S. Friedrich<sup>1,2</sup>, M. König<sup>1</sup>, A. Wicenec<sup>1,3</sup>

<sup>1</sup>Institut für Astronomie und Astrophysik, Universität Tübingen, 72076 Tübingen, Germany <sup>2</sup>Institut für Astronomie und Astrophysik, Universität Kiel, 24098 Kiel, Germany <sup>3</sup>European Southern Observatory, 85748 Garching, Germany

# ABSTRACT

In order to automatically identify intrinsic variable stars in the Tycho Epoch Photometry Annex B (TEP/B) we have tried various methods. The simplest one was to look for stars with high standard deviations or scatter in their  $V_T$  magnitudes. As an alternative approach, we have used different statistical moments like skewness and kurtosis to derive reasonable estimates on non-Gaussian variability in the Tycho light-curves. We also applied an epoch folding method to emphasize any occuring periodic structure.

Key words: Variables; Photometry; Tycho.

# 1. INTRODUCTION

The Tycho Epoch Photometry Annex B (TEP/B) contains single measurements for about 490 000 stars of the Tycho catalogue. The total number of records containing two color photometry sums up to more than 6 million, making it the largest homogeneous data base of stellar photometry in the world. The exploitation of this data base seems to be worthwhile to find intrinsic variability of stars. Due to the number of objects and measurements only an *automatic approach* is feasible. The problem of finding *intrinsic* variability of stars is twofold. The first step is to identify possible variability (e.g. enhanced scatter). The second step is to verify whether this is due to intrinsic variability of the star or to other disturbances of the measurements or to limitations of the model used to derive the expectations. In order to separate this spurious variability and intrinsic variability we first concentrated on finding periods in the measurements of stars. In order to test various methods we have used a set of known periodic variable stars and a set of confirmed non-variable and non-double standard stars. The results of these tests are presented in this paper.

# 2. SEARCH CRITERIA

In order to identify possible variable stars in the Tycho data stream we have tried various methods. One was to look for stars where the standard deviations or scatter of their  $V_T$  magnitudes exceed an empirically derived threshold, i.e. an outlier criteria. As an alternative approach, we have used different statistical moments like skewness and kurtosis to derive reasonable estimates on non-Gaussian variability in the Tycho light-curves.

## 2.1. Outliers

This method was carried out relative to the Tycho photometric precision, i.e. stars are considered to be 'potentially variable' if the variance of their observations exceeded the expected variance. The variances of the observations of stars were derived from the 15th and 85th percentiles of the actual magnitude distribution for all accepted transits. The expected variance was calculated empirically by a fit to the distribution of photometric standard errors against the number of observations of all standard stars and more than 50 000 randomly selected stars. The fit was compared to a similar sample consisting of 10 000 stars flagged as variable and more than 6000 known variable stars. The function used is:

$$F_{\rm err} = 4.6 \times 1.75^m \times (N - 40)^{-0.71}$$

where  $F_{\rm err}$  is the expected upper limit of the standard error of the mean magnitude with  $V_T$  magnitude mand N observations. This function is valid for  $N \geq 50$ at  $V_T = 6.5$  mag and  $N \geq 70$  at  $V_T = 10$  mag. In Figure 1 we show standard errors and the scatter for a sample of standard and variable stars in the magnitude range between 5–10 mag. Also plotted (left panel) are the upper limits for  $F_{\rm err}$  (gray scaled solid lines) for the magnitudes indicated. As it can be seen in Figure 1 there is not much difference between the sample of standard stars and variable stars!

<sup>\*</sup>based on observations made with the ESA Hipparcos satellite.



Figure 1. Standard error for a sample of standard (top) and variable (bottom) stars in the magnitude range between 5-10 mag.

## 2.2. Skewness and Kurtosis

and kurtosis are plotted together with the variance:

To investigate non-Gaussian variability higher statistical moments as skewness:

$$S = \sum_{i}^{N} \left( \frac{\left( V_T(i) - \bar{V_T} \right)}{\sigma} \right)^3 \sqrt{\frac{N}{15}}$$

and kurtosis:

$$K = \sum_{i}^{N} \frac{\left(V_{T}(i) - \bar{V_{T}}\right)^{4}}{v^{2}} \frac{1}{N}$$

were applied to a sample of standard and variable stars. In Figures 2 and 3 (left panels) the skewness

$$v = \sum_{i=1}^{N} \frac{\left(V_T(i) - \bar{V_T}\right)^2}{(N-1)}$$

and the average deviation:

$$D = \sum_{i}^{N} \frac{|V_T(i) - \bar{V_T}|}{N}$$

N is the number of transits and  $\sigma$  the standard error of the  $V_T$  magnitude. The right panels in these figures show the same quantities but a Gaussian distributed error was assumed instead of the real error. For the standard stars the left and the right panel should be similar.



Figure 2. Variance, average deviation, skewness and kurtosis (upper to lower) for the sample of standard stars (left) compared to a Gaussian distributed error (right).

# 3. EPOCH FOLDING

We also tried to identify possibly variable stars with an epoch folding technique, see Chadwick et al. (1985). First the light curves were folded with trial periods and each pulse profile tested against the mean value of  $V_T$ . The height of peaks in the resulting  $\chi^2$ -diagram is a measure for the existence of a period. Only peaks in the  $\chi^2$  diagram which are significantly higher  $(5-7\sigma)$  than the mean value of  $V_T$  were accepted. Subsequently the vicinity of these peaks were investigated to decide whether they are real or coincidental. A subsample of  $\delta$  Cephei stars in the magnitude range between 2–9.5 mag and the period range between 1 and 25 days was investigated in more detail. To improve the  $V_T$  the mean value of a group of transits was used. In the above magnitude and period range all 127  $\delta$  Cephei stars were found but only 43 of them with the right period within an error margin of 10 per cent. A similar result was found for RR Lyrae stars in the magnitude range between 4.5-9.5 mag and the period range between 0.1 and 1 day. Applied to a subsample of standard stars also a large number of 'variable' stars was found but their  $\chi^2$  values are slightly lower than those of the true variable stars.



Figure 3. Variance, average deviation, skewness and kurtosis for the sample of variable stars.



Figure 4.  $\chi^2$  diagram and light-curve (not folded) of  $\delta$ Cephei.

# 4. EXPLANATION OF THE POOR RESULTS

The results shown in the figures above lead to the conclusion (see below) that none of these methods

444



Figure 5. (a) Period search diagram of  $\delta$  Cephei for the original light-curve. Also drawn are the mean  $\chi^2$  and its standard deviation expected from simulations  $(3.6 \pm 1.6)$ . (b) Stochastic permutation of the measured magnitudes of  $\delta$  Cephei.

are able to separate true periodicity from fake periods produced by the windowing of the Tycho observations and other disturbances of the single measurements. To prove this point, we have simulated Tycho time series by a stochastic permutation of the measured magnitudes of  $\delta$  Cephei. The sampling pattern (199 data points, 797 day observation length) was not changed. Figure 5 shows again the  $\chi^2$  diagram of  $\delta$  Cephei (5a) together with a  $\chi^2$  diagram of one stochastic permutation (5b). The comparison of the two shows only slight differences. One of the most prominent differences is around the real period of  $\delta$  Cephei (5.366 days) but there are also others (e.g. between 6.1 and 6.2 days) which are at least as significant as the latter one. The problem with Tycho light-curves is *not* the estimation of magnitudes but the miserable sampling. This results in trial period diagrams which are substantially dominated by the observational window function. Even in the case of strictly periodic sources (as the variable star  $\delta$  Cephei) the period behavior is totally smeared out.

It should be clearly noted here that TEP/B is a data base containing invaluable information of stellar variability, but the method to find this information needs to be carefully adjusted to the peculiarities of the Tycho measurements. To illustrate the value of the Tycho Epoch Photometry Annex B, Figure 6 shows the folded light-curve of  $\delta$  Cephei.

#### 5. CONCLUSIONS

We have applied various search algorithms on samples of standard and variable stars in order to find criteria to unequivocally identify variability in Tycho light-curves. None of the applied methods showed convincing results. Epoch folding turned out to be the best method at least to identify variable stars. But period estimates must be taken with great cau-



Figure 6. Folded light-curve of  $\delta$  Cephei in the  $V_T$  band. The error bars correspond to the 1 sigma errors of the single measurements.

tion, since the  $\chi^2$  diagram is strongly disturbed by aliases.

Due to the poor sampling of the Tycho light-curves, all search statistics are severely hampered which makes it very difficult to automatically identify variable stars. One way to overcome these difficulties might be to filter the  $\chi^2$  diagram with the  $\chi^2$  dia-gram of the windowing function of the measurements. Further studies on the behavior of the differences of the  $\chi^2$  diagram of the original time series and the  $\chi^2$ diagram of a permutated time series will be carried out in order to evaluate the potential of such tests.

### ACKNOWLEDGMENTS

This work has been supported in parts by the Deutsche Agentur für Raumfahrt Angelegenheiten (DARA) under grants number 50 OO 9106.

# REFERENCES

Chadwick, P.M., Dowthwaite, J.C., Harrison, A.B., 1985, Nature, 317, 236-238