Astron. Astrophys. 258, 74-81 (1992)

ASTRONOMY
AND
ASTROPHYSICS

Reliability of the Hipparcos Input Catalogue tested by the 'First look'*

Reliability of the Hipparcos Input Catalogue

F. Crifo[1**] , **A. Gómez**[1**] , **F. Arenou**[1**] , **D. Morin**[1**] **and H. Schrijver**[2***]

[1] URA 335 du CNRS, Observatoire de Paris-Meudon, DASGAL, F-92195 Meudon, France
[2] SRON Space Research Utrecht, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands

Received April 1, 1991, accepted April 1, 1991

Abstract. The reliability of the Input Catalogue is one of the keys for the success of the Hipparcos mission. As the Input Catalogue is a catalogue made of very heterogeneous sources, many systematic reliability tests were conducted to eliminate errors on identifications, position, magnitude, etc.

In March 1991 half of the Input Catalogue had already been checked by the 'First look' facility developed by the FAST consortium at Utrecht. It is shown that less than 0.1 per cent of the 118 000 stars of the observing program will not be detected.

Key words: Star catalogue – catalogue reliability – Hipparcos

1. Introduction

The reliability of the Input Catalogue is one of the keys for the success of the Hipparcos mission. The Input Catalogue was made of very heterogeneous data sources, and throughout its construction great care was taken to ensure that as few as possible errors would be introduced, and that errors of older catalogues would be removed as much as possible. Soon after its completion the catalogue was used for satellite operation, and the main questions were: Would the stars be found? What would be the percentage of waste, both in observing time and loss of stars?

2. Reliability tests and criteria while preparing the catalogue

The INCA Database contains some 215 000 stars proposed by the astronomical community to be observed by Hipparcos (Gómez et al. 1989; Turon et al. 1991). For each star the following information is available:

- measurements compiled or obtained within the INCA Consortium on astrometric (positions and proper motions), photometric (magnitudes and colours) and spectroscopic (spectral type, etc) data, together with data on duplicity and variability (for details see Perryman & Turon 1989);
- identifiers from known catalogues or published lists. Many of these identifiers and data come from the SIMBAD Database of the Centre de Données Astronomiques de Strasbourg, along with the SIMBAD software for sampling and updating the database;
- identifiers introduced to identify each proposal (more than 200) and each star within a proposal.

Since the Input Catalogue was built in successive steps, after numerical simulations (Crézé et al. 1989; Turon et al. 1991), from the stars of the INCA Database, it was necessary to make a systematic search of possible errors for each of the proposed stars. Obviously, a database containing 2.3 million identifiers and 6.3 million measurements may also contain inconsistencies. A set of programmes was then developed, allowing detection of possible errors by the intercomparison of all the available data for each star of the Database. The principles of the systematic method used to check measurements and identifiers are rather general and may be adopted for any astronomical Database. The coherence tests performed can be summarized as follows:

2.1. 'Internal' consistency

(a) Consistency within measurements:
A comparison of the different measurements of a kind of data (e.g. positions, proper motions, magnitudes, spectral types, etc) with an adopted reference value, considered as the best, easily shows the 'outlying' values. Obviously, such a test can be applied when there exists more than two measurements for the considered data and this fact reduces its applicability.

(b) Consistency within identifiers:
This test is especially useful to verify if the stars with special properties, like stars belonging to catalogues or lists of variable, double, cluster, peculiar, . . . stars are well identified. It consists of checking the existence or not of the corresponding identifiers (e.g. the various identifiers of double and multiple stars, of variable stars, etc.). Its applicability depends on the number of identifiers introduced in the Database.

Send offprint requests to: F. Crifo
* Based on observations made with the Hipparcos satellite
** INCA Consortium
*** FAST Consortium

2.2. 'Crossed' consistency

When a reference value is adopted for each kind of data, one can test:
(a) Consistency between identifiers and the available data. This test is the most powerful tool to detect errors easily. It takes into account how the different catalogues and lists included in the Data Base were built:
- position catalogues arranged by increasing coordinates at the equinox and epoch of the catalogue (e.g. BD, CD, CPD, AGK, FK4, FK5, SAO, IRS, etc. catalogues);
- proper motion catalogues containing only stars with proper motions larger than a given value (e.g. Luyten's Catalogues);
- catalogues including stars up to a well established limiting magnitude (e.g. the Bright Star Catalogue, survey catalogues) or compiling stars of a given type (e.g. central stars of Planetary Nebulae, chemically peculiar stars, red dwarfs belonging to the Vyssotsvky lists, etc.). A large variety of tests is possible by comparing the chosen data (astrometric, photometric or spectroscopic) with the appropriate identifier.
(b) Consistency between the best adopted measurements. This test consists in the intercomparison of astrometric data (positions with proper motions), photometric data (magnitudes with colours), photometric with spectroscopic data (colour and spectral type), etc. A simple test is to analyse the consistency between the $(B-V)$ colour and the spectral type.

Examples of internal or crossed consistency may be found in Gómez et al. (1989), Figs. 3.12 to 3.16 and in Gómez (1988), Figs. 3 to 5. However all cases were not as obvious as these figures, and necessited more checks. In particular for some southern NLTT stars proper motions were recalculated, between a first epoch taken in CPC or YZ catalogues, and the newly measured INCA position as second epoch. As a result, some errors were found on SAO proper motions.
Example: SAO 181892. Components of p.m. in arcsec.y^{-1} :

$$\begin{aligned} \text{SAO} &: -0.001 \ +0.008 \\ \text{NLTT} &: +0.220 \ +0.031 \\ \text{INCA–YZ} &: +0.192 \ +0.029 \end{aligned}$$

The INCA position for this star was measured at Bordeaux on the ESO Quick Blue Survey plates (epoch 1977.131); the YZ epoch is 1933.36. The quickly determined (INCA-YZ) proper motion is in very good agreement with the NLTT, and the NLTT p.m. was finally adopted.

During the preparation of the Input Catalogue, the new or best selected data were prepared in the various specialized working groups, and then sent to Meudon for inclusion in the Database. Then the error lists (lists of inconsistencies) were established and corrected, mostly at Meudon. Also, all stars which could not be treated in the specialized groups were sent back to Meudon, where all these 'residuals' were handled with different methods, often star by star, and in cooperation with the specialized working groups. Numerous tests were applied, which don't pretend to be exhaustive: within the short period of the preparation of the catalogue, and given the available manpower, all the errors could not be found, but the essential ones were corrected. Clearly, systematic errors take less time to be corrected than individual errors! The success in detecting errors depended also on the data available for each star: obviously the task was easier for stars brighter than about $V = 9$ mag. For fainter stars it was necessary, in over a thousand cases, to check the data using plates or/and the available literature. Almost at the end of the preparation of the Input Catalogue, the HST Guide Star Catalogue (GSC) became available. The GSC positions and/or magnitudes for about thousand stars were used to improve the catalogue (see Crifo et al. 1991).

3. The 'First look' test of the catalogue using Hipparcos observations

One of the features of the FAST data reduction Consortium is the operation of the so-called 'First look' facility (located at Space Research Utrecht) intended to provide a possibility to quickly validate a fraction of the Hipparcos observations and to produce calibration data to be used by the other partners of FAST and also by ESOC (Schrijver 1989, 1991). This is done by performing a full scientific analysis of the selected data sets up to the great-circle level. From the start of the operations about 6.5 hours of observations per week have been processed in this way, corresponding to about 6 per cent of the useful observations.

Shortly after the beginning of the scientific operations it was realized that the results of this quick processing could also be used for a test of the validity of the Input Catalogue by verifying if star signals had been measured at the positions given by the catalogue and pointed to by the detector.

The sequences of photon counts from all stars in the data set are reduced to a set of 5 parameters characterizing the star signal observed during an 'observing frame' of 2.1333 s. The first of these 5 parameters represents the average signal from the observed star plus the background (consisting of sky background, straylight and detector dark counts). For the purpose of catalogue checking we have processed the obtained parameters by running a special program that labels all star transits in which the measured signal falls below a certain threshold.

We have chosen to select a single (magnitude independent) threshold based on the calibrated detector sensitivity and the measured background. This means that only stars that are imaged outside the detector instantaneous field of view (FWHM 38 arcsec), and thus having an error in position of more than about 20 arcsec are candidates for detection by this method.

The sensitivity of the Hipparcos detecting system has decreased from about 2 150 Hz (for a star of Hipparcos magnitude $H_p = 9$ and colour index 0.5 in the centre of the field of view) in December 1989 to about 1750 Hz in May 1991. The variations as function of the position in the field of view and of colour are of the order of 10 per cent. On the other hand the background is about 30 Hz in the most favourable conditions. We have selected 100 Hz as a reasonable threshold.

The list of all suspect transits obtained in this way is further checked by hand. Observations from stars fainter than $H_p = 11.5$ are not considered unless their signal is very low (near minimum background limit). Then it is verified if the signal is low for all transits of the star in the data set. Isolated lows can occur due to various causes: bad real-time satellite attitude, shutter manoeuvres near occultations, position of the star on the edge of the field of view. For all candidate stars, but in particular for stars having only one or two transits in the measuring period, it is investigated if one of these conditions occured. The final product of these verifications is a (biweekly) list of 'suspect stars'.

4. Stars not found with the 'First look' as indicators of the reliability of the Input Catalogue

In March 1991 (date of the last update of the operational version of the Input Catalogue), some 59 400 different objects had been already observed with the 'First look': just half of the Input Catalogue. Of these, only 102 fell into the category of 'suspect stars' (137 transits). A careful examination of these 102 objects shows that they were all errors in the Input Catalogue, and not detection or reduction problems. Therefore they are clear indicators of the reliability of the Input Catalogue. The types of errors may be listed as follows:
(1) no object at Hipparcos position, correct object in the vicinity ($\leq$ 1 arcmin): 39 objects, all components of double or multiple star. Five errors are due to incorrect proper motion;
(2) correct position, but object fainter than expected, and not variable: 23 objects;
(3) wrong star: 21 objects, the replacing one being fainter than the expected one, and thus not found; 15 of them are single stars;
(4) variable star, temporarily or definitively too faint: 11 objects. Most of them are very irregular variables, with important and unpredictable drops in brightness (like stars of RCB type). The AAVSO watch organized during the mission shows that these irregular variables were precisely in an extinction phase;
(5) others: 8 objects, mostly with unresolved identification problem. Such objects were deleted from Input Catalogue.

All these errors were corrected in the last update of the Input Catalogue used for satellite operation, but the errors to be found in the future will not be corrected in the Input Catalogue. A straightforward extrapolation shows that the printed version of the Input Catalogue will contain at most 100 empty targets which will be deleted from the final Hipparcos Catalogue (and hopefully less, owing to other corrections also included in the last update). From this brief analysis one may conclude that:
(1) In the Input Catalogue, data on double and multiple stars are statistically of lower quality than those on single stars. This problem was known in advance, and therefore a specially dedicated working group was set up in the INCA team. All the work performed on double stars by this working group and the other INCA groups has greatly improved the rather bad initial situation, by clarifying the geometry of many systems and introducing improved coordinates and magnitudes. The new CCDM catalogue, now being prepared at Bruxelles Observatory under the leadership of J. Dommanget, should be an enormous improvement. Before it is completed, the extract to be published in one of the Annexes of the Input Catalogue will already clarify many cases (Dommanget 1989). Nevertheless, position errors important enough to remove a star out of the instantaneous field of view of the satellite's detector (offset by more than about 20 arcsec) and produce an empty target are found here only for double stars. The latest single stars with uncertain coordinates were all corrected with the help of the Guide Star Catalogue (GSC) of the Hubble Space Telescope (see Crifo et al. 1991). Because double stars may not always be separated below 20 arcsec in the GSC, this method was not applied to double and multiple stars.
(2) Identification errors are at a very low level: only about 20 stars will be 'replaced' by a fainter one in the printed Input Catalogue. They should be deleted from the final Hipparcos Catalogue. However to this number one should also add the stars replaced by an equally bright one, or by a brighter one, which are harder to detect. Even so, we estimate that the number of identification errors remaining in the printed version of the Input Catalogue should not exceed 150, and some of them will be removed from the final Hipparcos Catalogue. Some of these errors, which are presently suspected but not yet clearly established, are indicated in the Notes of the printed version of Input Catalogue. The corresponding identification chart (if any, in the Annex), will be in error too.
(3) Magnitude errors, which may also result in loss of stars, are at a similar level. However some stars have a too high magnitude in the Input Catalogue and are of course detected without problem. We estimate that 300 to 350 stars have a magnitude error exceeding 1 mag. About 25 of them will produce an empty target and will be deleted from the final catalogue; the others will have a corrected magnitude, as observed by the satellite.

The above numbers of possible identification and magnitude errors were estimated from both the Utrecht 'First look' results, and from the comparison with the GSC performed at Geneva Observatory for the purpose of identification chart drawing.

5. Conclusion

The Input Catalogue, made in a relatively short time with rather severe requirements on stars not always clearly identified, has been, immediately after completion, submitted to the most severe test possible: the observation of each of the stars by the satellite. In contrast to what is the case with other catalogues, where erroneous stars, once recognized, are just left aside, for Hipparcos a star in error means lost observing time, degradation of the sphere solution, reduced yield of the scientific proposals. After almost 18 months of observation, the 'First look' system developed at Utrecht by the FAST Consortium clearly shows that only some 100 empty targets of the 118 000 (0.085 per cent) are to be found in the observing programme and in the printed version of Input Catalogue. Therefore the Input Catalogue may be considered as very reliable. Not more than 150 identification errors are expected. Most of these errors will be removed from the final Hipparcos Catalogue.

References

Crézé M., Nicolet B., Chareton M., 1989, ESA–SP 1111, II, 47
Crifo F., Grenon M., Jahreiss H., McLean B.J., 1991, Adv. Space Res., vol.11 (2), 137
Dommanget J.,1989, ESA–SP 1111, vol. II, 149
Gómez A., 1988, in Proc.'Sitges Coll.', Torra J.& Turon C.(eds), 33
Gómez A., Morin D., Arenou F., 1989, ESA–SP 1111,vol.II,23
Perryman M.A.C., Turon C., 1989, ESA–SP 1111, vol.II
Schrijver H., 1989, ESA–SP 1111, vol.II, 373
Schrijver H., 1991, Adv. Space Res., vol. 11 (2), 51
Turon C., Gómez A., Crifo F., Grenon M., 1989, ESA–SP 1111,vol.II,7
Turon C., Gómez A., Crifo F., Crézé M., Perryman M., Morin D., Arenou F., Nicolet B., Chareton M., Egret D., 1991, this issue (B1)

This article was processed by the author using Springer-Verlag TEX A&A macro package 1991.