lacroute (1981) for the French zones i.e.  $+31^{\circ} < \delta < -2^{\circ}$ . At the US Naval Observatory there exists a project to key-punch the whole sky; the project is progressing slowly, and is expected to be ready in the nineties. It might be that optical readers could be of great usefulness to speed up the project, but tests with optical readers at Strasbourg so far have not been very convincing.

If we ask the blunt question of what the AC could be used for right now, the answer is apparently easy -compare it to the CGS and get proper motions. From the paper by Fresneau (1985) one can find out the details; let me observe simply that the first step must be to convert the CGS star positions into  $\alpha$  and  $\delta$  in some of the existing reference systems -but this has not been done as yet. At the CDS, this is under way, but it is not simple. Preliminary results were reported by Didelon and Egret (1987).

The next logical question, to which nobody has given much thought lately is the use one will make of the proper motions. Until now proper motions were mostly used to detect nearby stars, because the surveys had picked out only the stars of largest proper motion, to say so the "tail" of the kinematic distribution. But if we can go to smaller proper motions we may be able to use the distribution of proper motion for stellar kinematics. This has not been a favorite subject for the last fifty years (to put it mildly) and the latest review paper on galactic structure (Bahcall, 1986) even quotes the exploitation of stellar kinematics as one of the important points for future work.

Having said that much on existing surveys, we may say a few words on a matter of general policy, which seems to have escaped general attention up to now. You know that many rapid automated measuring engines exist, of the PDS, GALAXY or COSMOS type. These machines scan Schmidt plates at great speed, and are able to provide positional and photometric data for several  $10^4$ - $10^6$  objects per plate. The fate of all these observations is rather obscure -it seems that once the data acquired are used (for instance for a statistics) the original data are destroyed, on the grounds that after all the plate can be remeasured.

This is true in a certain sense, but wrong in another: after all, measuring a plate represents an investment of time, energy and money, and furthermore, as Tritton has shown in this meeting, plates may degrade. So it would be worthwhile to look into the matter and see if the data acquired by automated plate measuring engines could be safeguarded somewhere. I made an inquiry in 1978 on the general policy of the large measuring engines and found out that no general policy existed. In 1987 I redid the enquiry; I wrote to about 20 centers and got 7 answers up to now. There seems still to be no generally accepted policy. This is surprising, but is in line with the general lack of attention to archival problems.

The exception is the Edinburgh Observatory. MacGillivray(1987) announces that they are scanning with the COSMOS all plates of the Southern Sky Atlas of the UK 1,2m Schmidt telescope in Australia(UKST). They expect to produce a catalogue down to  $B \sim 22^m$  and expect to do the same in the North, using the new Palomar sky survey. They expect to have the southern sky ( $b < -30^\circ$ ) ready in 1988. The whole material will be stored on optical discs. This project will be another milestone in our enterprise of "mapping the sky".  
 5Note added after the meeting: Humphreys at minnesota announces a similar project for the northern sky. See CDS Information Bulletin no. 33).

Let me devote finally a few words to another general problem, namely the needs of the non-astrometrists.

As you know, users of data centers are of two kinds -the specialist who in general wants original data, and the non-specialists who wants the best data available. The specialist (i.e. an astrometrists in our case) is often disenchanted of what he finds in a data bank -nothing else than what he knows already, and he knows too well about the quality of these data. On the other hand he finds with a certain condescendence that after all data banks are marginally useful because they can provide him the photometric data he had been looking for.

Data center people know this very well, and know moreover that the photometrist thinks about same, but for him the "marginally useful" data are the astrometric ones, and not the photometric ones, which he knows already and whose quality he doubts...

Such a tale is by no means a fiction, since we have heard this many times expressed, with varying degrees of feelings. The morale is clear: data centers must do something for both kinds of users; for the specialist, by furnishing him observational data- and for the non-specialists by providing him the "best" values recommended by some specialist. You can simply not answer to the photometrist looking for proper motion data: dear sir, you can get from us 121 astrometric catalogues -the answers you are looking for, are all there. The non-specialist will be unable to use the 121 catalogues; he wants just one answer per star and so in the end he will apply duly for the SAO, which probably satisfies his needs. This is the explanation for what we perceive very clearly at the CDS and what is summarized in table 4, which gives the number of orders received for astrometric catalogs.

TABLE 4

Catalogues requested  
from the CDS in 1984, 1985 and 1986

---

Total of catalogues existing 1987:	121
Total of catalogues requested:	119
Catalogues most frequently requested	
SAO	6 times
Perth 70	6
AGK 3	5
Cordoba Durchm. (-22° to -31°)	3
Others	
Washington Double Star Catalogue	6
Luyten $\mu > 0^m 5$	4

---

It is clear that out of one hundred catalogues, astronomers are interested in less than half a dozen: the SAO, Perth, AGK3. All other catalogues are only copied for other data centers where probably they are not much used either. Now it is unnecessary to ask your opinion about SAO, because there seems to exist a certain consensus that it is not the best catalogue which could have conceivably been constructed. But the general use made of it shows that for non-astrometrists it is the catalog "par excellence". I think this underlines very clearly the needs of the non-specialist for astrometric data, and it tells also that the astrometric community has not been very active in responding to these needs. If astrometrists feel that one could do better than SAO, then DO IT -but please do not tell that one has to wait until HIPPARCOS is reduced, or give some other date from here in twenty years. Users wish to have data NOW. This is fair because you would also not be very happy if somebody told you that you have to wait twenty years for the next catalogue of UBV photometry -which you happen to need.

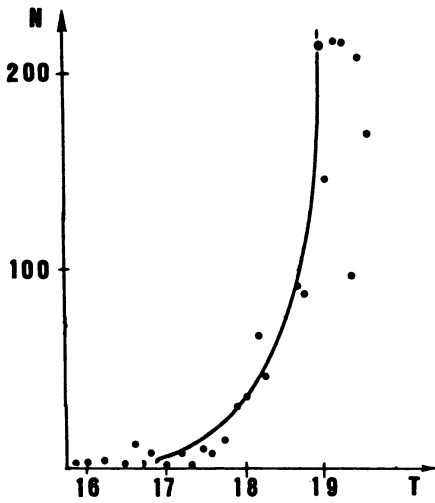


Figure 1. Number of catalogues produced by decade.  
 N = number T = epoch in centuries  
 The line represents the exponential curve given in the text.

#### BIBLIOGRAPHY

- Bahcall, J.N. (1986) *Annual Review of Astronomy and Astrophysics* 24, 577, Annual Reviews Inc., Palo Alto
- Didelon, P. & Egret, D. (1987) *Bull. Inform. CDS* 32,27
- Fresneau, A. (1985) *Astron. J.* 90, 892
- Jaschek, C. (1985) *Bull. Inform. CDS* 29, 43
- Lacroute, P. (1981) *Bull. Inform. CDS* 21, 2
- MacGillivray, H.T. (1987) Private Communication
- Rousseau, M. & Guibert, J. (1984) *Bull. Inform. CDS* 27, 43
- Schwan, H. (1986) *Highlights of Astronomy, Vol. XIX*, Ed. J.P. Swings p. 69, D. Reidel Publ. Co.
- Sevarlic, B., Teleki, G. & Szadeczky-Kardoss, G. (1978) *Publ. Dept. Astron. Belgrade* 7, 69 : *Epitome Fundamentorum Astronomiae, Part I*
- Teleki, G. (1985) *Bull. Obs. Belgrade* 135, 40



## Discussion:

**WESTERHOUT** I want to make it clear that Dr. Jaschek's statement that measuring the CdC plates might be a U.S. Naval Observatory program was entirely a trial balloon. USNO does not have such a program, but does encourage people with big fast machines to measure the CdC. The USNO has a big but not a fast measuring machine, which will be occupied for the next six years - measuring the Washington-New Zealand astrograph plates.

**CORBIN** I would like to make two comments in support of Dr. Jaschek's statements on converting the AC-X-Y measures to machine readable form.

1) I believe many measuring machines work more efficiently if the stars on the plate are first identified. The published X-Y's could be used for this.

2) Assuming that we may never measure all the AC plates, experience with the X-Y's can guide us as to which of the measurable zones are most in need of remeasurement.

**MULHOLLAND** It is not adequate to say that producers of astrometric positions should make them available. During the satellite program of which I was director for nearly ten years, it was necessary to determine positions of several thousand stars of magnitude 12-14, but they remain unavailable to others because of a lack of funding for the compilation, cross-checking, and editing of these data.

**EICHHORN** When around 1959 the SAOC was suggested by Whipple, I asked around Washington (where I was stationed then) for opinions. One of the most influential astrometrists in the area answered my question whether it would be worthwhile to put the catalogue out on the FK4 system when it would become available (then only the FK3 was published) and that it did not matter which system the catalogue was going to be on, because this was going to be such a bad catalogue that no one was ever going to use it.

**WORLEY** First there is a question of semantics involved in the word "catalogue". To me, a catalogue implies something global or universal (or semi-universal). Many of the "catalogues" in the astronomical literature are in fact just lists of observations. In the Washington Double Star Catalogue data base, containing now more than 2050 references, many are called catalogues, but are in fact not. Therefore, I think that the statistics you have presented on this subject require some revision. Second, I am disturbed that a prominent European journal's editors and board are apparently unaware of the proper mechanisms for the distribution and archiving of double star data, despite attempts to make this procedure clear.