

# Variable stars in the Tycho photometric observations

## I. Detection<sup>\*,\*\*</sup>

S. Piquard<sup>1</sup>, J.-L. Halbwegs<sup>1</sup>, C. Fabricius<sup>2</sup>, R. Geckeler<sup>3</sup>, C. Soubiran<sup>4</sup>, and A. Wicenec<sup>5</sup>

<sup>1</sup> Observatoire astronomique de Strasbourg, 11 rue de l'Université, 67000 Strasbourg, France

<sup>2</sup> Copenhagen University Observatory, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark

<sup>3</sup> Institute for Astrophysics, Technikerstrasse 25, 6020, Innsbruck, Austria

<sup>4</sup> Observatoire astronomique de Bordeaux, 33270 Floirac, France

<sup>5</sup> ESO, Karl-Schwarzschild-Strasse 2, 85748, Garching, Germany

Received 11 September 2000 / Accepted 24 April 2001

**Abstract.** We present an original way to detect variable stars in the Tycho photometric observations. A modified Kolmogorov-Smirnov statistical test has been constructed, based on a model of most of the satellite's parameters, leading to the dispersion in the measurements. This allows us to build a treatment taking into account truncated detections and censored measurements, and to search for variability in the faint part of the catalogue. The selection threshold of variable stars has been calibrated to minimize false alarm rate. Reliable results were thus obtained for stars as faint as 11 mag. One thousand and ninety-one suspected variable stars have been listed, 407 of which are already known in Hipparcos, GCVS or NSV, and 684 of which are suspected variable stars to be identified soon.

**Key words.** methods: statistical – stars: variables: general

## 1. Introduction

### 1.1. Variable stars surveys

The scientific interest in complete magnitude samples for many types of variable stars is considerable. A lot of variability types are used as distance candles, among them Cepheids, RR Lyræ, Mira,  $\delta$  Scuti, detached eclipsing binaries and supernovæ Ia (see for example Egret & Heck 1999), but variable stars are also useful in determining some fundamental parameters, and put constraints on the knowledge of stellar evolution or internal structure. For example, stellar masses, radii and luminosities can be derived from detached eclipsing binaries, and mass transfer processes can be constrained by studying semi-detached binaries. Contact binaries give constraints on theoretical models of double star evolution or the thermal evolution of such systems. We must understand the processes encountered in regular and irregular pulsations, as well as their implication on the internal structure and the metallicity evolution of these stars in order to construct models of population synthesis or galaxy evolution.

Unfortunately until some ten years ago samples of known variable stars were strongly inhomogeneous, and several variability types still have too few known members. This was emphasized by the variability search performed on Hipparcos data, when several variables of unfrequent types were found (as for example variable stars with period around 1 day, which are obviously rarely detected in on-ground surveys) (van Leeuwen 1997). The interest in discovering variable stars has grown in the last few years, as several large-scale surveys have discovered large numbers of new variable stars, in samples expected to be as complete as possible. These include the ASAS (Pojmański 1997, 1998) which is looking for transients in the southern hemisphere up to the 13th magnitude and ROTSE (Akerlof et al. 2000) which looks for variability in the northern hemisphere up to the 15th magnitude. The list is not exhaustive. The search for microlensing events also led to the discovery of variable stars by surveys such as OGLE (Udalski et al. 1992; Maceroni & Rucinski 1999, and references therein), MACHO (Alcock et al. 1999, and references therein), and EROS (de Laverny et al. 1998, and references therein).

New variable stars have also been discovered during the HIPPARCOS mission with  $V \leq 12.4$  mag (Eyer & Grenon 1997; van Leeuwen 1997). Among the 8237 new variable stars, 970 have been identified as periodic variables. As Hipparcos is complete up to  $V = 7.3$ – $9.0$  mag (depending on the galactic latitude), and Tycho is complete up to  $V = 10.9$  mag, new variable stars in the Tycho photometry can be expected.

---

Send offprint requests to: S. Piquard,  
e-mail: [piquard@astro.u-strasbg.fr](mailto:piquard@astro.u-strasbg.fr)

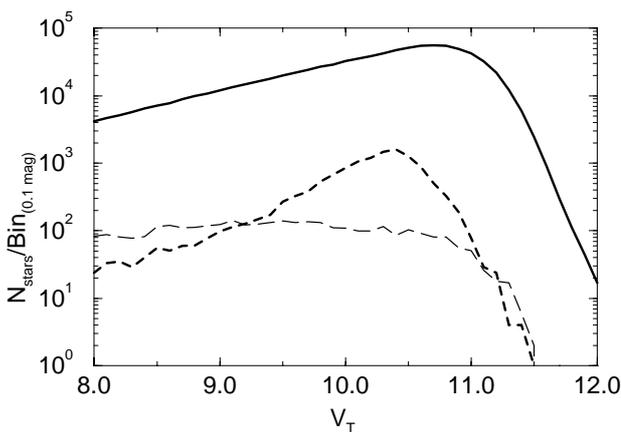
\* Based on observations made with the ESA Hipparcos satellite

\*\* The catalogue is only available in electronic form at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](http://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/373/576>

### 1.2. The variability flag in the Tycho catalogue

The Hipparcos star mapper was primarily devoted to determining the satellite’s attitude. The recording of the photon counts led to the Tycho catalogue of one million stars, for which an average of 130 measurements were obtained during the 3.5 years of the mission. The magnitude completeness of Tycho up to  $V_T = 10.9$ , and the time sampling of the measurements will enable the detection of new variable stars and improve the knowledge of those populations in the solar neighborhood (Mauder & Høg 1987). The Tycho photometric data were partly published in 1997. An earlier research into variable stars in the Tycho database was undertaken in 1997 through the 384 000 brightest stars of the catalogue (Großmann et al. 1997). The stars presenting a scatter in their measurements greater than a threshold were suspected to be variable stars and were flagged in the Tycho catalogue. Unfortunately, we notice that the proportion of suspected variable stars grows with magnitudes (Fig. 1), and those flags are said to show an unexpectedly large scatter in the measurements; this is not necessarily due to variability, but may arise from duplicity or disturbance caused by stars crossing the star mapper at the same time.

Friedrich et al. (1997) presented some comments on the establishment of these flags. They also tried several standard statistical methods and concluded that the usual spread, skewness or kurtosis of the light curve did not distinguish variable stars from constant stars. They also tried to detect variability by looking for periodicity in the light curves, but the poor sampling and the noise made automatic detection difficult.



**Fig. 1.** Variable stars distribution in the Tycho catalogue, as derived from the variability flag. The heavy-line represents the total number of star in the Tycho catalogue (849 806 stars having good astrometric solution). The thick dashed-line represents the number of stars flagged as “strong evidence of variability” (10 320 stars), and the thin long-dashed-line shows the stars either in the GCVS or in the NSV (3331 stars).

### 1.3. An overview of the Tycho photometric reduction

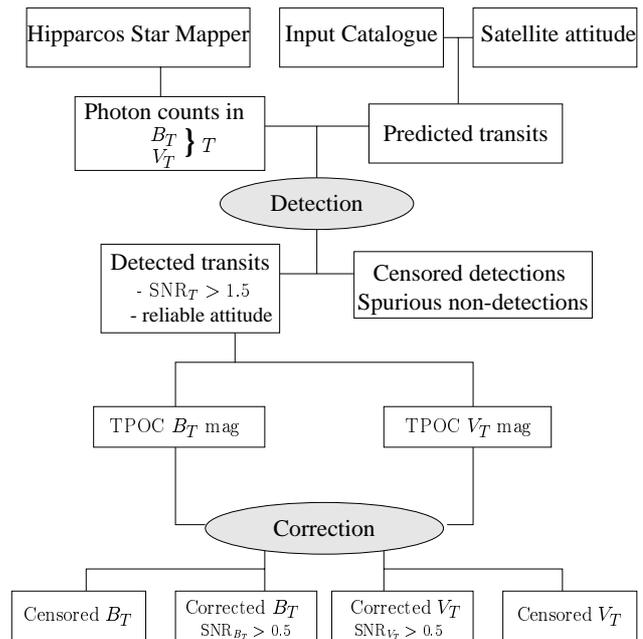
It is hard to deal with the photometric observations of the Tycho programme, since they were produced by a rather complicated process. The Tycho program is presented in ESA (1997). We summarize the main characteristics relating to the photometric reduction hereafter.

#### 1.3.1. The observations of the Hipparcos star mapper

The satellite observed two fields simultaneously, and both were projected and merged on the focal plane. As the satellite was precessing, the stars crossed the focal plane twice during a rotation, once for each of the two fields. The second crossing occurred 20 minutes after the first. The instrument used by Tycho was the star mapper, comprising two slit groups, the “inclined” and the “vertical”. When a star crossed the mapper, the inclined slits were crossed first and the vertical a few seconds later; however, the photons from both slit groups were recorded with the same device as for the fields. The light of the star was split in two photometric channels in the  $B_T$  and  $V_T$  bands, and the photon counts were recorded separately for each. Merging the counts in  $B_T$  and  $V_T$  defines a wide-band channel (the “T” channel).

#### 1.3.2. The photometric reduction

The successive steps and the outputs of the photometric reduction are briefly presented hereafter; they are also summarized in Fig. 2.



**Fig. 2.** The organisation of the photometric reduction, from the photon counts until the corrected magnitude measurements.

The transit of a star across a slit group was predicted from an input catalogue and from the satellite attitude. The transit was detected when a peak was found in the photon counts of the “T” channel with a signal-to-noise ratio (hereafter SNR) greater than about 1.5. The signals in the  $B_T$  and  $V_T$  channels were estimated and the magnitudes were derived. A *censored detection* was recorded when the transit was not detected. A censored detection usually means that the signal had a SNR below the threshold but in a few cases the non-detection actually arose from errors in the preliminary satellite attitude used in the detection process; the rate of these so-called *spurious non-detections* is about 6%. The photometric observations thus obtained were recorded in the Tycho Photometric Observation Catalogue (TPOC). Part of the TPOC was published with the Tycho catalogue: the Tycho Epoch Photometry Annexes.

When the mean magnitudes of the faint stars were calculated during the production of the Tycho catalogue it appeared that the magnitudes derived in the photometric reduction of the transits were too bright. The signal was systematically overestimated by a factor depending on the signal and on the background. A correction routine was then applied to the individual measurements when they were entered in the mean magnitude calculation software (however, the magnitude measurements in the TPOC and in the Tycho Epoch Photometry Annexes were kept unchanged). The estimated signal in the  $B_T$  or the  $V_T$  band may be very low after the correction, and it was assumed that a magnitude measurement was dubious when its corrected SNR was less than 0.5. When this happened the corrected magnitude was ignored, but we took account of the fact that the SNR was below the 0.5 threshold. The magnitude measurement was then called a *censored magnitude*.

The magnitude measurements, the non-detections, the rate of spurious non-detections, and the censored magnitudes were all taken into account in the *de-censoring* procedure which was used to calculate the mean magnitudes of the stars in the Tycho catalogue. Later a new reduction was performed on the basis of the photon counts and the censoring phenomena were avoided. The Tycho-2 catalogue was thus obtained (Høg et al. 2000a). The Tycho-2 magnitudes confirmed the overall validity of the original Tycho magnitudes, and, therefore, the validity of the de-censoring procedure.

In this paper the model used in the de-censoring was improved and it is used to review the problem of the detection of variable stars. A suitable statistical test was constructed, and applied to a selection of 752 115 stars fainter than  $V_T = 8$  mag and with good enough measurements. The choice of the selection criteria is described in Sect. 2. The statistical test is explained in Sect. 3. Section 4 presents our sample of suspected variable stars.

## 2. Preparation of the input data

The stars in the Tycho catalogue are not all equally accurate in their photometric measurements. Apart from the intrinsic magnitude of the stars, it depends on the ability of the reduction process to attribute one transit to one object and to derive magnitudes. We cannot expect to find relevant variable stars if we apply this detection method to stars whose measurements are contaminated, and give too many false alarms. Stars are then selected for processing; we retain those with reliably accurate photometry, and dubious transits are rejected, as explained hereafter.

### 2.1. Selection of the stars

#### 2.1.1. Selection of the stars with relevant photometric observations

The Tycho catalogue contains 1 009 938 stars fainter than  $V = 8$  mag, some of which have poor quality photometry and must be discarded. The catalogue provides some flags to describe the data; those allowing selection of stars with reliable photometry are summarized here.

- The T10 flag is put to “X” when the astrometric parameters are dubious. This may be due to a bad photometric quality, to a possible duplicity leading to a contamination of the measurements, or to a genuinely dubious astrometry. In this case, some measurements may be contaminated or may even belong to the other star; 159 198 stars having  $T10 = X$  have been rejected for this reason;
- The T36 flag indicates whether the given magnitude is a median value whether it arises from de-censoring, or whether it is an approximate magnitude. An approximate magnitude is provided only when the photometric data are too corrupted to derive a reliable one. The star was only retained when  $B_T$  and  $V_T$  could have been derived either by median value or by de-censored mean value arising from the de-censoring analysis ( $T36 = M$  or  $N$ ). This led to the rejection of 887 additional stars.

In this way all stars with identified probable photometric pollution were rejected.

Moreover, we used the part of the TPOC produced during the main processing of the Tycho data (see Høg 1997 and Høg et al. 2000b). This file differs from the Tycho Epoch Photometry Annexes as it contains more information about the reliability of the transit data. However, the only copy we obtained was partly illegible and some GSC zones were missing in the South hemisphere. These zones are gathered into four sets with the following numbers: 7523–7842, 8563–8642, 8883–9042 and 9283–9363. They contain 10.5% of the stars in the Tycho catalogue.

#### 2.1.2. Choice of the Tycho-2 mean magnitudes

To determine the variability of a star using a statistical test implies an assumption that the star is constant with

a given magnitude: This calls for an accurate estimation of the magnitude of the star. In practice, the magnitudes of the Tycho stars may be taken from the original Tycho catalogue or from the Tycho-2 catalogue already mentioned in Sect. 1.3.

The difference between taking Tycho-1 or Tycho-2 magnitudes as mean values has been tested on two samples, using the preliminary version of the test as described in Sect. 3.6. The first sample contains 326 Hipparcos variable stars having  $V > 9.5$  mag and amplitude larger than 0.06 mag. The second contains 1964 Hipparcos constant stars having  $V > 9.5$  mag. Using Tycho-1 magnitudes, 7.3% of variable stars were recovered and 1.1% of constant stars were wrongly detected as variables. Using the Tycho-2 magnitudes, 9.8% of variable stars were recovered when 0.1% of constant stars were wrongly detected as variables. The Tycho-2 mean magnitudes were chosen as they increase the detectability of variable stars and decrease the false alarms rate. We kept the remaining stars only when Tycho-2 magnitudes were available. This led to our sub-sample of 752 115 stars.

## 2.2. Selection of the photometric observations

As mentioned above, we prefer the TPOC to the Tycho Epoch Photometry Annexes, since it contains more information about the reliability of the transits. Moreover all transits close to the predicted crossing of one star are given with the star, and we must select the true detection of the star, if any. We examine the closest transit to each predicted position within  $1''$  of that predicted position (Høg in ESAb 1997). It is only selected if no transit is recorded for an interval of  $1''$  on both sides of the observed position, in order to avoid pollution by the afterglow from the photo-multiplier. All transits showing the following are rejected:

- cumulative background in the two bands larger than 20 counts by sample;
- no background estimate;
- height of the crossing on the slit system larger than  $\sim 20'$  ( $=1195''$ ) or closer than  $0.3'$  ( $=18''$ ) to the middle of the slit, to avoid misuse of the calibrating terms;
- magnitude indicated by a flag as having a bad goodness-of-fit after fitting the signal with the expected response of the slits groups<sup>1</sup>;
- indication that the transit belongs to more than one predicted group crossing, as it indicates that the transit may not belong to the star or may be disturbed<sup>1</sup>;
- jet firing during the record; this happens when adjusting the satellite scanning law and might affect the estimated signal amplitudes (Großmann 1997)<sup>1</sup>;
- updated transit velocity not consistent with the transit velocity interval used in detection<sup>1</sup>;
- parasite or background raised by bright star<sup>1</sup>.

<sup>1</sup> In the annexes these items are either not completely reported or absent.

Contrary to the calculation of the mean magnitudes with the de-censoring processing, the variability search is made solely on the transits containing valid detections. All the empty crossing predictions are thus rejected, because we cannot distinguish between the censored detections and the “spurious non-detections”, as explained in Sect. 1.3; assuming a constant rate of spurious non-detections would lead to an abnormally large rate of false alarms in the variability search. We prefer therefore to ignore all non-detections, and we only take account of the effect of the detection process on the statistical properties of the detected transits. As a matter of semantics, we will then consider transits affected by *truncation* rather than by *censoring*.

## 2.3. Correction and censoring of the magnitude measurements

When a transit is selected the magnitude measured in each channel is corrected with the routine already mentioned in Sect. 1.3. The SNR of the corrected magnitude is calculated afterwards; first the total photon count,  $C_M$ , is derived from the corrected magnitude,  $M_{\text{meas}}$ , and from the mean background  $b_M$ , expressed in count per sample:

$$C_M = \left( 10^{\frac{M_{\text{cal}} - M_{\text{meas}}}{2.5}} + b_M \right) / 0.15 \quad (1)$$

$M_{\text{cal}}$  is a calibration constant given in Table 8.1 of Großmann (Großmann 1997), and the coefficient 0.15 is used to translate the mean amplitude in total count, since a transit is recorded within  $1/0.15$  samples (Bässgen et al. 1992; Halbwachs 1997). The SNR of the measurement is then:

$$\text{SNR} = \frac{C_M - \frac{b_M}{0.15}}{\sqrt{C_M}}. \quad (2)$$

As in the calculation of the mean magnitudes of the Tycho catalogue, the magnitude measurement is censored when SNR is less than 0.5. Then, we just know that  $C_M$  is so small that  $\text{SNR} < 0.5$ .

## 2.4. Merging the measurements in multitransits

As it is reasonable to assume that the star does not vary between the crossings of the two windows, i.e. within twenty minutes, the four successive transits (two slits systems  $\times$  two windows) are gathered in a so-called *multitransit*. This decreases the noise in the data and also provides an opportunity to discard a few more disturbed transits: The transits are compared with each other before merging, and the mean intensity of three transits and the standard deviation  $\sigma$  are derived. If the intensity of the remaining transit is  $5\sigma$  or more higher than the mean of the other three, it is rejected on suspicion of parasites or spikes. If fewer than four transits are available the comparison is made in the same way for three or two transits.

The transits which are kept in the multitransit are then merged. In order to prepare the search of variability, we only need the total counts in the  $B_T$  and in the  $V_T$

channels. For that purpose, we total the counts derived in Eq. (1) for each transit. When  $M_{\text{meas}}$  is censored,  $C_M$  is replaced with its mean value  $\langle C_M \rangle$ , which is derived as in the de-censoring procedure: The statistical law of the photon count is derived, as explained in Sect. 3.1 hereafter, and  $\langle C_M \rangle$ , is calculated, taking account of the censoring condition  $\text{SNR} < 0.5$ . The photon counts of the multitransit are thus finally obtained for both channels. In order to avoid confusion with single transits, they are noted  $\widetilde{C}_{B_T}$  and  $\widetilde{C}_{V_T}$  hereafter.

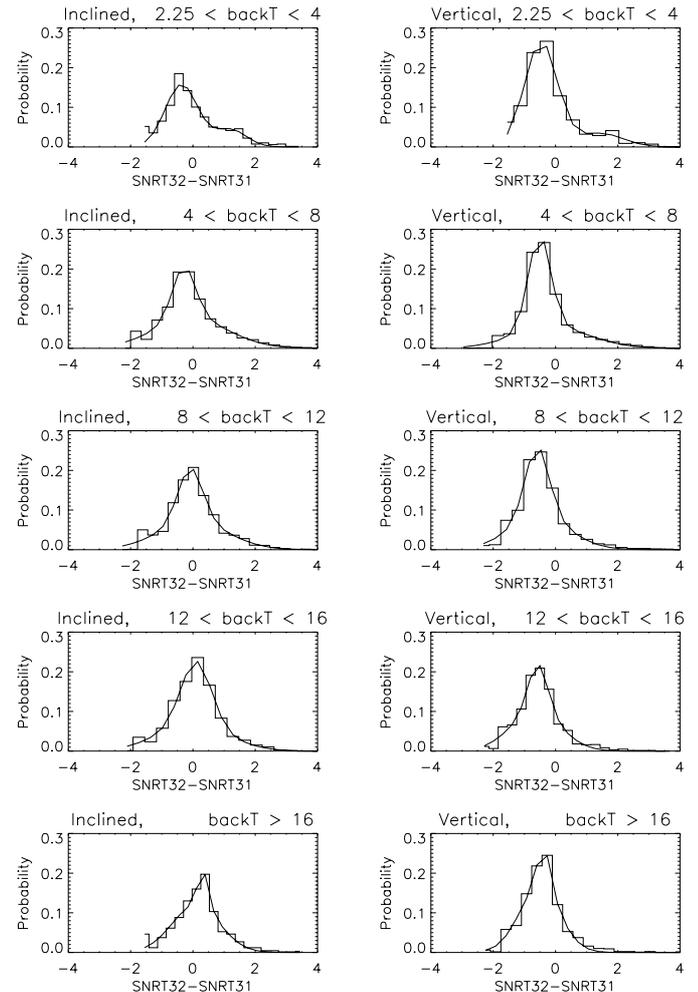
### 2.5. The detection threshold reviewed

As we saw in Sect. 1.3, when the SNR was less than about 1.5 in the “T” channel the transit was rejected. The distribution of the actual values of this detection threshold is important as we are seeking individual measurements near the Tycho detection limit. In reality the detection selection was based on a preliminary SNR estimation and the measurements used thereafter were estimated later by a more accurate process. If the former SNR estimate was not as close to 1.5 as we expected, the effect of the magnitude censoring will differ. In order to derive the statistical properties of the difference between the former SNR and the latter one, we select all transits with a preliminary estimation of SNR between 1.50 and 1.55<sup>2</sup>. The signal is corrected according to slit system and background value (Sect. 1.3), and the true SNR is calculated. The difference between the true SNR and the 1.5 value is then derived for each slit system and for different classes of background. We see that the detection threshold fluctuated around 1.5 with different Gauss laws, depending on the background and the slit system, with a standard deviation which may reach 1.4 (Fig. 3). Consequently the significance of the detection threshold must be taken into account statistically, as detection may occur with very faint SNR, but it may also be censored with SNR higher than 1.5.

### 3. Variability detection

For each multitransit, the counts in the channels  $B_T$  and  $V_T$  are available for the variability search. However, the count in the “T” band, which is simply  $\widetilde{C}_T = \widetilde{C}_{B_T} + \widetilde{C}_{V_T}$ , is a better parameter than  $\widetilde{C}_{B_T}$  and  $\widetilde{C}_{V_T}$  considered separately, the reason being that the standard deviation of the count is then decreased by  $\frac{\sqrt{\widetilde{C}_{B_T}} + \sqrt{\widetilde{C}_{V_T}}}{\sqrt{\widetilde{C}_{B_T} + \widetilde{C}_{V_T}}}$ . However, the “T” band is rather wide, and we expect a few variable stars are more likely to arise from fluctuations of the  $(B_T - V_T)$  colour index; this concerns stars such as semi-regular variables which have small amplitudes in bolometric magnitude, but which present important fluctuations of absorption lines in their spectra, making them vary in colour. The variability in  $(B_T - V_T)$  is therefore investigated too, by considering the ratio  $\widetilde{C}_{B_T}/\widetilde{C}_{V_T}$ . However,

<sup>2</sup> The preliminary SNR is not reported in the annexes.



**Fig. 3.** Probability distribution of the difference between the true SNR (SNRT32) and the SNR used in the detection process (SNRT31) when the latter is between 1.5 and 1.55, i.e. just above the detection threshold. The left-hand column refers to the transits of the inclined slit group, and the right-hand column to the vertical slit group. 5 ranges of background in the “T” channel are considered, from 2.25 to 20 counts per sample. The fits were used as statistical laws of the actual detection threshold.

the error is then much larger than for  $\widetilde{C}_T$ , and a poor efficiency is expected for this complementary search.

#### 3.1. The statistical laws of $\widetilde{C}_T$ and $\widetilde{C}_{B_T}/\widetilde{C}_{V_T}$

A mean number of photon counts  $\Lambda_M$  is expected for each given magnitude to be observed during a single transit. This mean number depends primarily on the magnitude of the star, which is assumed to be the Tycho-2 magnitude,  $M_{\text{TYC2}}$ , but also on the background and on the slit system on which it is recorded. For a given channel,  $\Lambda_M$  is given by the equation:

$$\Lambda_M = \left( 10^{\frac{M_{\text{cal}} - M_{\text{TYC2}}}{2.5}} + b_M \right) / 0.15. \quad (3)$$

This equation looks similar to Eq. (1), but it refers to an expectation of the measurement instead of the actual observation. If our model is correct, for each crossing of a slit group, detected or not, the actual photon count  $C_M$  obeys the Poisson law with parameter  $\Lambda_M$ . The final law of the multitransit counts is derived by transforming this original Poisson law for all the operations carried out in the data processing. In practice it is calculated by simulation: Raw counts are generated from the Poisson law, and they are introduced in routines describing the treatment of the actual data.

We first must give our attention to the truncation arising from the detection of the transit. For detected transits, the count  $C_T = C_{B_T} + C_{V_T}$  is above the threshold corresponding to:  $\text{SNR}_T = 1.5$  plus a random correction arising from the distribution in Fig. 3.

The count in ‘‘T’’ of a multitransit is obtained by adding the  $C_T$  values of all the detected transits. However, a detected transit is discarded when its mean intensity is  $5\sigma$  above the others, as explained in Sect. 2.4. However, when  $\text{SNR}_{B_T}$  or  $\text{SNR}_{V_T}$  is less than 0.5, the corresponding count  $C_{B_T}$  or  $C_{V_T}$  is replaced with the mean value from the censoring condition.

The statistical laws of the count  $\widetilde{C}_T$  and of the ratio  $\widetilde{C}_{B_T}/\widetilde{C}_{V_T}$  of the multitransit are the products of all these operations. It is worth noting that, apart from the magnitudes of the stars, these laws depend on parameters which are themselves variable. The background was dramatically rising when the satellite was crossing through the van Allen belts. Moreover background variations also occurred around a star when the other field of view was close to the ecliptic plane (zodiacal light) or pointed toward the Milky Way (Wicenc & van Leeuwen 1995). Finally, the calibration terms presented small variations during the mission. Consequently, the law of the photon counts is different from one multitransit to another.

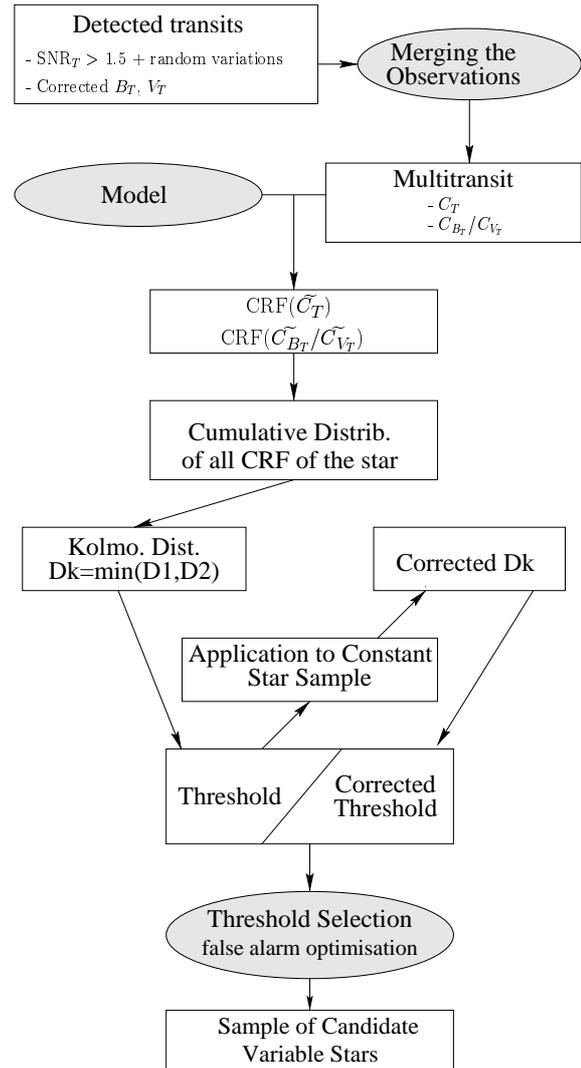
### 3.2. Choice of the method of Kolmogorov-Smirnov

The standard procedure for detecting variability in light-curves is to treat the distribution of the magnitude measurements, assuming an error law. With on-ground photometry, the error law of magnitudes follows a Gauss law, so that standard statistical tests may be used. This method could not be applied to Tycho measurements since the law of the observations is rather complicated and even variable, as shown in Sect. 3.1 above.

We decided to follow the approach of the Kolmogorov-Smirnov (KS) test (see Peacock 1983; Papoulis 1984). The KS test is based on the difference between the cumulative distribution of the observations, and that calculated from a model. It may be applied to any continuous statistical law. Hereafter, in Sect. 3.3, we use our own model to derive the Cumulative Relative Frequency of each of the actual measurements (hereafter CRF). The CRF will be used in place of the measurements themselves, and the test will be based on the cumulative distribution of the CRF. The

advantage of this approach is that the distribution of a CRF is simply the constant law, whatever the model used to derive the CRF.

In spite of its complexity, our model is not perfect and our results must be revised, as explained in Sect. 2.5. Afterward the selection threshold is optimised (Sect. 3.8), and the final selection of candidate variable stars is obtained. All the selection steps are summarized in Fig. 4, from the TPOC to the list of candidate variable stars.



**Fig. 4.** Organisation of the selection of candidate variable stars.

### 3.3. Calculation of the CRF of a multitransit

The CRF of the  $T$  magnitude of a multitransit is the probability that  $T$  is as faint as or fainter than the magnitude actually obtained. This is also the probability that  $\widetilde{C}_T$  is equal to or smaller than the sum of  $\widetilde{C}_{B_T}$  and  $\widetilde{C}_{V_T}$  which were actually obtained in Sect. 2.4. The calculation is done

by simulation: virtual transits are generated with the same observation conditions as the actual ones, as explained in Sect. 3.1; the proportion of virtual multitransits with  $\widetilde{C}_T$  smaller than the actual value is thus derived:

$$F(\widetilde{C}_T) = P(\widetilde{C}_{T\text{sim}} \leq \widetilde{C}_{T\text{obs}}). \quad (4)$$

At the same time, the CRF of the colour index is computed by comparing  $\widetilde{C}_{B_T}/\widetilde{C}_{V_T}$  to its actual value:

$$F\left(\left(\frac{\widetilde{C}_{B_T}}{\widetilde{C}_{V_T}}\right)\right) = P\left(\left(\frac{\widetilde{C}_{B_T}}{\widetilde{C}_{V_T}}\right)_{\text{sim}} \leq \left(\frac{\widetilde{C}_{B_T}}{\widetilde{C}_{V_T}}\right)_{\text{obs}}\right). \quad (5)$$

The observations of each multitransit are therefore translated in two CRF: one for the measurement in the “T”-band, and one for the colour index.

### 3.4. Kolmogorov distances

The standard version of the KS test is based on the uniform convergence theorem between two cumulative distributions (Darling 1957; Papoulis 1984), which states:

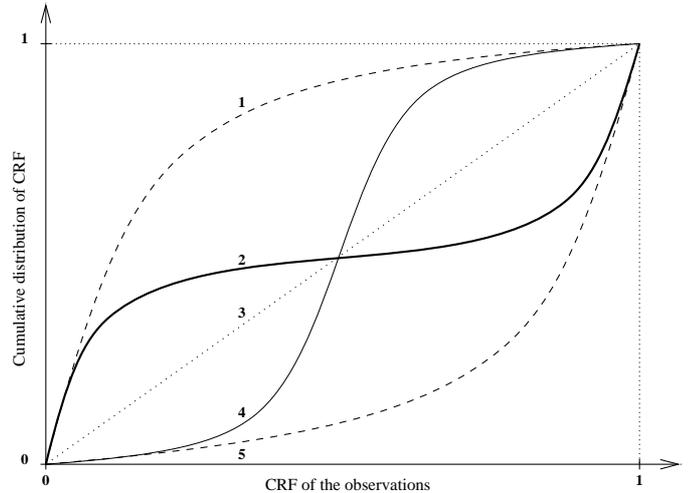
$$\forall \epsilon > 0, \exists N : [n \geq N] \Rightarrow \max_x |G_n(x) - G(x)| \leq \epsilon. \quad (6)$$

In other words, if the model is correct, the cumulative distribution of the observations,  $G_n$ , should not spread more than a limited distance  $D_{\text{KS}} = \max_x |G_n(x) - G(x)|$  from the theoretical one,  $G$ . The universality of the KS test arises from the fact that the distribution of  $D_{\text{KS}}$  does not depend on the model tested.

If the data acquisition model is correct and if the star is constant with the assumed magnitude, the CRF obtained in Eqs. (4) and (5) must follow a uniform law. The cumulative distribution of these CRF,  $G(F)$  is therefore compared with the cumulative distribution of the uniform law, using a Kolmogorov-Smirnov approach. Hereafter “CRF” refers only to the CRF of the observations, as derived from our model. To avoid confusion, the cumulative relative frequency of these CRF is called “cumulative distribution”.

In order to compare the cumulative distributions, we must adapt the method of determining our Kolmogorov’s distance calculation. We are looking for an unexpectedly wide difference between the two distributions, corresponding both to an excess of excessively bright measurements, i.e. with a low probability of being as bright as those observed, and an excess of extremely faint measurements, i.e. with a high probability of being as bright as those observed. We are not concerned with a mean difference. However, if there is a slight difference between the assumed mean magnitude and the actual one, the observed distribution will be wholly on or under the theoretical distribution, since the majority of the measurements will have an excess of low or high probability. These curves usually correspond to false alarms and are not relevant.

On the other hand, if the actual standard deviation of the observed distribution is lower than expected, the majority of the measurements will be too close to the assumed mean magnitudes, and the distribution of CRF will



**Fig. 5.** Cumulative distributions of CRF. A CRF close to 0 corresponds to a faint measurement (with low probability of being fainter than was observed), and a CRF close to 1 corresponds to a bright measurement (with high probability to be fainter than was observed). Lines (1) and (5) show the effect of an error in the mean magnitude, (1): the actual magnitude is fainter than assumed, (5): the actual magnitude is brighter than assumed. Line (3) shows the expected distribution when the magnitudes and the distribution of errors are as in the model. Line (2) shows the “N” template presenting excess of spread in the CRF distribution possibly due to variability, and line (4) shows the “S” template of a distribution less spread than expected.

show an excess of measurements which have a computed brightness probability close to 0.5. The observed distribution then will look like an “S” template, as shown in Fig. 5. This means that the star is probably constant and these cases should not be detected as variable.

In summary, only the “N” template curves correspond to significant variability in the data. Consequently, we must turn our null hypothesis into: “The cumulative distribution of CRF of the measurements does not spread more than a limited distance of the uniform distribution according to a “N” form template”. We define two Kolmogorov-like distances: the greater distance on the right side under the uniform distribution  $D_2$ , and the greater distance on the left side upon the uniform distribution  $D_1$ , with  $D_2$  to the left from  $D_1$ , as in Fig. 6:

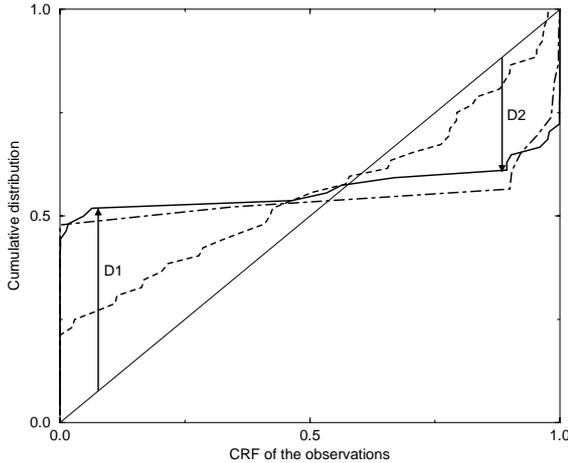
$$D_1 = \max(G(F) - G_U) \quad (7)$$

$$D_2 = \max(G_U - G(F)) \quad (8)$$

where  $G_U$  indicates the cumulative distribution for the uniform law. The distances are thus defined for the  $T$ -test as for the  $(B_T - V_T)$  test. We must stress that if the assumed and actual mean magnitudes differ slightly, this will generate a dissymmetry in the “N” template, favoring the great Kolmogorov’s distances. In order to avoid false alarms, we considered:

$$D = \min(D_1, D_2). \quad (9)$$

We verified that for the same theoretical false alarm rate, the minimum distance gives more variable stars and fewer



**Fig. 6.** Examples of cumulative distributions of CRF: the thick line shows the cumulative distribution for the known cepheid BG Lac (Tyc 3210-1636-1), the thin line shows the cumulative distribution for an uniform law, the dashed line represents an Hipparcos constant star (Tyc 2153-1005-1) and the dotted-dashed line shows the known eclipsing binary of Algol type X Tri (Tyc 1763-2733-1). The two distances  $D_1$  and  $D_2$  are shown (See text).

false alarms than the greater, using Hipparcos known variable stars of high amplitude and Hipparcos known constant stars. We looked at the behaviour of the “N” template curves for different variability types. Non-sinusoidal light-curves gave more or less symmetrical cumulative distributions, sometimes with several very faint CRF values, but remaining within the “N” template, as in Fig. 6.

### 3.5. Estimation of the rejection threshold

For each star, the rejection threshold of the model,  $t$ , is the probability that  $D$  will be even larger than the value actually obtained. A small  $t$  means that it was a priori unlikely for the star to have such a large  $D$  if the model is correct; it is then inferred that the model may be wrong, i.e. that the star could really be variable. However, when many stars are considered, some will have large  $D$  (and, consequently, small  $t$ ) just at random. Therefore, if we select all the stars having  $t$  smaller than a given threshold  $t_{\text{sel}}$  as candidate variables,  $t_{\text{sel}}$  is the proportion of stars which should be selected among those which are really constant. In other words,  $t_{\text{sel}}$  is the expected rate of false alarms, a parameter of major importance in any selection of particular objects.

For a given star,  $t$  depends on  $D$  and on the number of multitransits,  $n$ , but the dependence on  $n$  diminishes when  $D\sqrt{n}$  is used in place of  $D$ , especially when  $n$  is large. We therefore calculated tables providing  $t$  as a function of  $D\sqrt{n}$  for fixed values of  $n$  by simulation. The computation of each table was performed as follows: 1 000 000 sets of  $n$  virtual CRF were drawn from the constant law. For each set,  $D\sqrt{n}$  was derived as explained in Sect. 3.4 above. Finally,  $t$  was derived from the proportion of simulated values which exceeded  $D\sqrt{n}$ .

By interpolation between these tables,  $t$  may be derived for any values of  $n$  and  $D$ . We may now check the validity of the model: when a sample of constant stars is considered, the proportion of stars with  $t < t_{\text{sel}}$  should be exactly  $t_{\text{sel}}$ . In practice, we will see hereafter that the model is far from perfect, and the threshold must be revised, as explained in Sect. 3.7.

### 3.6. Preliminary results

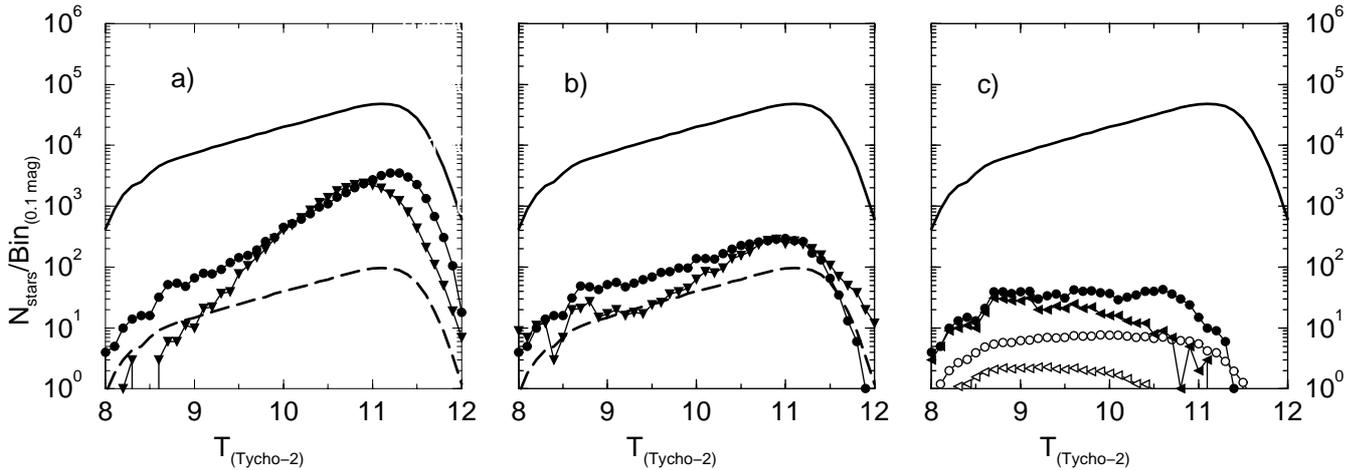
We applied the test above on the sample of 752 115 Tycho stars, with the selection threshold  $t_{\text{sel}} = 0.2\%$ . A preliminary sample of 12 731 candidate variable stars was thus obtained with the selection upon  $T$ , and 6 955 more candidate variable stars came with the selection upon  $(B_T - V_T)$ . The distribution of the candidates versus  $T$  is shown in Fig. 7a. It is noticeable that the rate rises abnormally with fainter  $T$  magnitudes. We also note that it rises for extreme  $(B_T - V_T)$ . This indicates that the estimated counts used in the calculation are more widespread than assumed with the Poisson law. To counter this effect, an a posteriori correction was applied to the distances  $D$  in order to obtain the right rejection thresholds.

### 3.7. The rejection threshold re-visited

Since our estimations of  $t$  seemed unreliable, we needed a sample of constant stars with every magnitudes in order to compare  $t$  to the actual rate of false alarms. This sample was used to find a way to obtain corrected values of  $t$  (Sect. 3.7.2), which were used to fix the selection criterion of the candidate variable stars (Sect. 3.8).

#### 3.7.1. The sample of known constant stars

During the Hipparcos reduction, the stars showing a constant light curve with dispersion smaller than the expected standard deviation were quoted as constant stars (micro-variability cannot be excluded, but, anyway, Tycho measurements are too inaccurate for being then altered). Consequently, from  $V = 8$  to  $V = 12$  mag, we have a sample of 27 055 Hipparcos constant stars. However, Hipparcos stars are usually quite bright, and we found very few faint stars among them. Fortunately, the Bordeaux CCD meridian circle, operating since January 1997, provides several measurements per star up to  $V = 17$ , between declination  $11^\circ$  and  $18^\circ$ , as part of the “Méridien 2000” program (Colin et al. 1998). A sample of stars with at least 6 measurements whose standard error was less than 0.03 mag were considered as constant stars. This led to a sample of  $\sim 4500$  stars from  $V = 7.8$  to  $V = 12.7$ , which was used to complete the Hipparcos constant star sample.



**Fig. 7.** Distribution of candidate variable stars using different selection criteria. **a)** Distribution of the candidate variable stars of the preliminary selection with  $t_{\text{sel}} = 0.2\%$ . The heavy line is the total number of stars, the dashed line is the expected number of false alarms, the lined circles show the candidate variable stars detected in  $T$  magnitude, and the lined triangles the candidate variable stars detected with  $(B_T - V_T)$ . **b)** Distribution of candidate variable stars after distance correction, with  $t_{\text{sel}} = 0.2\%$ . Same symbols as in **a)**. **c)** Distribution of candidate  $T$ -variable stars, selected on the basis of constant false candidate ratios. The heavy line indicates the total number of Tycho stars; the lines with full symbols show the numbers of selected stars with false candidate ratios of 20% (circles) and 10% (triangles); the lines with empty symbols indicate the corresponding false alarm numbers.

### 3.7.2. Correction of the distance $D$

The sample of constant stars is used to check if  $t$  is really the probability that a constant star will look variable. As inferred from the faint part of Fig. 7a, the proportion of false candidate variables with  $t < t_{\text{sel}}$  is always larger than  $t_{\text{sel}}$ . In order to straighten  $t$ , a correction of  $D$  is sought. It appears that reliable values of  $t$  are obtained when  $D$  is corrected with the formulae:

$$D_{T \text{ corr}} = \left( \sum_{i,j=0}^1 \alpha_{ij} \times T^i \times (B_T - V_T)^j \right) \times D_T \quad (10)$$

for the test on  $T$ , and

$$D_{(B_T - V_T) \text{ corr}} = \left( \sum_{i,j=0}^2 \alpha_{ij} \times T^i \times (B_T - V_T)^j \right) \times D_{(B_T - V_T)} \quad (11)$$

for the test on  $(B_T - V_T)$ . The  $\alpha_{ij}$  coefficients are taken from Table 1.

When  $t$  is derived from the corrected distances, the proportion of constant stars with  $t < t_{\text{sel}}$  is exactly  $t_{\text{sel}}$ .

**Table 1.** Coefficients  $\alpha_{ij}$  for corrected Kolmogorov distances in Eqs. (10) and (11).

		$j$		$j$		
		0	1	0	1	2
$i$	0	1.86	0.039	14.62	-10.11	5.68
	1	0.104	-0.004	-2.62	2.04	-1.16
	2	-	-	0.123	-0.10	0.06

We verified that this is true for any value of  $t_{\text{sel}}$ , and also for any value of  $n$ .

After correction of  $D$ , candidate variable stars with  $t < t_{\text{sel}} = 0.2\%$  are selected for the whole sample of Tycho stars. The distribution of the stars is in Fig. 7b. The rate of candidate variables is now almost constant until  $T \approx 11$  mag, but it decreases after.

### 3.8. The false alarm optimisation

The correction above leads to a false alarm rate which is related to the total number of constant stars. Since we can detect only variables with large amplitudes, the proportion of variables is rather small in our sample, and the number of false alarms is approximately  $t_{\text{sel}} \times N$ , where  $N$  is the number of stars in the tested sample. However, the number of candidate variables is  $n(t < t_{\text{sel}})$ , which is the sum of the false alarms and of the true variable stars. Therefore, the efficiency of our selection is not evaluated by  $t_{\text{sel}}$  alone, but by the ratio of false candidates among the selection:

$$r_{\text{fc}} = \frac{n(t < t_{\text{sel}})}{t_{\text{sel}} \times N}. \quad (12)$$

We want a selection with a constant – and hopefully small –  $r_{\text{fc}}$ , whatever the magnitude of the stars. This is obtained by making  $t_{\text{sel}}$  a function of  $T$ .

The selection was revised using several values of  $r_{\text{fc}}$ . The distributions of “ $T$ ” of the candidate variable stars are plotted in Fig. 7c for  $r_{\text{fc}} = 20$  and 10%. It appears in this figure that increasing  $r_{\text{fc}}$  from 10 to 20% adds only a few more stars brighter than 9 mag; therefore, these additional bright stars include a large proportion of false candidates, since  $r_{\text{fc}} = 10\%$  is sufficient to select the large amplitude

variables. At the opposite end, the number of stars increases a lot in range  $T \in ]9, 11]$  when  $r_{fc}$  is changed from 10 to 20%. That means that the faint variables with large amplitudes are then close to the detection limit. Since we want to select the faint variable stars, we made our final selection with the false candidate ratio  $r_{fc} = 20\%$  rather than 10%. This results in increasing the number of candidate variables from 487 to 954, but the number of actual variables should then be 763 instead of 438.

The selection thresholds for  $r_{fc} = 20\%$  are as follows :

$$t_{T \text{ sel}}(r_{fc} = 20\%) = 0.0134 - 0.0022 \times T + 9.33 \times 10^{-5} \times T^2 \quad (13)$$

and

$$t_{(B_T - V_T) \text{ sel}}(r_{fc} = 20\%) = 0.0103 - 0.0020 \times T + 9.96 \times 10^{-5} \times T^2. \quad (14)$$

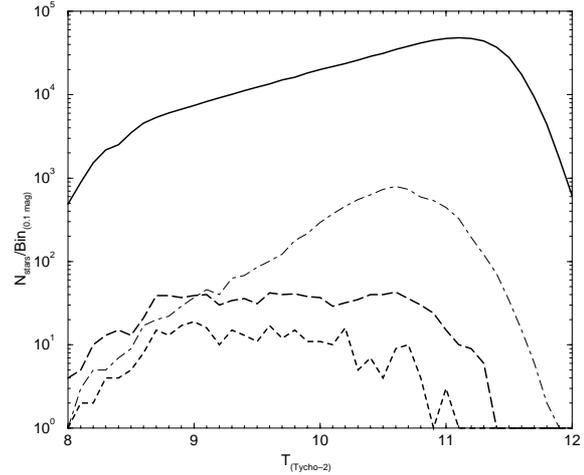
However, these thresholds were derived from fits which are valid only up to  $T \approx 11$  mag: It is easy to see in Fig. 7c that the number of candidates stars converges with the number of false candidates in the range  $T \in ]11, 11.5]$ . Improving the fits would be useless, however, since the range  $T > 11$  mag would then contain very few candidates, as the Tycho observations are far too inaccurate for variability detection among so faint stars. It is noteworthy that the limit of our selection coincides with the completeness limit of the Tycho catalogue.

## 4. Results

### 4.1. The final selection of candidate variables

With a false candidate ratio of 20% in  $T$ , 954 candidates were selected, including 191 false alarms; 374 stars are already known to be variable in Hipparcos, GCVS or NSV. Among the remaining 580 suspected variable stars, 22 are flagged as constant by Hipparcos, one is found to be variable by Hipparcos but with an amplitude smaller than 0.06 mag, 39 are not detected as variable by Hipparcos, and one is a constant from the Bordeaux CCD meridian circle. Five hundred and fifty-eight stars are likely to be identified in an incoming paper; the number of false alarms among them should be 128, since 63 stars were already pointed out by Hipparcos, leaving us with about 390 true new variable stars.

With a false candidate rate of 20% in  $(B_T - V_T)$ , a further 141 candidates were selected, 4 of which were already in the “T”-selection above. Merging both selections, we obtained a list of 1091 candidate variable stars, containing 379 already known in Hipparcos, GCVS or NSV. The 712 remaining candidates contain 28 Hipparcos constant stars, two of which were found to be variable by Hipparcos but with an amplitude smaller than 0.06 mag; 49 are not identified as variables by Hipparcos – again, one is a constant from the Bordeaux CCD meridian circle. Finally, we have a sample of 634 suspected new variable stars; the number of false alarms among them should be 138, since



**Fig. 8.** The new selection of variable stars compared with that in the Tycho catalogue. The thin dotted-dashed line gives the stars flagged as “strong evidence of variability” in the Tycho catalogue. The heavy long-dashed is the distribution of all the candidate variable stars in our sample, and in heavy short-dashed line the stars flagged as variables and recovered in our sample.

80 were already identified in Hipparcos. About 496 true new variable stars are expected.

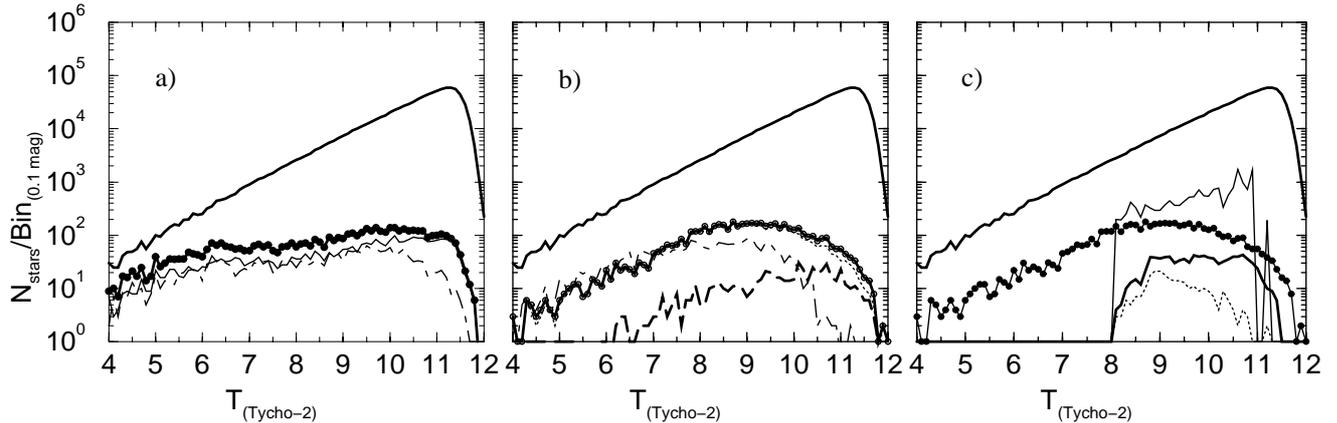
The intersection between our selection and the “variability flag” in the Tycho catalogue is shown in Fig. 8. Only 30% of our selection showed “strong evidence of variability”, and the two selections look very different for  $T > 9$  mag.

The distribution of variable stars previously known in GCVS or NSV catalogues is shown in Fig. 9a. It is noticeable that the fainter the magnitude, the lower the proportion of stars already known as variable.

### 4.2. Completeness of variability detection

The distributions of magnitudes of variable stars previously known in the Hipparcos catalogue are displayed in Fig. 9b, differentiating populations with the variation amplitude. It seems that the Hipparcos sample of variable stars with amplitude above 0.06 mag is complete for stars fainter than  $T = 8$  mag. This is due to the selection of the stars from the Hipparcos programme, which is complete up to  $V = 7.3-9$  mag, (the limit depends on the galactic latitude), but also to the fact that Hipparcos was fully efficient to detect variability with amplitude above 0.06 mag among stars brighter than 8 mag. The proportion of variables with amplitude larger than 0.06 mag, as derived from the Hipparcos stars with  $T$  between 4 and 8 mag is 5%. This proportion is to be compared with that among stars between 8 and 11 mag, which is derived hereafter.

The efficiency of our variability search may be estimated from the proportion of Hipparcos variable stars with amplitude larger than 0.06 magnitude which were recovered in our sample. The corrected distribution of variables was thus calculated (Fig. 9c). It appears that,



**Fig. 9.** Distribution of known variable stars in the Tycho Catalogue. The heavy line is the total number of Tycho stars in each window. **a)** The known variable stars in the Tycho Catalogue. The thin line shows the stars flagged as known in the GCVS, the dotted-dashed thin line corresponds to the NSV flag, and the circled heavy line shows the sum. **b)** Hipparcos variable stars distribution in the Tycho catalogue. The dashed line is the distribution of faint amplitude ( $< 0.06$  mag) Hipparcos variable stars. The dotted line gives the variable stars having an amplitude between 0.06 and 0.6 mag, and amplitudes larger than 0.6 mag are in the heavy dashed line. The circled line is for all amplitudes larger than 0.06 mag. **c)** Variable stars distribution in the Tycho catalogue. Our sample of candidate variable stars is in heavy line, and the circled line gives the Hipparcos known variables with amplitude larger than 0.06 mag. The dotted line shows the Hipparcos variables which were recovered in our selection, and the thin line is the distribution of variable stars extrapolated from the efficiency of our search.

for  $T$  between 8 and 11 mag, the proportion of variables is 5%, the same as among the stars brighter than 8 mag. This means that some 23 700 stars of the  $\sim 473\,800$  Tycho stars brighter than  $T = 11$  mag should be variable. Five thousand three hundred and thirty variable stars are already known to us from Hipparcos and this search should finally add 496 stars. There are still some 17 800 stars with amplitudes larger than 0.06 mag as yet undiscovered in the Tycho measurements.

Similarly, we estimate there are about 1500 stars with amplitudes larger than 0.6 mag. Some 550 of these are already quoted in Hipparcos, and a large number of our candidate variables are probably in this category of variability amplitude.

#### 4.3. Cross-identification with other catalogues of variable stars

None of the stars found by the ASAS project are common with our sample (Pojmański 1998), since most are in the missing zones of the Tycho Epoch Photometry Catalogue. Likewise, no stars found in the preliminary list of the ROTSE experiment were recovered, as most are too faint for Tycho.

Two stars in our sample are in the Woitas' list of Tycho variable stars, corrected by Kazarovets (Woitas 1997; Kazarovets 1999), and two more are in the list published during the Tycho reduction by Makarov et al. (1994). Five stars are common with the new Catalogue of  $\delta$  Scuti stars published by Rodríguez et al. (2000), and one is already mentioned in the Behlen Observatory Survey (Schmidt & Seth 1996). Koen & Schumann (1999) pointed out 70 stars using Tycho photometry, but 19 were flagged

**Table 2.** Stars cross identified with other catalogues.

<i>Tycho id.</i>			<i>other id.</i>			<i>Source</i>
2724	1872	1	<i>RZ</i>	<i>Peg</i>		Behlen
2549	677	1	2549	677		Makarov
1466	869	1	1466	869		Makarov
3038	566	1	<i>DU</i>	<i>Boo</i>		Kazarovets
1073	1391	1	20 15 53.6	+7 40 13.3		Kazarovets( <i>J2000</i> )
5495	421	1	5495	421		Koen
7079	334	1	7079	334		Koen
7850	1060	1	7850	1060		Koen
8442	1011	1	8442	1011		Koen
3600	2	1	<i>V1719</i>	<i>Cyg</i>		Rodríguez
3425	1038	1	<i>BE</i>	<i>Lyn</i>		Rodríguez
5926	1034	1	<i>RY</i>	<i>Lep</i>		Rodríguez
6638	779	1	<i>KZ</i>	<i>Hya</i>		Rodríguez
8428	217	1	<i>RS</i>	<i>Gru</i>		Rodríguez

with an "X" for dubious astrometry in the Tycho catalogue and are not treated in our search; 10 were brighter than our restricted magnitude region, 11 were in missing zones, and amongst the remaining 30, only four show a significant variability.

#### 4.4. Description of the candidate variable star catalogue

The 1091 resulting variable stars are available through the Vizier service of the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](ftp://cdsarc.u-strasbg.fr) (130.79.125.5) or via <http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/373/576>,

**Table 3.** Left-hand page of the table.

<i>Tycho - id.</i>			<i>Coordinates</i>		<i>Hip Type</i>		<i>N</i>	<i>B<sub>T</sub></i>	<i>V<sub>T</sub></i>
2549	677	1	213.99251657	34.43759537	69695	<i>I</i>	193	10.487	8.581
2554	1272	1	220.30577234	31.57213020	71802	<i>SRB</i>	104	9.542	8.174
2557	818	1	219.81605076	32.53952789	71644	<i>SRB</i>	200	9.832	8.378
2563	1338	1	230.34983708	31.36741257	75143	<i>M</i>	184	10.281	9.308
2583	823	1	245.14571892	34.62786865	80073		125	10.552	9.095

**Table 4.** Right-hand page of the table.

<i>Tycho - id.</i>			<i>t<sub>T</sub></i>	<i>t<sub>B<sub>T</sub>-V<sub>T</sub></sub></i>	<i>Am</i>	<i>H</i>	<i>G</i>	<i>V</i>	<i>D</i>	<i>An</i>	<i>name</i>	<i>Cross - id</i>
2549	677	1	0.0000000	0.7500000	2	<i>U</i>		<i>U</i>	<i>R</i>	<i>B</i>		<i>M</i>
2554	1272	1	0.0000000	0.7500000	3	<i>U</i>	<i>G</i>	<i>V</i>	<i>Y</i>	<i>A</i>	<i>RW Boo</i>	
2557	818	1	0.0000047	0.7500000	2	<i>U</i>	<i>G</i>	<i>V</i>	<i>Y</i>	<i>A</i>	<i>RV Boo</i>	
2563	1338	1	0.0000000	0.0010489	3	<i>P</i>	<i>G</i>	<i>V</i>	<i>Y</i>	<i>A</i>	<i>S CrB</i>	
2583	823	1	0.0000000	0.1351794	2	<i>U</i>		<i>V</i>	<i>Y</i>	<i>A</i>		

but a sample is provided in Tables 3 and 4. The content is as follows:

- The Tycho identifications, constructed with the GSC zone number, the number of the star in the zone, and the component number;
- The equatorial (ICRS) coordinates in degrees, for the epoch J1991.25, taken from the Tycho Catalogue;
- The Hipparcos number if any;
- Variability type identifications in Hipparcos if any (field P5 in the Variability Annex of the Hipparcos catalogue);
- Number of detected transits used in our variability search;
- The mean  $B_T$  magnitude from Tycho-2;
- The mean  $V_T$  magnitude from Tycho-2;
- $t_T$ : the probability for a constant star appearing as variable in the “T” channel. For stars with measurements less widely distributed than expected, the Kolmogorov distance is set to zero and the probability to an arbitrary value of 0.75;
- $t_{B_T-V_T}$ : idem for  $B_T - V_T$ ;
- $Am$  is the amplitude flag from Hipparcos (field H6). 1 means amplitude less than 0.06 magnitude, 2 is for amplitude between 0.06 and 0.6 magnitude, and 3 is for amplitude larger than 0.6 magnitude;
- The  $H$  flag notes the *Periodic* or *Unsolved* stars in the Hipparcos variability search (field H52);
- The  $G$  flag notes identification with GCVS variable stars, updated according to the General Catalogue of Variable Stars published in 1998;
- The  $V$  flag is the variability flag from the Tycho Catalogue (field T48).  $W$  notes weak indication of variability.  $U$  notes apparent variability probably due to duplicity and it is only set when the duplicity flag is set at  $R$  or  $S$ . The  $V$  flag is set for strong evidence of variability;
- The  $D$  flag is the duplicity flag from Tycho Catalogue (field T49).  $D$  denotes duplicity that is clearly indicated in the Tycho data,  $R$  for weak indication of duplicity,  $S$  for duplicity suspected but uncertain,  $Y$

- when no indication of duplicity was found, and  $Z$  when the star has not been investigated for duplicity;
- $An$  is the flag marking either Annex A or Annex B of the Tycho Epoch Photometry Annexes (field T50);
- GCVS name of the star if any;
- the cross-identification flag: the first name initial of the author of catalogue as given in Table 2.

A preliminary study of the sample already gave some new RR Lyræ,  $\delta$  Scuti, EB, EW, EA, SR (Andronov et al. 2000). A complete study will be published soon.

## 5. Conclusion

We present a new search of variable stars among the Tycho catalogue, which is devoted to the faint stars. In spite of difficulties arising from the complexity of the data acquisition process in the Tycho programme, we obtained a reliable selection of stars likely to be variable. Our research did not take account of the faintest Tycho stars, but our method is efficient up to the completeness limit of the Tycho catalogue.

Apart from several variable stars already identified, our list includes 496 variable stars previously undiscovered. This is about 7 times more than the selection of Koen & Schumann (1999). The variability types and the periods of these new variables will be derived in a forthcoming paper.

*Acknowledgements.* The authors would like to thank Photis Nobelis for fruitful discussions and Margaret Nelson for correcting the English. We thank the referee, Floor van Leeuwen, for his comments.

## References

- Akerlof, C., Amrose, S., Balsano, R., et al. 2000, *AJ*, 119, 1901  
 Alcock, C., Allsman, R. A., Alves, D. R., et al. 1999, *AJ*, 117, 920

- Andronov, I. L., Cuypers, J., & Piquard, S. 2000, in IAU Colloq. 176: The Impact of Large-Scale Surveys on Pulsating Star Research, ASP Conf. Ser., 203, 64
- Bässgen, G., Wicenc, A., Andreasen, G. K., et al. 1992, A&A, 258, L86
- Colin, J., Daigne, G., Ducourant, C., et al. 1998, in The Message of the Angles, ed. P. Brosche, W. Dick, O. Schwarz, & R. Wielen (Harri Deutsch, Frankfurt), vol. 3, 133
- Darling, D. 1957, *Annales of Mathematical Statistics*, 29, 823
- de Laverny, P., Beaulieu, J. P., Asplund, M., et al. 1998, A&A, 335, L93
- Egret, D. & Heck, A. 1999, PASP, 111, 123
- ESA SP-1200. 1997, The Hipparcos and Tycho Catalogues: Construction of the Tycho catalogue, vol. 4
- Eyer, L. & Grenon, M. 1997, ESA SP-402: Hipparcos – Venice '97, 402, 467
- Friedrich, S., König, M., & Wicenc, A. 1997, ESA SP-402: Hipparcos – Venice '97, 402, 441
- Großmann, V. 1997, in The Hipparcos and Tycho Catalogues: Construction of the Tycho catalogue, ed. ESA, vol. 4, ESA SP-1200, 123
- Großmann, V., Wicenc, A., Makarov, V., & Halbwachs, J.-L. 1997, in The Hipparcos and Tycho Catalogues: Construction of the Tycho catalogue, ed. ESA, vol. 4, ESA SP-1200, 267
- Halbwachs, J.-L. 1997, in The Hipparcos and Tycho Catalogues: Construction of the Tycho catalogue, ed. ESA, vol. 4, ESA SP-1200, 133
- Høg, E. 1997, in The Hipparcos and Tycho Catalogues: Construction of the Tycho catalogue, ed. ESA, vol. 4, ESA SP-1200, 15–30
- Høg, E., Fabricius, C., Makarov, V. V., et al. 2000a, A&A, 357, 367
- Høg, E., Fabricius, C., Makarov, V. V., et al. 2000b, A&A, 355, L27
- Kazarovets, E. V. 1999, *Information Bulletin on Variable Stars*, 4662
- Koen, C., & Schumann, R. 1999, MNRAS, 310, 618
- Maceroni, C., & Rucinski, S. M. 1999, AJ, 118, 1819
- Makarov, V., Bastian, U., Høg, E., Großmann, V., & Wicenc, A. 1994, *Information Bulletin on Variable Stars*, 4118
- Mauder, H., & Høg, E. 1987, A&A, 185, 349
- Papoulis, A. 1984, *Probability, random variables and stochastic processes* (New York: McGraw-Hill)
- Peacock, J. A. 1983, MNRAS, 202, 615
- Pojmański, G. 1997, *Acta Astron.*, 47, 467
- Pojmański, G. 1998, *Acta Astron.*, 48, 35
- Rodríguez, E., López-González, M. J., & López de Coca, P. 2000, A&AS, 144, 469
- Schmidt, E. G. & Seth, A. 1996, AJ, 112, 2769
- Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., & Mateo, M. 1992, *Acta Astron.*, 42, 253
- van Leeuwen, F. 1997, *Space Sci. Rev.*, 81, 201
- Wicenc, A., & van Leeuwen, F. 1995, A&A, 304, 160
- Woitak, J. 1997, *Information Bulletin on Variable Stars*, 4444