Monte Carlo Method for Stellar Distance and Magnitude Estimation with Parallax

Santanu Ganguly

Liverpoool John Moores University

Abstract

Conventional trigonometric parallax has long been used to calculate accurate distances to nearby stars. The regular nature of the parallactic motion of a star, caused by the Earth's orbit around the Sun, allows this motion to be decoupled from the intrinsic proper motion of the star in the heliocentric rest frame. Hence, the distance to the star can be calculated. The smallest parallax measureable from the ground is about 0.01 arcsec. For a given level of astrometric precision, this imposes a fundamental limit to the observable distance.

Absolute astrometry from the ground is fundamentally limited by atmospheric turbulence, gravitational and thermal flexure of telescopes throughout the night, and a limited field of view. Because of this, astrometry surveys are conducted in space [1, 2, 3].

This work investigates statistical methods for calibrating parallaxes to retrieve absolute magnitude of stars using Monte Carlo (MC) simulations [4]. An estimate for bias in the distances and absolute magnitudes derived from parallax measurements can best be made using MC models.

The MC code used here calculates first a synthetic sample of stars with a given 'true' absolute magnitude M (for example in the optical V band) and distribution of 'true' distances R drawn randomly according to given law. To each synthetic star at a true distance R the code assigns an apparent magnitude and a true parallax $p = 1/R$. Simulation of observations are achieved by assigning a random Gaussian error with a given σ which in turn provides the observed parallax p_0 which is different from the true parallax due to measurement error. In this model, the parallax

error σ can be chosen to be either constant or dependent on the apparent magnitude of the synthetic star – this follows the error-law of the parallaxes measured by the GAIA astrometric satellite (GAIA Data Release 1, 2016). Each value of p_0 gives the retrieved value R_0 , which in turn is different from the true distance R due to parallax errors. R_0 is used to determine M_0, the retrieved absolute magnitude of each synthetic star, which in turn is perturbed by a Gaussian photometric observational error σ_{phot}. All negative derived distances were ignored as the negativity is an indication that the parallax error was larger than the parallax itself.

Task 1

A sample of 400000 stars with a constant spatial distribution with a true absolute magnitude of 0.06 and a very small Gaussian dispersion of 0.001 mag weas ran through the simulation. The samples had a constant parallax error of 0.00005 arc and a small $\sigma_{phot} = 0.01$ mag without any magnitude limit.

i) All stars within $R_0 = 4$ Kpc were considered and the mean value and dispersion (standard deviation) were determined of the retrieved $(M_0)$and true absolute magnitudes (M) as follows (Table 1.):

Table 1

	True abs mag (M)	Retrieved abs mag (M_0)
Mean	0.599995093	0.626801532
Dispersion	0.001043615	0.2850161

There is a difference between values of true and retrieved absolute magnitudes due to parallax error. The dispersion for M_0 is also much larger than that of M, indicating the error due to parallax and perhaps also radial velocity data, proper motion and possibly presence of super-giants whose absolute magnitudes may skew the retrieved results.

ii) The sample was split into 4 different ranges according to retrieved distance $R_0 =$ 0-1000, 1001-2000, 2001-3000 and 3001-4000 and mean and dispersion calculated (Table 2):

Table 2

R_0	True abs mag (M)		Retrieved abs mag M_0)	
	Mean	Dispersion	Mean	Dispersion
1-1000	0.599973992	0.001015171	0.617426528	0.092589977
1001-2000	0.600002905	0.001052132	0.670734063	0.187905759
2001-3000	0.599989023	0.001045876	0.758506927	0.284561159
3001-4000	0.599995247	0.001040786	0.597096679	0.242858435

Generally speaking, as distances of stars grow longer, the error due to parallax, i.e dispersion of data for retrieved absolute magnitude should increase. The very small dispersion in the true absolute magnitude is expected and the large difference in dispersion between M_0 and M indicates that.

a) Histograms of the true number distribution of stars with distance and retrieved distribution of stars with distance were done with a bin size of 200 (Fig 1)

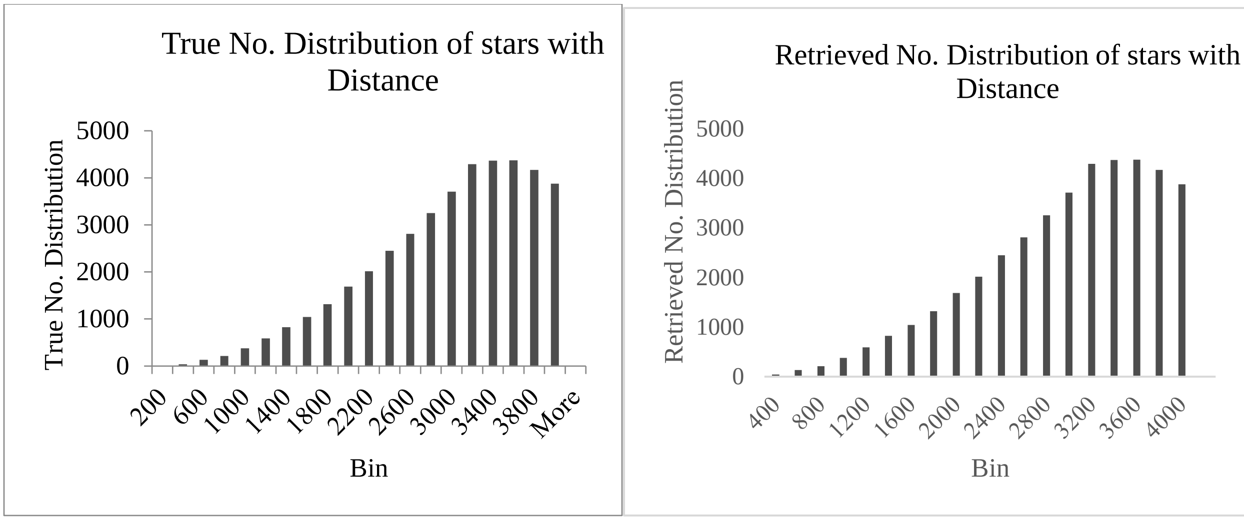


Fig. 1

The distribution of distances show a bias towards the higher distances. This is because of the bias introduced by the parallax errors.

b) A scatter plot of all individual retrieved absolute magnitudes as a function of the retrieved distances was produced for 41540 data points.

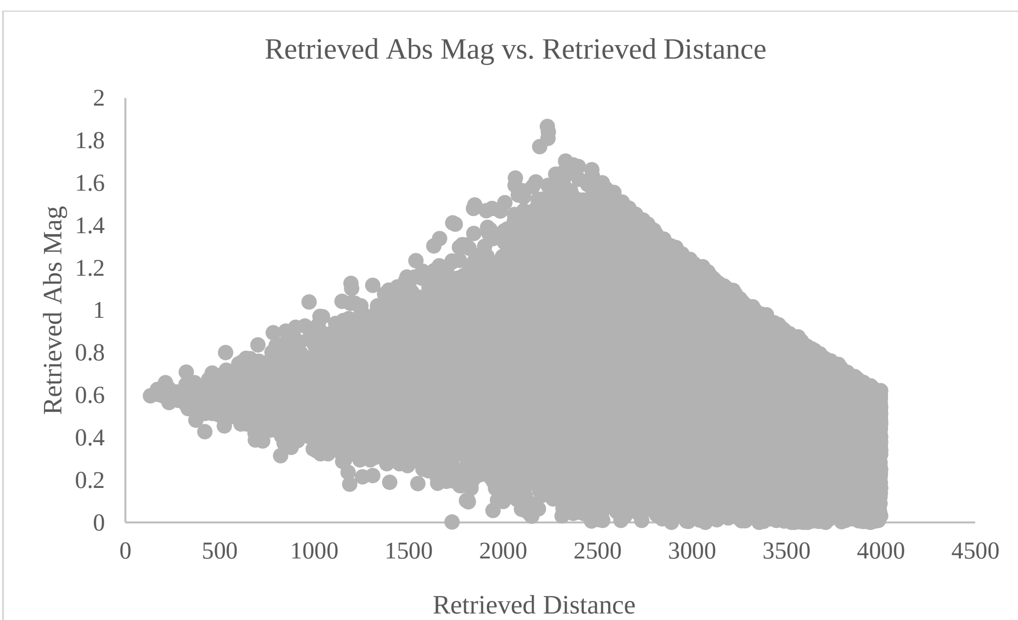


Fig.2

c) Histograms of the retrieved absolute magnitudes for $R_0 =$ 0-1000, 1001-2000, 2001-3000 and 3001-4000 were plotted

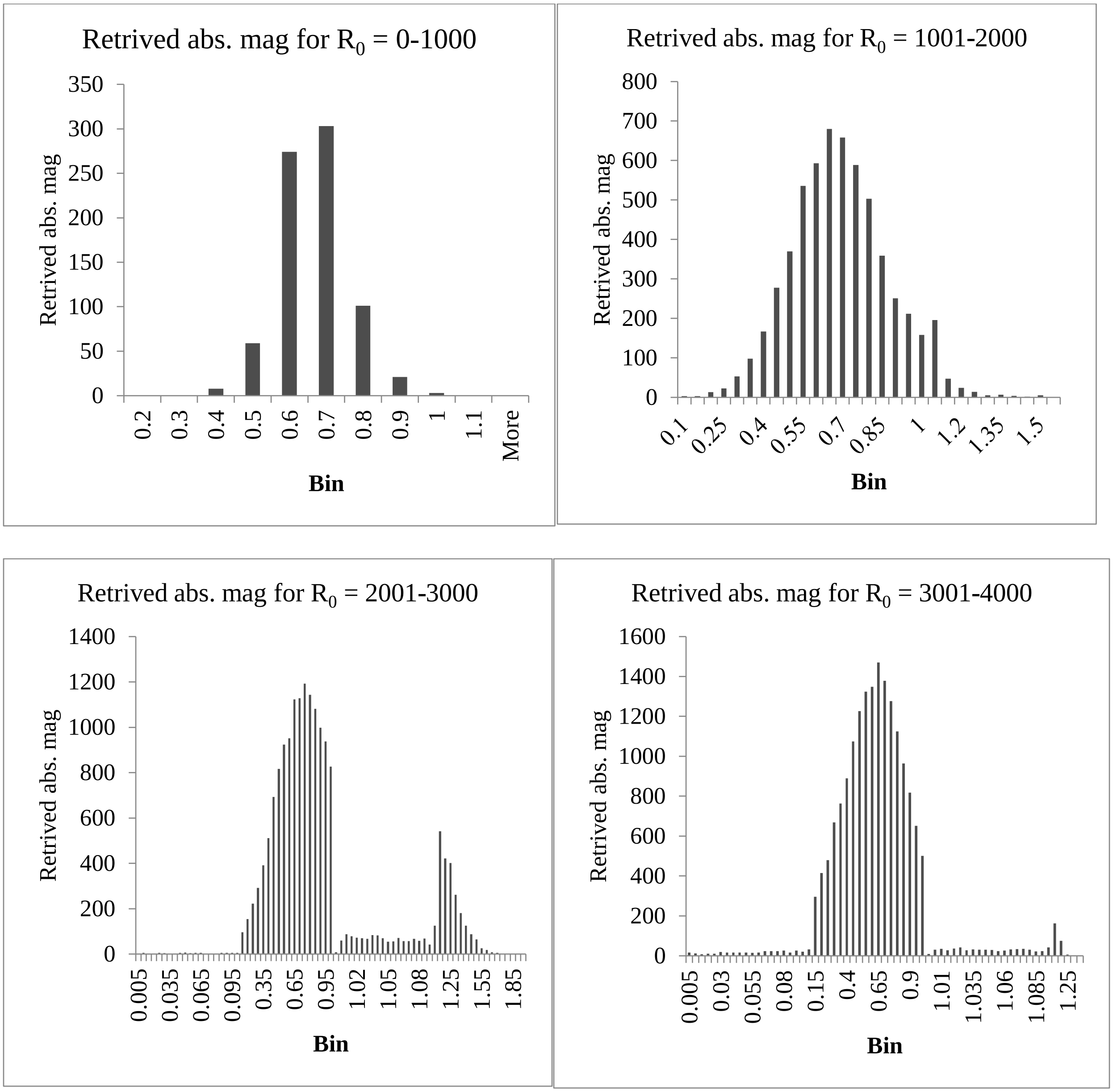


Fig. 3

The mean M (tables 1 & 2) is very similar in all the stellar distance ranges. However, as range of R_0 increases, dispersion also tends to increase and so does the difference between the values of M_0 in various ranges. The difference between M_0 and M indicates the parallax in measurement of M_0.

The reason for a less accurate measurement for more distant stars (and galaxies in particular) is because those distant objects remain fixed in position, whereas nearby objects shift position when viewed from our location on Earth, even when viewed from opposite sides of the Sun over six month intervals.

Bias in distances determined from parallaxes for the model population gets included as a bias in absolute magnitudes. Scatter plot of Fig. 2 shows that the bias is modified by the non-uniform density of the distribution and causes the parallax distances to be underestimated and the absolute magnitudes to be ~0.38 mag too faint for objects with estimated parallax errors of ~20% of the measured parallax.

For a constant parallax error, more distant tsars show largest migration to shorter and longer distances as is depicted in the histogram of $R_0 =$ 3001-4000 of Fig 3, where the distribution appears Gaussian.

Task 2

Task 1 was repeated *but* with a limit of 14 mag.

i) For all stars within $R_0 = 4$ Kpc the mean value and dispersion were determined of the retrieved (M_0)and true absolute magnitudes (M) (Table 3.):

Table 3

	True abs mag (M)	Retrieved abs mag (M_0)
Mean	0.599995755	0.749485497
Dispersion	0.001037097	0.302948438

Results and conclusions similar to that of Task 1(i), except the retrieved absolute magnitude show a mean higher by ~15% aided by an error higher by ~27%.

ii) The sample was split into 4 different ranges according to retrieved distance $R_0 =$ 0-1000, 1001-2000, 2001-3000 and 3001-4000 and mean and dispersion calculated. (Table 4):

Table 4

R_0	True abs mag (M)		Retrieved abs mag M_0)	
	Mean	Dispersion	Mean	Dispersion
1-1000	0.6	0.001020977	0.614595268	0.091393253
1001-2000	0.599972631	0.001043824	0.668107937	0.191843326
2001-3000	0.599997998	0.001028808	0.784864969	0.313074733
3001-4000	0.599998589	0.001041226	0.747759758	0.314185983

As expected, the values of true absolute magnitude are almost identical to those of Task 1(ii). However, the limit of 14 mag has added to the bias of parallax and the errors are generally larger by ~10%.

a) Histograms of the true number distribution of stars with distance and retrieved distribution of stars with distance were done with a bin size of 200 (Fig 4)

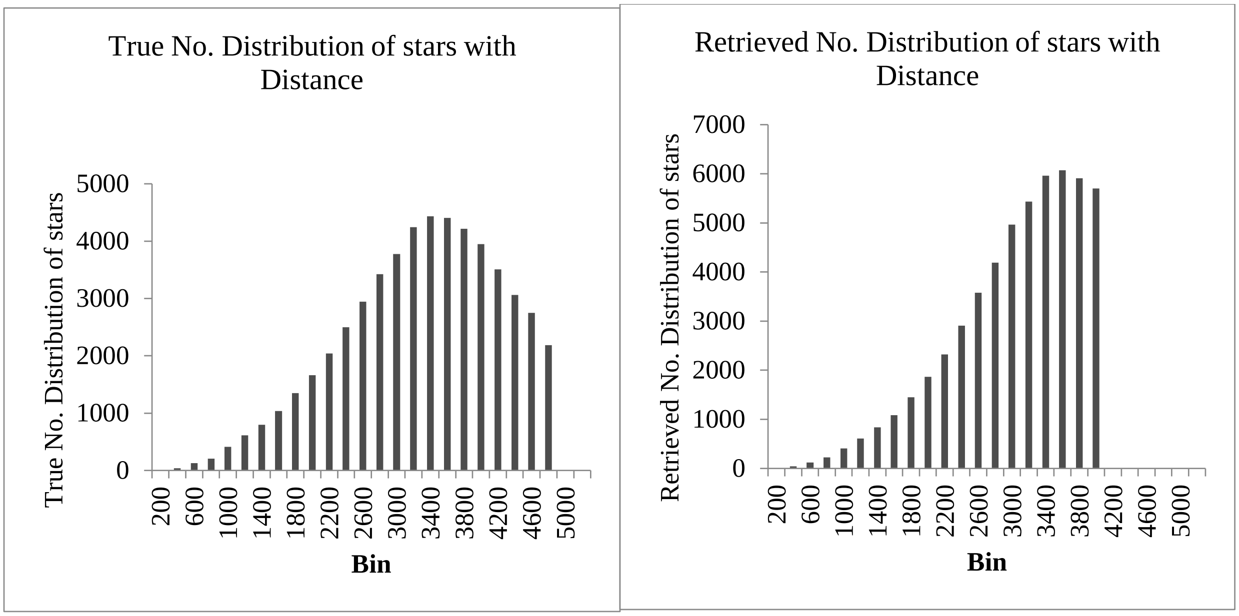


Fig. 4

The increased bias due to the imposed magnitude limit shows a higher gradient due to additional parallax error – more stars tend to migrate towards a clustered distribution.

b) A scatter plot of all individual retrieved absolute magnitudes as a function of the retrieved distances was produced for the data points.

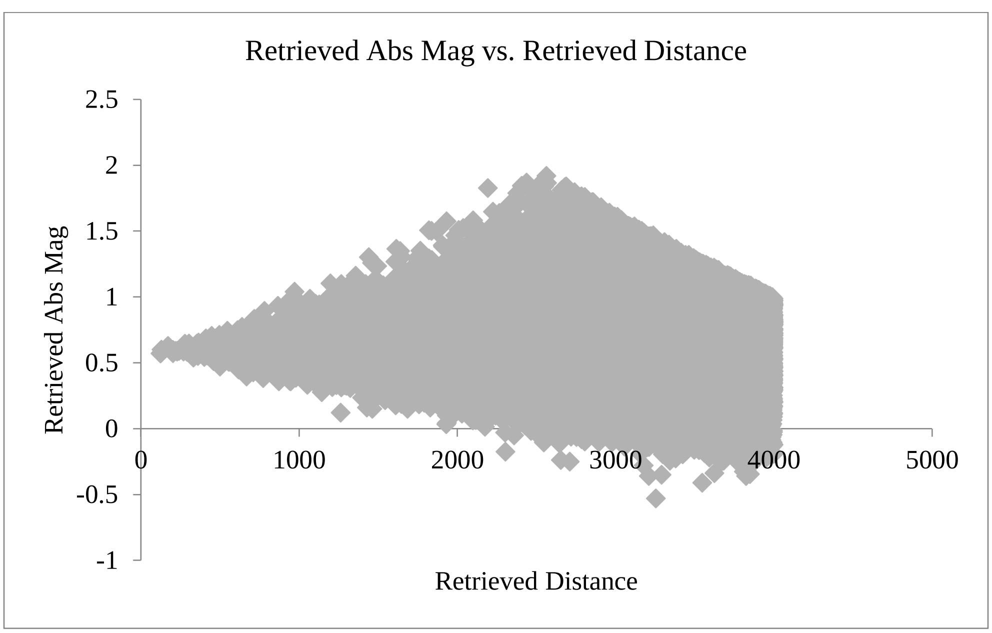
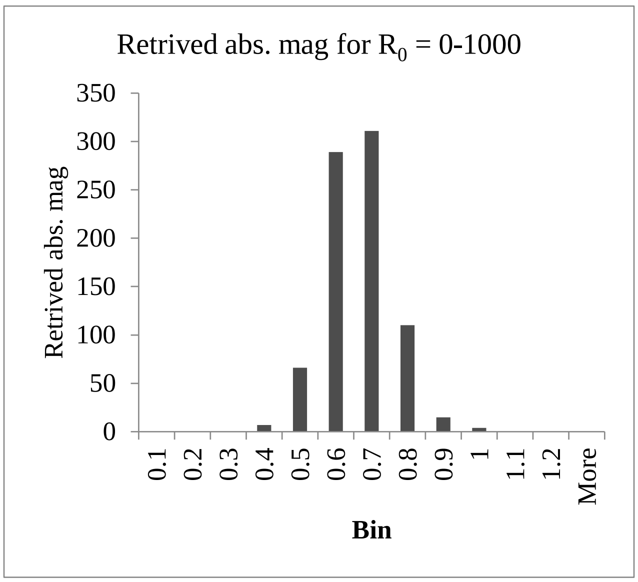
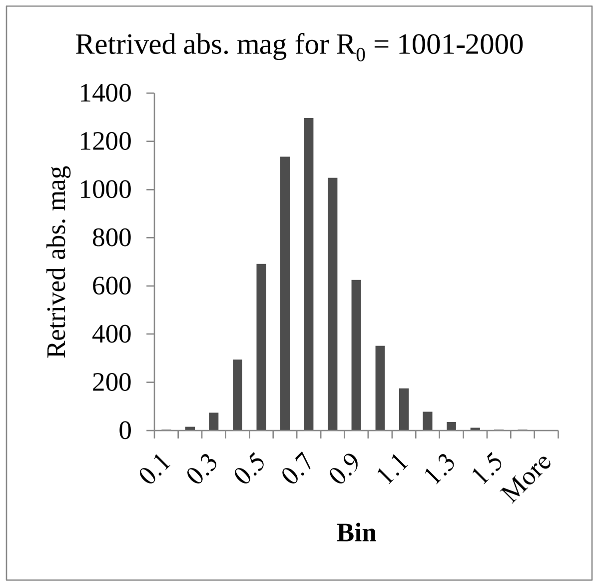


Fig.5

The scatter plot of Fig 5 as compared to that of Fig. 3 shows that there is an added selection of likely nearby objects, attempting to match our sample selection of the generally non uniform spatial distribution.

c) Histograms of the retrieved absolute magnitudes for $R_0 =$ 0-1000, 1001-2000, 2001-3000 and 3001-4000 were plotted





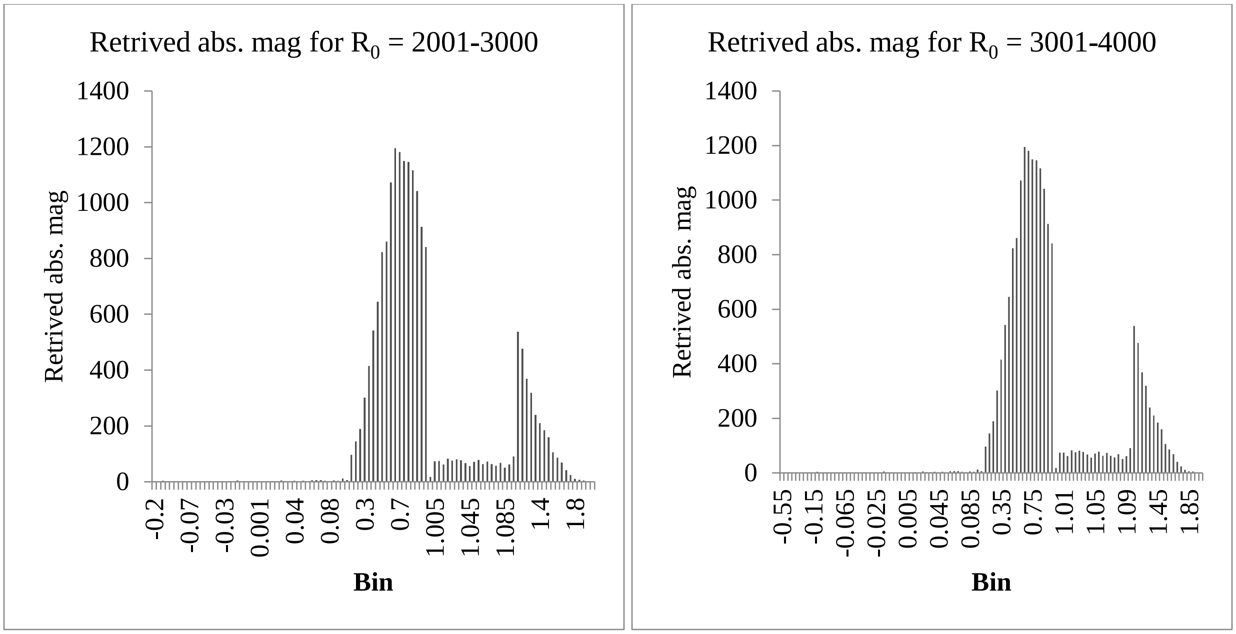


Fig. 6

The main difference between Fig 3 and Fig 6 are the added biases and lighter distribution density due to the magnitude limit.

Task 3

Task 1 was repeated *but* for the case where the spatial density of stars would decrease by R^2.

i) For all stars within $R_0 = 4$ Kpc the mean value and dispersion were determined of the retrieved $(M_0)$and true absolute magnitudes (M) (Table 5.):

Table 5

	True abs mag (M)	Retrieved abs mag (M_0)
Mean	0.599997928	0.682925493
Dispersion	0.001042508	0.291314958

Results and conclusions similar to that of Task 1(i).

ii) The sample was split into 4 different ranges according to retrieved distance $R_0 =$ 0-1000, 1001-2000, 2001-3000 and 3001-4000. (Table 6):

Table 6

R_0	True abs mag (M)		Retrieved abs mag M_0)	
	Mean	Dispersion	Mean	Dispersion
1-1000	0.600002441	0.001041375	0.604331526	0.064003748
1001-2000	0.600001521	0.001041278	0.634167888	0.171829241
2001-3000	0.599999128	0.001041781	0.695061629	0.297263576
3001-4000	0.599989396	0.001045321	0.78728638	0.43206919

As expected, the values of true absolute magnitude are almost identical to those of Task 1(ii). However, since the error is inversely proportional to R, the distribution proportional to R^2 has contributed to reduce errors by 3-5% below the 3001 mark has added to the bias of parallax and the errors are generally larger by ~10%. In the range 3001-4000 the errors have increased by over ~10% adding to the retrieved magnitude.

a) Histograms of the true number distribution of stars with distance and retrieved distribution of stars with distance were done with a bin size of 200 (Fig. 7)

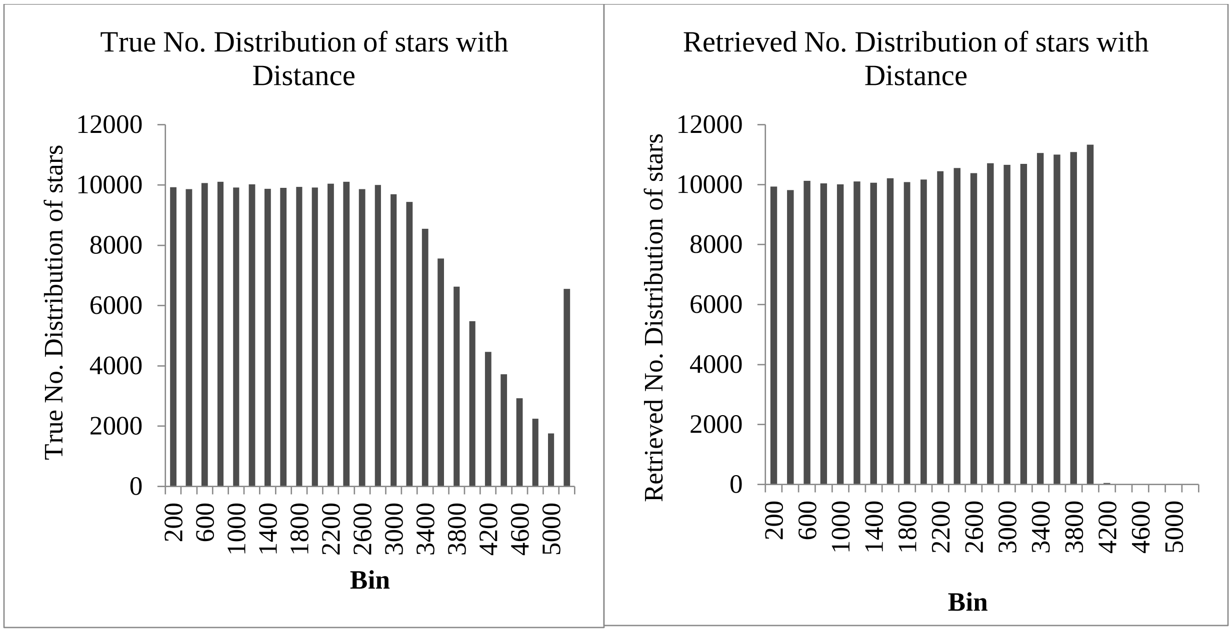


Fig. 7

The imposed inversely proportional bias sows an almost even distribution of retrieved magnitude which, in accordance to its mean values and parallax errors increases in gradient.

b) A scatter plot of all individual retrieved absolute magnitudes as a function of the retrieved distances was produced for 208645 data points.

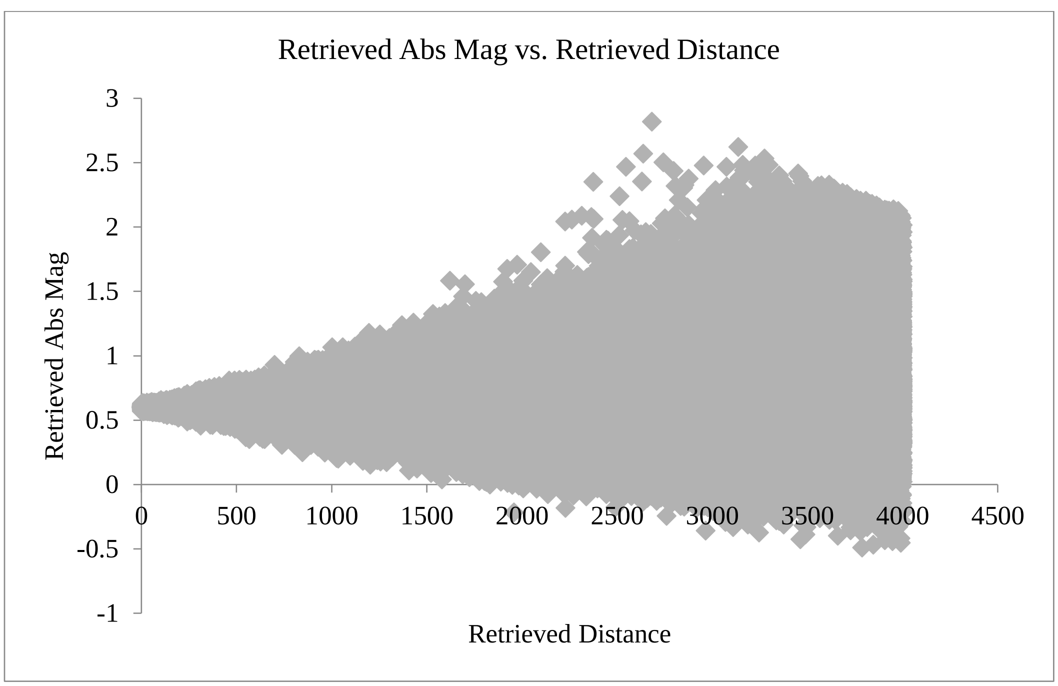
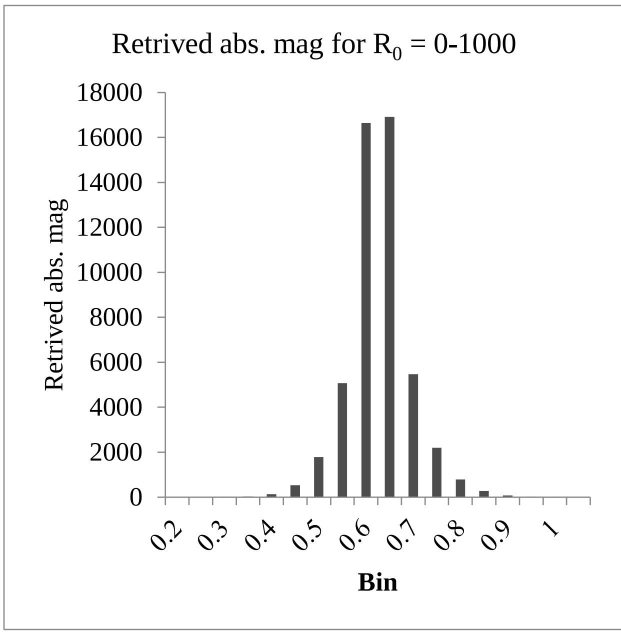
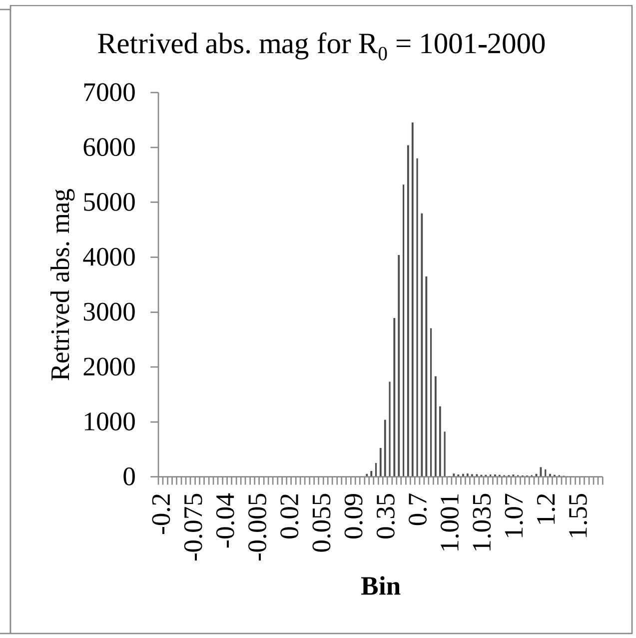


Fig.7

The scatter plot of Fig 7 as compared to that of Fig. 3 shows that there is an added selection of likely nearby objects.

c) Histograms of the retrieved absolute magnitudes for $R_0 =$ 0-1000, 1001-2000, 2001-3000 and 3001-4000 were plotted





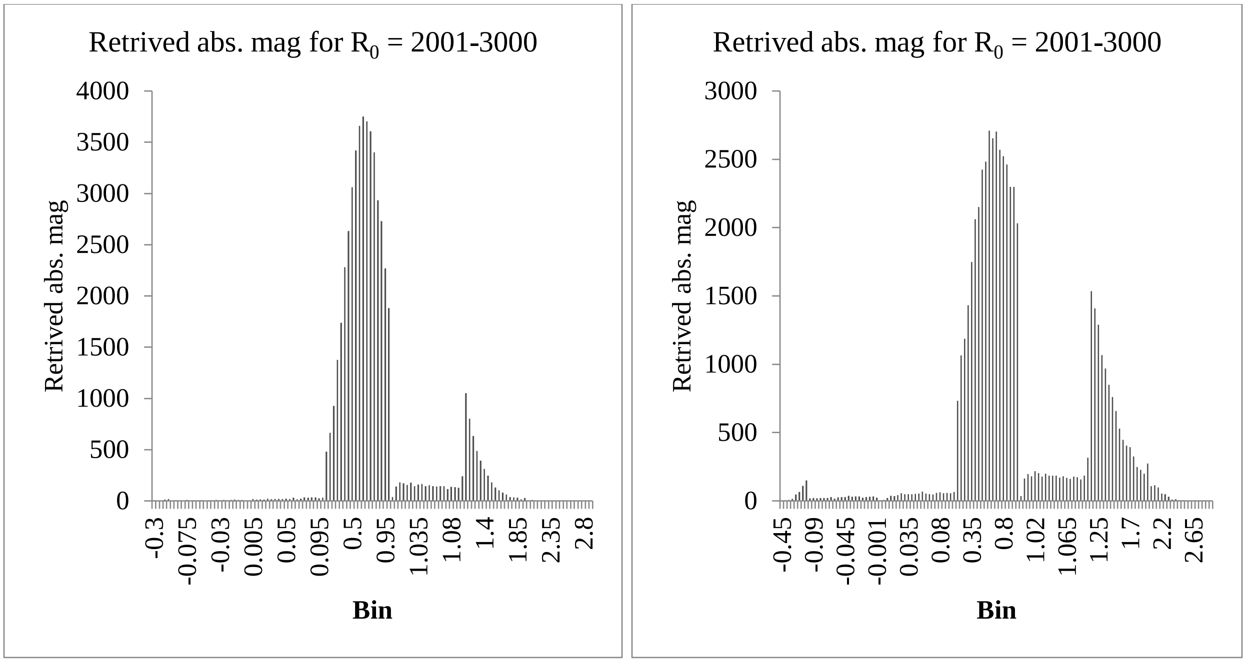


Fig. 8

The main difference between Fig 6 (Task 2) and Fig 8 are the distributions are sharper due to a inverse square spatial distribution.

Task 4

Task 1 was repeated *but* for a GAIA-like parallax error distribution.

i) For all stars within $R_0 = 4$ Kpc the mean value and dispersion were determined of the retrieved $(M_0)$and true absolute magnitudes (M) (Table 6.):

Table 6

	True abs mag (M)	Retrieved abs mag (M_0)
Mean	0.599996938	0.620290665
Dispersion	0.001037029	0.086053513

Results and conclusions similar to that of Task 1(i) except the error for M_0 is smaller by ~60% for the GAIA-like parallax.

ii) The sample was split into 4 different ranges according to retrieved distance $R_0 =$ 0-1000, 1001-2000, 2001-3000 and 3001-4000. (Table 7):

Table 7

R_0	True abs mag (M)		Retrieved abs mag M_0)	
	Mean	Dispersion	Mean	Dispersion
1-1000	0.600011364	0.001031632	0.600012626	0.01325821
1001-2000	0.599977891	0.001052463	0.600866435	0.023653642
2001-3000	0.599991296	0.001028333	0.608380997	0.05300515
3001-4000	0.600002807	0.001038637	0.630252337	0.104249444

As expected, the values of true absolute magnitude are almost identical to those of Task 1(ii). However, the GAIA-like error has contributed to reduce errors by ~75% below the 3001 mark. In the range 3001-4000 the errors have decreased by over ~50% as compared with Table 2.

a) Histograms of the true number distribution of stars with distance and retrieved distribution of stars with distance were done with a bin size of 200 (Fig. 9)

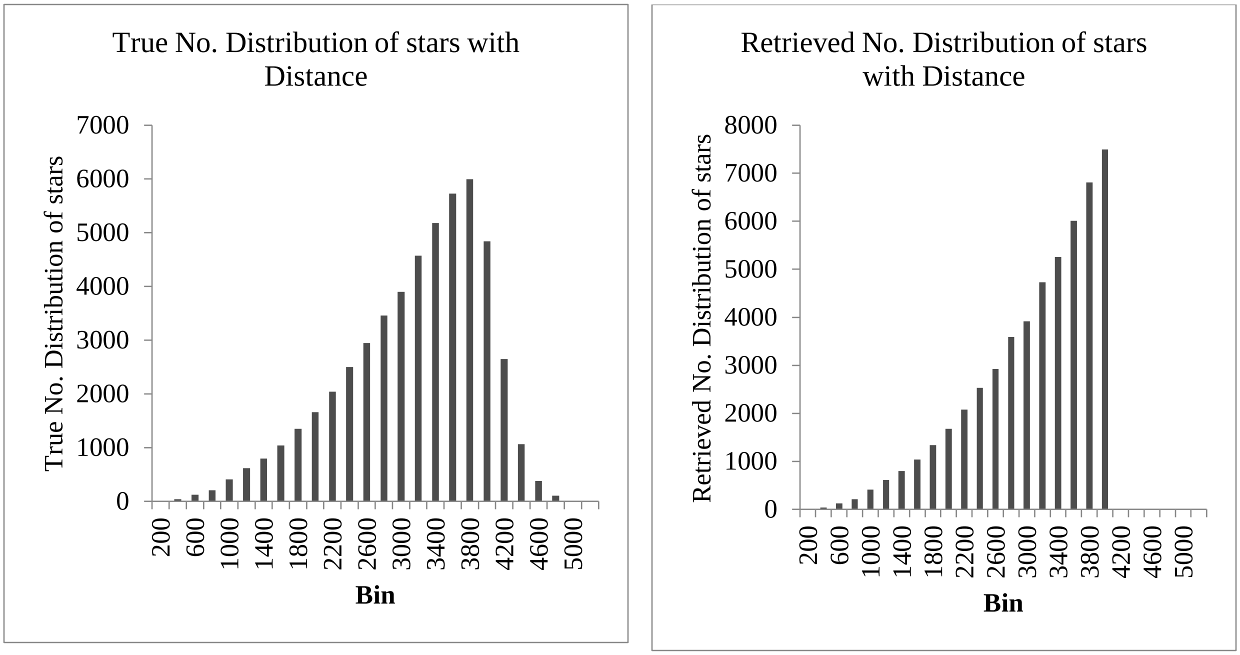


Fig. 9

The imposed GAIA-like bias shows a constant steep gradient for the retrieved magnitude in comparison with the weighted Gaussian nature of the true distribution.

b) A scatter plot of all individual retrieved absolute magnitudes as a function of the retrieved distances was produced for the data points.

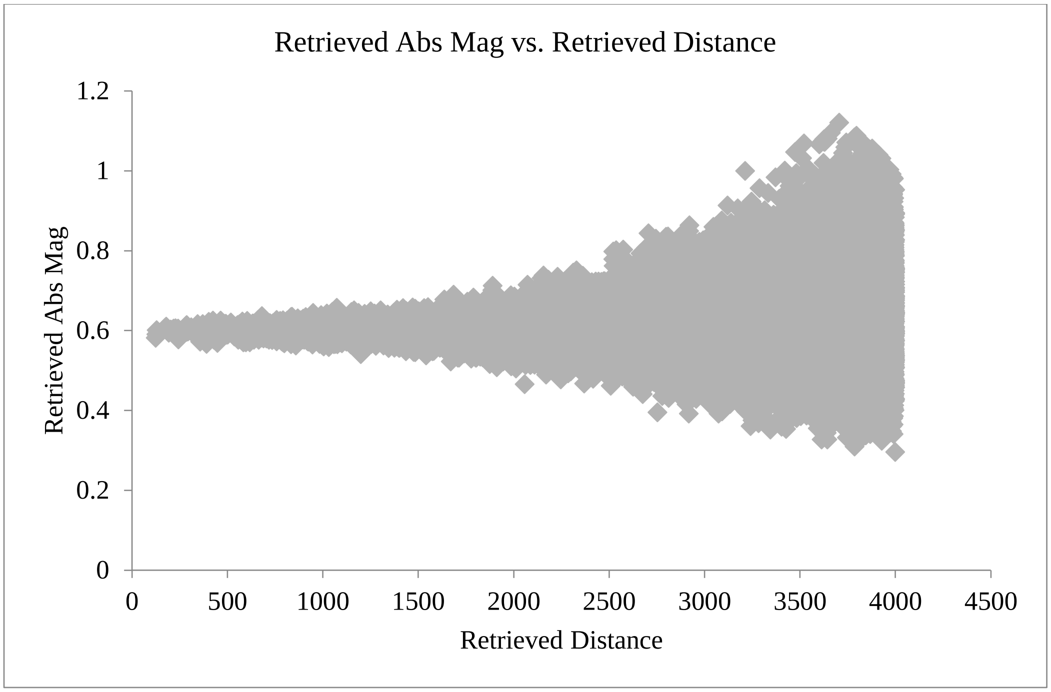


Fig.10

The scatter plot of Fig 10 as compared to that of Fig. 3 in Task 1 shows an almost identical distribution whereas comparison with those in Task 2 and Task 3 shows a clear bias induced distribution retrieved absolute magnitude.

c) Histograms of the retrieved absolute magnitudes for $R_0 =$ 0-1000, 1001-2000, 2001-3000 and 3001-4000 were plotted (Fig 11).

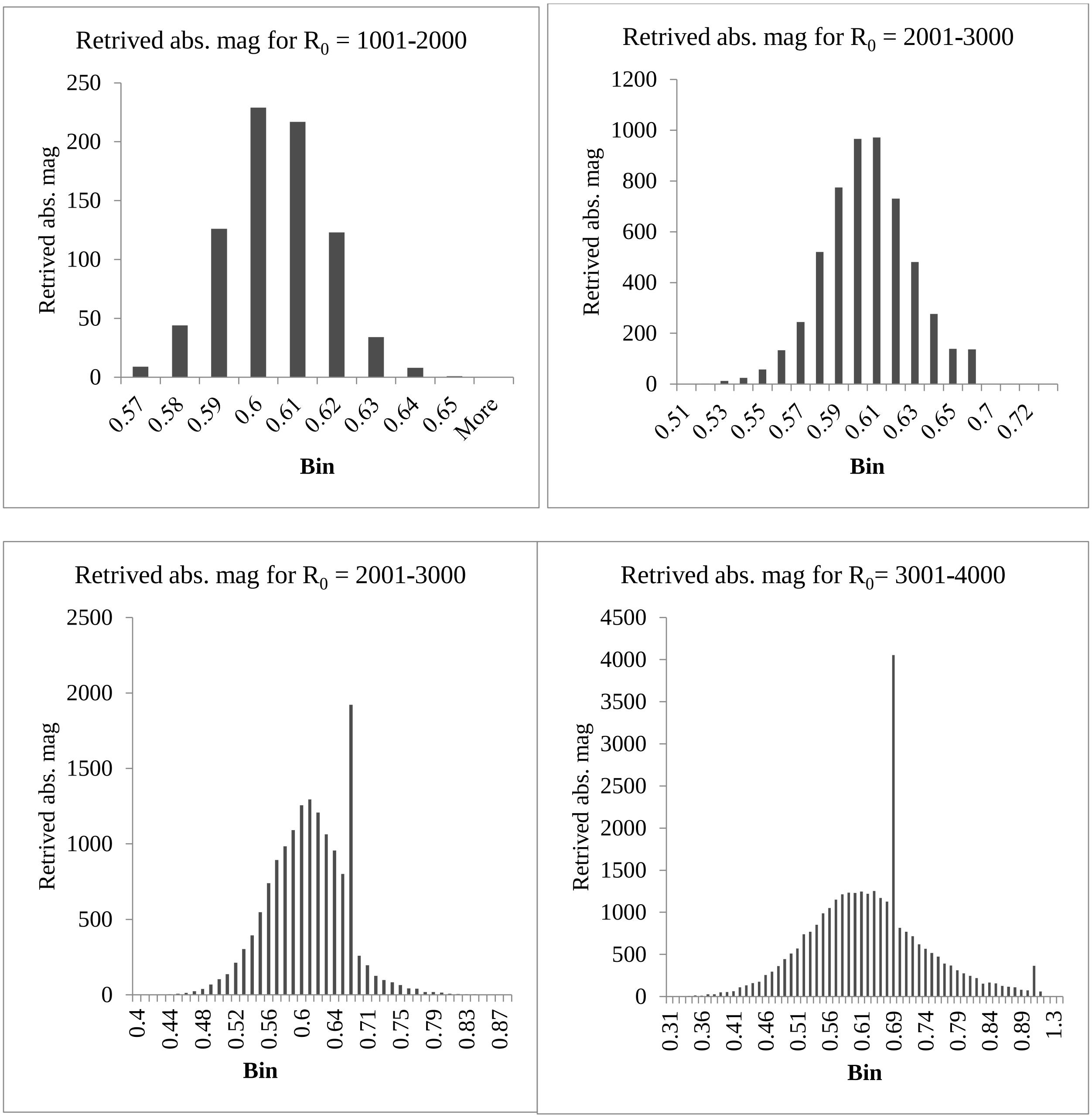


Fig. 11

The main difference between Tasks 1,2 and 3 and Fig 11 above are the distributions are all Gaussian due to a variable GAIA-like parallax error which is statistical in nature.

Conslusion.

A correction for the bias in Monte Carlo models to estimate the bias in the parallax-based distances is significant when the parallax errors are an appreciable fraction of the parallax values, but it becomes small or insignificant when the parallax errors are reduced. The comparison with models shows a good agreement. The good agreement implies that the dependency of the parallax errors on magnitude and latitude is in agreement with the expectations.

References

[1] Perryman, M. A. C. et al. (2001). GAIA: Composition, formation and evolution of the Galaxy. Astronomy and Astrophysics, v.369, p.339-363.

[2] Ratnatunga, K. U. & Casertano, S. (1991). Absolute magnitude calibration using trigonometric parallax - Incomplete, spectroscopic samples. Astronomical Journal (ISSN 0004-6256), vol. 101, March 1991, p. 1075-1088.

[3] Bailer-Jones, C.A.L. et al. (2018). Estimating Distance from Parallaxes. IV. Distances to 1.33 Billion Stars in Gaia Data Release 2. The Astronomical Journal, Volume 156, Number 2

[4] Gu J. et al.(2016). A Monte-Carlo Metod for Estimating Stellar Photometric Metallicity Distributions. The Astrophysical Journal, Volume 826, Number 1