

The extragalactic Cepheid distance bias: Numerical simulations

G. Paturel¹ and P. Teerikorpi²

¹ CRAL-Observatoire de Lyon, 69561 Saint-Genis Laval Cedex, France

² Tuorla Observatory, University of Turku, 21500 Piikkiö, Finland

Received 6 October 2003 / Accepted 27 November 2003

Abstract. We present numerical simulations for a selection bias in the extragalactic Cepheid method (as detected in Teerikorpi & Paturel 2002), which results in underestimated distances for host galaxies. The selection effect involves the Cepheid detection magnitude limits in V and I , the dispersion in $\langle M(P) \rangle$, and importantly, variation amplitude, observable upper period limit, and also extinction. It influences both the bias curve fitting method and the method where short-period Cepheids are removed. When V , I , and P -limits exist, one expects a typical bias pattern in the $\log H$ vs. M_{lim} diagram. A full bias correction is possible only for special favourable situations, but there always appears a nearly unbiased plateau. The simulated bias varies from galaxy-to-galaxy roughly as was revealed by the strong and puzzling observed trend for galaxies in TP02.

Key words. galaxies: general – cosmology

1. Introduction

The influence of selection on extragalactic distance measurements has been much studied for galaxy “standard candles”, e.g. in the TF-method. Sandage (1988) pointed out that Cepheids are also affected: only long periods may be fully sampled due to the magnitude limit (cf. Paturel et al. 1997; Lanoix et al. 1999). Because of the small scatter in the PL-relation, this source of error has been regarded as small and correctable in modern HST studies.

But Teerikorpi & Paturel (2002; TP02) found an unexpectedly large effect, when inspecting the Hubble parameter versus the deepness of the Cepheid sample in a galaxy. The deepness was measured by the absolute Cepheid magnitude limit $M_{\text{lim}} = m_{\text{lim}} - 5 \log V_c - 16.5 + 5 \log H/50$, where m_{lim} was derived from histograms of magnitudes for each galaxy and relative distances (V_c) were based on the Hubble law. In the $\log H$ vs. M_{lim} diagram a familiar bias pattern appeared: for deep samples, H remains about constant, while for bright M_{lim} , it starts to grow.

The same behaviour was seen in two sets of Cepheid distances: 1) the HST Key Project values (Freedman et al. 2001), for which shortest period cepheids were excluded in order to diminish the bias, and 2) those derived iteratively using a theoretical bias curve by Paturel et al. (2002a,b).

In order to explain this effect, a simple model was proposed in TP02. This bias differs from the usual selection affecting standard candles with dispersion σ . Though the scatter in the average magnitude $\langle M(P) \rangle$ for Cepheids is small at a period P , it is not this dispersion which basically defines the bias here as it does for the TF relation. A member of such a standard candle

class has a constant luminosity, but each Cepheid has a variation amplitude. When predicting the bias, one must distinguish between “Cepheid type” and classical “TF type” relations¹.

Let the Cepheid magnitude limit be m_{lim} and the amplitude ΔV . To determine $\langle m(P) \rangle$ requires the whole range $m(P) \pm \Delta V$ is observed. As $M(P)$ has a small scatter σ , an unbiased $\bar{m}(P)$ requires a somewhat wider range, to the magnitude $m(P) + \Delta V + \alpha\sigma$ ($\alpha \approx 1-2$). If the observational P -limit is too short or if long-period Cepheids are rare, this faint limit may never be reached.

It is instructive to make a diagram for real Cepheid measured galaxies, showing the M vs. $\log P$ relation and \pm its amplitude = 0.5 mag. Each galaxy is characterized by the limiting absolute magnitude and period (Fig. 1, with M_{lim} for $H = 56 \text{ km s}^{-1} \text{ Mpc}^{-1}$). The lower-left quadrant defines the region from which one was able to collect Cepheids. *If a galaxy is below the “PL + ΔV ” line, then one cannot measure the whole variation range.* Dots are examples of “unbiased” galaxies ($\log H < 1.8$) from Table 1 in TP02, open circles are “strongly biased” ($\log H > 1.9$). The unbiased ones lie well above the PL line, having a large triangle in which the PL relation is sampled also on its faint side. Notwithstanding the exact value of H , the biased ones have small areas allowing Cepheid detection, which means difficult distance determination.

In this Letter we report results of numerical simulations which test the selection model, now including also the I magnitude. We are interested in such questions as:

- What is the typical bias pattern one expects in the $\log H$ vs. M_{lim} diagram?

¹ We have here always in mind the Malmquist bias of the 2nd kind, in the terminology in Teerikorpi (1997): the systematic error in the average derived $\langle \mu_{\text{der}} \rangle$ when $\mu_{\text{true}} = \text{constant}$.

Send offprint requests to: G. Paturel,
e-mail: patu@obs.univ-lyon1.fr

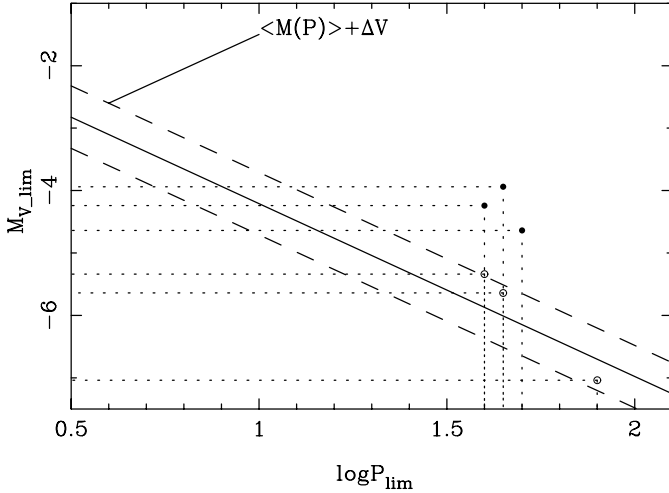


Fig. 1. Locations of some “unbiased” (dots) and “biased” (open circles) Cepheid-measured galaxies from Table 1 of TP02 on the M_{lim} vs. $\log P_{\text{lim}}$ diagram. The solid line is the PL relation. Dashed lines give the amplitude range, here assumed to be ± 0.5 mag. The lower-left quadrants show the regions from which Cepheids are allowed. Close to the upper dashed line, distance determination becomes increasingly difficult.

- On what conditions do bias correction methods give unbiased distances at all M_{lim} ?
- Does the normalized M_{lim} lead to a bias pattern with less scatter, like for real galaxies in TP02?
- What is the role of extinction?

2. Simulations: One-magnitude PL relation

In our numerical simulations we study what happens to distance determination, when one observes a galaxy containing thousands of Cepheids, affected by magnitude and period limits. One first chooses randomly a period normally distributed around $\log P = 1.2$ with dispersion = 0.5. Then the average magnitude comes from a normal distribution around $M = M(P)$, with dispersion σ . If a Cepheid at its faintest phase is dimmer than m_{lim} , it is removed from the sample. Thus we start with the available set of calibrator galaxies, fix their distances kinematically ($H = 56$ or $\log H = 1.75$), and use the apparent magnitude and period limits² from Table 1 of TP02, in order to create for each galaxy its observable Cepheid population.

There are two ways in use to treat the bias, previously thought to be small and depending basically on σ : 1) using iteratively the bias curve (Patrel et al. 2002a,b), or 2) removing short period Cepheids as suggested by Patrel et al. (1997) and Lanoix et al. (1999) and used in Freedman et al. (2001). We wish to check whether the corrections return, on the average, the input value of H . This was first done for one magnitude (V) PL relations. For simplicity, zero extinction was assumed.

Figure 2a shows a typical end result of simulation, when in the bias curve fitting the amplitude effect is ignored. Here σ is

² In TP02 P_{lim} was not the maximum observed period, but was indicative of into how long periods the bulk of the sample extends. In simulations we add 0.1 to the listed P_{lim} .

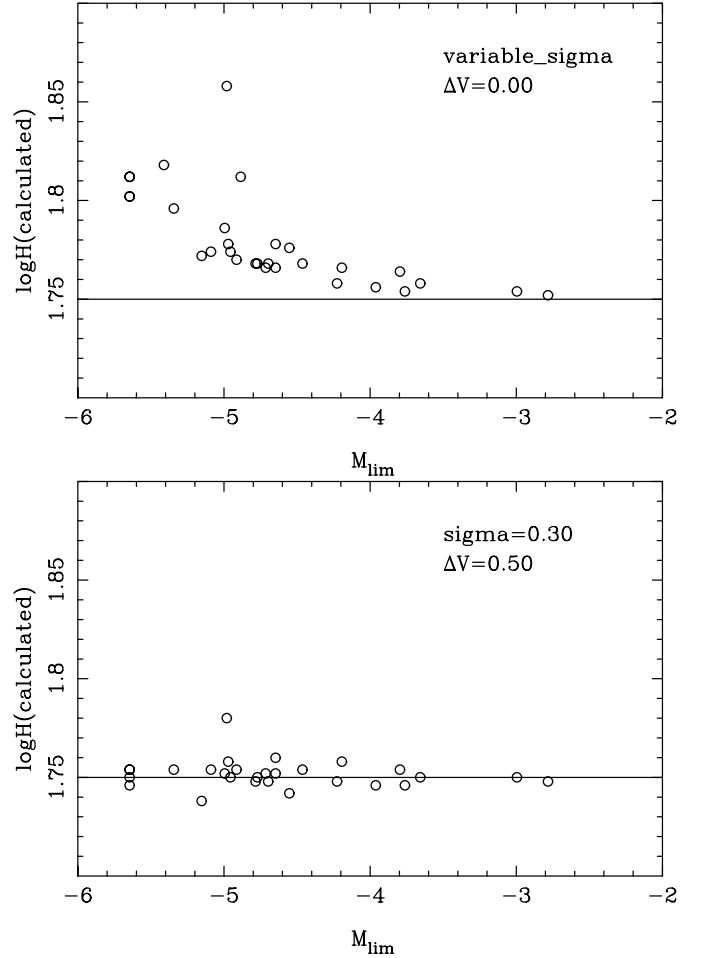


Fig. 2. One-magnitude PL relation: the simulated behaviour of the Hubble parameter versus the absolute magnitude limit for Cepheids in calibrator galaxies when distances were derived by bias curve fitting. **a)** The amplitude ΔV ignored in the bias curve and σ determined from each galaxy separately. **b)** The amplitude included in the bias curve and σ is fixed = 0.3 mag. The input $\log H = 1.75$ is shown by the horizontal line.

variable, i.e. it was derived for each galaxy individually from the observed dispersion. The bias pattern appears, but even at the horizontal plateau one underestimates distances (or the plateau is hardly reached).

The bias vanishes when one puts the amplitude ΔV into the magnitude limit, when the distance is derived using the bias curve ($m_{\text{lim}}(\text{effective}) = m_{\text{lim}} - \Delta V$) and fixes σ into its universal input value (Fig. 2b). For variable σ we still found a bias at bright magnitude limits, reflecting a too small σ derived from the biased part.

We conclude that the iterative bias curve fitting method works for the one-magnitude Cepheid relation, but only when the amplitude and dispersion are known.

3. Simulations: Two-magnitude PL relation

Things are more complicated when the Cepheids are measured in two magnitude systems. Thus, for the combined V, I method there remains a bias, which partly seems to be related to the

fact that now two magnitude limits are involved. A Cepheid may be accepted by its V magnitude, but is cut away because its I magnitude cannot be measured. The I limit is brighter than V_{lim} , while the V and I PL relations are numerically not widely apart from each other (we use: $M_V = -2.77 \log P - 1.44$, $M_I = -3.05 \log P - 1.81$). When we put $I_{\text{lim}} = V_{\text{lim}} - 1.5$ (based on the mean $V - I$ colour) when defining the I cut-off for each simulated galaxy, the bias is strong (Fig. 3a). Even the unbiased plateau at $\log H = 1.75$ is not reached.

In Figs. 3b,c we show examples in which the period limit is infinite. This affects significantly the end results. In Fig. 3c the ideal situation dreamed by an observer (a faint I -limit, a long P -limit) results in unbiased average distances all over the M_{lim} axis.

The manner in which we have in our previous papers derived distance moduli to calibrator galaxies roughly corresponds to that in Fig. 3a (here σ is fixed, while we had it variable, but this does not cause much difference). Though we cannot yet correct the two-magnitude bias exactly, at least we can conclude that the simulation well reproduces the trend seen in the data. The “unbiased plateau” gives close to the true input value of $\log H$, or slightly larger.

We also tested the “normalization” used in Fig. 1b of TP02 ($M_{\text{lim}}(\text{norm}) = M_{\text{lim}} + 2.77(\log P_{\text{lim}} - 1.6)$), and similarly as happened with observations it also resulted in a narrow relation for simulations (cf. Figs. 4 and 3a). The rationale of this second-order shift is seen from Fig. 3a, in which different symbols indicate different ranges of the limiting period (filled triangles are longest: $P_{\text{lim}} \geq 1.75$, asterixes are shortest: $P_{\text{lim}} \leq 1.65$, and open circles are intermediate. Shorter P_{lim} makes the bias start already at a fainter M_{lim} , which also broadens the bias pattern. That we found in TP02 a significantly narrower relation after this normalization, is evidence for the bias interpretation.

The HST Key Project distances agree rather well with the values derived by the TF-type bias curve fitting in Patrel et al. (2002a,b), though the latter result in somewhat larger values at large distances. Our above simulations are applicable to the latter treatment. As the resulting $\log P$ distribution for the observed Cepheid sample in our simulations is not far from what was obtained in the HST studies, we can also test that approach.

The HST values involve exclusion of some short period Cepheids for which the bias was expected to increase (Patrel et al. 1997; Lanoix et al. 1999). But this straightforward approach is also vulnerable to the present bias, as is shown by the simulation in Fig. 5. There we have used 30 longest period Cepheids for each galaxy. Roughly corresponding to the real practice, this results in a good unbiased plateau, but bright limit galaxies are clearly biased.

A plus-side of this unbiased plateau approach is that the plateau galaxies will be automatically corrected from extinction. But note that the average extinction between us and the whole Cepheid population in the host galaxy will brighten the effective limiting magnitude, hence increase the bias for the bright M_{lim} part of the diagram, *even if each observed Cepheid has been corrected for extinction*. This was explained in TP02 and we have confirmed it with simulations (open circles in Fig. 5: here Cepheids are viewed affected by reddening $\langle E(B - V) \rangle = 0.15$ and $\sigma_{B-V} = 0.05$. The bias

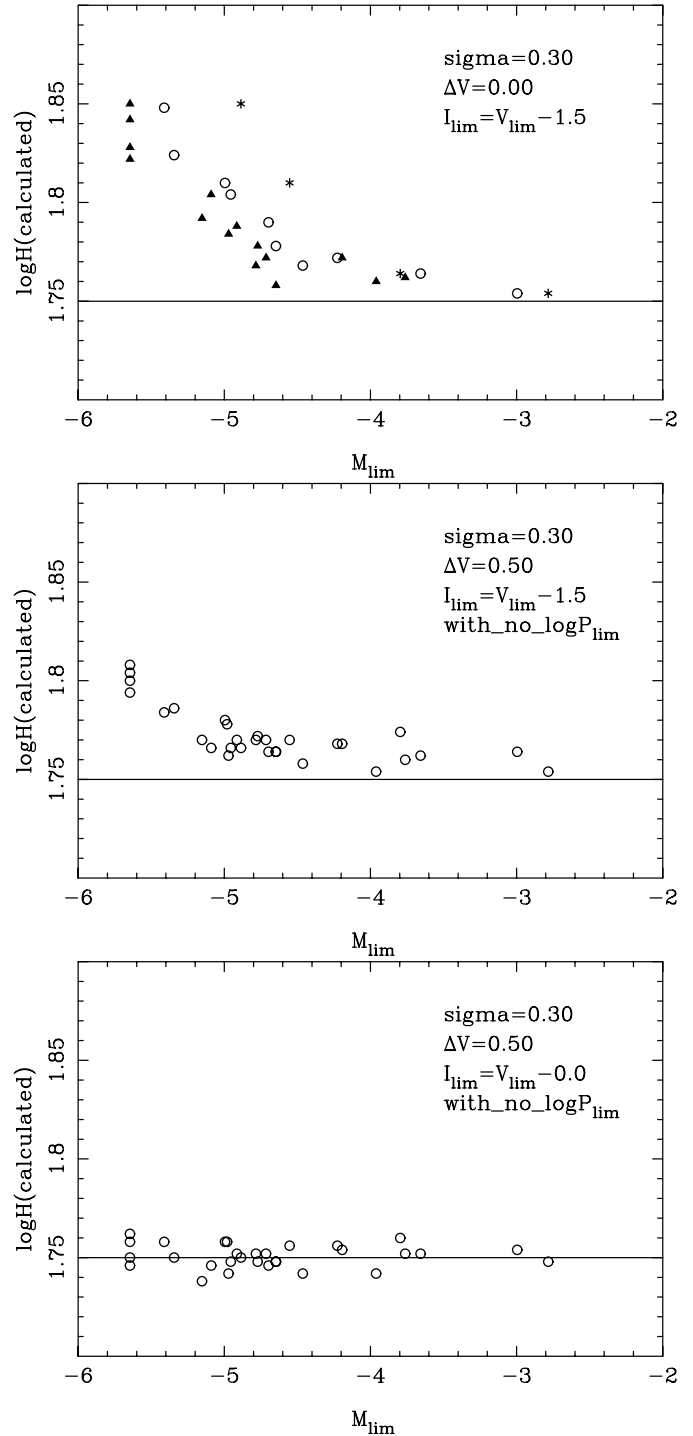


Fig. 3. V and I PL relations: the simulated behaviour of the Hubble parameter when the distances have been derived by the bias curve fitting method from V and I magnitudes with σ having the constant value 0.3 mag. **a)** $I_{\text{lim}} = V_{\text{lim}} - 1.5$, and the V amplitude ignored in the bias curve. **b)** As in **a)** but here the V amplitude has been put into the bias curve and the period limit is infinite. **c)** Here the I limit is made fainter. In **a)** different symbols mean different P_{lim} ranges (see the text).

pattern is roughly like in observations.) In our future application of these simulations to real Cepheids we will pay special attention to the extinction problem in the bias curve fitting method.

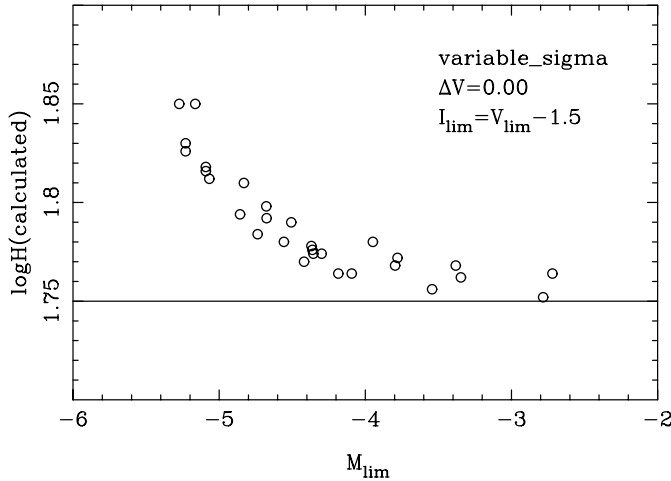


Fig. 4. V and I PL relations: the simulated behaviour of the Hubble parameter using the bias curve fitting method with σ variable. The limiting magnitude is here normalized as for the observations in Fig. 1b in TP02. Note the reduced scatter both in observations (TP02) and simulations (here).

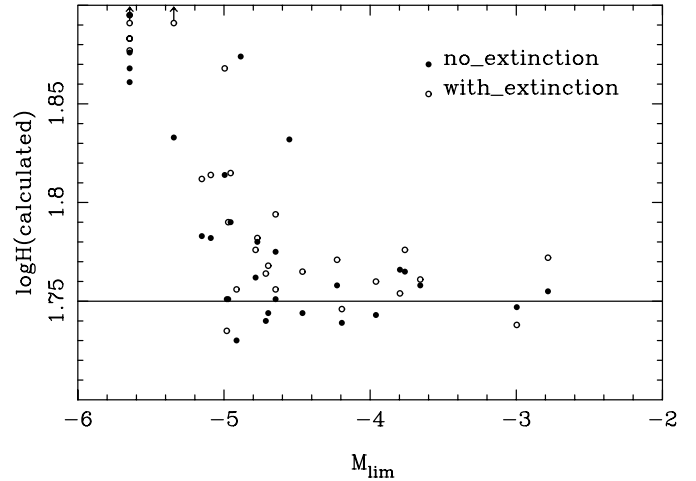


Fig. 5. The simulated behaviour of the Hubble parameter when one attempts to avoid the bias by excluding short-period Cepheids, leaving for each galaxy 30 longest period Cepheids. Dots: zero extinction; Open circles: the whole Cepheid population is affected by reddening, and each observed Cepheid is extinction corrected using the standard V, I procedure.

4. Conclusions

We summarize our results from numerical simulations:

- In practice, when V , I , and P -limits exist, one expects a typical bias pattern in the $\log H$ vs. M_{lim} diagram. A full bias correction is possible only for special favourable situations, but there always appears a nearly unbiased plateau.
- For the one-magnitude PL relation, the bias curve fitting gives unbiased distances at all M_{lim} , but only if the universal dispersion σ and amplitude ΔV are known and both are used in the bias calculation.
- The two-magnitude method, using the standard PL relations for V and I , is more sensitive to the upper limit P_{lim} and gives unbiased distances only if $P_{\text{lim}} = \infty$ and the I limit is faint enough.
- The extinction, which brightens the effective limiting magnitude and enhances the bias, is a problem to overcome in the analysis of real Cepheid samples.
- The normalization of M_{lim} , based on P_{lim} as in TP02, results in simulations in a bias pattern with less scatter, just as happened with the observations in TP02.

Our conclusion is that the strong trend in the Hubble parameter, as found in TP02, may really be due to the selection bias in Cepheid distance measurements. We wonder if we see here an old lesson for modern precision cosmology: the bias creeping into distance measurements has been a constant problem in astronomy.

As our next step, we will reanalyze our previous distance determinations, now understood to be partly biased, to the galaxies with available Cepheids in the light of these simulations and the results in TP02. We will then also discuss the implications on the cosmic distance scale (Patrel & Teerikorpi, in preparation). Independently of the present bias, the recent analysis by Tammann et al. (2003) of the non-uniqueness of the PL relation in different galaxies, anticipated in Patrel et al. (2002b), shows the distance scale problem is not yet finally settled.

Acknowledgements. We would like to thank the anonymous referee for useful comments. This study has been supported by the Academy of Finland (project “Fundamental questions of observational cosmology”).

References

- Freedman, W. L., Madore, B., Gibson, B. K., et al. 2001, *ApJ*, 553, 47
- Lanoix, P., Patrel, G., & Garnier, R. 1999, *ApJ*, 517, 188
- Patrel, G., Lanoix, P., Garnier, R., et al. 1997, in *Proc. ESA Symp. Hipparcos-Venice 97*, 629, ed. M. A. C. Perryman, & P. L. Bernacca, ESA SP-402
- Patrel, G., Theureau, G., Fouqué, P., et al. 2002, *A&A*, 383, 398
- Patrel, G., Teerikorpi, P., Theureau, G., et al. 2002, *A&A*, 389, 19
- Sandage, A. 1988, *PASP*, 100, 935
- Tammann, G., Sandage, A., & Reindl, B. 2003, *A&A*, 404, 423
- Teerikorpi, P. 1997, *ARA&A*, 35, 101
- Teerikorpi, P., & Patrel, G. 2002, *A&A*, 381, L37 (TP02)