Astron. Astrophys. 258, 149–156 (1992)

ASTRONOMY
AND
ASTROPHYSICS

23. Hipparcos photometry: NDAC reductions *

Dafydd W. Evans[1], Floor van Leeuwen[1], Margaret J. Penston[1], N. Ramamani[1] and E. Høg[2]

[1] Royal Greenwich Observatory, Madingley Road, Cambridge CB3 0EZ, United Kingdom
[2] Copenhagen University Observatory, Østervoldgade 3, DK-1350 Copenhagen K, Denmark

Received August 6, 1991; accepted January 18, 1992

Abstract. The Hipparcos mission will provide astrometry for about 120 000 stars. It will also give very accurate photometry for these stars. The resulting data base will by far surpass any existing photometric catalogue, both in accuracy and number of stars measured. This paper describes the calibrations and initial reductions of the photometry carried out by the Northern Data Analysis Consortium, one of the two data reduction groups analysing the data.

Key words: Hipparcos – data analysis – photometry – space vehicles and instruments

1. Introduction

The Hipparcos satellite has been successfully collecting data since the end of 1989. Although ending up in an elliptical transfer orbit, it seems quite likely that the main mission objectives will be achieved at the end of the mission. The current status of the satellite is dealt with by Perryman et al. (1992).

The main purpose of Hipparcos is to measure stellar positions very accurately. However Hipparcos can also measure magnitudes to millimagnitude accuracies. Using the counts from the main detector (the Image Dissector Tube or IDT) it is possible to calibrate these and produce magnitudes in a Hipparcos photometric system. Grenon, Mermilliod & Mermilliod (1992) describe how these Hipparcos magnitudes ($H_{\rm p}$) relate to other photometric systems. Very generally this passband covers both Johnson B and V.

Early investigations into the theoretical aspects of this work along with investigations using simulated data are contained in Mignard et al. (1987) and Mignard et al. (1989). The first paper contains details of work carried out by the FAST data reduction consortium, while the second refers to the work of both the data reduction consortia.

This paper describes the photometric calibration carried out at the Royal Greenwich Observatory (RGO) on the data from the IDT as part of the data processing for the Northern Data Analysis Consortium (NDAC). A general overview of all the NDAC reductions is given by Lindegren et al. (1992). Also covered in this paper are the various tests and comparisons that have been carried out in order to verify the procedure used.

Send offprint requests to: D. W. Evans

* Based on observations made with the ESA Hipparcos satellite

2. The data

The basic principle of the main detection system depends on the rotation of the satellite causing the images of stars from the two fields of view to pass over a grid situated at the focal plane of the telescope. Behind this lies the light gathering system, the image dissector tube. The details about the structure of Hipparcos can be found in Perryman & Hassan (1989).

As the satellite rotates, the image of a star being observed passes over the grid, thus modulating the signal received by the detector. On top of this signal there is the ever present background signal caused by detector noise, faint stars and the zodiacal light. The detector is sampled at a rate of 1200 Hz and an observation (also known as a frame transit), from which the magnitudes are calculated, consists of all the samples for a particular star within a 2.13 s period known as an observational frame. Since a number of different stars can be observed during an observational frame, the number of samples that make up a frame transit for a fully observed star can vary from 128 to 2560.

As the image of a star passes from one end of the grid to the other, an angular distance of $0\overset{\circ}{.}9$, it can be observed in up to 10 observational frames. After calibration has taken place, an average is calculated of all the frame transit magnitudes obtained from one complete crossing of the field of view or grid. This average is called a field transit.

A 5-parameter model is fitted to the frame transit data (Lindegren et al. 1992) and from this model two parameters, β_1 and β_2, can be used to estimate the magnitude of the star. β_1 is the average count rate and β_2 is the amplitude of the first harmonic in intensity. They are also referred to as the DC component and the AC component respectively. The corresponding magnitudes are referred to as H_1 and H_2.

In the case of single stars, β_1 is twice as accurate as β_2, but there is a possible disadvantage in that an estimate of the background count must be made. An error in this estimate could lead to small systematic effects. For double stars β_1 represents a combined intensity, while β_2 will depend on the orientation of the system with respect to the modulating grid and can be systematically fainter than β_1. In fact, this can be used to identify double stars (Söderhjelm et al. 1992).

In order to calibrate β_1 and β_2 a catalogue of photoelectrically determined magnitudes was produced and transformed onto the Hipparcos system (Grenon, Mermilliod & Mermilliod 1992). The colour information however comes from the main input catalogue. For the calibration stars this should be sufficiently accurate.

The counts (β) are converted into magnitudes using the equation

$$m = -2.5 \log_{10} \beta \, . \tag{1}$$

Since the accuracies of the counts are so high (much better than 10%), this procedure will not cause any bias larger than 0.001 mag.

3. The method

The calibration programme to convert the calculated β_1, β_2 to H_p is run on each dataset straight after the initial IDT data reduction has been carried out (see Lindegren et al. 1992). Separate solutions are carried out for both the preceding and following fields of view and for both the β_1 and β_2 counts.

The data is passed into the calibration routines through a preliminary rejection filter. The data must satisfy the following criteria:

1. The star is a calibration star i.e. classified as `'C'` in the Input Catalogue.
2. The colour of this star is between $-0.3 < (B-V)_T < 2.0$, where subscript T denotes the Tycho photometric system.
3. The star is not an unresolved double as classified by the Input Catalogue.
4. No error flags have been set for the data in the initial data reductions.
5. The star is fully observed i.e. visible in the frame for the full 2.13 s. This is done by rejecting all data with $\mathrm{abs}(G) > 0.88$, where G is the longitudinal grid co-ordinate which varies between ± 1 across the grid.
6. The data must be within about 2.5 mag of the expected value. This very broad filter improves the robustness of the solution.

The actual calibration is a least squares solution using the method of Householder Transformations (Bierman 1977). Since these use the observation equations, and not the normal equations, the weighting system used is inverse standard deviation (not variance). Thus for the magnitude transformed data the weights are given by

$$w = \left(\sigma_H^2 + (2.5 \log_{10} \mathrm{e} \times \sigma_\beta/\beta)^2\right)^{-\frac{1}{2}} , \tag{2}$$

where σ_H is the quoted standard deviation in the standard magnitude from the catalogue and σ_β the standard deviation in the measured count, β.

If the initial number of calibration observations (or frame transits) for any of the solutions is less than 150 the entire calibration is abandoned and the previous calibration must be used for that particular dataset.

The solution is calculated iteratively with an a posteriori four sigma filter being applied to the data to reduce the effect of outliers.

4. The model used

The model used in the least squares solution is linear and given by

$$-2.5 \log_{10} \beta - H_p = \sum_{i=1}^{N_{par}} a_i X_i \, , \tag{3}$$

where $H_p$is the ground based calibration magnitude in the Hipparcos system, N_{par} the number of parameters in the solution, a_i the parameters we solve for and X_i the variable (see Table 3).

Some of the terms that were introduced into the model are considered in the next two subsections.

4.1. The background term

Since β_1 is the sum of the source and background counts, I_s and I_b respectively (see Lindegren et al., 1992), it is important to take account of the background, especially for faint stars where it is a significant contribution to the counts. It should be stressed that the background does not affect the β_2 measurements and is not a term in its calibration.

The effect of including I_b in our calibrations can be seen in the following equations:

$$\log_{10} \beta_1 = \log_{10} I_s + \log_{10}\left(1 + \frac{I_b}{I_s}\right) \tag{4}$$

$$= \log_{10} I_s + \frac{1}{\ln 10}\frac{I_b}{I_s} + 0\left(\frac{I_b^2}{I_s^2}\right) . \tag{5}$$

for small values of I_b/I_s.

Therefore adding the effect of the background to Eq. (3) we get

$$\begin{aligned} -2.5 \log_{10} \beta_1 - H_p &= \sum a_i X_i - \frac{2.5}{\ln 10}\frac{I_b}{I_s} \\ &\sim \sum a_i X_i - \frac{2.5 I_b}{\ln 10} 10^{0.4(H_p + a_1)} \\ &\qquad \text{using } I_s \sim 10^{-0.4(H_p + a_1)} \\ &\sim \sum a_i X_i + a_9 \times 10^{0.4(H_p - 8)} \end{aligned} \tag{6}$$

where $a_9 = \frac{-2.5 I_b}{\ln 10} 10^{0.4(8+a_1)}$ and a_1 is the zero point term in the main model. $(H_p - 8)$ is used in order to have a convenient scaling.

Thus the additional term in our model to account for the background is $10^{0.4(H_p-8)}$ and the new coefficient is a_9. An example of the effect of the background term is shown in Fig. 1.

To reduce the effect of the second order term in this approximation an average value for the background has been measured using the provisional data and this is now subtracted from β_1 before the calibration is attempted. Thus what is being measured in the calibration is the additional background to that which has been estimated.

4.2. The grid position dependence

A calibration was first carried out with a simple model consisting of (1) a cubic in colour, (2) position on the grid as linear functions of G and H (where H is the grid co-ordinate perpendicular to G) along with their linear colour cross terms and (3) a background term i.e. terms a_1 to a_9 in Table 3. An example of the residuals, after this calibration, as a function of grid position is shown as a contour plot in Fig. 2.

Table 1. The positions of the reference points of the radial model given in grid co-ordinate units.

Radial distance	0.0	0.1	0.3	0.5	0.7	0.9	1.1	1.3
Variable name	R_0	R_1	R_2	R_3	R_4	R_5	R_6	R_7

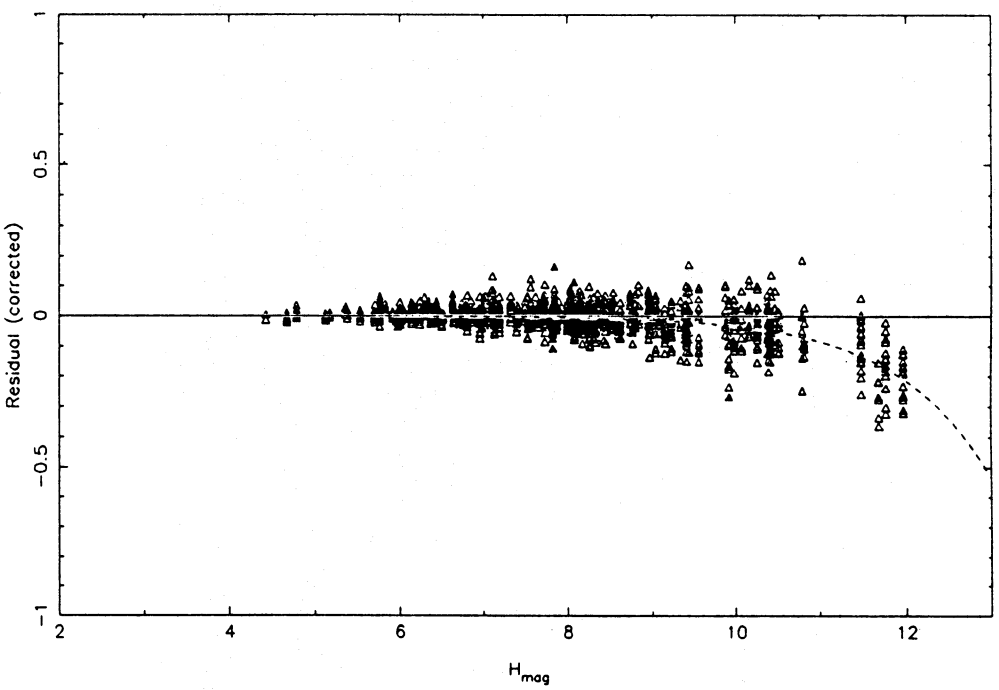


Fig. 1. This diagram shows the effect of the background term on β_1 for calibration stars in an example data set. The data points are residuals from a solution using all terms in Table 3 except the background term (a_9). Thus the need for this term is easily seen at the faint end where the dashed line shows the residual caused by not having a background term. This data comes from the preceding field of view of NDAC dataset DS4100

Table 2. Estimates on the size of the standard deviation caused by the colour dependence of the grid dependent error in units of 0.001 mag.

Field of View	Preceding β_1	Following β_1	Preceding β_2	Following β_2
Blue	3.4	3.5	6.8	5.4
Yellow	3.9	3.9	7.0	6.3
Red	4.9	4.8	7.3	5.6
All	3.4	3.3	6.4	4.8

Table 3. The variables used in the calibration

Parameter a_i	Variable X_i	Physical meaning
a_1	1	zero point
a_2	$(B-V-0.5)$	colour terms
a_3	$(B-V-0.5)^2$	
a_4	$(B-V-0.5)^3$	
a_5	G	grid abscissa
a_6	H	grid ordinate
a_7	$(B-V-0.5)\times G$	a_5 colour dependence
a_8	$(B-V-0.5)\times H$	a_6 colour dependence
a_9	$10^{0.4(H_p-8)}$	background term (see §4.1)
a_{10}	R_2	radial terms (see §4.2)
a_{11}	R_3	
a_{12}	R_4	
a_{13}	R_5	
a_{14}	R_6	
a_{15}	R_7	

A polynomial model to fit this radial dependence was deemed to be inappropriate. The model chosen for the calibration was a linear interpolation one in which each observation equation contains terms representing the relative fractions of the two nearest radial reference points. The positions of the reference points in grid co-ordinate units are given in Table 1.

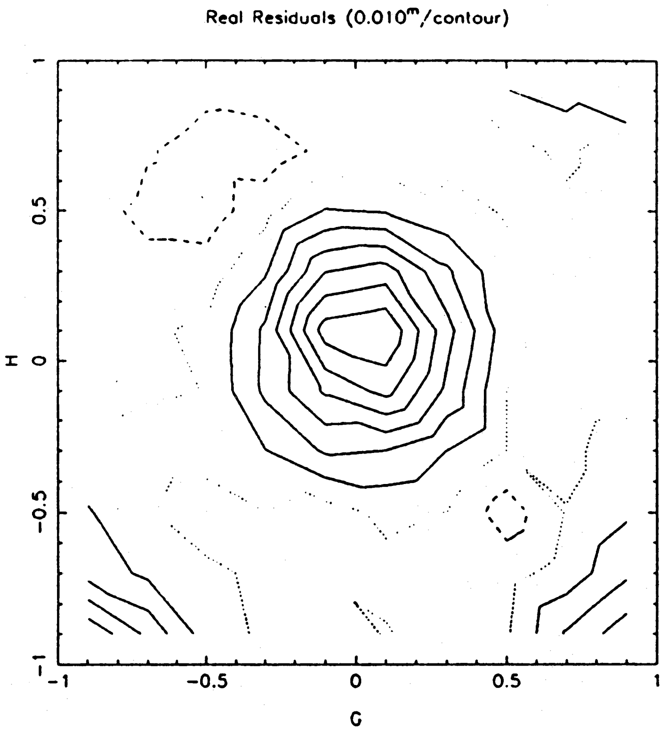


Fig. 2. This contour plot shows an example of the radial dependence of the calibrations after accounting for colour, linear gradient across the grid and the background term. The contours here are every 0.010 magnitude with the dotted line being zero, the dashed lines negative and the solid lines positive residuals

The central location for the radial dependence could be determined to the desired accuracy by eye. This was taken to be (0.03, 0.11) in grid co-ordinate units. Radial measurement is considered as having been taken from this point.

Although the model contains R_0 and R_1 points they are not solved but are set to zero. The reasons for this are that R_0 is effectively the zero point a_1 and a zero derivative at $R = 0$ is required which implies that R_1 has to be equal to R_0. Points outside $R = 1.3$ are calculated as an extension of the line between 1.1 and 1.3. Thus this radial model adds an extra 6 parameters to the overall calibration model.

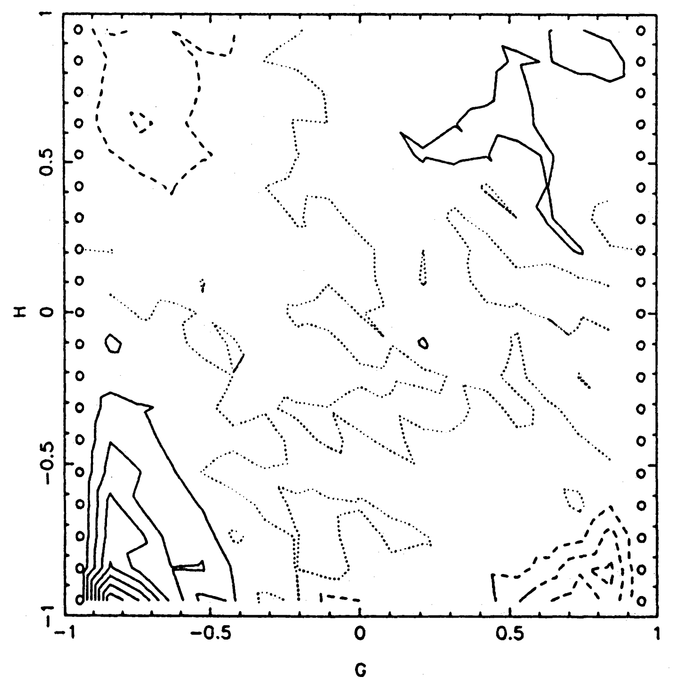


Fig. 3a. This contour plot shows an example of the remaining residuals of the calibrations after accounting for colour, gradient, background and radial terms for the provisional data taken *before* 11 December 1989. Note that these contours are every 0.010 magnitude with the dotted line being zero and the solid lines positive residuals. Another point to note is that these (and the following) contour plots are created from more elements (19 × 19) than the previous ones and that since no data from partially observed stars is used no values can be calculated for the extreme G values

The remaining residuals of all the calibrations as a function of G and H were collected in the form of a 19 × 19 array. Investigation of these arrays as contour plots as a function of time revealed that the remaining grid effects are small and that the pattern of the residuals remained constant except for a discontinuity that occurred between 11 December 1989 and 22 February 1990 probably associated with refocussing of the satellite that occurred on 3 and 12 January 1990. This change is illustrated in Fig. 3a (before) and Fig. 3b (after). Splitting the data into smaller time intervals produced no noticeable (apart from increased noise) departure from these overall plots.

The data was also split into three different colour bins. The three ranges defined at RGO were

1. $(B - V) < 0.45$ (Blue),
2. $0.45 < (B - V) < 0.95$ (Yellow),
3. $0.95 < (B - V)$ (Red).

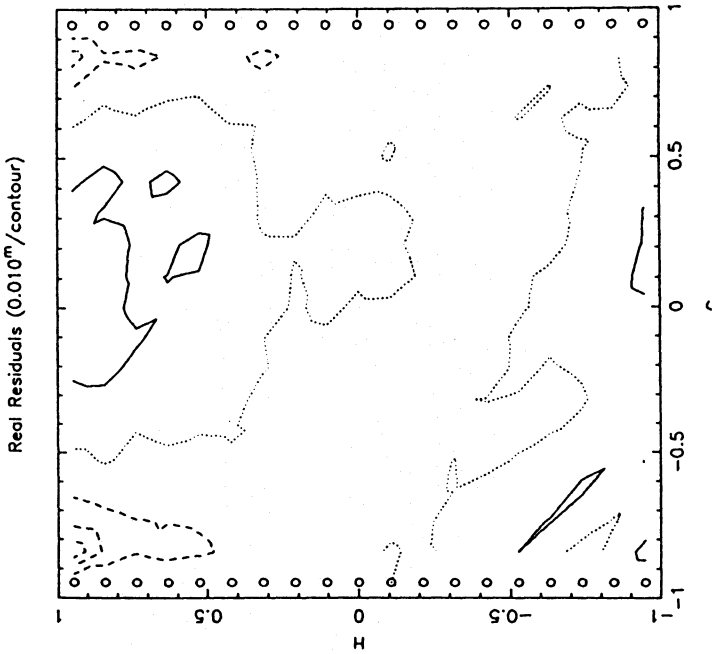


Fig. 3b. This contour plot is the same as Fig. 3a but that it is for the provisional data taken *after* 22 February 1990

These ranges were chosen so that approximately the same number of stars would go into each sample. Figure 4 shows an example of the results of this investigation. Each plot contains about 100 000 observations. This figure shows that the grid dependence remaining is slightly colour dependent.

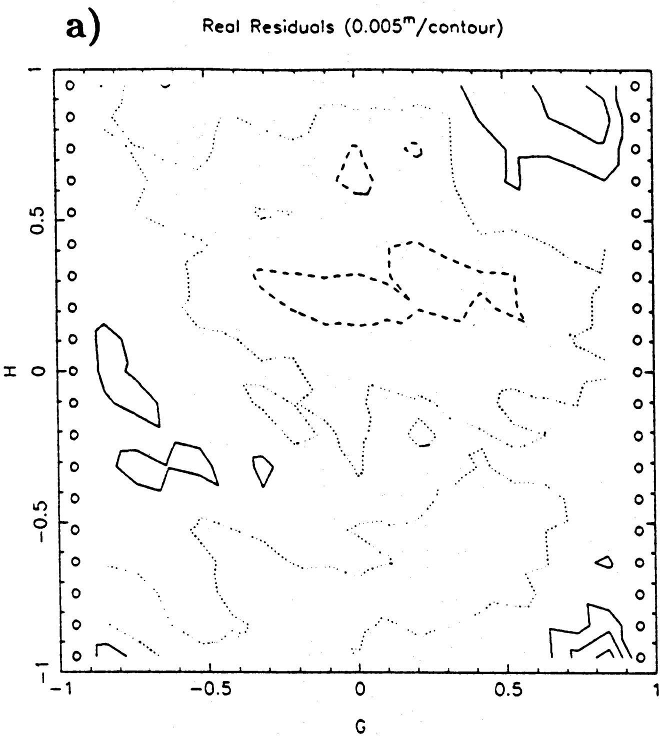
However, since observations for each star occur at different places across the grid we can treat these residuals as an additional random standard deviation to the photon counts. An estimate on the size of this standard deviation can be made by measuring the effective sigma of the array that has been used to produce the contour plots. Table 2 shows these values in millimagnitudes. For an eighth magnitude star the usual standard deviation on a β_1 measurement is about 0.03 mag so the above additional standard deviation is not really significant. Even for fourth magnitude stars it is only increasing the standard deviation by about 20%. It is therefore considered that this residual can be left uncorrected.

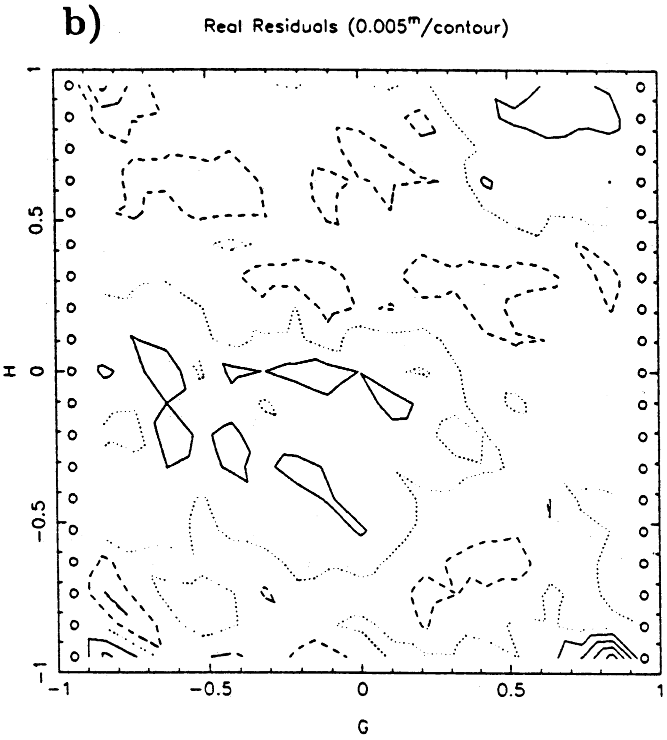
4.3. The current model

The model that is being used currently at the RGO is the 15 parameter model given in Table 3.

The standard deviation in the magnitudes calculated from the calibrated counts (β's) is almost all due to the standard deviation in the counts themselves. However there is some contribution from sources linked with the calibration itself.

One problem that might arise in the calibration comes from our lack of knowledge about the colour of the star and the existence of a colour term in the calibration. The solution adopted by NDAC is to use the star mapper updates for B_T and V_T where available to calculate the colour (see van Leeuwen et al. 1992). When the standard deviation from our updates is less than that of the input





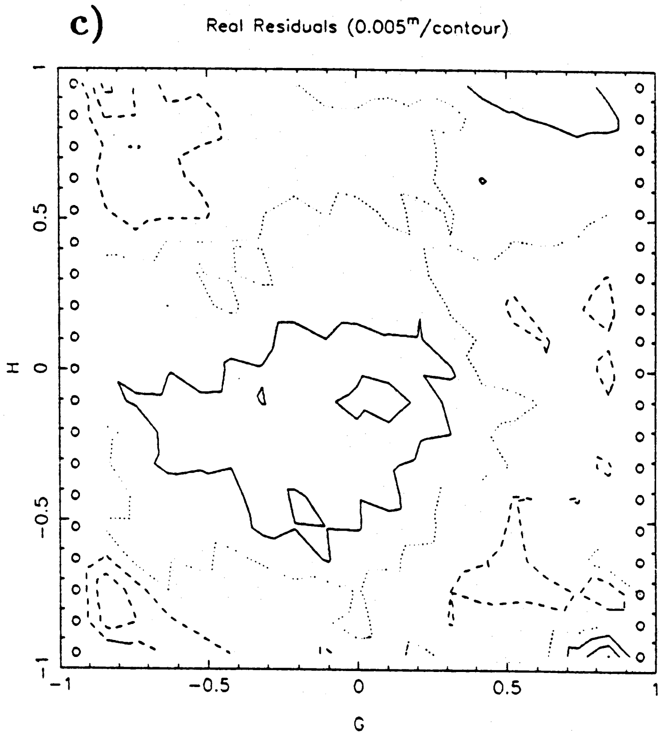


Fig. 4. These contour plots show the remaining residuals of the calibrations after accounting for all terms in Table 3 for stars in the colour ranges (a) $(B-V) < 0.45$ (Blue), (b) $0.45 < (B-V) < 0.95$ (Yellow) and (c) $0.95 < (B-V)$ (Red). Note that these contours are every **0.005** magnitude with the dotted line being zero, the solid lines positive and the dashed lines negative residuals

catalogue we use our updated colour in the conversion of counts to magnitudes.

The calibrations carried out at the RGO is on the **frame** transit data i.e. the individual observations. It is expected that the IDT photometry will be published in the form of **field** transits, an average of all a particular star's observations during a single crossing of the main detector grid. The field transits are calculated at the same time as the catalogue updates for IDT data are being carried out and are stored on optical disc for later use.

The standard deviation for these field transits is calculated without taking into consideration variability over the field transit (a period of about 20 s) and are simply the expected precisions calculated for each field transit. Typical values for an individual field transit standard deviation are given in Table 4.

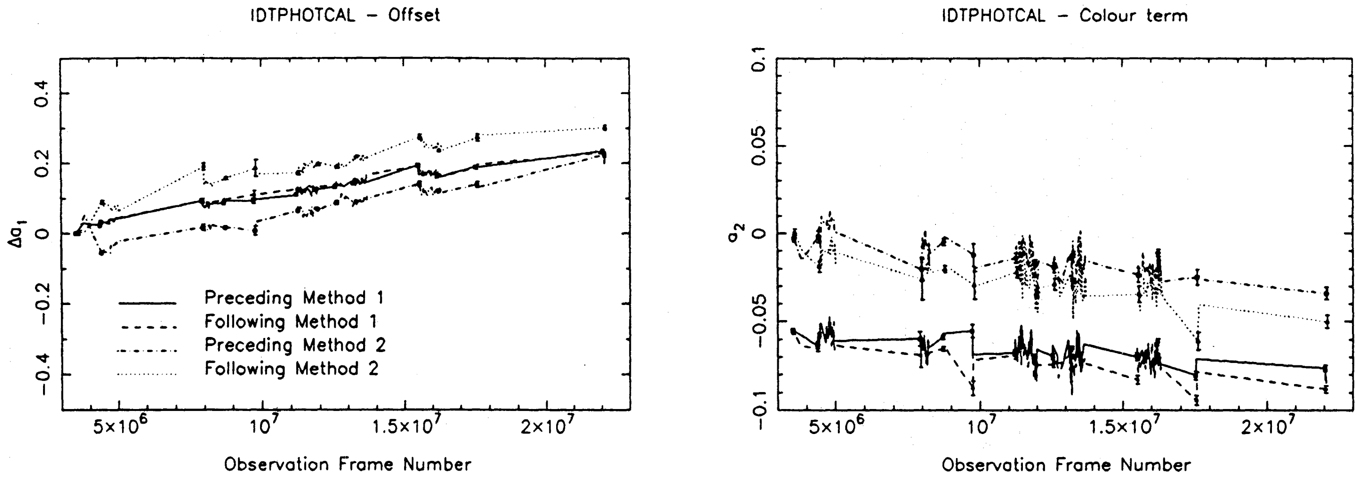


Fig. 5. These plots show the change of two of the IDT calibration coefficients (a_1 and a_2) over the period November 1989 – February 1991. In order to plot all four a_1 terms on the same plot the initial value has been subtracted from all data points hence they all start at 0.0. Method 1 refers to the β_1 calibrations and Method 2 to β_2. Note the strong time dependence on the a_1 term which is colour dependant hence the gradient seen in the a_2 plot

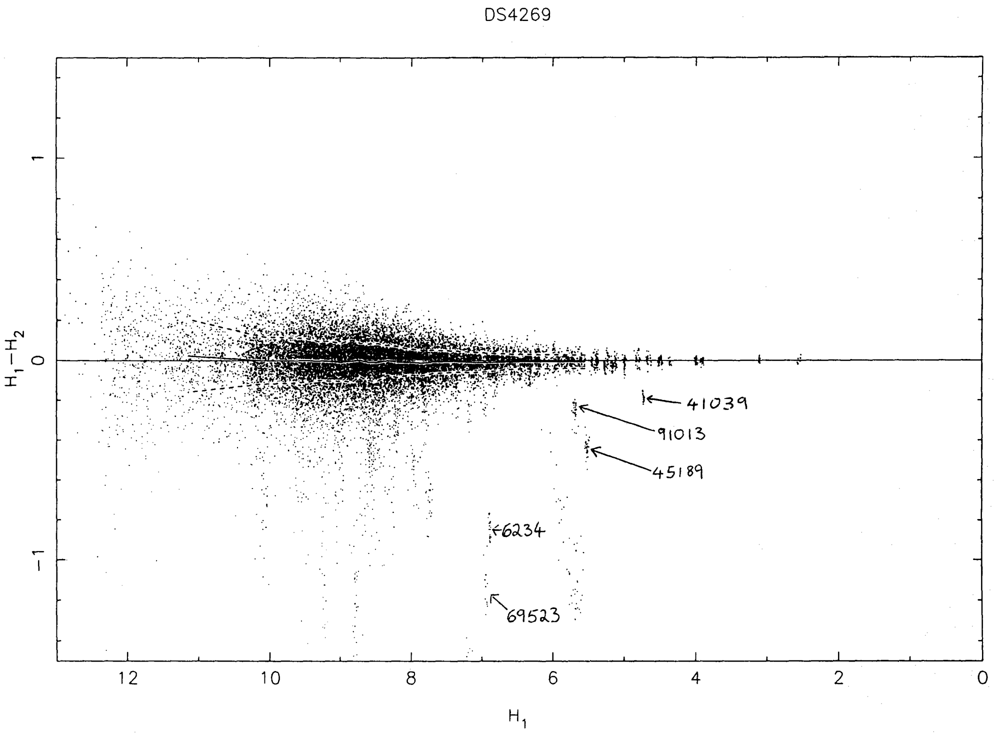


Fig. 6. This plot shows the differences between the individual measurements (frame transits) of H_1 and H_2 for all stars in dataset DS4269. A median (solid) and one sigma line (dotted) have been calculated. The median line is very close to the $H_1 - H_2 = 0$ line. Some data points that are far from the median line have been identified with INCA numbers. These stars are all classified as non-single by INCA

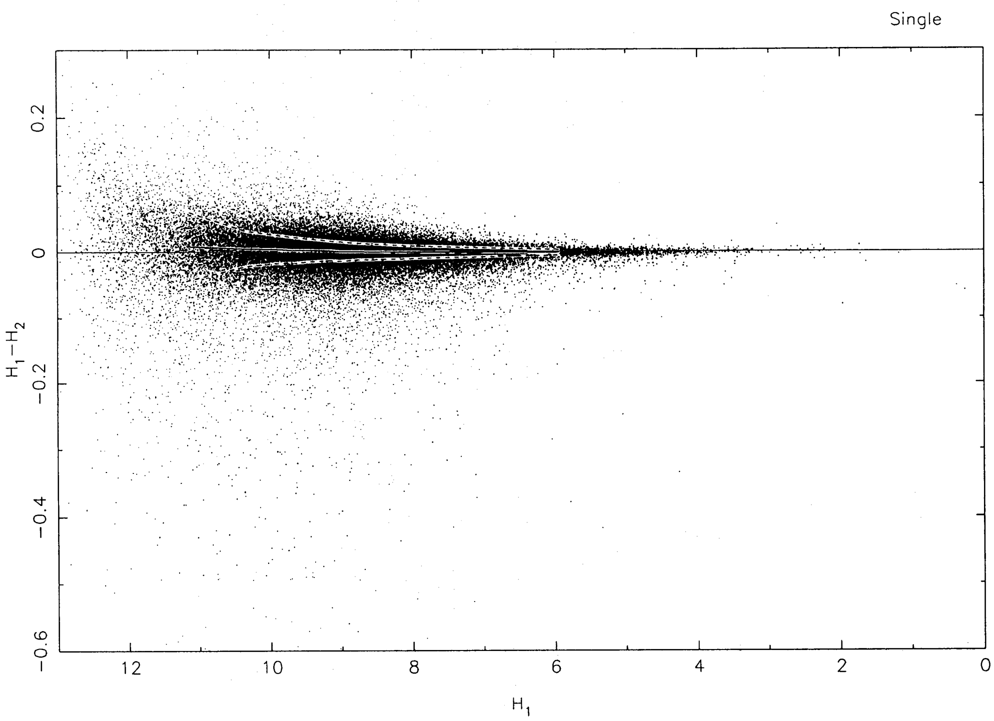
Table 4. Typical values for an individual field transit standard deviation in magnitudes.

H_p	4	5	6	7	8
H_1 s.d.	0.002	0.003	0.004	0.006	0.010
H_p	9	10	11	12	
H_1 s.d.	0.014	0.020	0.030	~0.040	

These standard deviations are directly comparable with the precisions quoted in van Leeuwen et al. (1992) since the calculation of a field transit magnitude is exactly the same as that of an updated magnitude in the working catalogue with about 6–8 observations (frame transits). The H_2 standard deviations are approximately twice the size of those of the H_1 field transits.

5. The variation of the calibration with time

Investigating how the calibration parameters change with time provides us with a good method of checking the stability of the calibrations. The 12 weeks of satellite data provided to the consortia for testing their software, usually referred to as the provisional tapes



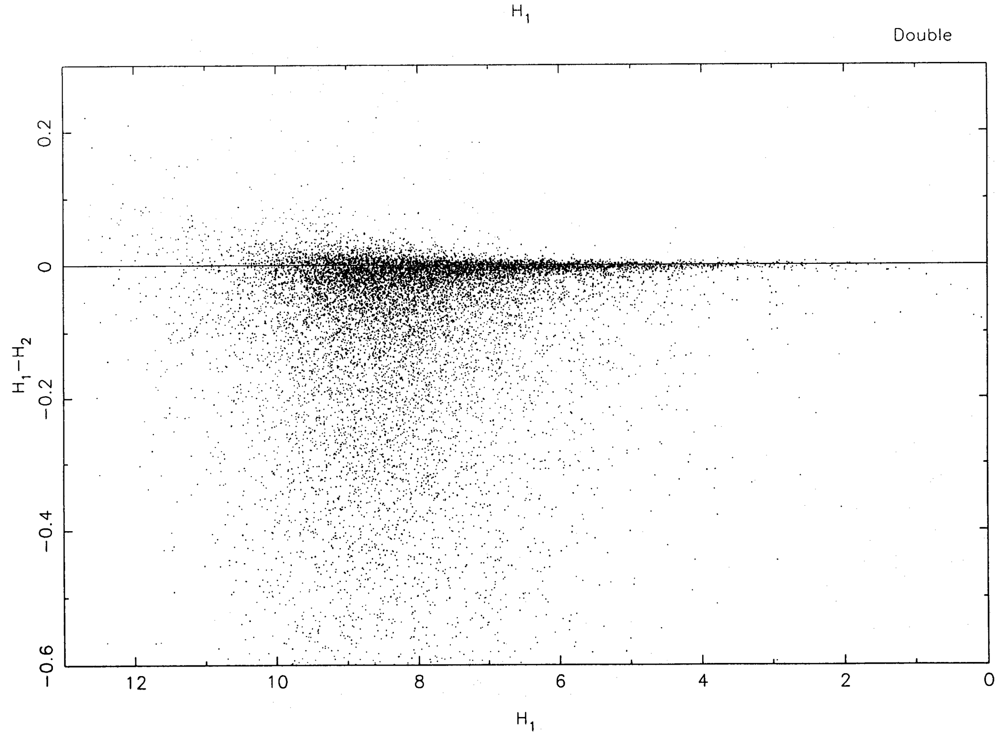


Fig. 7. These plots show similar data to that in Fig. 6. The difference is that the data here consists of the accumulated photometry of all the provisional tapes received at RGO. The scale has been changed to accommodate the reduction in the noise level. The top plot only contains data from stars classified by INCA as single. The lower plot is the corresponding one for non-single stars. No median or sigma lines were drawn for the lower plot due to its broad, asymmetric nature

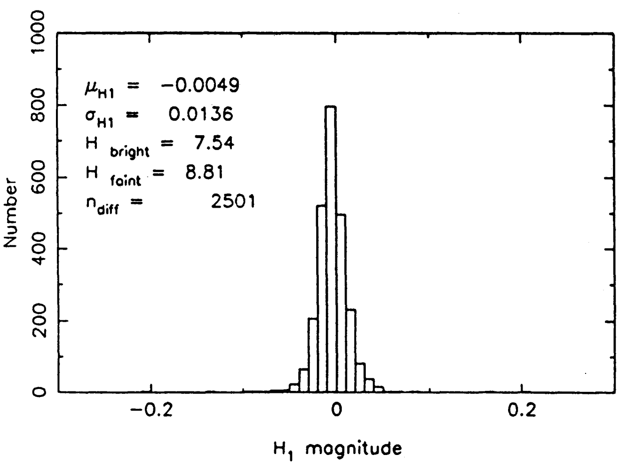
or data, span from November 1989 to February 1991 and should provide an indication of the large scale changes that have occurred during the first year of the mission. Figure 5 shows the development of two of the parameters over this period.

The data has large gaps in it which can cause inaccuracies in certain places. The reason for this is that after a large gap the running solution must be initialized otherwise a bias would be introduced due to the time dependence of the calibration. If a small dataset is the first one after the gap, then this calibration will be relatively inaccurate due to not having many observations. This explains the jumps and large standard deviation seen in some of the plots.

6. A comparison between the H_1 and H_2 results

A comparison between the results calculated from the β_1 and β_2 values for the programme stars is an independent check on the photometric calibrations. Figure 6 shows a plot of the differences between individual measurements (frame transits) of H_1 and H_2. There are a significant number of points in this plot that lie well below the median of the distribution. The vast majority of these data points belong to stars classified as unresolved doubles (see Söderhjelm et al. 1992). As explained in Sect. 2, this is to be expected.

This type of information will be used to identify "new" double stars. The scatter produced by using individual frame (or even field)



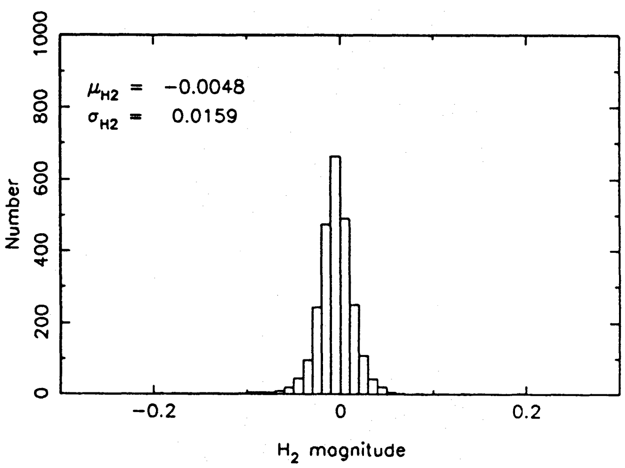


Fig. 8. These histograms show the H magnitude residuals (standard − observed) for all stars classified as C (confirmed standards) in the Input Catalogue in the magnitude range $7.54 < H_p < 8.81$. The plot on the left is the H_1 (dc) residuals and that on the right is for H_2 (ac). The μ and σ values quoted are medians and robust width estimates. The latter is equivalent to the σ of a Gaussian distribution and is less sensitive to the wings of a distribution than a simple rms calculation

transits can be reduced by using the accumulated data from the updated catalogue (van Leeuwen et al. 1992). This is shown in Fig. 7. The scale has been changed from that used in Fig. 6 due to the reduction of the random scatter and the plot has been split into two to show the different distributions for assumed single and known non-single stars. 90% of the points with $H_1 - H_2 < -0.2$ are classified as being non-single in the Input Catalogue.

It should be noted that the distributions will not look exactly like this with the production data since a more even spread of the scanning law will be available and problems which might arise with the provisional data due to the orientation of double systems with respect to the observing grid will disappear.

7. Internal tests

One of the most reliable tests that can be carried out on this calibration procedure is to see if the H_1 and H_2 updated values i.e. the weighted averages of all the calibrated data from the provisional tapes, agree with the standard values. Figure 8 shows the residual histograms for all stars classified as confirmed standards in the Input Catalogue. Two basic conclusions can be drawn from these plots. There is a small bias present and the widths are larger than expected. Investigations are still being carried out into these points.

8. Conclusions

This paper confirms the high potential expected from the photometry produced by the main mission of the Hipparcos satellite, both in terms of accuracy and the number of stars measured. Various tests have been carried out on the calibrations of the photometric measurements and are shown to confirm their validity, stability and accuracy. The high quality of the calibrated data will make it possible to identify many new double and variable stars.

Acknowledgements. The Hipparcos team at the Royal Greenwich Observatory is supported by the UK Science and Engineering Research Council. Thanks should also be given to Michel Grenon and François Mignard for interesting discussions.

References

Bierman, G.J., 1977, Factorization Methods for Discrete sequential Estimation, Volume 128 in Mathematics in Science and Engineering, Academic Press
Grenon M., 1992, A&A, this issue (paper 13)
van Leeuwen F., Evans D.W., Lindegren L., Penston M.J., Ramamani N., 1992, A&A, this issue (paper 19)
Lindegren L., Høg E., van Leeuwen F., Murray C.A., Evans D.W., Penston M.J., Perryman M.A.C., Petersen C., Ramamani N., Snijders M.A.J., Söderhjelm S., Andreasen G.K., Cruise A.M., Elton N., Lund N., Poder K., 1992, A&A, this issue (paper 3)
Mignard, F., Froeschlé, M., Falin, J.L., 1987, Third FAST Thinkshop: Processing of Scientific Data from the ESA Astrometry Satellite HIPPARCOS, eds. P.L.Bernacca & J.Kovalevsky, page 247
Mignard, F., Froeschlé, M., Falin, J.L., Andreasen, G.K., Høg, E., Grewing, M., Scales, D., 1989, ESA SP-1111, Vol. III, page 205
Perryman, M.A.C., Hassan, H. 1989, The Hipparcos mission, ESA SP-1111, Vol. I, The Hipparcos satellite
Perryman, M.A.C., Lindegren, L., Murray, C.A., Høg, E., Kovalevsky, J. 1989, The Hipparcos mission, ESA SP-1111, Vol. III, The Data Reductions
Perryman M.A.C., Høg E., Kovalevsky J., Lindegren L., Turon C., Bernacca P.L., Crézé M., Donati F., Grenon M., Grewing M., van Leeuwen F., van der Marel H., Murray C.A., Le Poole R.S., Schrijver H., 1992, A&A, this issue (paper 1)
Söderhjelm S., Evans D.W., van Leeuwen F., Lindegren L., 1992, A&A, this issue (paper 24)

This article was processed by the author using Springer-Verlag TEX A&A macro package 1992.