UTILISATION OF HIPPARCOS DATA FOR DISTANCE DETERMINATIONS: ERROR, BIAS AND ESTIMATION

X. Luri[1,2], **F. Arenou**[2]

[1]Departament d'Astronomia i Meteorologia UB, Avda. Diagonal 647, E08028, Barcelona, Spain
[2]Observatoire de Paris-Meudon, DASGAL, F92195 Meudon CEDEX, France

ABSTRACT

The Hipparcos astrometric data will allow the determination of a great number of stellar distances. To make a good use of their high accuracy the data have to be used with care to obtain unbiased distance estimations.

We discuss in this paper the effects of the observational errors and observational selection criteria on the estimation of distances and we analyse them from a mathematical point of view. Examples are presented using Monte Carlo simulated samples.

Key words: barium stars; luminosity calibration; stellar kinematics.

1. INTRODUCTION

The goal of the poster presented in the Venice symposium under this title was to illustrate the problems associated with the estimation of distances using trigonometric parallaxes. These difficulties were encountered and highlighted in several presentations and the need for some guidelines on the use of the Hipparcos data was made clear. To cover this need, the inclusion of a 'rough guide' to the use of Hipparcos data in these proceedings was decided (see Brown et al. 1997). Consequently, the contents of this contribution have been slightly changed to complement that work.

2. EFFECTS OF THE OBSERVATIONAL ERRORS IN DISTANCE ESTIMATES

To estimate the distances of a sample of stars using only its Hipparcos trigonometrical parallax π_H, the properties of the distance estimate used have to be carefully taken into account.

A usual estimate of distance is $R = \frac{1}{\pi_H}$, but this estimate is a biased one. If we take, for instance, a star at 200 pc whose parallax is observed at the mean Hipparcos precision $\sigma_\pi = 1$ mas. The error distribution of the observed parallax π_H (assumed to be Gaussian) around the *true* parallax π is depicted in Figure 1. The error distribution of the estimate $R = \frac{1}{\pi_H}$ is obtained using the law of transformation of *probability density functions* (p.d.f.), that is to say, by multiplying the p.d.f. of π_H by the Jacobian of the $\pi_H \rightarrow R$ transformation, $\mathcal{J} = R^2$, and is depicted in Figure 2.

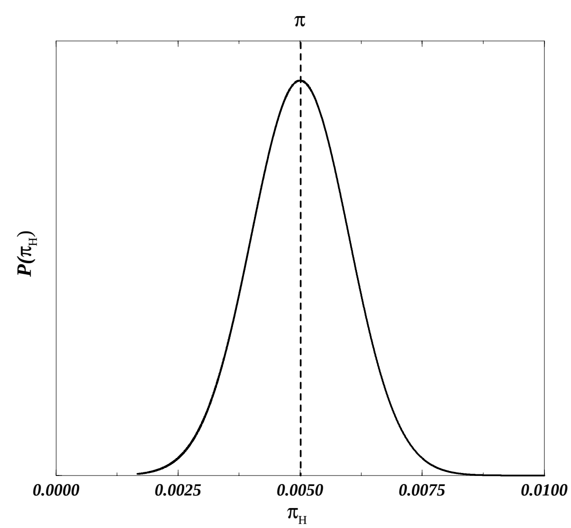

Figure 1. Distribution of the observational error in the observed parallax.

As a consequence, if $R = \frac{1}{\pi_H}$ is used to estimate stellar distances, an *a posteriori* correction is required. This correction, being of statistical nature, is only meaningful when *determined for* and *applied to* a given sample. See the next section and Brown et al. (1997) for more details.

3. SAMPLE SELECTION EFFECTS

Apart from the effects of the observational errors there are (at least) two more factors which must be taken into account when determining stellar distances for a sample of stars:

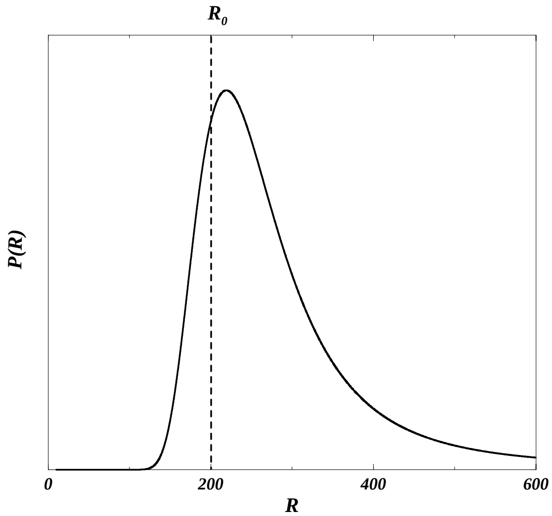

Figure 2. Distribution of the derived errors in $R = \frac{1}{\pi_H}$.

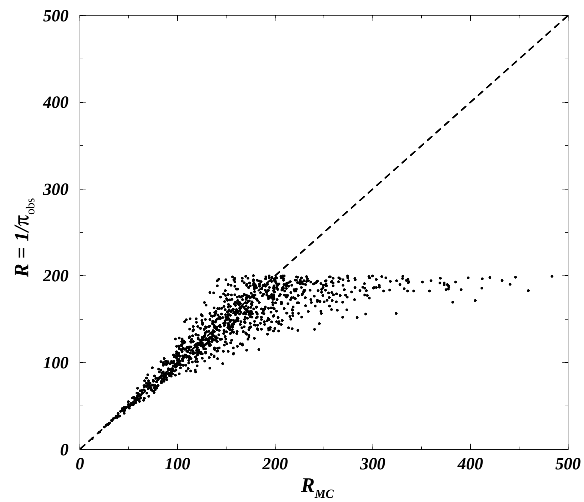

Figure 3. Simulated sample truncated at $\pi_H >\ 0.05$ *s.*

- the observational censorship of the sample;
- the spatial distribution of the sample.

Statistically unbiased distance estimates *for the sample as a whole* have to take into account these two factors, as well as the effects of the parallax observational errors described in the previous section.

To illustrate these effects we have generated Monte Carlo simulated samples to compare the generated distances R_{MC} (that is to say, the 'true' distances) with the values given by the $R = \frac{1}{\pi_H}$ estimate. The observational error in the trigonometric parallax has been fixed to 1 mas (Hipparcos mean value).

In Figure 3 the case of a sample limited to $\pi_H > 0.05$ sec is presented. Although the bias due to the observational error in the individual values of R is small ($\frac{\sigma_{\pi_H}}{\pi}$ is small for these stars, see Brown et al. 1997) on the average there is a strong bias in the estimated distances of the sample due to the truncation (see Figure 7).

On the other hand, for a sample limited in apparent magnitude, the effects are very different. In Figure 4 the case of a sample of stars with a Gaussian distribution of absolute magnitudes (mean $M_0 = 1.1$ mag and dispersion $\sigma_m = 0.4$ mag) limited to $m < 7.9$ mag is presented. In this case the bias goes in the opposite sense compared to the previous case (see Figure 7).

The spatial distribution of the sample also has an important influence on distance determinations from parallaxes. In the previous examples a homogeneous spatial distribution was used. If a cluster-like spatial distribution (an ellipsoidal space density at a mean distance of $R = 200$ pc) is combined with a relative error truncation $\frac{\sigma_{\pi_H}}{\pi_H} < 0.2$, the effects are very different, as shown in Figure 5. In this case, the mean bias is small, but there is a clear asymmetry effect: there are a fewer stars over the $R = R_{LM}$ line but they have higher errors than the ones under this line (see Figure 7). This asymmetry can have important effects when studying cluster sequences.

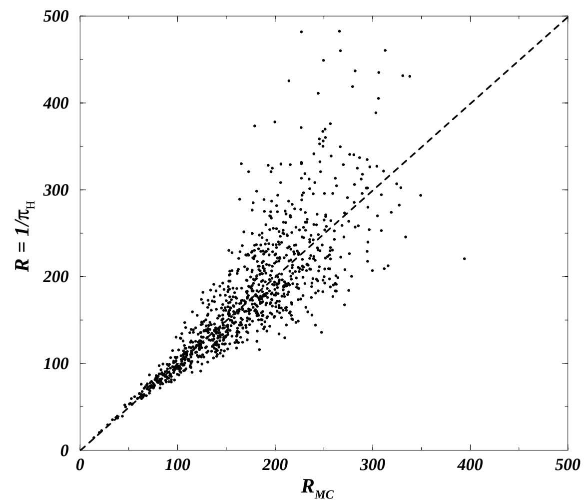

Figure 4. Simulated sample truncated in apparent magnitude.

These examples illustrate the difficulties of defining unbiased distance estimates for a sample of stars and underline the need for a *detailed analysis* of each specific case.

4. THE LM METHOD

The LM method (Luri, Mennessier et al. 1996) was developed to exploit the Hipparcos data to its full extent for luminosity calibrations and distance determinations.

Its results for several types of stars are presented in this issue:

- the luminosity calibration of the HR diagram revisited by Hipparcos (Gómez et al. 1997);

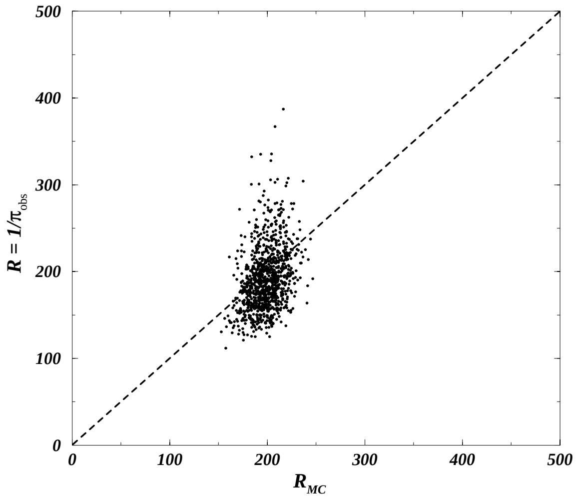

Figure 5. Simulated sample with a cluster-like spatial distribution and a $\frac{\sigma_{\pi H}}{\pi_H} < 0.2$ *truncation.*

- Barium stars: luminosity and kinematics from Hipparcos data (Luri et al. 1997);
- new aspect of Long-Period Variable stars from Hipparcos first results (Mennessier et al. 1997).

4.1. Use of All Available Data

The LM method uses all the astrometrical and photometric information provided by Hipparcos, as well as the available radial velocities and supplementary photometry.

4.2. Adaptation to the Sample

The LM method allows a detailed modelation of the characteristics of the luminosity, kinematical and spatial distributions of the sample. For instance, the galactic rotation and vertex deviation are included in the velocity distribution, and the spatial distribution used is an exponential disk.

4.3. Group Identification and Separation

An important feature of the LM method is its capability to identify and separate in the sample groups of stars with different luminosity, kinematics or spatial distribution characteristics. Separate results are given for each group identified, thus providing a much more meaningful information than a global result for the mixture of all of them.

4.4. Observational Errors

The knowledge of *individual* observational errors is included in the model so the estimations provided take into account its effects.

4.5. Observational Censorship

Any observational censorship used to define a sample is modelled and included in the method, so its effects are automatically taken into account and corrected.

4.6. Interstellar Extinction

The interstellar extinction is taken into account using the Arenou et al. (1992) model. Its effects in the apparent magnitudes as well as in the spatial distribution of the sample are automatically included and corrected.

5. LM DISTANCE ESTIMATES

The distance estimate defined by the LM method uses all the available information for a given star: trigonometric parallax π_H, apparent magnitude V, position (α, δ), proper motions (μ_α, μ_δ), radial velocity v_r as well as any other available physical parameters, as photometric indices C_i.

Once the parameters $\vec{\theta}$ defining the luminosity, kinematical and spatial distributions of a sample have been determined by the LM method, we can define the *a posteriori* probability of the star being at a distance R given $\vec{\theta}$ and the above cited observational parameters:

$$P(R \mid \vec{\theta}, V, \pi_H, \alpha, \delta, \mu_\alpha, \mu_\delta, v_r, [C_i]) \tag{1}$$

From this probability we can calculate the expected value of R, that can be used as an estimate of the true distance:

$$R_{LM} = \int R\ P(R \mid \vec{\theta}, V, \pi_H, \alpha, \delta, \mu_\alpha, \mu_\delta, v_r, [C_i]) dR$$

As this distance estimate uses a maximum of information it is very efficient. Its detailed properties are related to the properties of the sample and the observational errors. In the case of Hipparcos it can extend the range of reliable distances beyond the limit of usefulness of trigonometric parallaxes.

In Figure 6 we present an example using a Monte Carlo simulation: we have simulated a sample of A0 stars ($M_0 = 1.1$ mag and $\sigma_M = 0.4$ mag, with a young kinematics and a scale height $Z_0 = 50$ pc) with Hipparcos-like observational errors ($\sigma_\pi = \sigma_{\mu_\alpha} = \sigma_{\mu_\delta} = 1$ mas and $\sigma_m = 0.01$ mag) and an apparent magnitude limit at $m = 7.9$ mag. In Figure 7 the relative errors in distance of this estimate can be compared with the previous cases. This comparison shows that the asymmetry has been greatly reduced, giving an almost unbiased estimate.

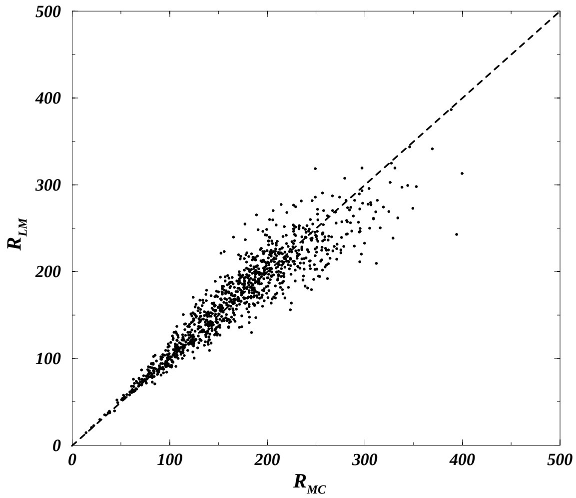

Figure 6. LM distance estimates for a simulated sample of a magnitude-limited sample of A0 V stars.

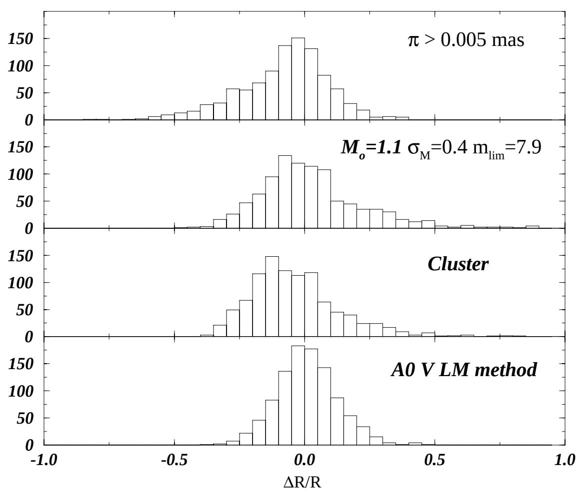

Figure 7. Histograms of relative errors in distance estimates.

REFERENCES

Brown, A.G.A., Arenou, F., Leeuwen, F. Van, Lindgren, L., Luri, X., 1997, ESA SP–402, this volume

Gómez, A.E., Luri, X., Mennessier, M.O., Torra, J., Figueras, F., 1997, ESA SP–402, this volume

Luri, X., Mennessier, M.O., Torra, J., Figueras, F., 1996, A&AS 117, 405

Luri, X., Gómez, A.E., Mennessier, M.O., Figueras, F., Torra, J., Grenier S., 1997, ESA SP–402, this volume

Mennessier, M.O., Mattei, J.A., Luri, X., 1997, ESA SP–402, this volume