Astron. Astrophys. 304, 52-60 (1995)

ASTRONOMY
AND
ASTROPHYSICS

Zero-point and external errors of Hipparcos parallaxes*

F. Arenou[1], **L. Lindegren**[2], **M. Frœschlé**[3], **A.E. Gómez**[1], **C. Turon**[1], **M.A.C. Perryman**[4], **and R. Wielen**[5]

[1] URA 335 du CNRS - DASGAL, Observatoire de Paris-Meudon, F-92195 Meudon, France.
[2] Lund Observatory, Box 43, S-22101 Lund, Sweden
[3] Observatoire de la Côte d'Azur - CERGA, URA 1360 du CNRS, F-06130 Grasse, France
[4] Astrophysics Division, European Space Agency, ESTEC, NL-2200 AG Noordwijk, The Netherlands
[5] Astronomisches Rechen-Institut, Mönchhofstr. 12-14, W-6900 Heidelberg, Germany

Received Dec 30, 1994; accepted Sep 6, 1995

Abstract. A good knowledge of both accuracy and precision of the Hipparcos parallaxes is one of the keys for their future scientific use. For this purpose, the Hipparcos preliminary parallaxes, as obtained after the processing of the first 30 months of Hipparcos data, are compared to various ground-based parallax estimates, using astrometric, photometric and spectroscopic data.

In order to find unbiased values of the global zero-point and of external errors, a new maximum-likelihood algorithm has been built, taking into account the censorships of the observed data. Applying the method to a sample of distant stars, it is shown that the global zero-point error of the Hipparcos preliminary parallaxes should be smaller than 0.1 mas and that the external errors are unlikely to be underestimated by more than about 5%.

Key words: Hipparcos – stars: distances – astrometry – methods: data analysis

1. Introduction

One of the most significant impacts of the Hipparcos mission is to measure the trigonometric parallaxes of a large number of stars to an accuracy of some 1 to 5 mas (Perryman et al., 1989). In contrast with parallaxes obtained with ground-based programmes, the Hipparcos parallaxes should be absolute (Lindegren, 1992). However, a (small) global zero-point shift may exist due to periodic basic angle variations of the satellite beam-combining mirror (Lindegren et al., 1992). If a global zero-point error exists, even if it is small, it must be found and shown to be independent of the astrometric and photometric data of the stars: positions, parallaxes, proper motions, apparent magnitudes and colours.

Using the first 30 months of Hipparcos data, the aim of this paper is twofold: to obtain unbiased estimates of the global zero-point and of the external error of the Hipparcos preliminary parallaxes.

Firstly, the Hipparcos preliminary parallaxes are compared with external parallax determinations available from various ground-based data sources: trigonometric, spectroscopic and photometric parallax as well as distance moduli of open clusters and of the Magellanic Clouds. The Hipparcos data are described in Sect. 2 and the comparisons are given in Sect. 3. Beyond the global comparisons, spectroscopic and photometric parallaxes of distant stars allow to obtain a first estimate of the global-zero point and of the external error.

Secondly, an algorithm based on maximum-likelihood estimation is used and described in Sect. 4. This method takes into account the fact that the samples suffer from selection biases and makes use of the available astrometric and photometric data. Moreover, the algorithm allows to check the quality of the fit between the adopted model and the observations and, also, to detect the outliers. As a result of this procedure unbiased estimates of the global zero-point and of the external error and their corresponding standard errors are obtained.

Finally, in Sect. 5 we analyse whether the errors on the parallaxes vary with distance, position, proper motion and photometry of the stars.

The results obtained, based on the first 30 months among the 37 months data of the whole Hipparcos mission, foresee the high quality of the final Hipparcos parallaxes.

2. The Hipparcos preliminary parallaxes

The parallax catalogue used in this paper (called H30 in what follows) is the union of the two Data Reduction Consortia (FAST and NDAC) sphere solutions for the first 30 months of Hipparcos data. The construction of this intermediate Hipparcos astrometric catalogue is described in Kovalevsky et al. (1995). For about 89 % of the stars contained in H30, the parallaxes, together with their formal

Send offprint requests to: Frederic.Arenou@obspm.fr

* Based on observations made with the ESA Hipparcos satellite.

variances, are the mean values of FAST and NDAC results; for the remaining stars, the parallaxes come either from FAST or NDAC.

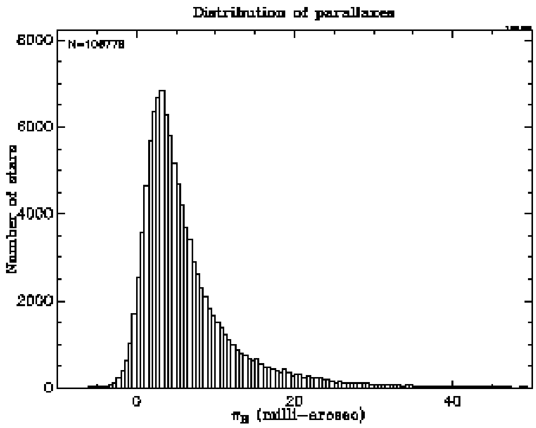

Fig. 1. Distribution of the H30 Hipparcos preliminary parallaxes (mas)

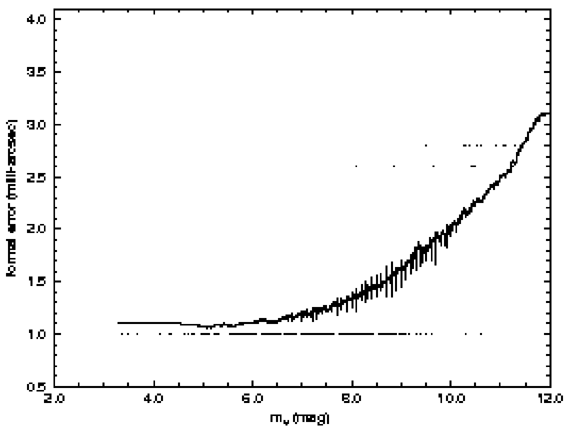

Fig. 2. H30 parallax formal errors (mas) versus apparent V–magnitude; solid line: running average over 500 points

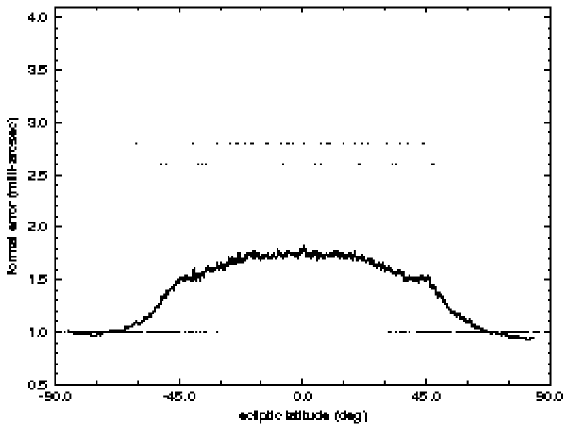

Fig. 3. H30 parallax formal errors (mas) versus ecliptic latitude; solid line: running average over 500 points

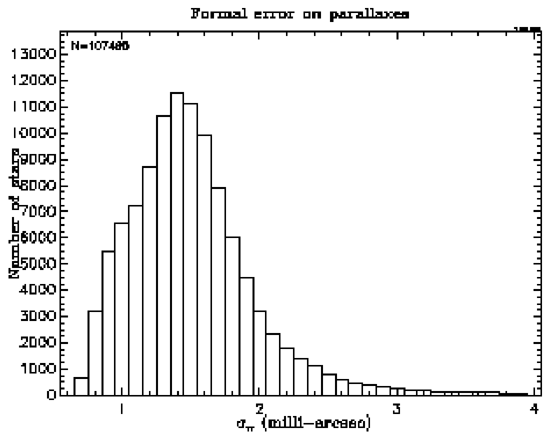

Fig. 4. Distribution of the formal errors on the H30 parallaxes (mas)

The distribution of the 107 495 Hipparcos preliminary parallaxes is given in Fig. 1. Half of the stars are closer than about 220 pc. One impressive point is the small fraction of stars (≈ 4 %) with negative parallaxes which implies that the parallaxes are of high precision.

Due to photon noise and to the scanning law of the satellite, the formal standard errors on the parallaxes ($\sigma_{\rm H}$) vary both with magnitude (Fig. 2) and ecliptic latitude (Fig. 3). The distribution of the formal errors is shown in Fig. 4, the mode of the distribution being at 1.4 mas.

In what follows, the true parallax is noted π, $\pi_{\rm H}$ denotes the Hipparcos preliminary parallax and $z = \langle \pi_{\rm H} - \pi \rangle$ is the global zero-point systematic error on the Hipparcos preliminary parallaxes.

The external error ($\sigma_{\rm ext}$) is the result of the contributions of the formal errors, as computed by the Data Reduction Consortia, and of the possible errors arising from an incomplete modelling in the reduction processes. Given that $\sigma_{\rm H}$ varies between 0.6 and 4 mas (see Fig. 4), it is more appropriate to study the unit-weight error $k = \langle \frac{\sigma_{\rm ext}}{\sigma_{\rm H}} \rangle$ instead of studying $\sigma_{\rm ext}$. In the best case we should get $z = 0$ and $k = 1$.

3. Global comparison of the Hipparcos preliminary parallaxes with external estimations

The programme stars observed by Hipparcos are contained in the Hipparcos Input Catalogue (Turon et al. 1992b); a comprehensive description of the Catalogue contents may be found in Turon et al. (1992a). Cross-identifications, spectroscopic and photometric data used in this paper come mainly from the Hipparcos Input Catalogue. In this section the Hipparcos preliminary parallaxes are compared to various ground-based parallax estimations. The distribution of the differences between these different sources of parallaxes is analysed. As the width of the distribution (equal to the RMS error if the distribution is normal) depends on the individual errors of both the Hipparcos and the ground-based parallaxes, the obtained results give only an upper limit of the external error of the Hipparcos preliminary parallaxes. However, as can be seen below, estimates z and k may be found in some cases.

3.1. Trigonometric parallaxes

In order to compare ground-based trigonometric parallaxes to the Hipparcos preliminary parallaxes, the General Catalogue of Trigonometric Stellar Parallaxes (GCTSP) (van Altena et al. 1991) has been used. About 5 800 GCTSPstars are contained in the Hipparcos Input Catalogue, 4 400 stars remain in H30.

The distribution of the differences between the Hipparcos preliminary parallaxes and the GCTSP parallaxes is plotted in Fig. 5 for the stars closer than 25 pc (about 700 stars). The median of the distribution is -0.25 ± 0.42 mas and the width is 8.90 ± 0.46 mas. These statistics were obtained using the quantiles of the distribution, in this way less weight is given to the outliers under the hypothesis of a normal distribution. The Gaussian distribution having the corresponding mean and standard deviation is plotted in Fig. 5. It does not mean that the distribution is supposed to be Gaussian which is obviously not the case (the formal errors on the differences range from 1.5 to 24 mas), but this plot allows to get a glance at the skewness and the kurtosis of the distribution. The median value is not significantly different from 0, but the width principally reflects the contribution of the errors of the GCTSP parallaxes themselves. If the Hipparcos preliminary parallaxes are compared with the GCTSP parallaxes using about 3 700 stars farther than 25 pc, a small bias appears: the median of the differences becomes -2.65 ± 0.23 mas.

3.2. Parallaxes of stars in open clusters

For open clusters far enough, all stars of one cluster may be assumed to be placed at the same distance. Taking into account the uncertainty on the distance moduli of clusters (Mermilliod, 1993), a formal error on the parallaxes of cluster stars better than 10% is obtained. Due to this high precision, parallaxes of stars in clusters are well suited to obtain an estimation of the Hipparcos preliminary parallax zero-point and of the external error provided that all the considered stars are true members of the clusters.

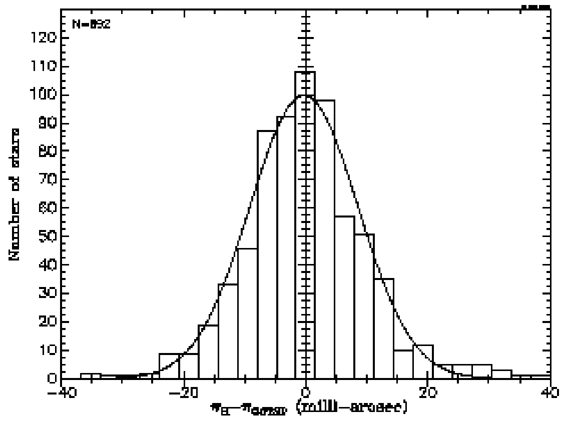

Fig. 5. Distribution of the parallax differences (Hipparcos − GCTSP) (mas) for stars closer than 25 pc

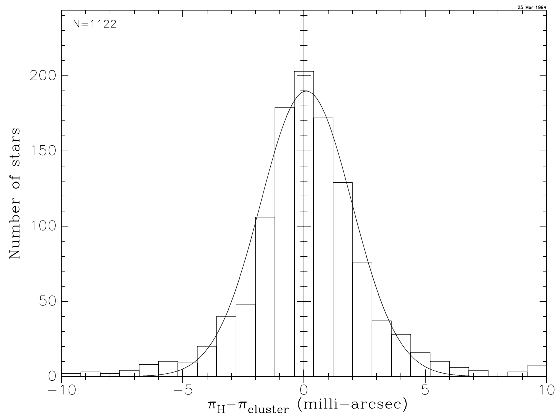

Fig. 6. Distribution of the parallax differences (Hipparcos − cluster) (mas)

First, the Hipparcos Input Catalogue stars with a cluster identifier were selected and then, the suspected non-members were suppressed using the BDA cluster data base (Mermilliod, 1992). For the 1 300 remaining stars, the parallax of the cluster, computed from the distance moduli quoted by Lyngå (1987), was assigned to each star.

The difference between Hipparcos preliminary parallaxes and the cluster parallaxes is shown in Fig. 6, the median and the width are respectively 0.09 ± 0.07 mas and 1.89 ± 0.08 mas. The long tail of the distribution is probably due to non-members (although known non-members were suppressed, the sample still contains stars which are not physically members of the cluster). The result obtained is consistent with $z = 0$ and gives an upper limit of the external error. In Fig. 6, the Gaussian distribution having a mean and a standard deviation equal to the obtained median and width values respectively, is also shown.

The results quoted before concern the global comparison using the observed stars in all the clusters together; a comparison cluster by cluster was also performed. All clusters (34 clusters) with at least 5 observed stars were considered. The proximity of these stars on the sky allows to test the behaviour of Hipparcos preliminary parallaxes on a small sky zone. The precision on the mean parallax of a group of adjacent stars is, due to correlation between data, about $\frac{\sigma_\pi}{n^{0.35}}$ instead of $\frac{\sigma_\pi}{\sqrt{n}}$ (Lindegren, 1989). The obtained difference between the mean H30 parallax for each cluster and the known parallax of the cluster is within 2 times the standard deviation of this difference for all the clusters with the exception of two of them. In these last cases, there could possibly exist bad membership identifications.

3.3. Magellanic clouds stars

The Magellanic clouds are so far away that their star parallaxes (≈ 0.017 and 0.022 mas for the Small and Large cloud respectively) cannot be determined by Hipparcos. However, 46 stars were observed by the satellite in order to get an upper limit on their proper motions. The measured parallaxes (close to 0) could be used to obtain directly the measurement error of the Hipparcos preliminary parallaxes. The result of the comparison is given in Fig. 7. The mean value: -0.16 ± 0.26 mas is not significantly different from 0 and the standard deviation of the distribution is 1.72 ± 0.18 mas in excellent agreement with the formal errors on the parallaxes. The Gaussian distribution corresponding to the mean and standard deviation values is also plotted in Fig. 7.

3.4. Dynamical parallaxes

There are 369 dynamical parallaxes in the Hipparcos Input Catalogue compiled by Dommanget & Nys (1982). Their relative error is assumed to be smaller than about 20%.

The distribution of the difference between H30 parallaxes and dynamical parallaxes, after rejecting 6 obvious outliers, is shown in Fig. 8. The median value is -0.05 ± 0.25 mas and the width 2.63 ± 0.28 mas. This result indicates that there are no important systematic effects on both parallaxes and that the width of the distribution is consistent with the expected value taking into account the errors on both dynamical and Hipparcos preliminary parallaxes.

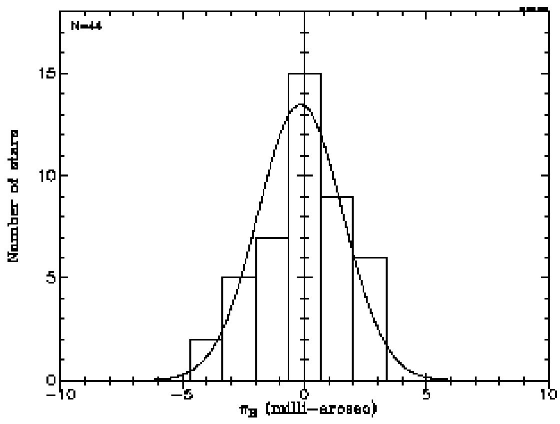

Fig. 7. Distribution of the Hipparcos preliminary parallaxes of stars in the Magellanic clouds (mas)

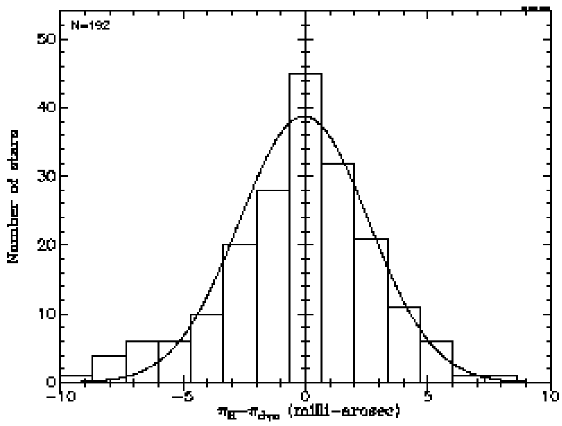

Fig. 8. Distribution of the parallax differences (Hipparcos − dynamical) (mas)

3.5. Spectroscopic and photometric parallaxes

More than half of the stars in the Hipparcos Input Catalogue have a spectral type and luminosity class, allowing to obtain an estimate of their distance through the use of spectroscopic absolute magnitude calibrations.

Apart from problems arising from the inhomogeneous aspect of visual spectral classification, the mean absolute magnitude calibrated for spectral type groups suffers from different drawbacks: a) it is not always clear whether or not the existing absolute magnitude calibrations have taken the Malmquist bias into account, b) if no criterion related to the age of the stars is taken into account, a main-sequence width of about 2 magnitudes for the earlier-type stars may be expected (Jaschek & Mermilliod 1984), and the mean value of the absolute magnitude may lie almost anywhere in this interval.

In order to obtain a spectroscopic parallax, the absolute magnitude and intrinsic colour calibrations from Schmidt- Kaler (1982) were chosen because they cover almost the whole HR diagram; the interstellar extinction was estimated with $A_V \approx (3.3+0.28(B-V)_0+0.04E(B-V))E(B-V)$ (Schmidt-Kaler 1982). A spectroscopic parallax ($\pi_{\rm S}$) was then obtained for about 54 000 stars with a relative error of about 25%.

Besides the spectroscopic parallaxes, a better galactic distance indicator is obtained, also for a large number of stars, using the photometric absolute magnitudes. The *uvby*-β photometry was chosen because of the high number of measurements available (Hauck & Mermilliod 1990) and because it is well suited to compute the parallaxes of distant stars observed by Hipparcos, essentially dwarfs, giants and supergiants of types B, A and F.

Reddening and visual absolute magnitudes were obtained (Arenou 1993), using various photometric calibrations corresponding to different groups in the HR diagram, among others: Arellano Ferro (1990), Balona et al. (1984), Crawford (1975, 1978, 1979), Figueras et al. (1991), Gray (1991), Guthrie (1987), Hilditch et al. (1983), Moon (1985), Olsen (1988) and Zhang (1983). The apparent visual magnitude was corrected from interstellar visual extinction using $A_V \approx 4.3E(b-y)$ (Crawford & Mandwewala 1976). Finally, a photometric parallax ($\pi_{\rm P}$) was obtained for about 12 000 stars, with a relative error better than 20%.

As the true parallax is not known, one way to study the errors on the measured Hipparcos preliminary parallaxes is to make use of distant stars, for which the true parallaxes are ≈ 0; in this case, the values of the Hipparcos preliminary parallaxes are only due to the measurement errors. However, if we want to put into light a zero-point error smaller than 0.1 mas, stars farther than about 10 kpc must be selected, but there are very few of them in the Hipparcos Input Catalogue. An alternative solution consists in using less distant stars, selected by some distance criteria (for instance stars with $\pi_{\rm S} < 2$ mas), and in computing $\langle\pi_{\rm H}-\pi_{\rm S}\rangle$.

However, due to the non-uniform distribution of the parallaxes and to their measurement error, if a sample is truncated in observed parallax, the computed mean parallax of a sample is a biased estimate of the true mean parallax of the stars in this sample. For instance, if only stars with $\pi_{\rm S} < 2$ mas are kept, the average difference $\langle\pi_{\rm H}-\pi_{\rm S}\rangle$ is biased by at least 0.3 mas. In order to reduce the bias, the distant stars may be selected with the help of one parallax estimate and the other parallax estimate may be compared to the Hipparcos preliminary parallaxes. In particular, an estimate of the global zero-point of Hipparcos preliminary parallaxes was obtained with $\langle\pi_{\rm H}-\pi_{\rm P}\rangle$ for stars with $\pi_{\rm S} < 2$ mas. The drawback of this method is that less stars remained at our disposal and that, although reduced, a smaller bias still remained, since $\pi_{\rm P}-\pi$ is correlated with $\pi_{\rm S}-\pi$, through the common use of the observed apparent magnitude and the log-normal law of the errors on $\pi_{\rm P}$ and $\pi_{\rm S}$.

In our sample, the stars known to have a variability > 0.2 mag were excluded, together with stars with a joint photometry and spectroscopic peculiar stars; this was achieved using the informations contained in the Hipparcos Input Catalogue, and some stars with unknown duplicity or variability could remain in this sample. Firstly the stars with a photometric parallax smaller than 2 mas were selected, and the Hipparcos preliminary parallaxes were compared to the spectroscopic parallaxes with a parallax relative error better than 25% (Fig. 9). The median and width are -0.02 ± 0.07 mas and 1.51 ± 0.08 mas, respectively. The distribution is not symmetric, probably due to giant stars wrongly classified as dwarfs. In contrast, the distribution of the residuals $\pi_{\rm H}-\pi_{\rm P}$ for stars with $\pi_{\rm S} < 2$ mas and a relative error on the photometric parallax smaller than 20% (Fig. 10) has a smaller width (1.26 ± 0.07 mas) and few (3) outliers. In this case, the contribution of the random errors on the photometric parallaxes should be smaller than 0.4 mas and, therefore, the width of the distribution of $\pi_{\rm H}-\pi_{\rm P}$ becomes a realistic approximation of the true width of the H30 parallax errors. Among the comparisons between the Hipparcos preliminary parallaxes and various ground-based parallax estimates, the use of the photometric parallaxes is the most promising, because of their precisions and of the number of the concerned stars. This allows us to obtain a reliable estimate of the Hipparcos preliminary parallax zero-point: the weighted mean of the differences $\pi_{\rm H}-\pi_{\rm P}$ is $z = 0.08 \pm 0.05$ mas. The normalised residuals, $(\pi_{\rm H}-\pi_{\rm P})$ divided by formal

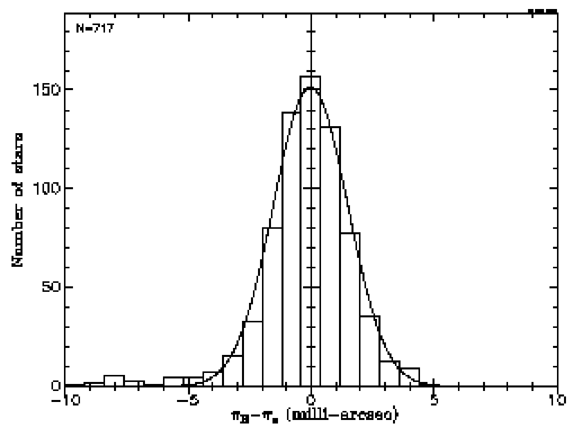

Fig. 9. Distribution of the parallax differences (Hipparcos − spectroscopic) for stars with photometric parallax < 2 mas and relative error on spectroscopic parallax < 25%

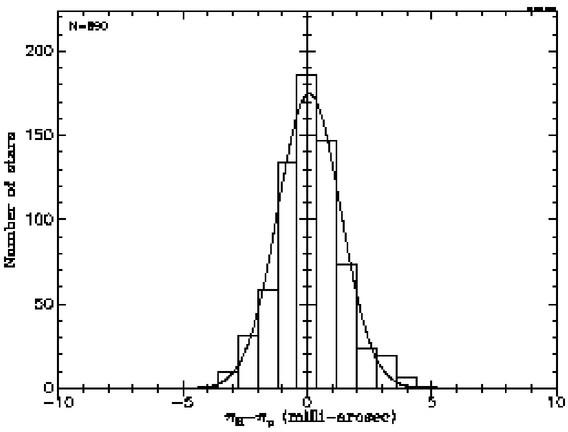

Fig. 10. Distribution of the parallax differences (Hipparcos − photometric) for stars with spectroscopic parallax < 2 mas and relative error on photometric parallax $< 20\%$

error, may also be studied. Using a Kolmogorov-Smirnov test, the hypothesis that their distribution is Gaussian (0,1) is not rejected at the 5%-significance level. The width of the normalised residuals is $k = 1.01 \pm 0.05$. These results suggest that the errors on the Hipparcos preliminary parallaxes are normally distributed with mean 0 and variance equal to the square of the formal error, as computed by the Data Reduction Consortia.

4. A maximum-likelihood algorithm

It was mentioned in Sect. 3.5 that the computation of $z = \langle \pi_{\rm H} - \pi_{\rm P} \rangle$ for stars with $\pi_{\rm S} < 2$ could lead to a small bias. In order to cope with this problem, and get an unbiased value of z and k, a more sophisticated method is presented below, which takes explicitly into account the various censorships on our sample.

4.1. The method

The *uvby*-β photometric calibration described above was used in order to obtain an absolute magnitude M_V and an interstellar visual extinction A_V, and the apparent distance moduli $t = m_V - M_V - A_V$ was computed.

For each star, the conditional probability density function (pdf) to observe the Hipparcos preliminary parallax $\pi_{\rm H}$ given its apparent distance modulus t, its galactic latitude b, and the unknown parameters z and k is:

$$f(\pi_{\rm H}|t,b,z,k) = \frac{g(\pi_{\rm H},t,b|z,k)}{h(t,b|z,k)} = \frac{g(\pi_{\rm H},t,b|z,k)}{\int_{-\infty}^{+\infty} g(\pi_{\rm H},t,b|z,k)d\pi_{\rm H}} \quad (1)$$

where $g(.)$ may be expressed as a marginal pdf:

$$g(\pi_{\rm H},t,b|z,k) = \int_0^{+\infty} q(\pi_{\rm H},t,b|\pi,z,k)p_4(\pi)d\pi \quad (2)$$

$q(.)$ being the product of the independent pdfs:

$$q(\pi_{\rm H},t,b|\pi,z,k) = p_1(\pi_{\rm H}|\pi,k,z).p_2(t|\pi).p_3(b|\pi) \quad (3)$$

where the conditional pdfs $p_1...p_4$ are determined below.

It has already been noticed that the normality hypothesis for the errors on the Hipparcos preliminary parallaxes may be admitted. Noting $\mathcal{G}_x(\mu,\sigma)$ a Gaussian pdf of variable x with mean μ and variance σ^2, and taking explicitly into account the fact that the observed parallax $\pi_{\rm H}$ may have been censored if it was outside the interval $[\pi_{\rm H}^-, \pi_{\rm H}^+]$, the two unknown parameters z and k are introduced into the conditional pdf of the errors:

$$p_1(\pi_{\rm H}|\pi,k,z) = \left| \begin{array}{l} \frac{\mathcal{G}_{\pi_{\rm H}}(\pi+z,k\sigma_{\rm H})}{\int_{\pi_{\rm H}^-}^{\pi_{\rm H}^+} \mathcal{G}_{\pi_{\rm H}}(\pi+z,k\sigma_{\rm H})d\pi_{\rm H}} \text{ if } \pi_{\rm H} \in [\pi_{\rm H}^-, \pi_{\rm H}^+] \\ 0 \text{ otherwise} \end{array} \right. \quad (4)$$

The error on the calibrated absolute magnitude M_V is supposed to be Gaussian $\mathcal{G}_{M_V}({M_V}', \sigma_M)$ around the true absolute magnitude ${M_V}'$, and the errors on the apparent magnitude and the interstellar extinction are also taken as following the normal laws $\mathcal{G}_{m_V}({m_V}', \sigma_m)$ and $\mathcal{G}_{A_V}({A_V}', \sigma_{A_V})$, where the prime denotes the true quantities. As a consequence, the conditional pdf of t is also Gaussian around $t' = {m_V}' - {M_V}' - {A_V}' = -5\log\pi - 5$, the variance being $\sigma_t^2 = {\sigma_m}^2 + {\sigma_M}^2 + {\sigma_A}^2$. If a censorship on the observed distance moduli outside $[t^-, t^+]$ is also introduced, we have:

$$p_2(t|\pi) = \left| \begin{array}{l} \frac{\mathcal{G}_t(-5\log\pi-5,\sigma_t)}{\int_{t^-}^{t^+} \mathcal{G}_t(-5\log\pi-5,\sigma_t)dt} \text{ if } t \in [t^-, t^+] \\ 0 \text{ otherwise} \end{array} \right. \quad (5)$$

In the last term of Eq. (2) we have to compute the product $p_3(b|\pi)p_4(\pi) = p(b,\pi)$. If we admit that the space distribution is independent of the galactic longitude, we obtain:

$$p(b,\pi) \propto p(r,l,b)|\frac{\partial r}{\partial \pi}| = \frac{1}{\pi^2}p(r,l,b) = \frac{1}{\pi^2}p(X,Y,Z)|J|$$

where $J = r^2 \cos b$ is the Jacobian of the transformation from the heliocentric Cartesian coordinates (X,Y,Z) to the corresponding spherical coordinates (r,l,b). Using a realistic model of the Galaxy, the Z-space distribution was assumed to be an exponential law with mean scale height h_Z: $p(X,Y,Z) \propto \frac{1}{2h_Z} e^{-\frac{|Z|}{h_Z}}$. Finally, we obtain:

$$p_3(b|\pi)p_4(\pi) \propto \frac{\cos b}{2h_Z\pi^4} e^{-\frac{|\sin b|}{\pi h_Z}} \quad (6)$$

The maximum-likelihood estimator (MLE) of (k, z) is the one which maximises the log-likelihood of our n-sample:

$$\ln\ \mathcal{L} = \sum_{i=1}^{n} \ln f(\pi_{\mathrm{H}i}|t_i, b_i, z, k) \tag{7}$$

It provides the largest probability of observing the Hipparcos parallaxes given the photometric and astrometric properties of each star. This estimator, which is asymptotically unbiased, is found numerically using the equations (1) to (7). The formal errors on the obtained parameters and the correlations between them are also found numerically using the inverse of the Fisher information matrix.

This method will also be used to calibrate the absolute magnitude from the final Hipparcos parallaxes. It follows, partly, the method developed by Ratnatunga & Casertano (1991) and takes into account, in a very efficient way, the Lutz & Kelker (1973) and Malmquist (1936) biases.

The considered sample contains distant stars and, for the present purpose, the existence of a kinematical bias was neglected. However our general model may also include in Eq. (3) a kinematical pdf:

$p_5(\mu_\alpha, \mu_\delta, V_R|\pi, \overline{U}, \overline{V}, \overline{W}, \sigma_{UU}, \sigma_{VV}, \sigma_{WW}, \sigma_{UV}, \sigma_{UW}, \sigma_{VW})$

which has not been used here.

4.2. Fit

Once the unknowns z, k are found, the expectation of $f(\pi_{\mathrm{H}}|t, b, z, k)$ may be considered as a predicted "observed" parallax. The star's parallax residual is defined by

$$\delta_i = \pi_{\mathrm{H}i} - \int_{-\infty}^{+\infty} \pi_{\mathrm{H}} f(\pi_{\mathrm{H}}|t_i, b_i, z, k) d\pi_{\mathrm{H}} \tag{8}$$

in order to study the quality of the fit and to reject the outliers. If the above adopted model is correct, the normalised residuals $\frac{\delta_i}{\sigma_{\delta_i}}$ should have a distribution with mean 0 and variance 1, and should be independent from the observed quantities (t, b). Assuming as the null hypothesis that the residual and the corresponding observed quantity are uncorrelated, the independence is tested using a Kendall's τ.

4.3. Outliers

Since the analytical form of the distribution of the normalised residuals is not obvious, the search for outliers (i.e. stars with observed quantities not consistent with the adopted model) is done by simulation: several hundred samples are drawn randomly, the MLE of (k, z) is found for each sample and the normalised residual is computed for each star in the sample. Only the extreme values of the normalised residuals are considered, the smallest and the largest ones, sorted from the simulated samples. A one-sided 97.5%-confidence interval is then obtained for both the smallest and the largest normalised residuals. Coming back to our real sample, a star is considered to be an outlier if its residual is outside of the corresponding 95%-confidence interval. Such outliers are excluded one by one, and the whole algorithm is rerun, until no outlier remains.

4.4. Simulations

The simulations use the fact that the *a posteriori* pdf is $s(\pi|t, b) \propto p_2(t|\pi).p_3(b|\pi).p_4(\pi)$. For each star i, a random π_i is drawn from the distribution $s(\pi_i|t_i, b_i)$, and then an "observed" parallax $\pi_{\mathrm{O}i}$ is drawn from $p_1(\pi_{\mathrm{O}i}|\pi_i, k, z)$. Apart from giving a confidence interval on the normalised residuals, these simulations also allow to verify the formal errors on the (z, k) estimates obtained by the maximum-likelihood algorithm.

4.5. Application to the Hipparcos preliminary parallaxes

The method described above was applied to the Hipparcos preliminary parallaxes, using the absolute magnitudes obtained with the *uvby*-β photometric calibrations, keeping all the stars with $\sigma_t < 0.35$. A photometric censorship was applied to this sample: only the distant stars were selected in order to minimise the effect of a possible zero-point error on the absolute magnitudes and the effect of their random errors. On the other hand, the stars must be numerous enough to get the best accuracy on the global zero-point. We kept only the stars with a distance modulus $8.5 < t < 14.5$ ($0.126 < \pi_{\mathrm{P}} < 2$ mas). No censorship was applied to the Hipparcos preliminary parallaxes; the final sample contains 487 stars. The scale height of the selected stars, mostly B- and A-type stars, is expected to be $h_Z \approx 100$ pc.

The final result for z and k with their corresponding formal errors is:

$z = -0.02 \pm 0.06$ mas and $k = 1.014 \pm 0.034$

with a correlation coefficient -0.19 between them. The quoted error bars may be considered as lower limits because the H30 parallaxes used in the estimations are given in tenth of mas. In any case, the global zero-point z is clearly smaller than the initial mission specifications (0.1 mas), and the true external errors are unlikely to be underestimated by more than about 5%.

One thousand simulations were done with $(z, k) = (-0.02, 1.014)$ and showed that the algorithm recovers unbiased estimates of the input parameters. The corresponding 95%-confidence interval of the normalised residuals was [-3.77,4.01], outside of which a star was considered to be an outlier: no outliers were found in our sample.

The normalised residuals were also found to be independent, at a 5%-significance level, from t and b (Fig. 11), their mean (-0.025 ± 0.015) and standard deviation

(0.992 ± 0.010) being not statistically different from 0 and 1, respectively. The fit may then be considered as satisfactory.

Finally, a simple test was done to check the quality of the results: a Gaussian random error $\mathcal{G}_\epsilon(0.5, 0.5\sigma_{\mathrm{H}i})$ mas was added to each H30 parallax $\pi_{\mathrm{H}i}$. Applying the algorithm, we found $z = 0.47 \pm 0.07$ mas (0.5-0.02=0.48 mas expected) and $k = 1.129 \pm 0.04$ ($\sqrt{1.014^2 + .5^2} = 1.131$ expected) which is in remarkable agreement.

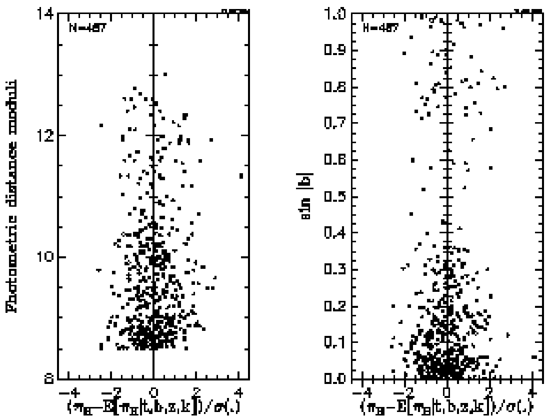

Fig. 11. Normalised residuals versus distance moduli (a) and galactic latitude (b)

5. Variation of the parallax errors

In this section, we study whether the errors on the Hipparcos parallaxes depend or not on the astrometric and photometric data of the stars.

From the comparisons done in Sect. 3, the zero-point error is not significantly different from 0 at a 5%-significance level. As the comparisons use the full range of the parallaxes of the stars in the Hipparcos Input Catalogue, the parallax errors do not likely depend on the parallax itself: neither for distant stars (Sect. 3.3, 3.5) nor for nearby (Sect. 3.1) or intermediate (Sect. 3.2) stars.

In what follows, only distant stars were used, selected as in Sect. 3.5 ($\pi_{\mathrm{S}} < 2$ mas), for which the astrometric and photometric data come from an external source: the Hipparcos Input Catalogue. The final sample contains about 700 stars.

Instead of using the Hipparcos parallax errors themselves, we study the variation of the normalised errors $\frac{\pi_{\mathrm{H}} - \pi_{\mathrm{P}}}{\sqrt{\sigma_{\mathrm{H}}{}^2 + \sigma^2_{\pi_{\mathrm{P}}}}}$ (defined at the end of Sect. 3.5) as a function of positions, proper motions, magnitudes and colours.

Figures 12 to 17 show the variation of the normalised errors with ecliptic longitude, ecliptic latitude, proper motion in right ascension, proper motion in declination, V apparent magnitude and $(B-V)$ colour, respectively. In each figure, the solid lines give the running average over 100 points (around 0, formal error ≈ 0.1) and the running standard deviation (around 1, formal error ≈ 0.07): they show the variation of z and k, respectively. A Kendall's

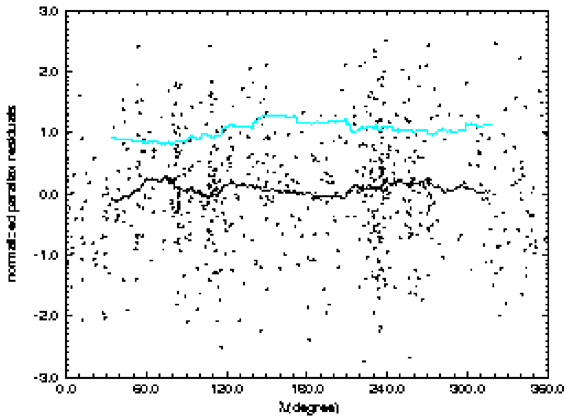

Fig. 12. Parallax normalised errors versus ecliptic longitude

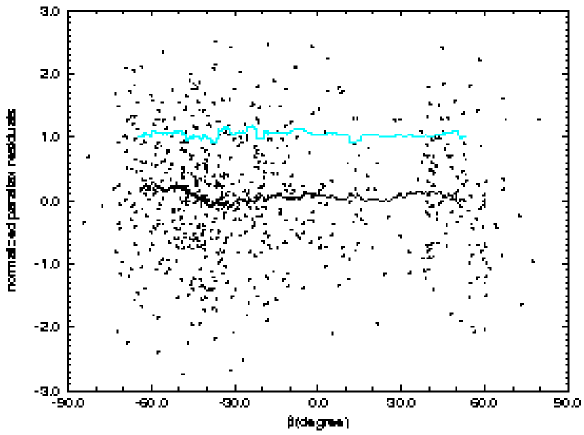

Fig. 13. Parallax normalised errors versus ecliptic latitude

τ test was applied in order to accept or to reject the following null hypothesis: the normalised errors and the astrometric or photometric data are independent. At a 5%-significance level, the null hypothesis was rejected only in the case of the variation of the normalised errors with the proper motion in right ascension and with the V-magnitude; however Fig. 16 shows a trend for the stars with a proper motion modulus in declination larger than 0.01 arcsec/y. There is no sensible variation of the errors with positions (Fig. 12, 13) and $(B-V)$ colour (Fig. 17). It is difficult to interpret whether the problems come from the determination of the astrometric parameters or from the $\pi_{\rm P}$ (stars wrongly classified, not detected variable or non-single stars, etc). In any case, the effect on z is smaller than 0.5 mas. For k the effect is small and it can be concluded that, with few exceptions, the formal errors are good estimates of the external errors.

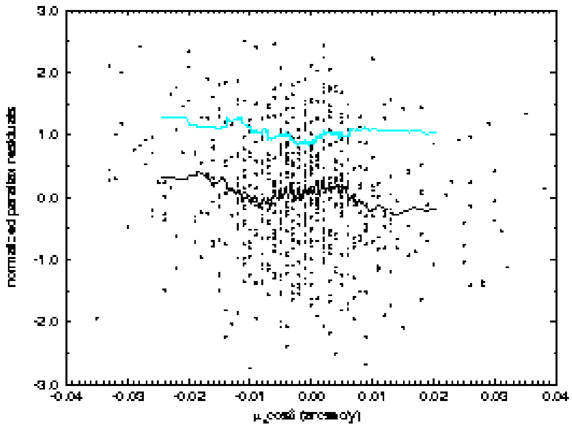

Fig. 14. Parallax normalised errors versus proper motion in right ascension

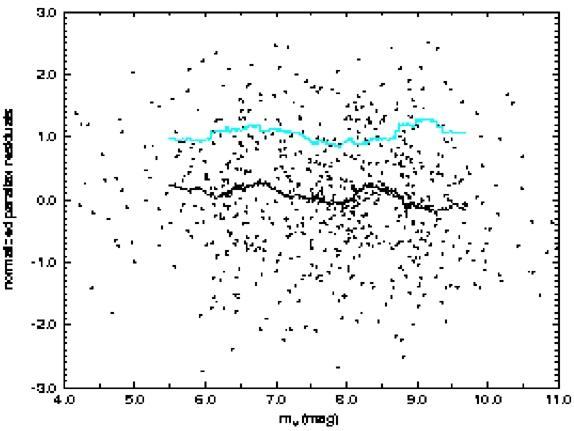

Fig. 15. Parallax normalised errors versus visual magnitude

6. Conclusion

The statistical properties of the Hipparcos preliminary parallaxes and of their errors were studied in order to obtain unbiased estimates of the global zero-point and of

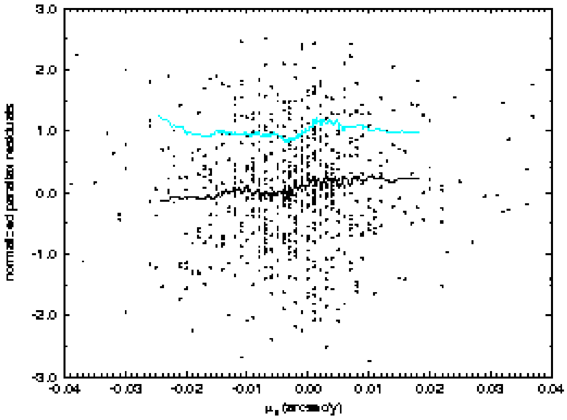

Fig. 16. Parallax normalised errors versus proper motion in declination

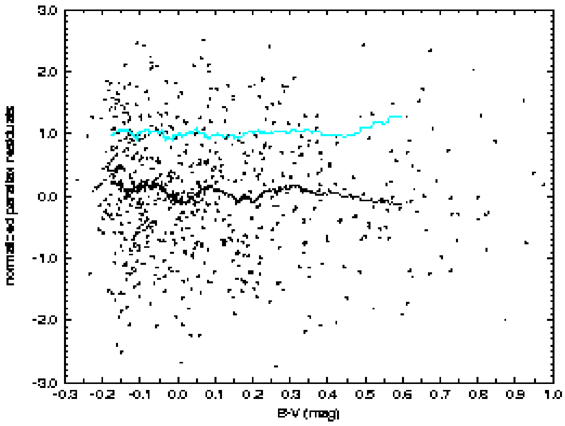

Fig. 17. Parallax normalised errors versus colour index

the external errors of the Hipparcos parallaxes. The parallaxes were compared to external parallax determinations available from various ground-based data sources: trigonometric, spectroscopic and photometric parallaxes, distance moduli of open clusters and of the Magellanic Clouds. In each case, the global zero-point shift of the Hipparcos preliminary parallaxes was not statistically different from 0.

A new, complete algorithm based on the maximum-likelihood of the conditional probability to observe the Hipparcos parallax of a star, given its magnitude, colour and galactic coordinates has been built and applied to a sample of distant stars. Various censorships are explicitly taken into account in the model and the detection of possible outliers is implemented. The algorithm allowed to obtain unbiased estimates of the global zero-point and of the external errors of trigonometric parallaxes together with their formal errors. It may be emphasised that the algorithm is firstly intended to calibrate the absolute magnitudes as a function of colour, however the most precise estimate of a star parallax, given its astrometric, photometric and kinematical data may also be obtained. The model was based on the following assumptions: exponential Z-distribution and Gaussian photometric errors, although other properties (e-g kinematical) could also have been introduced.

The obtained results, based on the first 30 months data, give strong indication that the Hipparcos preliminary parallaxes are free from any systematic errors (i.e. up to 0.1 mas), and that, for most of the stars, the true external errors are unlikely to be underestimated by more than about 5%. An improvement of the results could perhaps be obtained, using the new information on variability and multiplicity given by Hipparcos in order to reject stars with doubtful ground-based data.

In view of the results described above, it seems very likely that the final Hipparcos parallaxes will be both accurate and precise, beyond the original mission goals.

References

Arellano Ferro, A., Parrao, L., 1990, *Astron. Astrophys.* **239**, 205.

Arenou, F., 1993, *Contribution à la validation statistique des données D'Hipparcos: Catalogue d'Entrée et données préliminaires*, PHD thesis, Observatoire de Paris.

Balona, L.A., Shobbrook, R.R., 1984, *Mon. Not. R. Astron. Soc.* **211**, 375.

Crawford, D. L., 1975, *Astron. J.* **80**, 955.

Crawford, D. L., Mandwewala, N., 1976, *Publ. Astron. Soc. Pac.* **88**, 917.

Crawford, D. L., 1978, *Astron. J.* **83**, 48.

Crawford, D. L., 1979, *Astron. J.* **84**, 1858.

Dommanget, J., Nys, O., 1982, *Second Catalogue d'Ephémérides*, Comm. Obs. Royal de Belgique, Série B, n° 124.

Figueras, F., Torra, J., Jordi, C., 1991, *Astron. Astrophys. Suppl. Ser.* **87**, 319.

Gray, R.O., 1991, *Astron. Astrophys. Suppl. Ser.* **252**, 237.

Guthrie, B.N.G., 1987, *Mon. Not. R. Astron. Soc.* **226**, 361.

Hauck, B., Mermilliod, M., 1990, *Astron. Astrophys. Suppl. Ser.* **86**, 107.

Hilditch, R.W., Hill, G., Barnes, J.V., 1983, *Mon. Not. R. Astron. Soc.* **204**, 241.

Jaschek, C., Mermilliod, J. C., 1984, *Astron. Astrophys.* **137**, 358.

Kovalevsky, J. et al., 1995, A&A, this volume.

Lindegren, L., 1989, **The Hipparcos Mission**, ESA–SP 1111, Vol. III, 311.

Lindegren, L., 1992, *Space sciences with particular emphasis on High Energy Astrophysics* Satellite Symposium 3, Munich 30 March 1992, ESA SP-349.

Lindegren, L., van Leeuwen F., Petersen, C., Perryman M. A. C., Söderhjelm, S., 1992, *Astron. Astrophys.* **258**, 136.

Lingå, G., 1987, *Catalogue of open cluster data*, available through CDS, Strasbourg, France.

Lutz, T. E., Kelker, D. H., 1973, *Publ. Astron. Soc. Pac.* **85**, 573.

Malmquist, K.G., 1936, *Meddel. Stockholm Obs.* **26**.

Mermilliod, J-C., 1992, *Astronomy from large databases II*, Haguenau, ed. Murtagh F. & Heck A., 373.

Mermilliod, J-C., 1993, private communication.

Moon, T.T., 1985, *Communications from the University of London Observatory* n° **78**.

Perryman, M. A. C., et al. 1989, **The Hipparcos Mission**, ESA–SP 1111, Vol. I-III.

Ratnatunga, K. U., Casertano, S., 1991, *Astron. J.* **101**, 1075.

Schmidt-Kaler, T., 1982, in *Landolt-Börnstein*, Vol. VI/2b, K. Schaifers & H. H. Voigt eds., p. 1.

Turon C., Gómez A., Crifo F., Crézé M., Perryman M. A. C., Morin D., Arenou F., Nicolet B., Chareton M., Egret D., 1992, The Hipparcos Input Catalogue: I. Star selection, *Astron. Astrophys.* **258**, 74.

Turon C., Crézé M., Egret D., Gómez A.E., Grenon M., Jahreiß H., Réquième Y., Argue A.N., Bec-Borsenberger A., Dommanget J., Mennessier M.O., Arenou F., Chareton M., Crifo F., Mermilliod J.C., Morin D., Nicolet B., Nys O., Prévot L., Rousseau M., Perryman M.A.C., 1992, 1992, **The Hipparcos Input Catalogue**, ESA–SP 1136, 7 volumes.

van Altena, W. F., Truen-Liang Lee, J., Hoffleit, E. D., 1991, *The General Catalogue of Trigonometric Stellar Parallaxes*, CD-ROM preliminary version.

Zhang, E.H., 1983, *Astron. J.* **88**, 825.

This article was processed by the author using Springer-Verlag LaTeX A&A style file *L-AA* version 3.