A&A manuscript no.
(will be inserted by hand later)

Your thesaurus codes are:
11(05.01.1; 03.20.4; 04.03.1; 03.13.2)

ASTRONOMY AND ASTROPHYSICS
March 13, 2000

Construction and verification of the Tycho-2 Catalogue ⋆

E. Høg[1], C. Fabricius[1], V.V. Makarov[1], U. Bastian[2], P. Schwekendiek[2], A. Wicenec[3], S. Urban[4], T. Corbin[4], and G. Wycoff[4]

[1] Copenhagen University Observatory, Juliane Maries Vej 30, DK-2100 Copenhagen Ø, Denmark
[2] Astronomisches Rechen-Institut, Mönchhofstrasse 12–14, D-69120 Heidelberg, Germany
[3] European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany
[4] United States Naval Observatory, 3450 Massachusetts Ave N.W., Washington DC 20392-5420, USA

Received 28 Dec. 1999 ; Accepted 7 March 2000

Abstract. The construction of the Tycho-2 Catalogue is described. The catalogue contains astrometric positions and proper motions as well as two-colour photometric data for 2.5 million stars. The Tycho-2 positions and magnitudes are based on precisely the same observations as the original Tycho Catalogue (hereafter Tycho-1) collected by the sky mapper of the ESA Hipparcos satellite, but Tycho-2 is much bigger and slightly more precise, owing to a more advanced reduction technique. Components of double stars with separations down to 0.8 arcsec are included. Proper motions precise to about 2.5 mas/yr are given as derived from a comparison with the Astrographic Catalogue and 143 other astrometric catalogues, all reduced to the Hipparcos celestial coordinate system. Tycho-2 supersedes in most applications Tycho-1, and the ACT and TRC catalogues based on Tycho-1.

Key words: Astrometry – Techniques: photometric – Catalogs – Methods: data analysis

1. Introduction

The ESA astrometric satellite, Hipparcos, produced a wealth of astronomical observational data of unprecedented quality during its 3.5 year long in-orbit operation. Its two measuring instruments, the main modulating grid and the auxiliary star mapper, were used by the Hipparcos scientific consortia to produce two major products, the Hipparcos (HIP) and Tycho-1 Catalogues. They were published along with complementary annexes comprising astrometric and photometric information on over 1 million stars (ESA 1997), hereafter referred to as ESA97. Volume 4 of ESA97 describes the Tycho-1 Catalogue construction and is referred to as EVol4 hereafter. The Tycho-2 Catalogue is presented by Høg et al. (2000a, b).

The main instrument of the satellite observed in a predetermined way only small patches of the sky for the 120 000 brightest stars and selected stars of particular astrophysical interest (Perryman & Hassan 1989, Perryman et al. 1995). The operation of the star mapper was however of a survey type. In the 3.5 years of the mission, the Tycho instrument swept over the entire sky in such a way that each spot was scanned many times in different directions. This resulted in a telemetry data flow transmitted to the ground, representing chiefly long continuous time series of photon counts from two spectral passbands, B_T and V_T.

Generally, the entire area of the sky was observed by the star mapper. The number of significant observations was however limited by the moderate size of the telescope, the sensitivity of the detector and the relatively high sky background. The star mapper was not designed to observe fainter stars than the main grid instrument, serving basically the purpose of satellite attitude determination. It was only owing to the thoroughness of the observation and data analysis processes with the star mapper that led to a catalogue of nearly ten times as many stars as in the main Hipparcos Catalogue.

In the original Tycho-1 project astrometric and photometric data for 1 052 000 stars to a limit of V_T =11.5 mag were derived. Magnitudes in the two passbands, B_T and V_T, were obtained for the brightest 99 per cent of stars. The median precision in photometry was 0.012 mag for V_T of the stars with $V_T < 9$ mag, and 0.06 mag for the whole catalogue. A precision (median internal standard error) of 7 mas was achieved in astrometry (positions, annual proper motions and parallaxes) for the bright stars, and 25 mas for all stars. Double stars with separation larger than 2 arcsec were resolved, and duplicity down to 0.4 arcsec was sometimes detected.

The processing of Tycho observations for the Tycho-1 Catalogue was based on detections of individual transits over a slit group (cf. Sect. 2) above a certain signal-to-noise ratio (i.e. 1.5 or 1.8).

⋆ Based on observations made with the ESA Hipparcos astrometry satellite. *Send offprint requests to:* E. Høg : erik@astro.ku.dk

An improvement of the limiting magnitude of the Tycho-1 Catalogue by about 0.4 mag could have been obtained by superposition of the photon counts for each star from two consecutive transits in the two fields of view of the instrument. This was however beyond the available capabilities with respect to software development and computing facilities when discussed in 1991.

For the Tycho-2 Catalogue a much better solution, namely *a photon superposition for the whole mission* has been realised. This idea was proposed in principle by Høg et al. (1982), but had also to be abandoned at that time. The photon superposition gives better astrometric and photometric values for the fainter half of the original one million stars. A major reason for the improvement is that an astrometric estimation based on very few photons does not achieve the Cramér-Rao limit (see Yoshizawa et al. 1985). An estimation based on all photons from many transits comes much closer to that limit. Furthermore, many transits of the faint stars were below the limit of signal-to-noise ratio for the detection and were thus completely lost. Figures for the number of lost Tycho transits are given in Sect. 20.4 of EVol4.

The Tycho-2 Catalogue is the major result of a new reduction of the original Tycho observations (Høg et al. 1999) which improves the astrometric and photometric precision of the stars in the original Tycho-1 Catalogue (EVol4 and Høg et al. 1997) and increases the number of stars by a factor of about 2.5. Such improvements have become possible with the much better computer facilities available today, using the Hipparcos and Tycho Catalogues and the calibration data from the first Tycho reduction, and by combining several new catalogues into a Tycho-2 Input Catalogue.

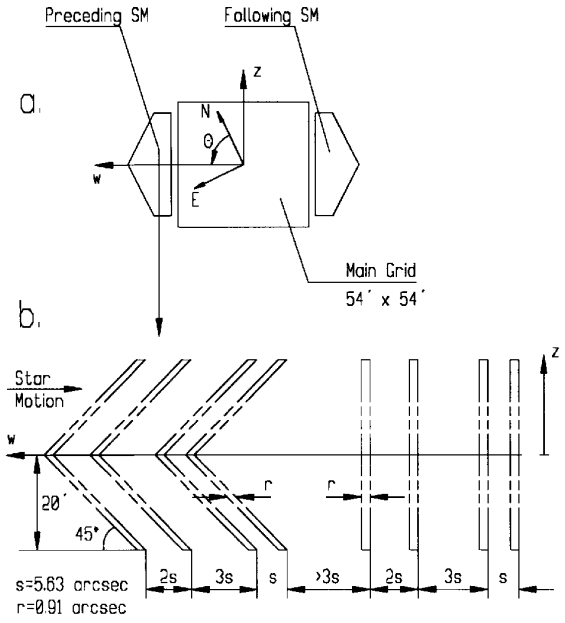

Fig. 1. The slit systems at the focal plane of the Hipparcos telescope. The telescope superposes two fields on the sky separated by 58 deg. (a) Arrangement of the star mappers and the main grid. Only the preceding star mapper was used during the mission, the following one being provided for the sake of redundancy. (b) The star mapper. The group of 'vertical' slits of the star mapper is perpendicular to the motion of the stars, while the 'chevron' slits are inclined by 45 deg. The light from the whole star mapper area is divided by a dichroic beam splitter onto two photomultiplier tubes which count the photons simultaneously in two bands: B_T and V_T.

2. Observations

The ESA satellite Hipparcos carried out astrometric and photometric observations from 1989 to 1993. This paper is concerned with the Tycho instrument of the satellite, which served an auxiliary role for the attitude determination, but it was also used to derive precise astrometric and two-colour photometric results for 1 million stars. The instrument contained two groups of slits in the focal plane of the telescope, 4 inclined and 4 vertical, as shown in Fig. 1. The 8 slits of 40 arcmin length across the scan direction continuously scanned the sky while the satellite rotated around its spin axis.

Two photomultiplier tubes, one for each passband, register the combined light from all the 8 slits. Two fields of view, 58° apart, are superposed in the focal plane. A transit of a star across the four inclined slits of the star mapper, and then a few seconds later across the four vertical slits, generated 4+4 separated peaks at known distances in the photon count records. The positions of the peaks contain astrometric information on the instantaneous position of the star, and the height or amplitude of the peaks above the sky background provides photometric information on the brightness of the star.

Because each of the slits covered 0.9 arcsec by 2400 arcsec, it happened quite often that more than one star crossed one of the slits at nearly the same time. If the star under study was badly affected by such a disturbing star the transit in question was rejected. If the disturbance was smaller and from a known star, the disturbing signal was subtracted. Such subtraction played an important role in the Tycho-2 processing, but was not possible in the Tycho-1 reduction where only transit rejection was performed.

3. Data reduction

The Tycho-2 processing of all raw counts is based on the available satellite attitude, available astrometric and photometric calibration parameters from the first processing, and an input catalogue of 4.35 million stars.

3.1. Input catalogue

Construction of the Tycho-2 Input Catalogue (TIC2) was one of the basic stages of the whole processing. Two major requirements to an input catalogue should be met:

- accuracy, inasmuch as the star to be observed should be secured within about 2 arcsec of the predicted position due to the size of extracted and treated count samples;
- completeness, so that all stars on the sky should be included for which a significant result could be obtained.

The Tycho-2 Input Catalogue was constructed by including stars from a hierarchy of large catalogues. In doing so, the first requirement was relatively easy to meet, as the accuracy of positions was generally much better than 2 arcsec. The photometric quality of large photographic catalogues is however poor, and gross errors in magnitudes are quite frequent. Early numerical simulations and analytical estimations showed that the expected results for stars fainter than $V = 12.2$ and $B = 12.7$ mag would hardly make sense with a total of 100 observations and an average background of 3–4 counts/sample in the B and V channels, respectively. The total number of observations is in fact strongly position dependent (cf. ESA97, Vol. 1, p. 403), rising sometimes above 400, and the background value is both position and time dependent (Wicenec & van Leeuwen 1995). The Tycho-2 faint limiting magnitude is thus not constant, varying in the range 11.8 to 12.5 mag in V. Due to the uncertainty of both the Tycho-2 faint limit and of the magnitudes from the source catalogue, objects with substantially fainter magnitudes had to be selected. On the other hand, many millions of uselessly faint stars would burden too much the computing capabilities and break our time schedule. As a result of a trade-off, a limit of $V \approx 12.5$ mag was finally adopted for inclusion into the input catalogue.

Most of the stars were obtained from the Hipparcos and Tycho Catalogues giving 1.1 million stars, the STARNET (Röser 1995) giving 2.4 million, the Guide Star Catalog version 1.2 (Röser et al. 1997) giving 0.5 million and the USNO-A version 1.0 (Monet et al. 1996) giving 0.3 million stars. While Hipparcos and STARNET provide proper motions of sufficient precision, the Tycho proper motions needed improvement and the other catalogues contain no proper motions at all. For Tycho stars precise proper motions were taken from a preliminary version of the Tycho Reference Catalogue (Høg et al. 1998) and the ACT (Urban et al. 1998b). Both these catalogues combine the Tycho positions with positions from the Astrographic Catalogue to derive the proper motions. For the remaining stars proper motions could be constructed for 0.3 million stars by combining GSC and USNO-A positions with sufficiently large epoch differences. In the end 0.5 million stars were left without proper motions. For the Hipparcos stars the parallax was also included. We reckoned that neglecting proper motions for the 0.5 million mostly faint stars could hardly affect the final positional accuracy. The proper motion should have been about 30 mas/yr to produce a discernible smearing of the star image at the faint end over the mission duration. In order not to miss these potentially important stars, we constructed 90 000 'proper motion entries' mostly from GSC entries not matching in position the nearest USNO-A1.0 entries of a different epoch, where a significant motion might be guessed. The result of this effort was modest with only a few hundred of these stars accepted in Tycho-2. Most of the finally adopted stars resolved this way, have in fact moderate proper motions and/or can be found in the NLTT catalogue, implying good completeness of the existing proper motion surveys in the given magnitude range.

The positions in TIC2 generally have an accuracy about 0.25 arcsec at the observation epoch, but deviations up to 2 arcsec do occur and were acceptable in the processing. Blended images of double stars in the photographic surveys (STARNET, GSC and USNO-A1.0) were by necessity entered in TIC2 with a photocentre position. When the separation exceeds 4–5 arcsec both components are normally lost to Tycho-2.

3.2. Prediction of transits

The aim of the prediction is to compute expected times of slit crossings of the input catalogue objects in terms of on-board clock units using the following information:

- *apparent* position of the objects as function of time;
- so-called on-ground geometric calibration of the instrument;
- orbital position and velocity of the satellite as functions of time;
- attitude of the satellite as a function of time.

As is seen from the list, the bulk of the raw data (observed photon counts) was not involved in the prediction computations, therefore this procedure was relatively fast. The input data listed above were readily available from the first processing, and essentially the same algorithm was used as described in EVol4. The largest task was to set up a new interface with the new input catalogue. The final on-ground, most precise attitude was used instead of the on-board 'real-time' attitude used largely in the first Tycho reduction (see Chapt. 11 of EVol4).

In computing apparent positions of the stars, input catalogue proper motions and parallaxes were used, whenever available. The parallaxes were known only for the HIP stars. We estimated that a neglected parallax of above 30 mas can cause a considerable smearing of the star image. It can have some effect on the astrometric position in case of non-uniform distribution of observations over a year, and a systematic change of estimated magnitudes in any case.

A large number of prediction files resulted from the prediction runs, ordered chronologically. Each record in a file contained full information on a slit crossing of one of the 4.3 million objects, including e.g. the star number in the input catalogue, the expected time of the event in units of on-board clock, and the same in the Terrestrial Dynamic Time scale in units of 0.0001 yr, predicted geometric field coordinates of the crossing, parameters of the scan motion, input catalogue magnitudes and assorted flags required in the consecutive reductions.

Unfortunately, the so-called 're-doing prediction' software which had to be adopted from the first Tycho processing, had a hidden bug which weakened the astrometric results for stars brighter than about $V_T = 8$ mag, but for the bulk of fainter stars had no noticeable effect.

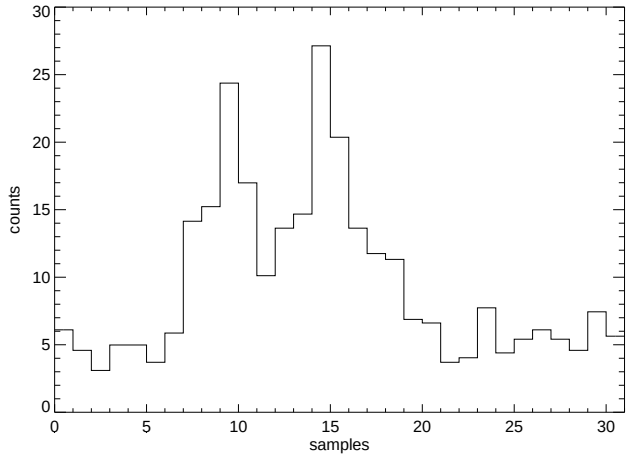

Fig. 2. An elementary observation, i.e. a series of 31 samples in the V_T passband, selected for a predicted transit time (within the central sample). The star is double, therefore two peaks are seen. A sample in the plot is actually the average of four original samples observed at the four slits (see Sect. 3.3).

3.3. Identification of transits and astrometric calibrations

The original observations (80 Gb of telemetry data) were provided as large sets of continuous series of 6400 photon counts each, given strictly in time sequence as they were measured by the photomultipliers. These data had to be matched and combined with the output prediction files, so that photon counts corresponding to each transit could be found in the raw data, and be extracted and stored in a data base of direct access. This stage was called 'photon counts identification', and the resulting data set 'identified counts data base' (ICDB). Technically, the task of identification was for each predicted transit of an input catalogue object to find a series of photon counts in the raw data, centered on the predicted time, and to record it in the data base. Astrometric corrections were introduced at this stage, viz. time dependant large-scale calibration corrections and z-coordinate dependant medium-scale irregularities. The corrections were applied to the predicted transit time, expressed in terms of the on-board clock. This was done in much the same way as in the first Tycho reductions, (see EVol4, Ch. 7). The so-called reprocessing geometric calibration parameters were used, consistent with the prediction software version and the attitude files. The variations and scatter of the 11 reprocessing calibration parameters with time, describing the instrumental distortions, rotations and shifts, reflect degradation of the prediction data with respect to the main Tycho-1 processing (cf. Figs. 7.2 and 7.3 of EVol4) presumably due to the hidden software error. The tables of the medium-scale irregularities, describing geometric imperfections of the star mapper slits were also used in the same way as before. These were assumed to be constant in time.

The observed photon counts were compressed into single-byte numbers. Before any processing could be done, they had to be decoded into 2-byte integer numbers by special decoding tables. The next step was a folding of the decoded series of counts (6400 in each) with a sliding filter. Each transit of a star across the 4 slits (vertical or inclined) of the star mapper generated 4 peaks in the photon counts, which had to be folded or convolved into one, referred to a certain fiducial line. The non-linear filtering technique was adopted as described in Chapter 4 of EVol4. The non-linearity of the filter was a sophistication of a simple 4-point sliding filter designed to cope with side lobes and cosmic hits in the data. By the folding, the resulting counts were no longer integer numbers, and the intrinsic photon shot noise decreased a factor of 2. It is noted also, that due to the width of the filter, some additional loss of data occurred after folding at the edges of disrupted series.

For each estimated transit time of a star a photon sample was sought which contained the predicted moment in its interval of 1/600 s. On both sides of this central sample, 15 continuous samples were selected, making a series of 31 samples. This series spans 8.7 arcsec on the sky with the vertical slits, and 6.2 arcsec with the inclined slits. Thus, for each predicted transit two series of counts were selected and stored in the ICDB, for the B_T and V_T channels separately. An example of an elementary series of counts is shown in Fig. 2.

The basic astrometric information is contained in the shift between the actual position of the peak and the predicted value. Knowing the geometry of observations, a correction to the input catalogue position and, in principle, to the proper motion and parallax can be computed from these shifts. Contrary to the first Tycho processing, the shift was not determined for each observation, but all observations were treated together to derive the two-dimensional positional correction. Hence, an important simplification and improvement of the data process-

ing scheme was achieved in that detections of transits were no longer required.

3.4. Disturbed transits and parasite rejection

A considerable improvement was achieved in the way disturbed transits were treated in the Tycho-2 processing. As the total size of the star mapper slits projected on the sky amounts to $2.7 \cdot 10^{-3}$ deg^2, the probability that more than one star crosses the slits nearly simultaneously is not negligible. In that case, the predicted star signal would be disturbed by another overlapping signal. Each observation can be affected differently due to the changing scan directions. The probability of disturbances is of course higher at multiple systems or dense stellar aggregates, but even a single isolated star can occasionally be disturbed from the other field of view. Since the input catalogue is fairly complete with bright stars, almost all disturbing bright star transits can be accurately predicted. In the Tycho-1 processing, a disturbed observation was simply rejected if an interfering star brighter than a certain limiting magnitude was present within a certain distance. A similar rejection, called *parasite rejection*, was implemented again. If there was at least one disturbing star brighter than 9.4 mag in V_T transiting within 3.4 arcsec interval of along-scan motion, centered on the predicted moment, the observation was not used in the following reductions (but still recorded in the ICDB).

The majority of disturbing signals were fainter than 9.4 mag or outside the rejection interval. If neglected, they can deteriorate the astrometric and photometric estimation. It was decided to *subtract* such disturbing signals if the stars producing them have sufficiently accurate magnitudes and positions in the input catalogue. The height of the disturbing signal can be calculated from the magnitudes in the catalogue, assuming constant brightness, and the relative location from the astrometric position. Technically, the relative position is fully specified by the difference in the predicted (corrected) transit times of the star in question and the disturbing star. Up to 3 disturbing transits could be recorded within an interval of $31 + 8$ samples (11 arcsec) centered on the given predicted time, from the stars with brightest input catalogue magnitudes. If both fields of view are at galactic latitude $b = 0$ there is a probability of only about 0.001 for having more than 3 disturbances in the interval. At star densities twice as high the probability is about 0.01. About one million stars from the Tycho-1 Catalogue, having accurate B_T and V_T magnitudes and positions, were eligible for the recording in the ICDB. Other stars, presumably much fainter, were not considered as possible interferences.

3.5. Recording in the ICDB

For each transit of a given programme star across one of the two groups of slits of the star mapper the following data were recorded in the ICDB:

- input catalogue star number;
- predicted time of transit, precise to 0.1 day, in Julian days;
- various flags describing the status of the observation and the status of the star in the input catalogue;
- phase, i.e. the relative position of the predicted transit time within the duration of the central count in the sample;
- position angle of the instantaneous scan direction, counted from north through east;
- instantaneous scan velocity relative to the nominal one (168.75 arcsec/s);
- data for a maximum of three interfering 'parasitic' stars, such as the relative transit time, the input catalogue B_T and V_T magnitudes and some flags;
- so-called 'format' sky background in the two passbands separately;
- two series of 31 photon counts each for the two passbands, centered on the predicted transit time.

3.6. Mapping technique

Each series of photon counts representing a transit of a star can be thought of as a one-dimensional image of the small patch of the sky around the predicted instantaneous position. This one-dimensional image results from a convolution of the real two-dimensional image with the infinitely long slit response function. In fact, each transit is a combined image of two widely separated patches, as the telescope points in two viewing directions simultaneously. As the spot is observed in many different orientations throughout the mission, the contribution of the other field of view has a rather random character, just adding to the sky background and increasing the frequency of interfering star passages.

Let $c_{ij}, j = 1, \ldots, 31$ be the ith series of photon counts for a star in the input catalogue. In the tangential plane to the instantaneous predicted position of the star a local rectilinear coordinate system is defined so that the x axis points in the east direction, and the y axis in the north direction. It is noted that the predicted position, and thus the origin of the local coordinate system moves on the sky with respect to an inertial reference system as governed by the proper motion and parallax in the input catalogue. The geometry of the observation is then fully described by the three numbers:

- the phase ϕ_i, i.e. the position of the origin within the central sample $j = 16$;
- the position angle of the slit normal p_{ni}, readily obtained from the position angle p_{si} of the scan direc-

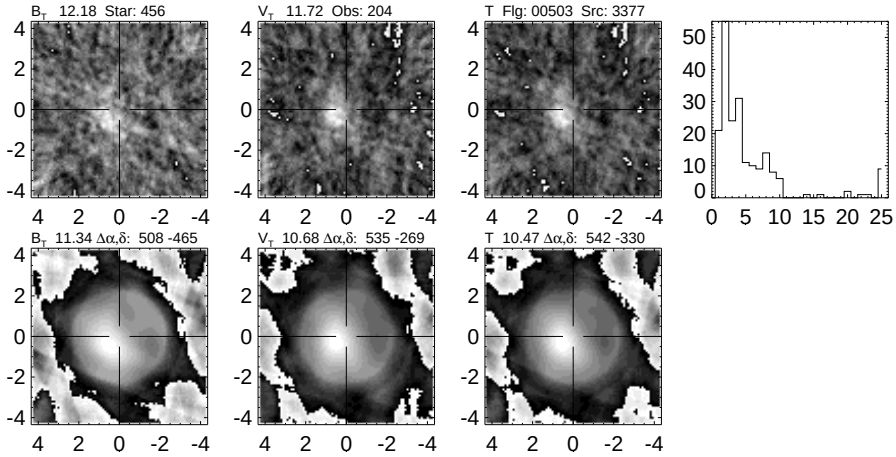

Fig. 3. Digital maps of the counts for the star #456 in the passbands B_T, V_T and the combined $T = B_T + V_T$, and a histogram of the counts in T (upper row). The lower row shows the corresponding cross-correlation maps, i.e. the maps constructed from the original counts convolved with the appropriate slit response functions for each pixel in the map. The star is a typical moderately faint object in the input catalogue taken from STARNET. A noticeable offset is seen from the input catalogue position at the centre of each map.

tion and relevant slit group flags given in the identified counts record;

– the scaling factor of the instantaneous scan velocity s_{vi} with respect to the nominal value.

Then at any point (x, y) of the map the corresponding intensity is

$$I_i(x, y) = c_{ik}, \tag{1}$$

where

$$k = \left\langle \frac{\sqrt{f_{\rm slit} + 1}}{0.28125\, s_{vi}} \boldsymbol{r} \cdot \boldsymbol{d}_i + 16 + \phi_i \right\rangle, \tag{2}$$

where $\langle q \rangle$ denotes the integer part of q, $f_{\rm slit} = 0$ for transits across the vertical slits and 1 for the inclined, 0.28125 arcsec is the sampling interval at the nominal scan velocity of 168.75 arcsec/s, and the vectors $\boldsymbol{r} = (x, y)$ and $\boldsymbol{d}_i = (\sin p_{ni}, \cos p_{ni})$.

Having computed such intensities at a fine grid of points, one would obtain a map for one transit, which would consist of 31 parallel strips of different brightness. The width of the map would be 6 arcsec for inclined slit transits and 8.4 for vertical. The next step is to combine all observations of the star into one map, in either of the passbands or in both of them together. This can be done by simply adding up all intensities:

$$I_{\rm tot}(x, y) = \sum_i I_i(x, y), \tag{3}$$

or by averaging with weights in an attempt to reduce the noise in the resulting image. When a few transits of different position angles are co-added, the star would appear on the map as a concentration of light of somewhat radial pattern. When many observations are combined, the bright spot assumes a round appearance, more like a photographic plate or CCD star image. Fig. 3 shows a map for a star in B_T, V_T and the combined light (so-called T passband), where a simple weighting of individual transits was done, inversely proportional to the highest counts.

In principle, such combined digital maps can be used for two-colour photometry and astrometry with the well-tested techniques for two-dimensional images. That would however be quite awkward to do on the 4 million target stars, as the maps do not contain any new information compared to the original transit samples, and computing them is a time-consuming task. Nothing can be gained in terms of solution accuracy by the finer sampling of the already sampled photon counts, nor by projecting the intrinsically one-dimensional images onto a two-dimensional map.

Yet, the mapping technique proved to be very useful for visual inspection of difficult and special cases, e.g. double stars and high proper motion candidates. It was extensively used for testing the solution methods and algorithms, specially with regard to their robustness and efficiency on heavily disturbed or scarce observations.

4. Astrometry and photometry of single stars

Single stars constitute the majority of the Tycho-2 Catalogue. Any star having a separate entry in the input catalogue was treated as a single object, no matter if there was a close companion. Since all components of HIP double and multiple systems had separate entries, even com-

ponents of binaries as tight as 0.1 arcsec were originally treated as single objects.

4.1. The estimation strategy

The cornerstone difference between the first and second Tycho processings was in the way the general characteristics of the stars were estimated. In the first Tycho reductions, each series of counts of the kind shown in Fig. 2 was treated separately, and it produced as a result a separate astrometric observation equation. All linearized equations for a given star were then weighted and solved in a least-squares adjustment. Similarly, the amplitude of each transiting star image was estimated, and the mean magnitude was derived from a set of collected amplitude estimations.

In the second Tycho processing, however, all count series for a given star were first added with appropriate weights, or 'stacked' on top of each other, and then both astrometric and photometric estimates were derived from the accumulated signal. (In reality, this process must be tried for a series of corrections to the assumed star position.) At first glance, the resulting precision should not be dependent on the sequence of steps, that is whether the estimation is performed first and then the averaging, or the reverse. According to the simplistic error propagation law, the gain in precision is always about $\sqrt{N_{\rm obs}}$, where $N_{\rm obs}$ is the total number of observations. The important fact is that this law is not valid at small signal-to-noise ratios (SNR) of ≈ 1. It is empirically known that observational data with SNR below 3.0 to 3.5 are statistically poor, and traditional ground-based astrometry did not use data at this noise level. With satellite observations, it is quite normal that one has to observe objects at the limit of detectability. As this was the case with Tycho, a substantial effort was spent on the theoretical development of the error model for faint signals, accompanied by numerical simulations.

Yoshizawa et al. (1985) showed that the Maximum Likelihood (ML) estimator for astrometry, which is known to be the optimal one, approaches the Cramér-Rao lower bound, specified by the inverse of the Fischer information matrix, only for 'large samples', that is at high SNRs. With 'poor samples', the performance of the ML estimator, as well as other estimators in use, deviates from the Cramér-Rao limit further and further with decreasing SNR. The authors found a rather good approximation for the ML estimation variance by using a Taylor expansion of the inverse of the information matrix.

This approach was improved and developed by Makarov & Høg (1995), which led to a numerically justified astrometric error model used throughout the Tycho-1 reductions. Very generally, the variance from one observation can be represented in this model by

$$\sigma^2 = k_1 A^{-1} + k_2 (\mathrm{SNR})^{-2} + k_3 (\mathrm{SNR})^{-4} + k_4 (\mathrm{SNR})^{-6}, \quad (4)$$

where A is the signal. For SNR $\gg 1$ the last two terms are negligible, and the formula becomes an approximation for the Cramér-Rao lower bound. At SNR ≈ 1 the high-order terms begin to dominate. In that case, increasing the SNR in the above formula a factor of $\sqrt{N}$ by stacking N observations before doing any actual estimation gives a much higher gain in precision than a factor of $\sqrt{N}$. For extremely faint signals, the gain may be dramatic in the sense that the uncertainty of each estimation, taken separately, can be infinitely large, while stacked data can still produce useful information.

This fact is illustrated in Fig. 4, where a randomly chosen single series of counts is shown of a faint Hipparcos star, and a stacked image from 133 observations. We neglected the fact that the predicted position of the star can be anywhere within the central sample duration in one transit. Estimating the position and magnitude of the star from the single observation seems pointless. Without knowing that the star *must* be at the centre, one of the noise peaks at the edge of the series could be picked up. The presence of the star is quite obvious in the accumulated data, and a reasonable unambiguous estimation can be done.

The change in the data reduction strategy is the basic reason for the increase in the number of stars and improvement of precision at faint magnitudes with respect to the Tycho-1 Catalogue. It may be noted, that, even though realised long ago, the present stacking technique could not have been implemented in the first Tycho reductions, being at the time too demanding in terms of computer facilities.

4.2. Background determination

According to the adopted observation model, the photon counts c_{ij}, $j = 1, \ldots, 31$ in the series number i for a given star, are composed of two elements, the background B_i and the star signal of amplitude A_i:

$$c_{ij} = A_i P_i(\tau_j - \tau_c) + B_i, \quad (5)$$

where P_i is the slit response function (that is the response to a transit of a point-like source), τ_c is the true position of the star within the given series, τ_j is the position of the j-th count. The positions are expressed in units of on-board time. The chosen normalization of the slit response function is

$$\sum_{j=1}^{31} P_i(\tau_j - \tau_c) = 1, \quad (6)$$

so that the observed amplitude A_i is interpreted as the integral signal (illuminance) of the given star, and can easily be calibrated into magnitudes. An important simplifying assumption for the Tycho-2 processing (as well as for Tycho-1) is that the background is constant within the single observation series, i.e., it does not vary significantly

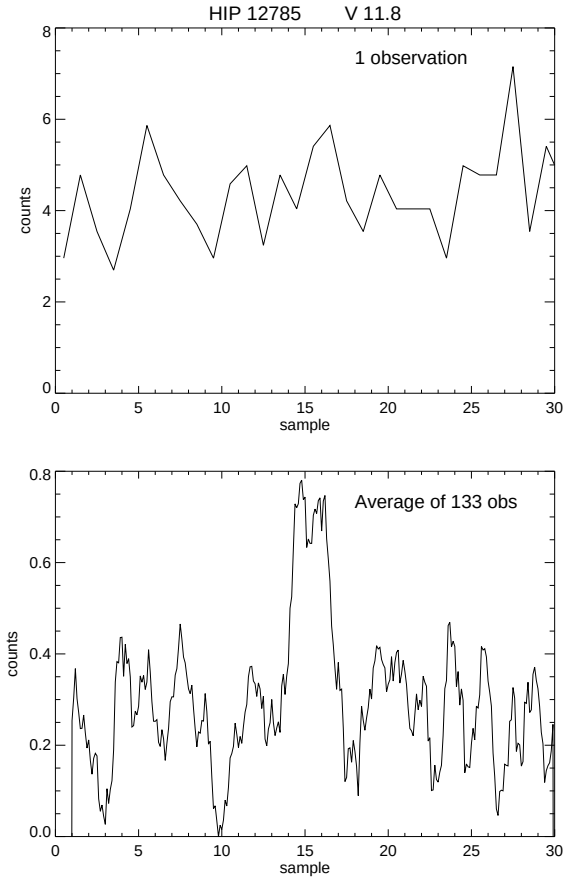

Fig. 4. Observations of the faint star HIP 12785, $V = 11.8$ mag. Upper plot: a randomly selected series of 31 samples, i.e. one observation. Lower plot: the average counts from 133 observations of this star.

in the nearest vicinity of the star. For the vast majority of stars this assumption is absolutely valid. The constant background value B_i must be determined and subtracted from the counts, prior to making any photometric estimation. Astrometric estimation, however, can in principle be independent of the background determination, given the above assumption.

The background value is usually determined as either the mean or the median of a selected series of counts, preferably undisturbed by stars or other objects' signals. This is never known with certainty, and the presence of an undetected object raising the determined value, can not be precluded. The median, being more indifferent towards a few enhanced counts, performs better than the mean. We chose to use the 31 counts recorded in the ICDB for background determination, thus inferring this value from the nearest surrounding of each star. It is recalled that in Tycho-1, a much longer series of counts was typically used (6400 counts for the format background), and sometimes a smaller, but still rather extensive series of 192 samples when a non-uniformity had been noticed. The use of a longer stretch of samples gives a better precision, but may be systematically inaccurate in the presence of local background enhancements. Since it is not the background precision that defines the final photometric performance for faint stars, but rather the Poisson shot noise in the counts, the loss of precision due to the background estimation was reckoned acceptable.

After numerous numerical simulations and photometry test runs, the following algorithm was adopted. Firstly, the median of the 31 collected sample values, $c_{\rm med}$, is computed. This value, if taken as a background estimate, would be biased by the presence of the star signal, leading to systematically too faint magnitudes. On the other hand, the star signal contribution is negligibly small for the majority of faint stars and the median would be good enough as a background estimation. We therefore iterate the estimation by computing the median among the counts which are smaller than $c_{\rm med} + 2.7\sqrt{c_{\rm med}}$. The second term accounts for the surplus of high counts well above the expected Poisson noise, caused by the presence of the star in question or possible disturbing stars. An obvious flaw of this method is that around the brightest stars, it still produces too bright background estimations.

4.3. Maximum cross-correlation estimator

When chosing among the various methods and techniques for image location estimation, one has to take into account their theoretical performances in a given problem, as well as practical merits and reliability. The theory of astrometric estimation was much advanced at the preparatory stage of the Hipparcos project. An extensive simulation was made e.g. in connection with the Tycho data reduction. The performance of different kinds of estimators, such as maximum likelihood (ML), least squares (LS), trimmed median (TM) and maximum cross-correlation (MCC) were assessed (Yoshizawa et al. 1985). It was found that for high signal to noise observations, when the star is comfortably bright with respect to the background level all these methods provide a very similar performance, and theoretically the ML estimator is equivalent to the lower Cramér-Rao bound. This is no longer the case when the signal to noise ratio is so small that the star signal is hidden in the photon noise pattern. Since every effort had to be done to push the faint limit of the Tycho catalogue as far as possible, the choice of estimator became quite important. Finally a simple linear filtering technique resembling to some extent the trimmed median, was adopted for the astrometric estimation in the first Tycho processing, and a maximum likelihood method for the photometric reduction.

In the present second processing, the idea of a simple linear filtering was abandoned. The greatest practical advantage of these methods, i.e. the computational quickness was not an issue any more since much more powerful computing facilities had become available. Further-

more, some practical drawbacks of the antisymmetric sliding filter used for Tycho-1 could not be disregarded, especially the multiplicity of solutions for each observation and a considerable loss in angular resolution of double stars (Makarov et al. (1994) private comm. TDAC-TD242).

The ML and LS methods, being essentially non-linear, are based on an analysis of photon counts residuals. That implies that the astrometric unknowns (coordinates of the star) and the photometric ones (B_T and V_T magnitudes) are solved in one run, which could hamper the solutions for many faint stars burdened by frequent interferences. For that reason the choice was eventually confined to the more robust TM and MCC estimators which can be used directly for the astrometric problem. In order to judge between the two, a rather detailed Monte-Carlo simulation was undertaken.

The simulations were carried out with the following simplifying assumptions: observations in only the V_T channel, only on vertical slits, 100 crossings per star in random scan directions, constant background. The knowledge of the background, attitude parameters and instrumental calibration was assumed to be perfect. For each of the magnitudes $V_T = 8$, 9, 10, 11, 12 and 13, a hundred such stars were simulated. Thus, 10 000 observations were simulated for each magnitude. The Poisson shot noise was the only one simulated in the counts. The TM filter was cut off at 0.8 arcsec from the centre. A cut-off at 1.0 arcsec was also tried, but found to be slightly inferior.

In order to compare the performance of the TM and MCC astrometric estimators, we set up a simulation software in IDL, which was later used as a prototype for the real reductions code. Positions of simulated star images were estimated for each series of counts and compared to the true positions. A robust estimate of the standard error, i.e. the half difference of the 85 and 15 percentile on the distribution was derived. A comparison of the coordinate errors showed exactly the same performance of the two estimators at magnitudes brighter or equal to 10. At magnitudes 11 and 12, the MCC estimation produced a 6% more precise result, but, interestingly, for extremely faint stars (12.5 and 13.0 mag), the precision of the TM estimation was better by 4% and 12%. We decided in favour of the MCC estimator from this experiment, since the majority of our objects are 11 to 12 mag in V_T. Another reason for choosing the MCC as an astrometric estimator was that it should obviously be more robust to random disturbancies. If some of the observations are partly overlapping with other star signals, the disturbance will have a smaller effect on the estimated position than in the case of TM filtering, thanks to the downweighting of the counts at the wings of the light distribution, when the observed series is convolved with the slit response function. Another advantage of the MCC method is that it takes into account in a natural way the asymmetry of the slit response functions (especially for inclined slits). Finally, we found the MCC estimation to be faster in computing.

The cross-correlation of an observed series c_{ij} is a convolution with the expected slit response function:

$$C_i(\tau) = \sum_{j=1}^{31} c_{ij}\, P_i(\tau_j - \tau). \tag{7}$$

The MCC location estimation is then the position τ at which the estimator function $C_i(\tau)$ has its maximum. If assumed in the image model (5) that the background is constant within the series, and the function $P(\tau_j - \tau_c)$ is finite (i.e., the whole star image is entirely contained in the series), the MCC estimator becomes equivalent to the LS estimator in its effect. The practical advantage of the MCC method is that it does not require a knowledge of the amplitude A_i, which allows to use it before and independently of a photometric estimation.

The core idea of the Tycho-2 processing was in fact to perform astrometric and photometric estimations on the combined and weighted observations rather than on each observation separately. The combined MCC estimator is therefore

$$C_{\rm tot}(\boldsymbol{r}) = \sum_i w_i\, C_i(\boldsymbol{r}\cdot\boldsymbol{d}_i/v_i), \tag{8}$$

where $v_i = s_{vi} v_{\rm nom}$ is the instantaneous scan velocity, w_i is the weight of ith observation and the parameters $\boldsymbol{r}$, $\boldsymbol{d}_i$, s_{vi} and $v_{\rm nom}$ are the same as in Eq. 2. The estimated position vector $\boldsymbol{r}_{\rm e}$ where $C_{\rm tot}(\boldsymbol{r}_{\rm e})$ obtains its maximum gives an offset of the star in tangential coordinates with respect to the input catalogue position.

4.4. Downhill simplex minimization

The central estimation problem is to find the maximum of the total cross-correlation function, using Eqs. 7 and 8 and provided the weights are known. Among the variety of numerical methods dedicated to such problems, the Downhill Simplex Algorithm was chosen (Press et al. 1992). It is conceptually simple and well-suited for minimization of 2D functions (the task can be converted into a minimization problem by inverting the sign of $C_{\rm tot}$). The procedure begins with placing three initial points within the area where the function is defined, and thus forming a triangle (a simplex in 2D). The function values are computed at the points and compared with each other. Governed by certain rules, the three points are relocated in the course of following iterations, so that the triangle can undergo reflections, expansions and contractions. The process can be visualised as a triangle of constantly varying form and size, crawling across the area until it finds the position of the minimum, and infinitely shrinks around it. The procedure is terminated when the size of the triangle becomes smaller than a certain limit, 2 mas in our case.

Although not the fastest method, it has some appealing properties. It does not involve gradients estimations,

which in our case are impossible. We also found it very robust when applied to faint stars data, where the $C_{\rm tot}$ function may have secondary maxima and valley-like maxima.

4.5. Iterative cleaning and weighting of observations

Despite all the efforts with interference (parasite) recording, rejection and subtraction, some fraction of the collected observations may be corrupted by unrecorded interferences from other stars, or simply by wild distortions of unknown origin. Generally, such bad observations may constitute only a few percent of the data, yet they were recognized as a major threat to the result, especially for fainter stars in dense regions or in multiple systems. Therefore, special care was taken to find and eliminate the most disturbed observations for each star.

The degree of disturbance of each series c_{ij} can be evaluated from the counts themselves. Once the background and the amplitude are known, the observation model (Eq. 5) can be subtracted from the observed counts. The residuals should mostly represent Poisson noise, so the χ^2 statistics over the whole set of observations should have a Poisson distribution with 29 degrees of freedom. An unexpectedly high value of χ^2 indicates a high degree of disturbance, and the observation should be rejected. Having on average 120 observations per star, we computed the 30, 50, 70, 80 and 90th percentiles of the χ^2 distribution. For each star, the 80th percentile was compared with a limiting value $\chi^2_{\rm lim}$, and all observations with χ^2 values exceeding *the largest* of the 80th percentile and the $\chi^2_{\rm lim}$ were dropped. The limiting value was chosen such, that the majority of undisturbed stars had their 80th percentiles well below it. Thus for each star, a certain fraction of observations were eliminated from the solution, but never more than 20 percent. The actual rejection levels, along with the other percentiles for each of the two channels, were recorded in the solution log files. They were used later on in the final selection of solutions for the catalogue. Solutions with too high χ^2 percentiles and rejection levels were dropped as unreliable. It is noted that a gross variability of the star can cause high χ^2 values.

This cleaning procedure implied an iterative processing, because at the start the amplitudes of the star are not known and the χ^2 values can not be computed. A zeroth (initial) iteration is required, without any cleaning of the observaton set. Another reason for repetitions of the whole estimation process is that the proper weighting of observations is also based on amplitude estimations. Thus, one full iteration included:

- star position estimation by the downhill simplex method on the combined $C_{\rm tot}$ function;
- photometric estimation in B_T and V_T;
- cleaning of the observation set by χ^2 statistics;
- calculation of weights and formal errors.

Three such iterations were performed for each star.

Two different kinds of weighting of individual observations were used for photometric and astrometric estimations, insofar as the error models were different (Sect. 4.6). In both cases, however the weights were inversely proportional to the expected variances of the estimations, assuming a pure Poisson noise in the data. The weighting of observations was, of course, done separately in B_T and V_T, but for the astrometric estimation in the next iteration, the two cross-correlation functions were combined, again with appropriate weighting.

4.6. The signal to noise ratio, formal errors and covariances

The positions discussed here are those derived from the satellite observations, not the mean positions also given in the catalogue. The standard errors of positions and magnitudes, as they are given in the catalogue were computed solely from analytical error models for individual observations. They were not corrected for unit weight errors representing the real observed scatter because the residuals of individual observations were not available in the Tycho-2 method. The errors are therefore called *formal* errors. The photometric error model was adopted unchanged from the Tycho-1 processing, cf. p. 125 in EVol4. This model is actually identical to the ML approximation, Eq. 4, limited to the first two terms. Given the errors of individual observations, the formal standard errors of the mean magnitudes for each object are computed by proper weighting. The real magnitude errors in the catalogue are probably 50 % larger than the formal standard errors at magnitudes fainter than 10, as shown in Fig. 6b.

Unlike the first Tycho reductions, the same type of error model was used for astrometric error estimation, limited to only two terms in the ML approximation. The higher order terms were neglected due to the anticipated rise of the signal to noise ratio. The parameters of the model were recalibrated by a comparison with high-precision Hipparcos positions. For faint stars, observations on the inclined slits and in the B_T channel received as a result a relatively higher weight than they had in the previous reductions. A full covariance matrix for each star was computed from these formal errors, combining them with individual weights, the B_T and V_T weights and slit position angles. The correlation coefficients between the α and δ position components are given in the catalogue, but they are of limited use for the general user, since the more accurate mean positions should be used in most applications, and they have uncorrelated components.

The signal to noise ratio, f, was computed for each star in the course of astrometric/photometric treatment, based on the co-added and weighted observations. Contrary to the first Tycho processing, such a ratio was never

computed for individual transits. In the framework of the MCC method, the signal to noise ratio is derived as

$$f = \frac{\sum_i w_i C_i - \sum_i w_i B_i}{\sqrt{\sum_i w_i^2 \sum_j \mathrm{var}[c_{ij}] P_{ij}^2 + \sum_i w_i^2 \mathrm{var}[B_i]}}, \qquad (9)$$

where $P_{ij} = P_i(\tau_j - \tau_e)$, τ_e is the estimated position $\boldsymbol{r}_e$ projected onto the scan direction, and $\mathrm{var}[c_{ij}]$ is readily derived from Eq. 7.

The signal to noise ratio was the main quantity to discriminate stars which were too faint or too poorly observed, or simply were spurious false entries in the input catalogue. Typically, only solutions with $f > 2.5$ were accepted for the catalogue. Sometimes, stars with f down to 2.3 were however accepted if we had confidence in their reality. The signal to noise ratio was not the only criterion for acceptance. We planted a large number of 'empty spots' in the input catalogue, i.e. randomly distributed positions on the sky with no objects. These fake target stars passed the same reduction procedures as normal stars. They were useful afterwards to estimate the rate of false detections with specific limits on f. Curiously, some empty spots produced detections with f up to 5 due to severe interferences. Therefore, employing more selection criteria was deemed necessary to damp the false detection rate. Finally, the set of criteria was an intricate system of f, χ^2 values, offset from the expected position and the 12 points discussed in the following section.

4.7. The 12-points method

We define 12 points at regular intervals on a circle of radius 0.8 arcsec centered on the estimated position $\boldsymbol{r}_e$. The cross-correlation value C_{tot} is computed at each of the 12 points (Eq. 8). By definition, all these values are smaller than $C_{\mathrm{tot}}(\boldsymbol{r}_e)$. If the number of observations for the star is sufficiently large, and the distribution of slit position angles is even, the 12-points values must be all about $\frac{3}{5}C_{\mathrm{tot}}(\boldsymbol{r}_e)$, representing the nominal sharpness of the cross-correlated star image. In practice, the distribution of slit directions is never quite uniform, and some modulation of the 12 values as a function of direction occurs, corresponding to elongation or distortion of the star image on a map like Fig. 3. It is therefore necessary to correct the 12 values for the directional non-uniformity of observation by normalizing them onto unit weight. Such corrected values proved indispensable as duplicity indicators (cf. Sect. 5) and for the rejection of false or unreliable objects. For instance, a false detection generated by chance intersections of bright parasitic transits, was often marked by 12 points with too low or too high median values, or a wild modulation with position angle.

4.8. Photometric calibration

The photometric calibration, i.e. the transformation of the observed signal (C_{tot} in our case) to stellar magnitudes had to be done almost from scratch. The well-established and detailed calibration of the first Tycho processing (see Chapter 8 in EVol4) could not be used due to a very different estimation method and, possibly, different adopted slit response functions. A simpler version of photometric calibration was used. It includes zero-point calibration constants and a universal time dependent trend, representing the flux sensitivity of the photometers and its change with time. The trend was approximated by a linear function (constant degradation of the photocathodes), as estimated from the Tycho-1 calibration. The zero-point corrections were determined anew from large samples of individual observations of photometric standard stars. The colour and slit abscissa-dependent terms were neglected.

It turned out later that Tycho-2 magnitude zeropoint was still magnitude-dependent, as described in Sect. 7.1, and a calibration refinement was carried out.

4.9. Special corrections

Because only 31 counts were extracted around the predicted transit time for a star, only stars closer than about 2.5 arcsec from the TIC2 position had an acceptable coverage of their signal by the slit response function. Some stars had, by chance, two TIC2 entries typically differing by one or two arcseconds. Such two entries, where the star had different offsets from the predicted position, should lead to the same position and the same magnitudes for the star. By comparing the actual results for entries with a large offset with entries with a small offset, an offset dependent calibration was established which was applied to all stars. The magnitudes were corrected by up to 0.07 mag and the positions by up to 0.15 arcsec.

5. Double star and photocentre solutions

For the double star processing, about 120 000 candidate stars were selected. Different criteria were used for the selection in an attempt to include as many resolvable pairs as possible. At the same time, the number of candidates was limited by the available computing capacities and time. The major selections were made in the following three groups.

Group 1. About 6 000 entries of TIC2 constituted the first selection, for which no routine single star solution had been achieved due to parasite rejection of all the observed transits. That could happen when for a given entry there was another entry in the input catalogue within 6 samples and brighter than $V_T = 9.4$ mag, as recorded in the catalogue. Then all the projected distances in time between the two entries crossing the slits were below 6 samples (1.7 arcsec in the scan direction), and consequently,

no observations could survive the parasite rejection. The brighter component of such a pair, however, could have a normal routine solution if its companion was fainter than 9.4 mag. In case of successful double star solution, the resolved entries should replace their single star solution counterparts, in that the latter were usually impaired by the very uncertain magnitudes in the input catalogue. For the sake of consistency, rejected double star solutions should never be used for the output catalogue construction, even though some of the brighter components may seem reasonable. Hence, a special treatment was required for unresolved close pairs in order to retain them in the output catalogue, which was essentially a single star solution, but without the parasite rejection and subtraction.

The vast majority of the stars in group 1 were narrow doubles from HIP, since only this catalogue of all used for the TIC2 construction, has a sufficient angular resolution, viz. down to 0.1 arcsec. There were also a handful of resolved Tycho-1 binaries and spurious or redundant entries from other catalogues. The group 1 reduction resulted in 1 500 resolved objects in Tycho-2.

Group 2. This selection comprised objects, for which a significant routine solution had been obtained, but which showed a prominent 'duplicity signature' in the 12 points. The 12 points, as described in Sect. 4.7, are the cross-correlation values computed at 12 regularly spaced points on a circle of 0.8 arcsec radius centred on the estimated position of the star. The routine single star solution provides a photocentre position for a visual double, therefore the character of the duplicity signature in the 12 points depends on the separation and magnitude difference of the pair. When the separation is not too big, and the components have comparable magnitudes, the photocentre takes a position near the middle, and there are two maxima in the 12 points as a function of position angle, at approximately opposite directions. This type of modulation can be approximated by a $c_0+c_{2c}\cos(2p_{ni})+c_{2s}\sin(2p_{ni})$ function. When the separation is not much wider than 0.8 arcsec and the secondary component is considerably fainter than the primary, there is only one hump in the 12 points in the direction of the secondary, since the photocentre position is much closer to the primary. This type of modulation is better approximated by $c_0 + c_{1c}\cos p_{ni} + c_{1s}\sin p_{ni}$. A harmonic analysis of the 12 points, restricted to the terms of second order would provide 5 coefficients, from which information on the amplitude of modulation and position angle of the system could be gleaned.

Unfortunately, this conceptually simple method proved vulnerable to the confusing effect of bright interfering signals and side lobes. A more intricate but empirically well-tested *merit* function was used instead, that combined the amplitude of the 12 points, their median value and the signal to noise ratio of the single star solution with the measured degree of the image elongation, as derived from the same 12 points. The cross-correlation map of each star was approximated with a two-dimensional ellipsoidal function by a least squares fit on the 12 points, and the degree of elongation was estimated as the ratio of the semiaxes. Obviously, this selection gave preference to components of similar magnitude separated by about 0.8 arcsec. The type of the merit function and its limiting value was carefully tested on Hipparcos double and single stars in order to increase the success rate on the sample. The selection gave by far the largest sample of 65 000 candidates, with a partial overlap with the Group 1.

Group 3. This group comprised stars selected from a compilation of known Tycho-1 visual doubles (Halbwachs 1994, private comm.) with separations 0.4 to 2.5 arcsec. It gave 13 500 very good candidates, but more than half of the solutions were not included in the catalogue, the separations being below 0.8 arcsec.

A few more groups based on various criteria, were selected (e.g. stars with large positional differences between Tycho-1 and Tycho-2, which could be an indication of duplicity), but proved less successful. The majority of solutions accepted for the catalogue were in the first three groups. All the candidates were treated by the same ad hoc double star programme, irrespective of the way they had been selected. The programme is a sophistication of the single star algorithm, based on the same maximum cross-correlation method, the simplex downhill algorithm for astrometric estimation, similar χ^2 iterative cleaning, weighting of observations and the photometric calibrations. But instead of a straightforward search for the maximum of a 2D cross-correlation function, the algorithm should be able to find two, partly overlapping images of the components, and unambiguously resolve them astrometrically and photometrically. The iterative method of repeated estimation and subtraction, developed for this task descends from the more general cleaning algorithm of image processing. Each iteration in this method includes two principal steps:

- astrometric and photometric estimation on the whole set of (cleaned) observations;
- subtraction of the star image from each series of counts with the just estimated position and B_T, V_T magnitudes (cleaning).

The first such iteration gives a photocentre solution, with a position closer to the brighter component but probably somewhat shifted towards the fainter one. The estimated magnitude is likely to be too bright, since part of the signal from the secondary is assigned to the primary. After the first subtraction, an asymmetric dip would be seen on the digital map at the place of the primary, and a displaced hump at the position of the secondary. The first magnitude estimation for the secondary is systematically too faint. As iterations proceed, the estimated brightness of the secondary is growing and the primary is waning, until convergence is reached. The required number of iterations depends largely on the magnitude difference between the components. The worst case for this method is two

equally bright stars at small separation, when the residual signal after the first subtraction is so symmetric around the photocentre, that it prevents the algorithm from rapid convergence. The number of iterations could not be chosen arbitrarily large, since it is approximately equal to the factor of computational time increase. It was found experimentally that a good convergence was achieved after 12 iterations for the majority of solutions at separations greater than 0.8 arcsec.

Each astrometric/photometric estimation for each of the components was almost identical to a full-scaled single star solution, including the background estimation, χ^2 iterative cleaning and weighting, and auxiliary parameters calculation, viz. the 12 points, χ^2 distribution percentiles and signal to noise ratio. The major difference was that no parasite rejection or subtraction was carried out, for obvious reasons. The auxiliary parameters were used for the final selection of double star solutions to be included in the catalogue. Generally, much more rigorous acceptance criteria were used than for single stars, following the policy that a result as reliable as possible should be given in the catalogue. A careful analysis of the Tycho-2 solutions was performed by a comparison with well-established Hipparcos double star solutions. The success rate and reliability were checked on a sample of a few thousand single Hipparcos stars. The rate of false resolutions for single Hipparcos stars was estimated at less than 0.05 %. Good quality Tycho-2 solutions were often obtained for separations above 0.43 arcsec and, surprisingly, realistic results were sometimes derived at separations as close as 0.25 arcsec. For the Tycho-2 Catalogue the safer limit of 0.8 arcsec was however adopted, drastically reducing the number of resolved doubles. These rejected doubles, instead underwent a photocentre treatment.

5.1. Photocentre solutions

Quite a few stars were resolved in the input catalogue, i.e. the components had separate entries, but they failed to be resolved in the Tycho-2 processing, chiefly because the separations were below 0.8 arcsec. Since the routine single star solution for them was not good enough, they obtained a special *photocentre* solution, which was a copy of the single star solution, except that the parasite rejection and subtraction was switched off. This allowed them to enter the catalogue with positions corresponding roughly to the centres of light, and magnitudes combining in a rather undetermined way the magnitudes of the components. These solutions, a total of 8884, are marked with a special flag in the catalogue and should be used with caution, but the proper motions can still be valuable.

6. Catalogue construction

The final catalogue contains positions and magnitudes for 2 539 913 stars and for 96 % of the stars proper motions have been derived as described in Sect. 8. For all stars we give the *observed* position at the epoch of the Tycho observations, which is close to 1991.5 and slightly different in right ascension and declination. When a proper motion is derived we also obtain a *mean* position. The mean position is a weighted mean for the catalogues contributing to the proper motion determination. The corresponding mean epochs for RA and Dec are given in the catalogue. The mean position has been brought rigorously to epoch 2000.0 by the computed proper motion and is given in the catalogue. More than two catalogues were often used to determine the mean position and proper motion of a star, and bringing an *observed* Tycho-2 position to the epoch 2000.0 can therefore give a position slightly different from the mean position.

Due to the data compression used in the ICDB, stars brighter than about $B_T = 2.1$ or $V_T = 1.9$ could not be treated properly and are excluded from the catalogue. These stars are all in the Hipparcos Catalogue and we could never hope to improve their astrometry. For the convenience of the user, stars from the Hipparcos and Tycho-1 Catalogues which for one or the other reason are not included in the Tycho-2 Catalogue, have been listed in supplementary catalogues. The first supplement contains 17 588 good quality stars while the second supplement contains 1146 Tycho-1 stars which are either false or heavily disturbed by brighter stars. Any star from Tycho-1 with an astrometric quality flag 9 which was not detected in Tycho-2 processing was excluded from both supplements.

6.1. Identification of Hipparcos stars

The catalogue provides cross references to stars in the Hipparcos Catalogue and its annex of double and multiple stars with component solutions. In general a distance limit of 0.8 arsec was used for the identification, but in some 50 cases distances up to 10 arcsec were accepted after a close examination. The record is HIP 67593 which has a very poor solution in the Hipparcos Catalogue. It may not be fully meaningful to give such poor references, but it reduces the number of ghosts in the Tycho-2 Supplement.

As explained in Sect. 3.4 the signal from a neighbouring, disturbing star was subtracted, if possible, in the data reductions. A Hipparcos star which was subtracted in the reduction of some TIC2 entry, cannot be identified with the resulting Tycho-2 entry and must go to the Supplement. This happens only rarely, because close Hipparcos doubles normally underwent a special photocentre treatment where neighbouring stars were not subtracted and the Tycho-2 star is then identified with both Hipparcos stars.

Four Hipparcos stars fall outside the normal picture. In two cases three Hipparcos stars are identified with one Tycho-2 star: TYC 3534–00302–1 = HIP 90284 ABC and TYC 3732–01030–1 = HIP 21730 ABC. A few cases where two Tycho-2 stars had the same Hipparcos identification

were mostly resolved by deleting one of the Tycho-2 stars from the catalogue, leaving only HIP 67626 (a VIM) identified with TYC 7291–1949–1 and also with TYC 7291–1949–2. Finally the star HIP 114923, which is erroneously resolved in HIP, is identified with TYC 7513–1039–1 without any CCDM identifiers.

6.2. Identification of Tycho-1 stars

The purpose of the identification of Tycho-1 stars in Tycho-2 is partly to set a flag in Tycho-2 indicating a Tycho-1 star and partly to ensure a consistent assignment of TYC numbers to the Tycho-2 stars (see Sect. 6.3).

As for Hipparcos, a general distance limit of 0.8 arcsec was used for identification but 2.4 arcsec was allowed for the Tycho-1 quality 9 stars. The limit of 0.8 arcsec is of course somewhat arbitrary and many stars are close to this limit. The 0.8 arcsec was chosen because this is the chosen resolution limit for double stars in Tycho-2 and because many of the limiting cases turned out to be double stars where we simply picked different components in Tycho-1 and Tycho-2.

When one Tycho-1 star can be identified with two Tycho-2 stars, it is generally a resolved Tycho-1 star with a Tycho-2 double star solution. In these cases the brighter Tycho-2 component is chosen and it gets the TYC3 component number from Tycho-1 while the fainter component gets the first available TYC3 and is not flagged as a Tycho-1 star.

6.3. Assignment of TYC numbers

For Tycho-2 we have chosen to use the same system of star numbers as in Tycho-1 which is basically the same as used in GSC, the Guide Star Catalog (Jenkner et al. 1990), but with a component number appended.

Each Tycho-2 entry has a set of three identification numbers: TYC1, TYC2 and TYC3. TYC1 is the GSC region number, TYC2 is the running number within the region and TYC3 is a Tycho specific component number. Double star components will either have two different running numbers or the same running number and different component numbers. For Tycho-2 we have tried to be as consistent as possible with Tycho-1, so TYC numbers from Tycho-1 have only been used in Tycho-2 in the same sense as in Tycho-1.

In about 4500 cases, new GSC running numbers were constructed for Tycho-2 with the kind assistance of Dr. Daniel Egret and the Guide Star Catalog team.

6.4. The proximity indicator

When selecting stars from a catalogue like Tycho-2, it is often useful to know if there are other stars present in the neighbourhood. For this purpose we give a proximity indicator for each star, which is simply the distance to the nearest neighbour in the main catalogue or in the main supplement. Stars in the second supplement, of presumably false Tycho-1 stars, were ignored.

7. Verification

We describe in this section a few general verifications and checks that were made on the astrometric and photometric results. Both accuracy and precision of the data in the Tycho-2 catalogue are assessed, and reliability of the formal errors is discussed. General statistical properties, such as angular resolution of double stars and the number of false entries are also addressed. Values for the completeness of the catalogue to a limiting magnitude, and the number density of stars on the sky are given by Høg et al. (2000a, b).

7.1. Photometric accuracy, precision

The photometric accuracy of the Tycho-2 magnitudes can be assessed by means of a large set of photometric standards of good quality. Such a set must include a sufficient number of stars in the magnitude interval $V = 10$ to 12 mag, which are the vast majority of our catalogue entries. It is not easy to find suitable photometric sets of standards from ground-based observations. Eventually, we used two sets of photometric measurements as a check for our photometry, as explained below. Both of them were used for the final correction of the magnitudes, so they do not really provide an external source of verification.

The most important set is the 17 785 standard stars, compiled for the purpose of photometric calibration in the Tycho-1 reductions (EVol4, Ch. 8)[1]. The precise ground-based magnitudes of the standard stars were transformed into Tycho B_T and V_T magnitudes, which were made available to us (Halbwachs 1998, private comm.). The standard stars were indirectly used for the photometric calibration, i.e. the slit-dependent zero-points determination, as explained in Sect. 4.8. They were also used for the final correction of the magnitudes. Due to the different method of background determination and probably different slit response functions used, there was a slight systematic change in the derived median $B_T - B_{\rm stand}$ and $V_T - V_{\rm stand}$ as functions of magnitude at bright and medium magnitudes (i.e. $B_T < 11.5$ and $V_T < 11.0$). At fainter magnitudes, an increasing bias was detected. The source of the bias is most likely in the intrinsic property of the maximum correlation estimator, which inevitably picks up a position where the signal achieves its maximum, thus effectively increasing the flux estimated from noisy data. It was however not possible to derive the bias reliably from the comparison with the standards because of the unfortunate lack of faint stars in the sample. The

[1] The number of stars in common with the set in Tycho-1 was slightly smaller.

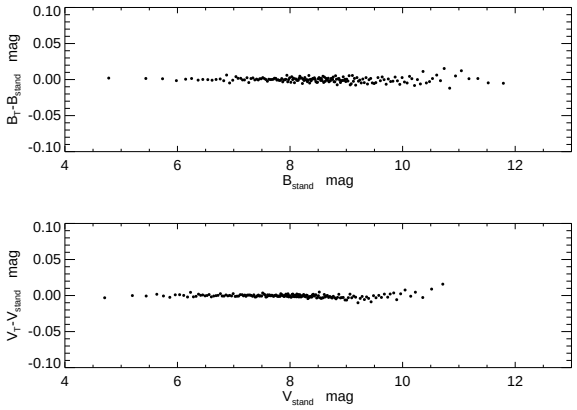

Fig. 5. Systematic differences between the Tycho-2 B_T and V_T magnitudes and the standard magnitudes transformed into the Tycho photometric system. Each dot is the median of 100 difference values in one bin.

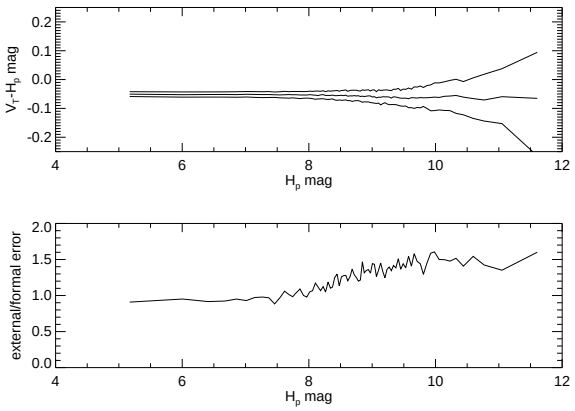

Fig. 6. The percentiles 15, 50 (the median) and 85 of the differences between the observed Tycho-2 and Hipparcos magnitudes (upper panel). The ratio of the robust external standard error, as derived from the percentiles 85 and 15, and the combined formal standard error in the two catalogues (lower panel).

sample is basically limited to 11.8 mag in B_T and 10.7 in V_T. The comparison with the standard stars was therefore used only to correct the Tycho magnitudes in the specified range of bright to medium magnitudes.

It was found that smooth spline approximations could be found for the functions $B_T - B_{\rm stand}$ vs $B_{\rm stand}$ and $V_T - V_{\rm stand}$ vs $V_{\rm stand}$. These approximated corrections were then subtracted from each magnitude in the catalogue. The corrected magnitudes (Fig. 5) show a satisfactory agreement with the standards.

A truly external check on Tycho-2 magnitudes is provided by a set of stars at the North Galactic Pole and in SA 80, 81 and 132 observed by Knude (1986, 1993, 1998 private comm.) with high accuracy in *uvby*. Only early type dwarf stars were used in order to keep a well defined transformation from the Strömgren system to Tycho. The *uvby* photometric data were first transformed to Johnson B and V using transformations by Jønch-Sørensen (1998, private comm.) and then transformed to B_T and V_T using the formulae in Sect. 1.3 of Vol. 1 of ESA97. The corrections for the faint end bias were thus obtained and applied to all faint magnitudes in the catalogue. Fortunately, the bias was not as dramatic as in the first Tycho processing, and it never approached the slope of -1, where any aposteriori correction would make no sense. A simple spline approximation was used again.

As a conclusion, the Tycho-2 photometric system has been verified with regard to systematic errors to represent very well the Hipparcos/Tycho-1 photometric system and the available set of ground-based standards. In the range of magnitudes fainter than $V \approx 11.5$, an additional set of highly accurate ground-based observations has been used to correct the faint end bias. It is believed that Tycho-2 magnitudes are systematically correct down to magnitudes 13 and 12 in B_T and V_T, respectively. Fainter magnitudes are not trustworthy, and a very faint magnitude in the catalogue should be understood to mean that the star was not detected in that colour.

In order to verify the standard errors of the Tycho-2 magnitudes, we compare the latter with Hipparcos Hp magnitudes in Fig. 6. Although the Hp magnitudes are completely uncorrelated with the Tycho measurements, the comparison is hampered by the great difference in the passbands in question. The Hp passband is much broader than either of the B_T or V_T, which brings a significant dependence of the comparison on the spectral type of common stars. By definition, the V_T passband is much closer to Hp than the B_T, but the transformation is not well known. Therefore, Fig. 6 (lower plot) gives the upper bound for the ratio "external error / formal error", the differences having been enhanced by the colour effects.

The upper plot in Fig. 6 shows the 15th, 50th and 85th percentiles of the differences $Hp - V_T$ for some 50 000 common stars in the colour range 0.2 to 0.8 mag in $B_T - V_T$. The bias after all the previously described corrections, is confirmed to be within 0.02 mag, as seen from the median difference. The robust estimation of the external error, corresponding to 1σ for a Gaussian distribution, is the difference between the 85th and 15th percentiles divided by 2. It reaches about 0.15 mag at magnitude 11.5. The ratio of the robust external error and the standard error (quadratically combined from the Tycho-2 and Hipparcos formal errors) in the lower part of Fig. 6 is close to 1 for stars brighter than magnitude 7.5 and then gradually rises to 1.5 at magnitude 10 and fainter. It is a satisfactory result, inasmuch as the external error is expected to be somewhat larger than the formal error, which leaves some sources of additional noise unaccounted, e.g. disturbances from other stars and positional errors.

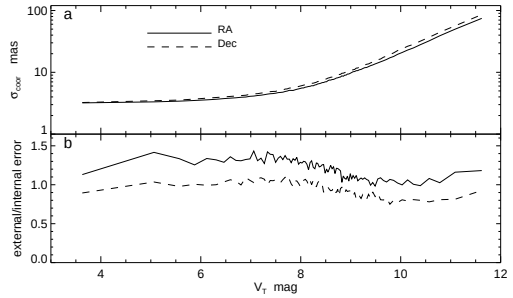

Fig. 7. (a): Internal positional standard errors for $\alpha \cos\delta$ and δ as given in the Tycho-2 catalogue. (b): The ratio of external to internal errors. It is typically 1.1 for RA and 0.8 for Dec for faint stars, i.e., the internal errors of positions are too conservative. For bright stars, $V_T < 9$ mag the ratio is about 1.1.

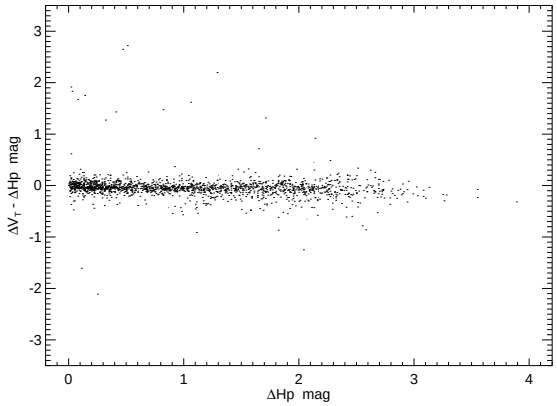

Fig. 8. The scatter of observed differences in V_T magnitudes between the components of close double stars, as compared to Hipparcos Hp magnitude differences. The plot shows 1752 Tycho-2 pairs in common with Hipparcos, with separations less than 2.5 arcsec.

7.2. Accuracy and precision of the observed positions

Unlike the first Tycho processing, the Hipparcos astrometric data provide an external reference to the Tycho-2 positions. The Hipparcos astrometry was not used in Tycho-2 reductions in any other way than for general calibrations of the instrument and attitude determinations. The errors of Tycho-2 coordinates as given in the catalogue are shown in Fig. 7a. The external errors have been measured as half the differences between the 85th and 15th percentiles of the residuals of 1000 stars per magnitude bin. The ratio of external to internal errors is shown in Fig. 7b. The errors of Hipparcos positions can be neglected at faint magnitudes. The figure confirms the anticipated improvement of positional accuracy at faint magnitudes with regard to Tycho-1. For example, the Tycho-2 external errors at $V_T = 10.5$ mag are about 32 mas in RA and 28 mas in Dec, according to Figs. 7a and b, to be compared with approximately 25 mas internal and 38 mas external standard error in both coordinates in Tycho-1.

The rotation between the Hipparcos and Tycho-2 reference systems was determined by the least squares method and found negligibly small. The systematic positional errors were checked and corrected in the same way as in the first processing (Sect. 11 in EVol4). The mean positional differences in cells of 6 by 6 deg^2 showed a rather smooth pattern rarely exceeding 2 mas. After a correction by means of a spherical harmonics representation, the typical systematic difference became well below 1 mas.

7.3. Double star solutions

The special double star processing, as described in Sect. 5, resulted in the inclusion in Tycho-2 of 7 500 resolved pairs with separations from 0.8 to 2.5 arcsec. Stricter criteria than for single stars were imposed to accept double star solutions for the final catalogue in an attempt to reduce the number of blunders. For instance, all solutions with separations less than 0.8 arcsec were rejected. The main hazards of the Tycho-2 double star solution include contamination of the signal by side lobes and interferences from proximate brighter stars and possible variability of the components. As the following very general analysis shows, an overall high quality of the results was achieved in Tycho-2, although a small number of mis-hits could not be avoided.

To verify Tycho-2 photometry for doubles is not a straightforward task. Existing ground-based observations in the separation range of interest are either scarce and limited to bright stars (Corbally 1984), or are not matching our data in precision. The Hipparcos double star data are superior to ours, and provide a truly external check, but the main instrument performed photometric measurements in just one very broad passband, Hp. The Hp passband nearly includes both B_T and V_T, and transformations to either of these are not known. Undoubtedly, the transformations will be strongly dependent on the spectral type. We compare directly the differences in V_T and Hp in Fig. 8, bearing in mind the additional scatter due to the colour variations. The figure shows that the vast majority of magnitude differences are in good agreement. There are however about 1.5 % gross errors, where the differences deviate more than 0.5 mag. Most probably, these are Tycho-2 errors, and their origin is not known. At ΔHp above 2.2 mag the scatter becomes larger due to increased observation errors and big colour variations.

It is simpler to check the relative astrometry for double stars, since Hipparcos provides an external reference. Figure 9 compares separations between the components of 1877 double stars as found in Tycho-2 with those in HIP. Only 'good' quality solutions in HIP were included (which

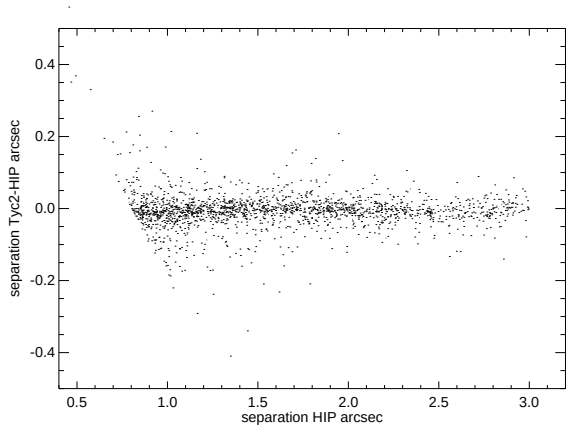

Fig. 9. Comparison of Tycho-2 and Hipparcos separations for 1877 doubles in common, with flag 'A' in H61 as a quality of solution.

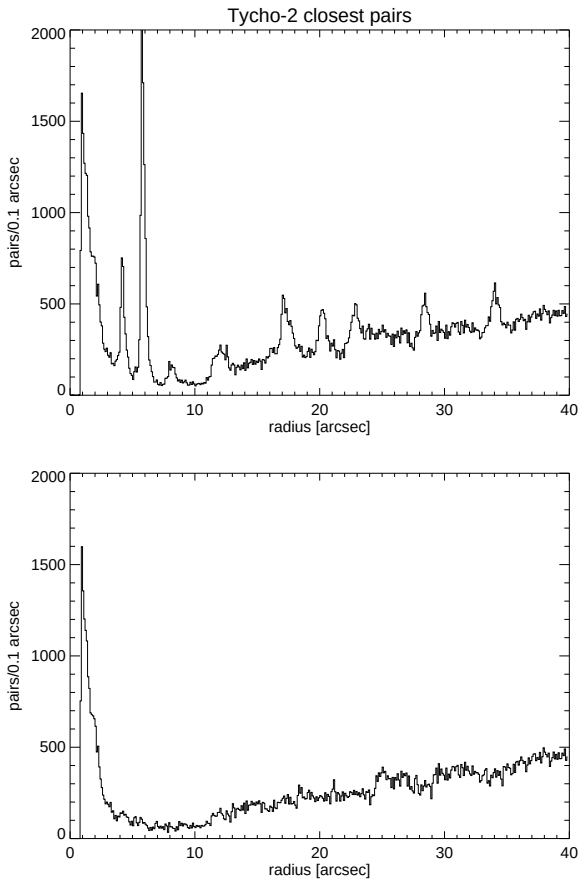

Fig. 10. Distribution of distances to the nearest star in the Tycho-2 Catalogue, before (a) and after (b) deletion of spurious entries caused by slit response side lobes (see Sect. 7.4).

actually eliminates a few extreme outliers). We know that the formal astrometric standard errors for close doubles in Tycho-2 are too optimistic, but the overall precision of the relative astrometry seems to be better than 50 mas. There are gross errors again, where the Tycho-2 and HIP data are discordant. Up to a few percent of erroneous solutions in HIP can not be precluded, as was demonstrated by a few authors successfully resolving transit data for abut 200 problematic stars (Fabricius & Makarov 2000, Falin & Mignard 1999, Söderhjelm 1999), but they must be rare among the quality 'A' solutions. It is provisionally concluded that there is an unknown source of occasional gross errors in the Tycho-2 double star method that should be born in mind by a user. An inspection of these bad solutions shows that they are not always associated with extremely small separations or large magnitude differences. Perhaps, a large-amplitude variability can cause a confusion in a Tycho-2 double star solution.

7.4. General statistical properties

The nearest-neighbour distance statistics, i.e. the distribution of the distances to the nearest neighbour for each star is a powerful tool to study clustering properties of a given set of objects. For a random (Poisson) distribution of star positions, a theoretical expected distribution function can be computed and compared with the observed histogram. This has been done on different samples of stars, e.g. at selected star formation regions (Gomez et al, 1993), and around the North Galactic Pole (Bahcall & Soneira, 1981). The tests usually show a shift of the observed median nearest-neighbour distance from the expected $0.47/\sqrt{\theta}$, where θ is the average number density over a region of a size Ω, $\theta = N/\Omega$, and N is the total number of objects in the region. The shift towards smaller distances stems from an excess of small-separation pairs due to physical binarity of stars, and perhaps other clustering phenomena.

We have used the nearest-neighbour distribution of all Tycho-2 stars for the purpose of verification and cleaning the catalogue of false entries. Fig. 10, upper panel, is the distribution for the originally compiled catalogue. The resolved physical pairs are the main constituents of the highest peak at around 1 arcsec, close to the angular resolution limit of 0.8 arcsec. The secondary peaks at 4.1, 5.8, 8.2, etc., are caused by sidelobes of the instrumental slit response and diffraction rings around bright stars in combination with the non-linear numerical filtering. The peaks appear at the separations of $n \times 5.8$ arcsec for the vertical slits and $n \times 4.1$ for the inclined slits ($n = 1...6$). In a small interval around these positions an excess number of faint companions were found which were then eliminated by application of suitable limits of Δm, i.e. entries at a critical distance from a much brighter star were deleted.

Using the mapping technique, we visually checked a large number of spurious pairs and found an efficient set of

criteria, based on the SNRs of the components and the χ^2 statistics, enabling a rather clean deletion of false objects. The lower panel of Fig. 10 shows the nearest-neighbour distance distribution for the whole catalogue after the final cleaning.

Of course, a relatively small number of false, i.e. non-existing, objects may remain in the catalogue, and a number of real stars affected by some disturbances have been deleted by the cleaning. An additional thousand entries were eliminated from a similar visual check of suspicious cases. For instance, whenever an entry with faint magnitudes was found for a star from the GSC or USNO A1.0 input catalogues with much brighter expected magnitudes, a digital map was constructed and inspected. More often than not, the entry appeared obviously wrong. Such false entries in big numbers were inherited from these two photographic catalogues, where spikes from overexposed star images had often been taken for star images.

8. Proper motions and mean positions

The proper motions in Tycho-2 were computed by combining the Tycho-2 positions with those from earlier transit circle and photographic programs. Note that this is somewhat different than the procedure used for ACT (Urban et al. 1998b) and TRC (Høg et al. 1998), where the proper motions were computed from essentially two positions: from Tycho-1 at epoch 1991.25 and from the Astrographic Catalogue (AC) at epoch about 1905. The Tycho-2 proper motions were derived using 143 transit circle and photographic catalogues in addition to the AC and Tycho-2 positions. Including these extra catalogues, generally observed at epochs between the AC and Tycho-2, allows for a stronger determination of proper motions, greater investigation of individual stars' errors, and a higher confidence of identification especially in the case of high proper motion stars. A total of 870 384 stars contained positions from catalogues other than the AC and Tycho-2, i.e. 35 per cent of all Tycho-2 stars. The full list of catalogues included in the proper motions is given in the Guide to Tycho-2 (Høg et al. 2000b).

Computing the Tycho-2 proper motions was essentially performed in five separate tasks: 1) putting all ground-based catalogues on the system defined by Hipparcos, including a complete re-reduction of the Astrographic Catalogue data; 2) computing weights for the individual catalogues; 3) computing proper motions and errors via weighted least-squares; 4) removing discordant observations; 5) repeating steps 3 and 4.

8.1. Ground-based catalogues to the Hipparcos System

The International Astronomical Union has designated the Hipparcos Catalogue as the optical realization of the International Celestial Reference System, ICRS, (IAU 1998). Therefore, all catalogues being used to compute proper motions (*source catalogues*) should closely follow Hipparcos; systematic errors with respect to positions, magnitudes, and colours should be minimized.

For all source catalogues with the exception of the AC, the procedure to convert to the Hipparcos system started with identifying all stars in common to each source catalogue and Hipparcos. Differences between an source catalogue and the Hipparcos positions at the source catalogue's epoch were computed using the Hipparcos proper motions. At this point, only the high quality HIP stars were used; those stars with G, V, X, or O codes in the HIP Field H59 were excluded.

Next, systematic errors based on right ascension and declination were removed. For each star in the catalogue (the 'central star'), identification of the nearby stars (in position space) in common with HIP was made. Weights were computed for each catalogue position of a HIP star according to how close it was to the central star; those closest received the most weight. The weighted mean residual was then computed and this value was applied to the central star. This process was continued for every star in the source catalogue.

The area surrounding the central star with which to bring in the HIP data changed from one catalogue to another. Ideally, one would like to perform this very locally. However, the area used was a function of the density and precision of each source catalogue. In general, the catalogues resulting from astrograph observations had a radius of 2 degrees around the central star. The lower density transit circle catalogues usually had 5 degrees in declination, and 30 minutes of right ascension.

The same procedure was used to remove systematic errors with respect to colour and magnitude. Weighted, positional residuals were computed for every star using nearby (in position-colour space, position-magnitude space, and colour-magnitude space) observations of HIP stars. These residuals were applied to the source catalogue positions. Following this, each catalogue was investigated to ensure the differences between it and HIP could be explained with random errors only. At this point, the catalogue was said to be on the system defined by HIP.

It was decided not to use the AC 2000 (Urban et al. 1998a), but to re-reduce the AC plate measures using a new reference catalogue and using the photometry from Tycho-2. The former would include many new catalogues reduced directly to HIP, thereby making the plate reductions more accurate. The latter would facilitate the removal of systematic errors based on magnitude, which has been a problem in the AC since its inception. The details regarding this reduction are beyond the scope of this paper; they will be given elsewhere (Urban et al. 2000, in preparation).

8.2. Weighting and combining the data

Due to the varying precision of the data used to compute the proper motions, weighting of the catalogue positions is necessary. Many of the source catalogues do not contain error information, so this must be computed by using well-observed stars; the process is explained below. Where the number of observations per catalogue position coordinate exists, weights per position coordinate are computed from the overall catalogue accuracy and the number of observations going into each position. Where the number of observations per catalogue position does not exist, the weights per position coordinate are simply a function of the catalogue standard error only.

Weights were determined for most catalogues in the following manner. First, all stars with 5 or more catalogue positions were identified. Next, unweighted mean positions and proper motions were determined, excluding the positions from the catalogue for which weights are being determined. Then, the residuals for each star were determined. Following the removal of gross outliers, the variance of the residuals was then computed. The weight of a position coordinate is the inverse of its variance. In actuality, some transit circle catalogues had a very large spread of the number of observations per star. Therefore, where possible, weights as function of catalogue standard error and the number of observations per position coordinate are used.

For a few catalogues, the weights were determined differently. Weights for positions from Tycho-2, Second Cape Photographic Catalogue (Zacharias et al. 1999) and the Twin Astrograph Catalogue (Zacharias & Zacharias 1999) were derived directly from each stars' internal standard error in each coordinate given in those catalogues. For the AC positions, the weights were a function of magnitude, declination zone, and number of observations. The weights for each catalogue are given in the Guide to Tycho-2.

With the catalogue positions all on the system defined by Hipparcos, and the appropriate weights associated with each, positions for each star are combined in a weighted least-squares adjustment using the techniques detailed by Corbin (1977).

The superior ability of the Tycho instrument to measure individual components of close binary systems where transit circles and astrographs measure centre-of-light, created a unique problem. It was decided to try to compute proper motions for these systems by combining the photocentre positions found in the astrograph catalogues with *computed* photocentres of the Tycho pairs. Since photocentres are dependent on the colour of the stars, the Tycho blue magnitudes were used to compute the Tycho photocentre; only astrograph catalogues were used for the early epoch positions since photographic plates are generally blue sensitive. The shift in the photocentre from the position of the brightest component of a close pair is computed as:

$$\Delta = \frac{\rho}{2 \times 2.5^{\Delta \mathrm{mag}}} \tag{10}$$

where ρ is the separation of the two stars and Δmag is the difference in the Tycho-2 blue magnitudes. Once the photocentre position is computed, the calculation of the proper motion is the same as for single stars.

8.3. Computation of standard errors

Two techniques were used to determine the standard errors associated with each mean position and proper motion; the *scatter method* and the *model method*. The scatter method computes the standard errors using the residuals of the source catalogues' positions from that of the computed mean position and motion. For example, if w_i is the weight of right ascension a_i, and a is the right ascension of the star computed at epoch of a_i, then the variance of the mean value of a is:

$$\sigma_a^2 = \frac{\sum w_i (a_i - a)^2}{(n-2)\sum w_i} \tag{11}$$

and its square root is the standard error. In reality, all computations are made using position unit vectors (xyz), not equatorial coordinates. The formulation for the standard errors of the weighted mean values in right ascension times cosδ, declination, and proper motion in both are

$$\sigma_\alpha = \left[\frac{\sum(\Delta y_i \overline{x} - \Delta x_i \overline{y})^2 w_i}{(1-\overline{z}^2)(n-2)\sum w_i}\right]^{\frac{1}{2}} \tag{12}$$

$$\sigma_\delta = \left[\frac{\sum(\Delta z_i^2 w_i)}{(1-\overline{z}^2)(n-2)\sum w_i}\right]^{\frac{1}{2}} \tag{13}$$

$$\sigma_{\mu_\alpha} = \left[\frac{\sum(\Delta y_i \overline{x} - \Delta x_i \overline{y})^2 w_i}{(1-\overline{z}^2)(n-2)\sum w_i E_i^2}\right]^{\frac{1}{2}} \tag{14}$$

$$\sigma_{\mu_\delta} = \left[\frac{\sum(\Delta z_i^2 w_i)}{(1-\overline{z}^2)(n-2)\sum w_i E_i^2}\right]^{\frac{1}{2}} \tag{15}$$

respectively. The values Δx, Δy, Δz are the residuals from the mean of x, y, and z respectively. E_i is the difference in epoch of the i^{th} position from the mean epoch. The value n is the number of catalogue positions for the star.

If all observations, by chance, happen to give very small residuals, an erroneously low value for σ will result. Therefore, another technique, the *model method*, to determine standard errors was also employed. The model method uses a pre-determined estimate of individual catalogue position errors to compute a standard error. The equations for errors in right ascension and proper motion in right ascension are

$$\sigma_\alpha = \left[\frac{1}{\sum w_i^2}\sum w_i^2 \sigma_{\alpha_i}^2\right]^{\frac{1}{2}} \tag{16}$$

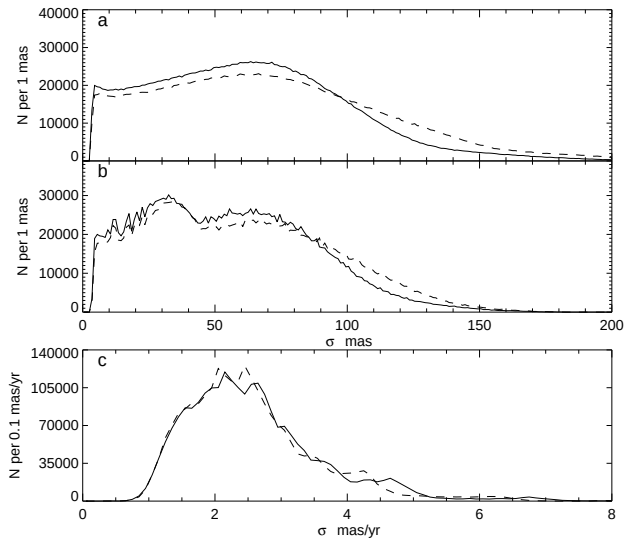

Fig. 11. Distributions of internal astrometric standard errors, solid line for $\alpha\cos\delta$, dashed line for δ: (a) Tycho-2 position as observed by the Tycho experiment about the epoch 1991; (b) mean position at mean epoch from Tycho-2 and ground-based observations; (c) proper motion from the same observations.

$$\sigma_{\mu_\alpha} = \left[\left(\frac{1}{\Sigma E_i^2}\right)^2 \frac{1}{\sum w_i^2} \sum w_i^2 E_i^2 \sigma_{\alpha_i}^2\right]^{\frac{1}{2}} \quad (17)$$

where σ_{α_i} is the computed standard error in right ascension of the i^{th} position and E_i is the epoch difference of the i^{th} positions from the weighted mean epoch. Errors in declination are similar.

Several statistics are used to determine which observations are outliers. Details can be found in the Guide to Tycho-2. Removal of outliers is performed, and the process of determining proper motions is repeated. Please note that the Tycho-2 position is never removed in this process.

9. Accuracy of Tycho-2 astrometry

The distribution of internal astrometric standard errors in the Tycho-2 Catalogue is shown in Fig. 11.

The external standard errors of the observed Tycho-2 positions of individual stars were derived by comparison with the Hipparcos positions as explained in Sect. 7.2. A similar procedure for the proper motions of individual stars will not give truly external errors because the Hipparcos proper motions have been used in the reduction of the old ground-based catalogues used for the Tycho-2 proper motions. Since no other comparison standards of sufficient quality exist we have however compared with Hipparcos.

Stars in Hipparcos with a 'Double/Multiple System' flag were excluded from the comparisons. Results for the common stars were collected in 1.0 magnitude intervals from $V_T = 7.0 - 12.0$. Computed were the standard deviation, σ_d, of the differences HIP–Tycho-2 and the rms values of the internal errors of the proper motions given in the two catalogues, σ_H and σ_T. (Note that σ_T is from the model method.) The expected cosmic error of proper motions, σ_C, for stars of median magnitude V_T was computed from

$$\sigma_C = 2.13\ 10^{-.20(V_T-4.9)} \quad (18)$$

This formula is in accordance with the results obtained by Wielen et al. (1998) from an analysis of the differences in proper motions of FK5 and HIP which gave $\sigma_C = 2.13$ mas/yr for the FK5 stars of median magnitude $V_T = 4.9$ mag. The dependence on magnitude would be correct if the distribution of absolute magnitudes of stars were the same for all apparent magnitudes. This is not true, but the formula should be adequate for the present application because the cosmic error becomes almost negligible compared to the other errors at magnitude 8 and fainter.

The cosmic error is uncorrelated with the Hipparcos or Tycho-2 proper motions, but these two latter may be correlated with the correlation coefficient ρ, hence in general

$$\sigma_d^2 = \sigma_H^2 + \sigma_T^2 - 2\rho\,\sigma_H\,\sigma_T + \sigma_C^2 \quad (19)$$

If the proper motions are uncorrelated ($\rho = 0$) and the internal standard errors of both catalogues are equal to the external we would expect the ratio $R = \sigma_d/\sqrt{\sigma_H^2 + \sigma_T^2 + \sigma_C^2}$ to be unity. Results are shown in Table 1 where the value N is the number of stars used to compute the statistics.

A value of R much smaller than 1.0 could indicate that the internal errors are larger than the external (which nobody would expect) and/or that $\rho > 0$. The vast majority of bright and accurate Hipparcos stars determine the systems of the other catalogues. We can therefore expect some correlation ($\rho > 0$) between Hipparcos and Tycho-2 proper motions for these stars. But there should be little correlation with the fainter and less accurate Hipparcos proper motions. This is expected because magnitude equations were never determined on a plate-by-plate basis, but using a large set of plates (usually an entire zone). The magnitude equation was then applied to the x, y data prior to solving for the plate parameters.

Since R-values appear to be about 1.0 and never significantly larger we conclude that the external standard errors of proper motion in Tycho-2 are approximately equal to the internal. The columns σ_H and σ_T in Table 1 thus give reliable measures of the proper motion accuracy in the Hipparcos and Tycho-2 Catalogues, showing that the two catalogues are equally accurate for stars fainter than $V_T = 9$. A comparison of the accuracy for brighter stars is pending.

Table 1. Proper motions compared with Hipparcos values. For each interval of magnitudes a line is given separately for RA and Dec. Unit for σ values is mas/yr.

V_T	σ_d	σ_H	σ_T	σ_C	R	N
<7.0	1.17	.76	1.05	–	–	11115
	1.18	.60	1.07	–	–	11121
7.0 - 8	1.35	.88	1.20	.64	.83	20507
	1.32	.71	1.21	.64	.86	20514
8.0 - 9	1.48	1.06	1.24	.40	.89	35572
	1.43	.86	1.25	.40	.91	35583
9.0 - 10	1.70	1.33	1.38	.26	.88	22733
	1.63	1.08	1.38	.26	.92	22791
10.0 - 11	2.49	1.84	1.84	.16	.95	6345
	2.31	1.46	1.82	.16	.99	6392
11.0 - 12	3.84	2.61	2.66	.10	1.03	1494
	3.64	2.09	2.62	.10	1.09	1497
> 12.0	4.64	3.41	3.32	–	.98	135
	4.81	2.73	3.25	–	1.13	134

The systematic errors of positions and proper motions have been studied by considering the average differences between Hipparcos and Tycho-2 values for the common stars in areas of typically $6° \times 6°$.

The observed Tycho-2 positions showed no significant differences to Hipparcos, i.e. well below 1 mas, as a result of the systematic corrections mentioned in Sect. 7.2. The proper motions showed a median length of the difference vectors of 0.21 mas/yr when all 119 740 common stars were included. This is satisfactorily small, but it is a statistically significant difference when the standard errors of both catalogues are taken into account. The maximum mean difference for the stars in any of the areas is less than 1 mas/yr and there is no conspicuous systematics in the directions of the vectors. It should be noted that larger systematic differences on smaller angular scales, e.g. about the size of the Astrographic Catalogue photographic plates ($2° \times 2°$) cannot be precluded from this investigation. Such small scale errors could possibly be reduced by a block-adjustment as proposed by Kuzmin et al. (1999).

Systematic errors of proper motions with a dependence on magnitude were investigated by use of stars in common to Hipparcos. Differences of the 39 664 stars fainter than $V_T = 9.0$ and the 11 611 stars fainter than 10.0 mag were computed. The median difference was 0.47 and 0.85 mas/yr, respectively, which are statistically insignificant due to random errors of the catalogues.

Since we consider the systematic errors of Hipparcos astrometry to be negligible in this context, it is concluded that the systematic errors of Tycho-2 proper motions are less than about 0.5 mas/yr on angular scales of 6 deg or more.

10. Conclusions and future plans

The Tycho-2 processing emphasized the optimal treatment of single stars, which constituted the vast majority of targets observed. However, an effort was made to detect and measure double stars. This resulted in the measurement of 7 500 pairs with separations between 0.8 and 2.5 arcsec, while the lower limit of resolution in Tycho-1 was about 2 arcsec. The processing also gave insight into still unexploited possibilities. A special double star processing where the Tycho-2 results are used as input would give detection and measurement of closer pairs and of presently unknown fainter companions of many thousand stars. It has also become clear that a special variable star processing should be applied because thousands of new variables can be detected, many more than can be detected from the published photometric annexes to the Tycho-1 Catalogue. These surveys of double and variable stars using the original Tycho observations and the Tycho-2 Catalogue as basis are planned at Copenhagen as the *Tycho-3 project*. The surveys would be complete and uniform over the sky to an extent which could only be superseded by a future astrometric satellite.

Acknowledgements

This work was supported by the Velux Foundation of 1981 and the Danish Space Board.

We are grateful to Dr. S. Röser, Astronomisches Rechen-Institut in Heidelberg for permission to use the STARNET catalogue as part of the Tycho-2 Input Catalogue, and to Dr. B. McLean, Space Telescope Science Institute in Baltimore for making the GSC 1.2 available to us for the same purpose. We wish to thank Dr. D. Egret of the Strasbourg Astronomical Observatory and the Guide Star Catalog team of the Space Telescope Science Institute for assistance in supplying additional GSC running numbers for the Tycho-2 stars.

References

Bahcall J.N., Soneira R.M., 1981, ApJ 246, 112

Corbally J.C., 1984, ApJS 55, 657

Corbin T.E., 1977, The Proper Motion System of the AGK3R, PhD dissertation, Leander McCormick Observatory, University of Virginia.

ESA, 1997, The Hipparcos and Tycho Catalogues, ESA SP-1200, Vol. 1–17, referred to in this text as *ESA97*. Volume 4 is referred to as *EVol4*

Fabricius C., Makarov V.V., 2000, submitted to A&A, Hipparcos astrometry for 257 stars using Tycho-2 data

Falin J.L., Mignard F., 1999, A&AS 135, 231

Gomez M., Hartmann L., Kenyon S.J., Hewett R., 1993, AJ 105, 1927

Høg E., Jaschek C., Lindegren L., 1982, ESA SP-177, 21

Høg E., Bässgen G., Bastian U., et al., 1997, A&A 323, L57

Høg E., Kuzmin A., Bastian U., et al., 1998, A&A 335, L65

Høg E., Fabricius C., Makarov V.V., et al., 1999, in *Modern Astrometry and Astrodynamics*, R. Dvorak, H.F. Haupt & K. Wodnar (eds.), Vienna Obs. 43

Høg E., Fabricius C., Makarov V.V., et al., 2000a, A&A, The Tycho-2 Catalogue of the 2.5 Million Brightest Stars, in press

Høg E., Fabricius C., Makarov V.V., et al., 2000b, The Tycho-2 Catalogue on CD-ROM, including Guide to Tycho-2, Copenhagen University Observatory

IAU, 1998, IAU Transactions, Volume 23B, resolution B2.

Jenkner H., Lasker B.M., Sturch C.R., et al., 1990, AJ 99, 2081

Knude J., 1986, A&AS 63, 313

Knude J., 1993, A&AS 98, 213

Kuzmin A., Høg E., Bastian U., et al., 1999, A&AS 136, 491

Lindegren L., Perryman M.A.C., 1996, A&AS 116, 579

Makarov V.V., Høg E., 1995, in IAU Symposium No. 166. E. Høg & P.K. Seidelmann (eds.), p. 372

Monet D., Bird A., Canzian B., et al. 1996, USNO-A V1.0, A Catalog of Astrometric Standards, U.S. Naval Observatory

Perryman M.A.C., Hassan H., 1989, *The Hipparcos mission*, Vol. 1, *The Hipparcos Satellite*. ESA Publications Division, ESA SP-1111

Perryman M.A.C., Lindegren L., Kovalevsky J., et al., 1995, A&A 304, 69

Press W.H., Teukolsky S.A., Vettering W.T., Flannery B.P., 1992, *Numerical Recipes in Fortran*, Cambridge Univ. Press

Röser S., 1995, In: IAU Symposium No. 172. S. Ferraz-Mello, B. Morando, J.-E. Arlot (eds.) Kluwer, Dordrecht, 1996, p. 481

Röser S., Morrison J., Bucciarelli B., Lasker B., McLean B., 1997, In: IAU Symposium No. 179, Eds. B.J. McLean, D.A. Golombek, J.E. Hayes, H.E. Payne. Kluwer, Dordrecht, 1998, p. 420

Söderhjelm S., 1999, A&A 341,121

Urban S.E., Corbin T.E., Wycoff, G.L, et al., 1998a, AJ 115, 1212

Urban S.E., Corbin T.E., Wycoff G.L., 1998b, AJ 115, 2161

Wicenec A., van Leeuwen F., 1995, A&A 304, 160

Wielen R., Schwan H., Dettbarn C., Jahreiss H., Lenhardt H., 1998, in *The message of the Angles. Astrometry 1798–1998*, P. Brosche, W.R. Dick, O. Schwarz, R. Wielen (eds.), Proceedings of the International Spring Meeting of the Astronomische Gesellschaft, held in Gotha, 11–15 May, 1998, p. 123

Yoshizawa M., Andreasen G.K., Høg E., 1985, A&A 147, 227

Zacharias N., Zacharias M. I., 1999, AJ 118, 2503

Zacharias N., Zacharias M. I., de Vegt C., 1999, AJ 117, 2895