On the good use of Gaia parallaxes

May 2018

X. Luri, ICCUB/IEEC

Basic message:

No pain, no gain

or

There ain't not such a thing as a free lunch

In scientific terms:

To extract all the possible information from parallaxes requires effort and good methodologies

I – Know thy data

Gaia astrometry

The astrometric processing is detailed in:

"Gaia Data Release 2: The astrometric solution", Lindegren & et al. (2018)

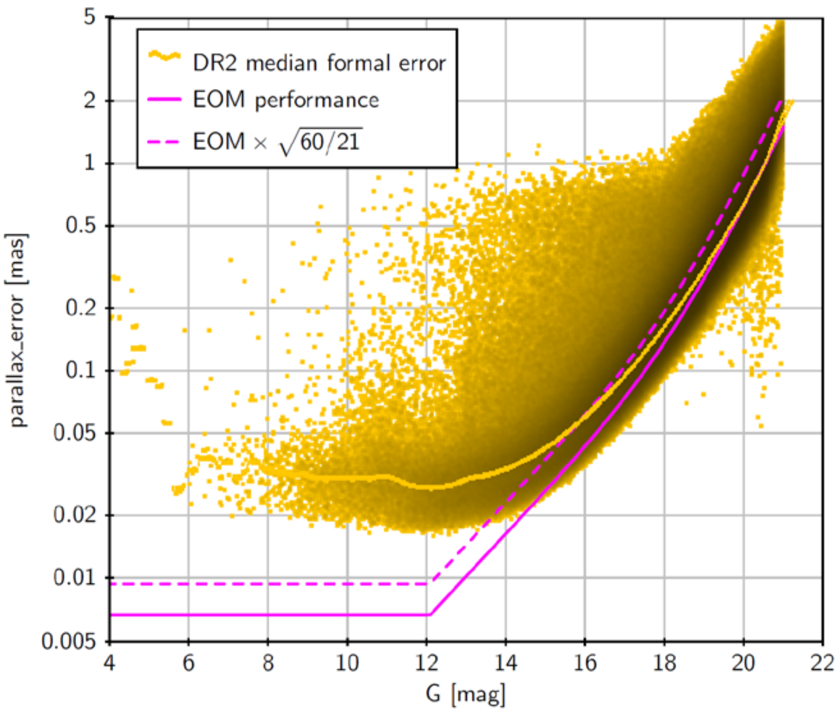
I encourage you to familiarise yourselves with the contents of that paper in order to understand the strengths and also the weaknesses of the published astrometry

Random errors

The random errors in Gaia parallaxes are characterized by the catalogue published uncertainties.

Typically:

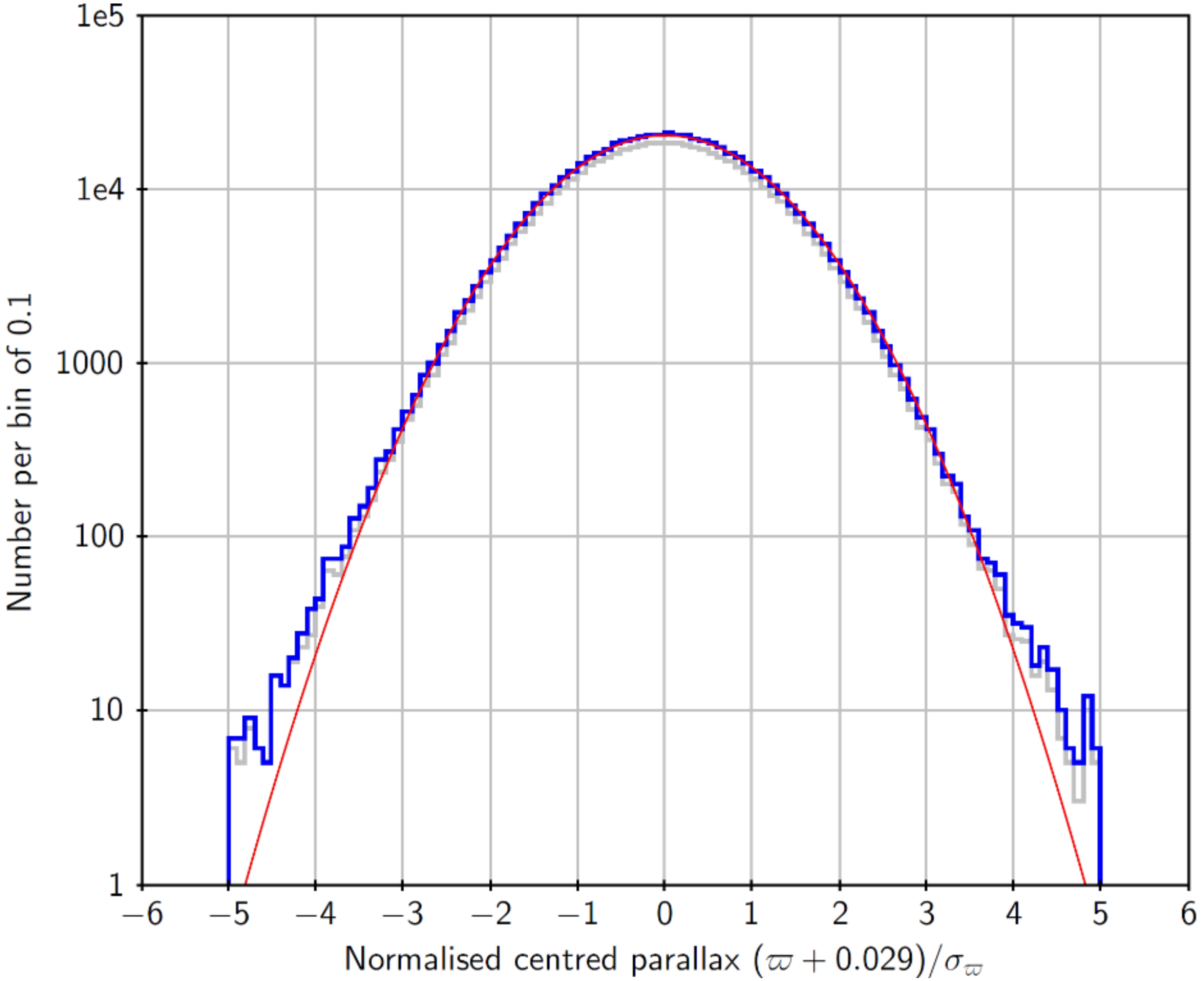
- Around 0.04 milliarcsecond for sources G<14 mag
- Around 0.1 mas for sources with a G ~17 mag
- Of the order of 0.7 mas at the faint end, around 20 mag



The astrometric uncertainties provided in Gaia DR2 have been derived from the formal errors computed in the astrometric processing.

Unlike for Gaia DR1, the parallax uncertainties have not been calibrated externally, i.e., they are known, as an ensemble, to be underestimated by 8–12 percent for faint sources (G >16 mag) outside the Galactic plane and by up to 30 percent for bright stars (G <12 mag).

Random error distribution is very approximately normal

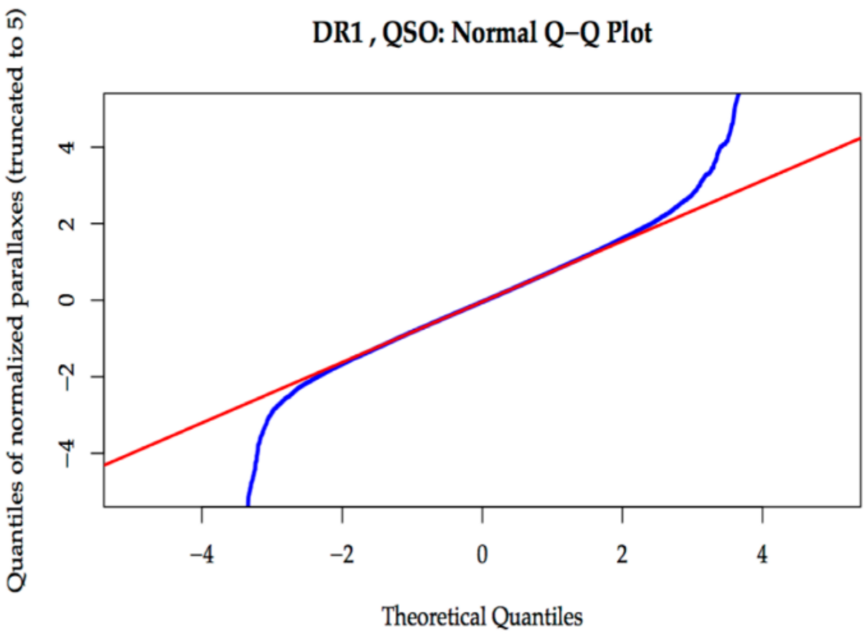


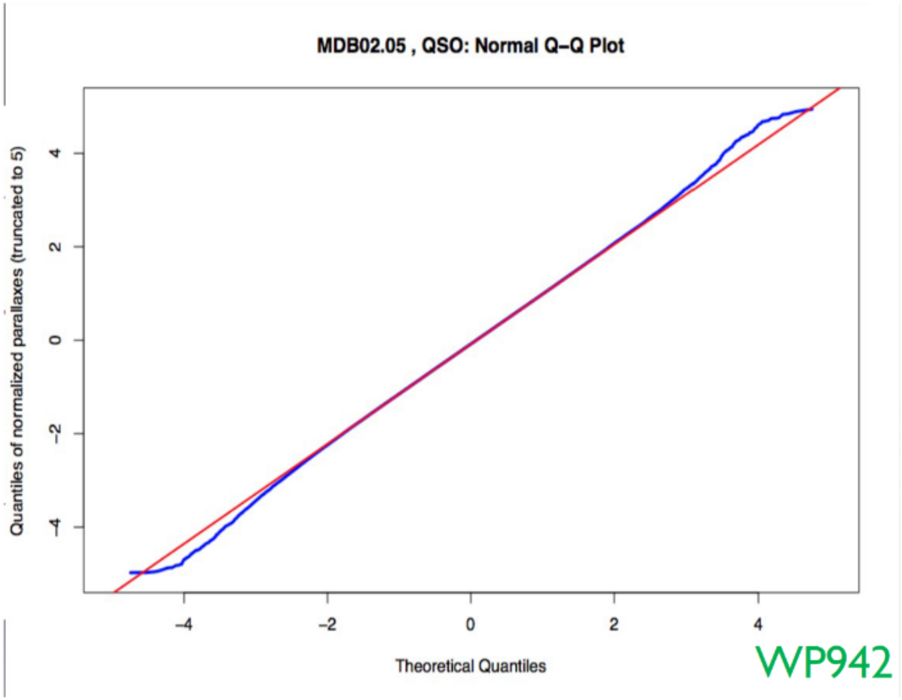
557,000
AllWISE QSO

Normalized & centered parallax distribution

DR1

DR2


DR1 , QSO: Normal Q–Q Plot
Quantiles of normalized parallaxes (truncated to 5)
−4
−2
0
2
4
Theoretical Quantiles


MDB02.05 , QSO: Normal Q–Q Plot
Quantiles of normalized parallaxes (truncated to 5)
−4
−2
0
2
4
Theoretical Quantiles
WP942

Beware of outliers

As in any (very) large catalogue outliers are present.

- For instance, there is a small number of sources with (unrealistic) very large positive and very large negative parallaxes (possibly x-match issues).

- *There are 59 stars in dr2int6 closer than Proxima…*

- Also, in some cases outliers have unrealistically high confidence levels (e.g., at the 100 level).

Systematic errors

Both the design of the spacecraft and the design and implementation of the data processing software and algorithms aim to prevent biases or systematic effects in the astrometry.

However, systematic errors, at low levels, nonetheless exist in Gaia DR2 (see Arenou & et al. 2018; Lindegren & et al. 2018).

Zero point

Average parallax zero-point shift of about -30 μas in the sense Gaia minus external data.

WP944

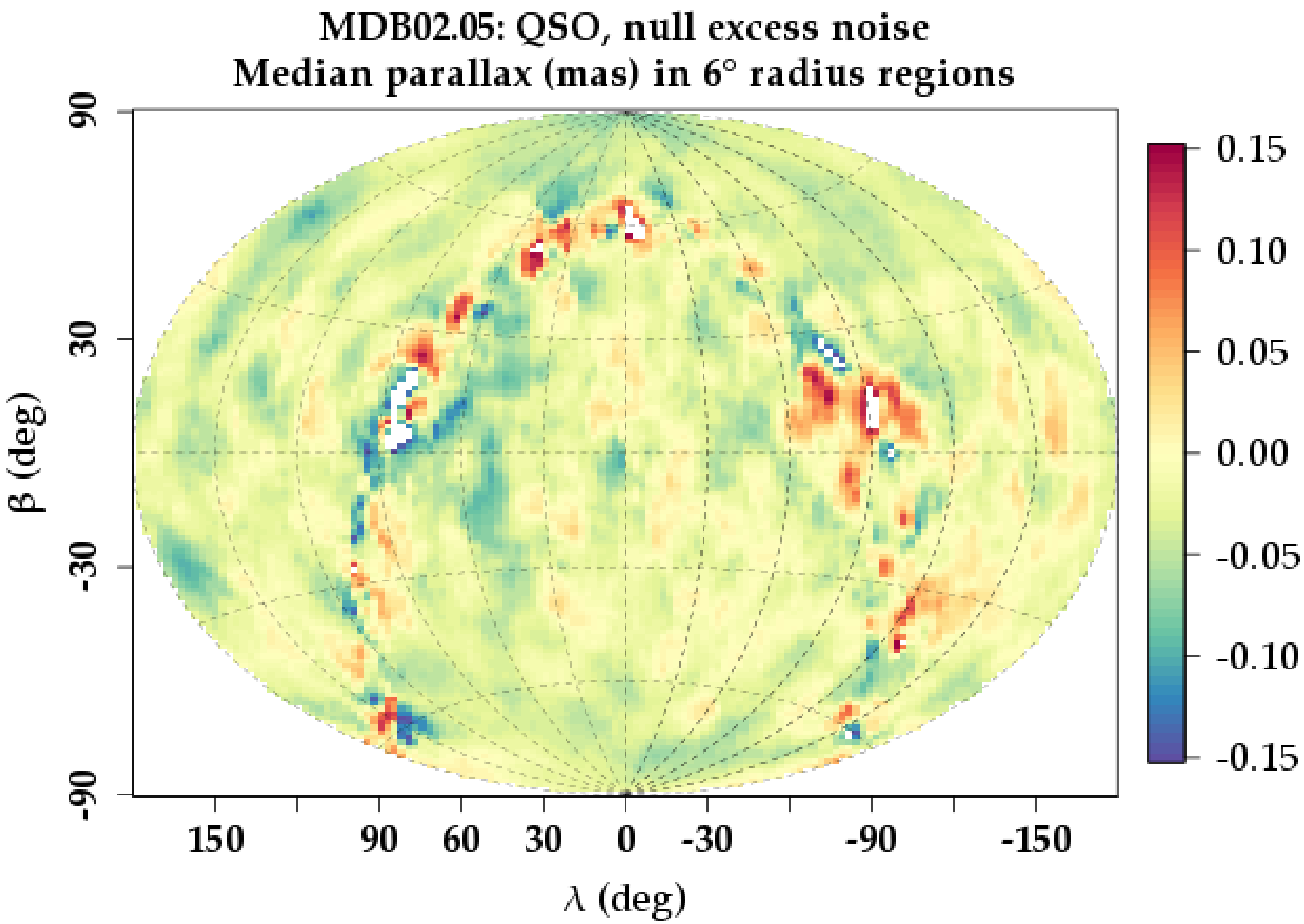
Catalogue	Outliers	ϖ residuals	ϖ sd factor (b)	Correlations
Hipparcos	0.3%	-0.118 ± 0.003	1.25 ± 0.003	mag, col, pscol, coord...
RRLyrae	26 / 1603	-0.049 ± 0.002	1.37 ± 0.02	mag, col, pscol, coord...
APOGEE	107 / 5374	-0.049 ± 0.002	1.41±0.01	mag, col, pscol, coord, ...
LAMOST	14 / 198	-0.041 ± 0.005	1.47±0.08	NA
SEGUE Kg	4 / 3244	-0.042 ± 0.002	1.07±0.01	col, pos
LMC plxpm	32997 sel.	-0.028 ± 0.0008	1.098 ± 0.004	mag, col, pscol, coord, ...
LMC Vr	9 / 317	-0.040 ± 0.001	1.46 ± 0.06	-
SMC plxpm	18936 sel.	-0.022 ± 0.0004	1.043 ± 0.005	mag, col, pscol, coord, ...
SMC Vr	12 / 115	-0.037 ± 0.002	1.3 ± 0.09	-
Draco	0 / 424	-0.047 ± 0.008	1.07 ± 0.04	-
Ursa Minor	0 / 78	-0.054 ± 0.008	0.98 ± 0.08	-
Sculptor	4 / 1252	-0.030 ± 0.006	1.10 ± 0.02	-
Sextans	1 / 375	-0.09 ± 0.02	1.07 ± 0.04	pscol
Carina	0 / 858	-0.021 ± 0.007	1.04 ± 0.03	-
Crater2	0 / 63	-0.05 ± 0.03	0.95 ± 0.09	-
Fornax	11 / 2612	-0.052 ± 0.004	1.21 ± 0.03	mag, col
LeoII	0 / 123	0.05 ± 0.05	1.0 ± 0.06	-
LeoI	2 / 292	-0.25 ± 0.07	1.30 ± 0.05	pscol
Phoenix	0 / 80	0.092 ± 0.08	1.08 ± 0.09	-
all dSph	18 / 6206	-0.044 ± 0.003	1.12 ± 0.01	mag, col, pscol, coord
ICRF2	3 / 2327	-0.039 ± 0.003	1.13 ± 0.02	mag
LQRF	25 / 77958	-0.0320 ± 0.0009	1.077 ± 0.003	mag, coord...
RFC2016cnoU	4 / 3478	-0.038 ± 0.003	1.12 ± 0.01	mag

Arenou & et al. 2018

Regional effects

Significant spatial correlations between stars, up to 0.04 mas in parallax exist on both small (<1º) and intermediate (<20º) angular scales.

As a result, averaging parallaxes over small regions of the sky, for instance in an open cluster, in the Magellanic Clouds, or for the Galactic Center, will not reduce the uncertainty on the mean below the ~0.1 mas level.



Peak-to-peak differences between small regions (5º radius) can go up to 0.13 mas

On the good use of Gaia parallaxes

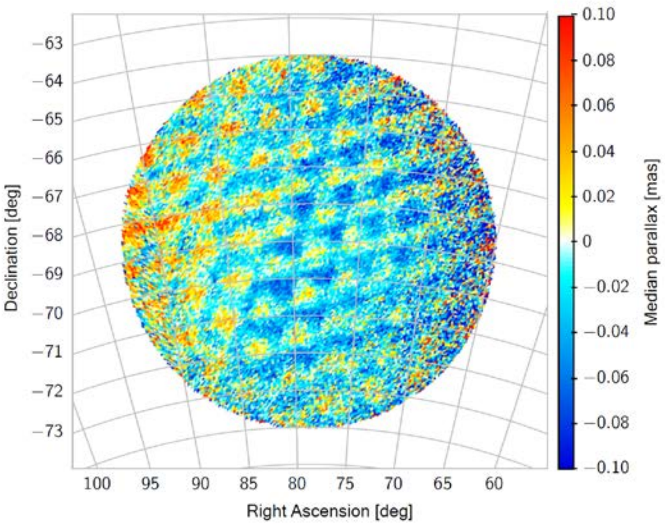
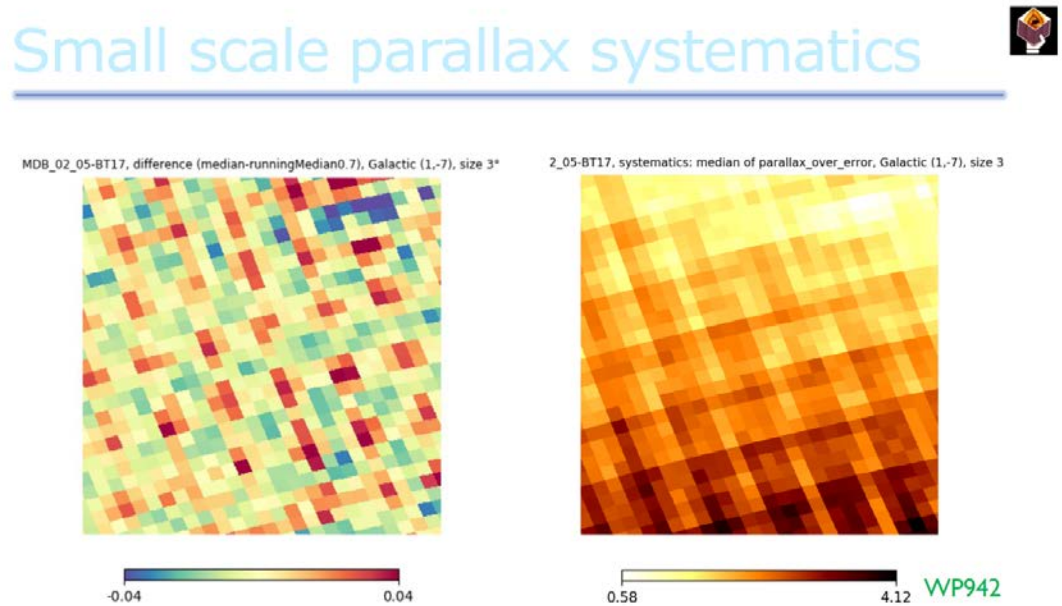


Fig. 13. Map of the median parallaxes for a sample of sources in th LMC area, showing small-scale variations of the parallax zero point. Median values are calculated in cells of about 0.057×0.057 deg^2.



From Lindegren et al. (2018)

Table 4. Summary of estimated systematics for faint sources ($G \gtrsim 16$ mag).

Angular scale	Parallax [μas]	Proper motion [μas yr^{-1}]	Remark	Reference
global	−29	15	offset (zero point or spin)	Fig. 6 and Mignard et al. (2018)
∼ 90 deg	5	18	RMS value	Fig. 7 and Mignard et al. (2018)
∼ 14–20 deg	17	28	RMS value	Eqs. (16) and (18)
< 1 deg	43	66	RMS value	Figs. 14 and 15

Notes. Columns 2 and 3 give the estimated offset or RMS level of systematics in parallax and proper motion based on the analysis described in the reference. The RMS values should be interpreted as the noise floor when averaging many sources at the given angular scale. For bright sources (especially $G \lesssim 13$ mag) the systematics may be significantly larger; see for example Fig. 4.

Summary of errors

$$\epsilon_{\varpi} \rightarrow \sigma_{random} + \sigma_{systematic} + z_{point}$$

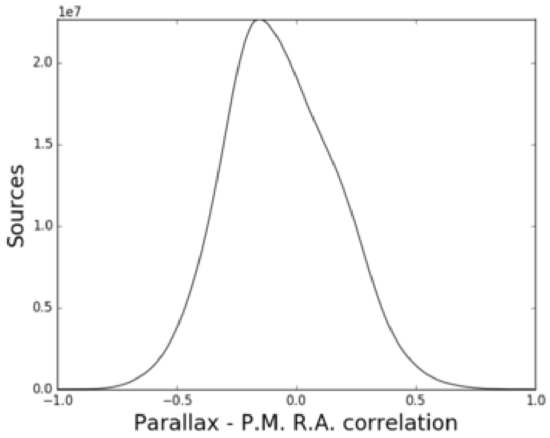
Gaussian, with some outliers. Given in the catalogue but can be underestimated.

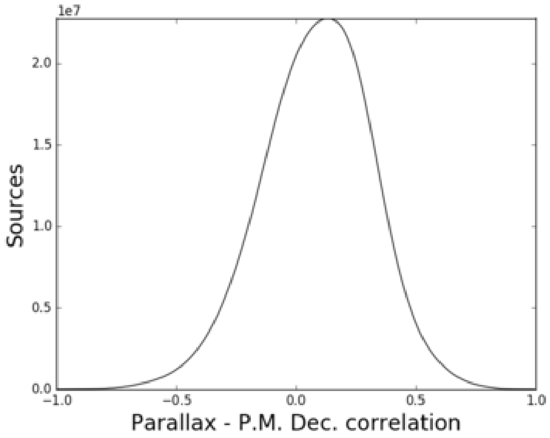
Regional, not gaussian, precision floor ~0.1mas

Global ~ -0.030mas

Correlations

The five astrometric parameters of Gaia objects are correlated. Correlation matrix is provided in the catalogue; it is important to take into account when jointly using $(\alpha,\delta,\mu_\alpha,\mu_\delta,\varpi)$





II – Distances and parallaxes

Important:

Unless you are working for DPAC, most likely you are not interested in the parallaxes "per se", but in some derived quantity like distances or absolute magnitudes.

This is not as straightforward as it may seem.

Inverting distance

A naive approach to distance estimation is

$$r = \frac{1}{\varpi}$$

Let's take a closer look.

True vs. observed

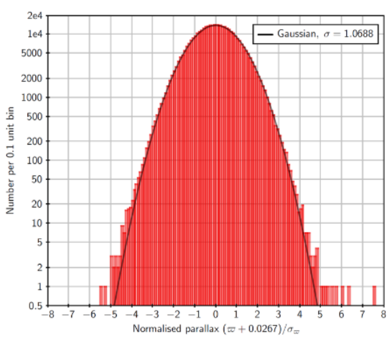
We do not have access to the **true** parallax, but to the **observed** parallax, and:

$$r = \frac{1}{\varpi_{True}}$$

What happens if we use ϖ instead of ϖ_{True} in the equation above? Not trivial…

The parallax as a random variable

In the rest of the talk I will assume that the observed parallax ϖ can be treated as a random variable distributed normally around the true parallax ϖ_{True}



Normalised, centred parallaxes $(\varpi + 0.0267)/\sigma_\varpi$

Mean = −0.0010
Median = −0.0000

Std = 1.0688
RSE = 1.0639

$$p(\varpi \mid \varpi_{\mathrm{True}}) = \frac{1}{\sigma_\varpi \sqrt{2\pi}} \exp\left(-\frac{(\varpi - \varpi_{\mathrm{True}})^2}{2\sigma_\varpi^2}\right)$$

(ignoring systematic errors)

Now let's imagine that in order to estimate the distance of a star we follow the usual reciepe and use the naive distance estimate:

$$\rho \equiv \frac{1}{\varpi}$$

Let's explore the consequences

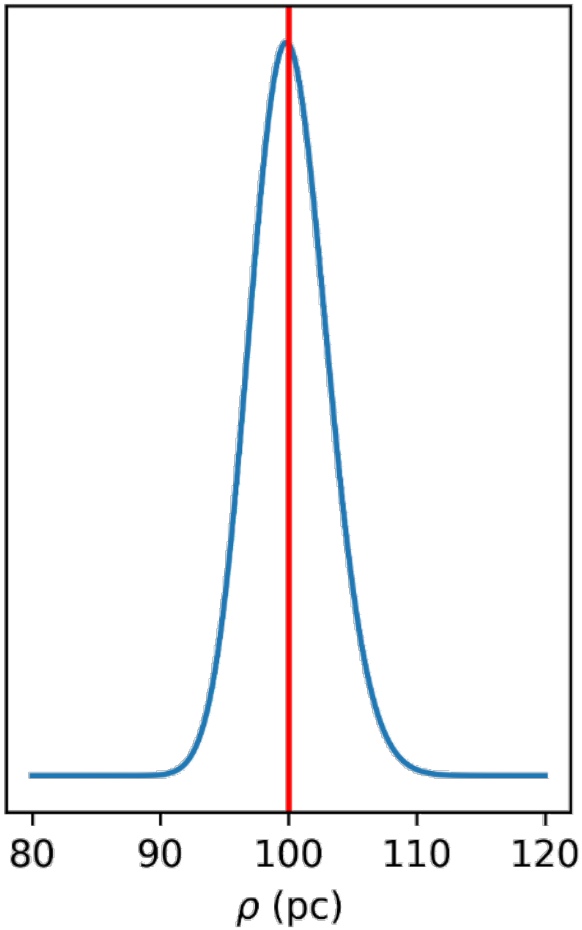
(notice that I am not calling this "r")

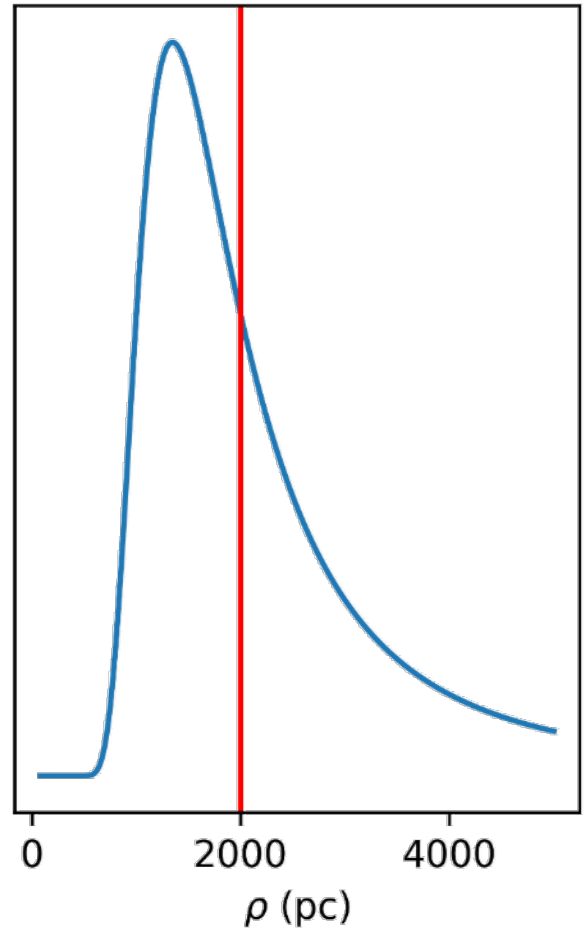
If ϖ is a random variable normally distributed around ϖ_{True} we can derive the PDF of ρ from it.

$$p(\rho \mid \varpi_{\text{True}}) = p(\varpi = 1/\rho \mid \varpi_{\text{True}}) \cdot \left| \frac{d\varpi}{d\rho} \right|$$

$$= \frac{1}{\rho^2 \sigma_\varpi \sqrt{2\pi}} \exp\left(-\frac{(1/\rho - \varpi_{\text{True}})^2}{2\sigma_\varpi^2} \right)$$

Two extreme cases





r= 100pc σ_{ϖ}= 0.3mas

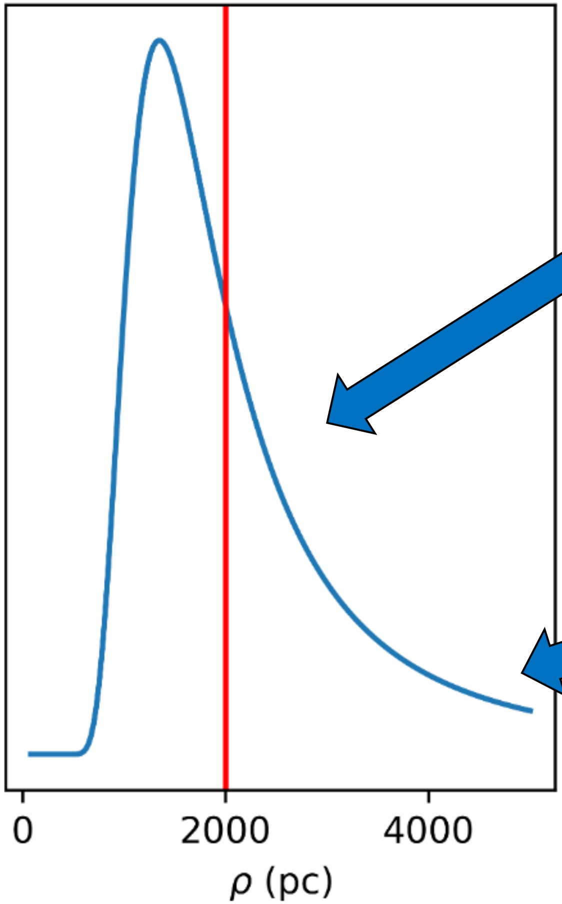
r= 2000pc σ_{ϖ}= 0.3mas

The mode does not coincide with the true value of r

Asymmetrical distribution

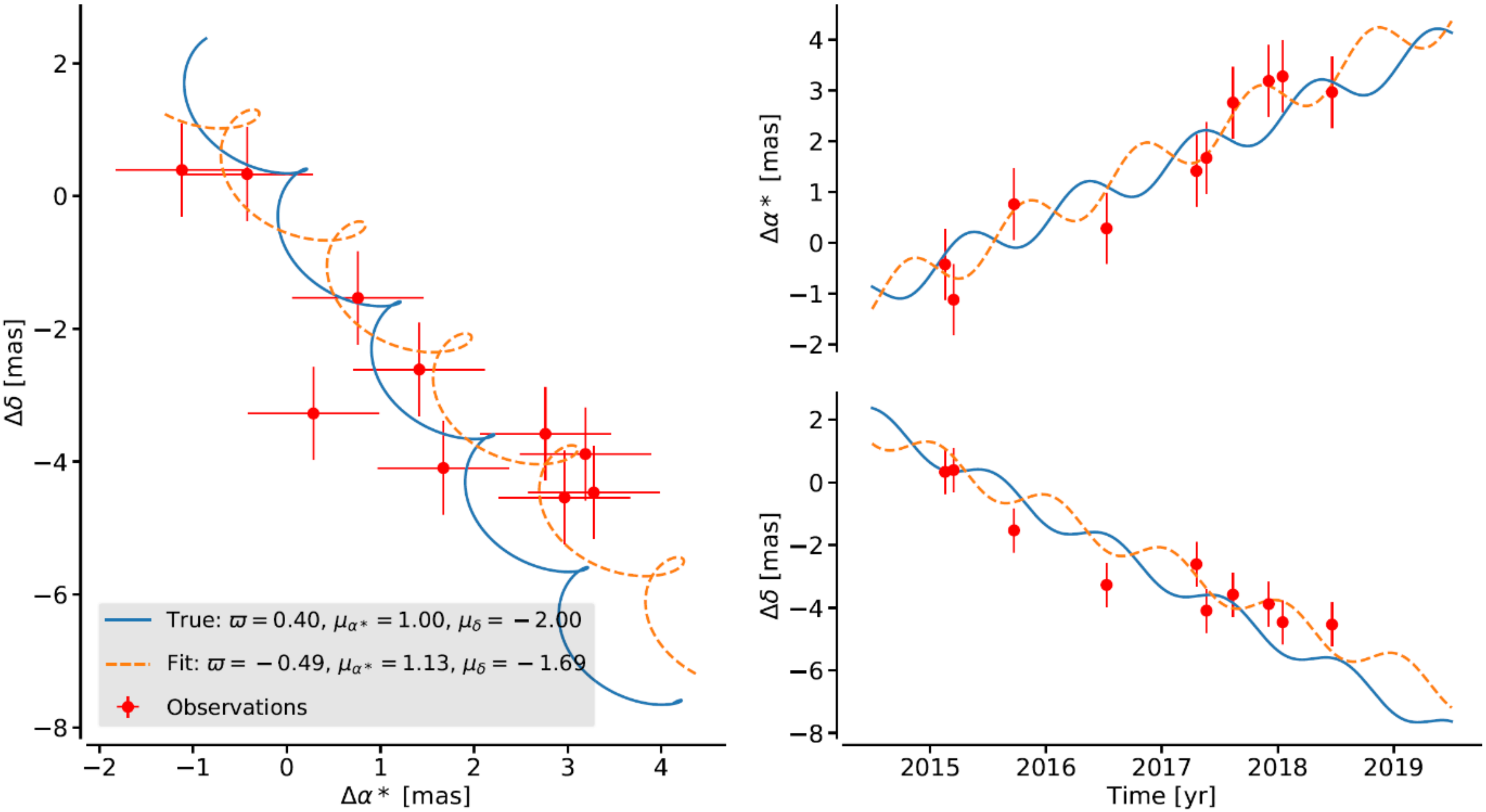
Negative values

Very long positive tail



Negative parallaxes

A normal distribution of errors can lead to **negative parallaxes**. This is an expected result of the Gaia data processing, based on a fitting of the five (six) astrometric parameters to the path of the stars in the sky.



True: $\varpi = 0.40$, $\mu_{\alpha*} = 1.00$, $\mu_\delta = -2.00$
Fit: $\varpi = -0.49$, $\mu_{\alpha*} = 1.13$, $\mu_\delta = -1.69$
Observations
Δα* [mas]
Δδ [mas]
Time [yr]

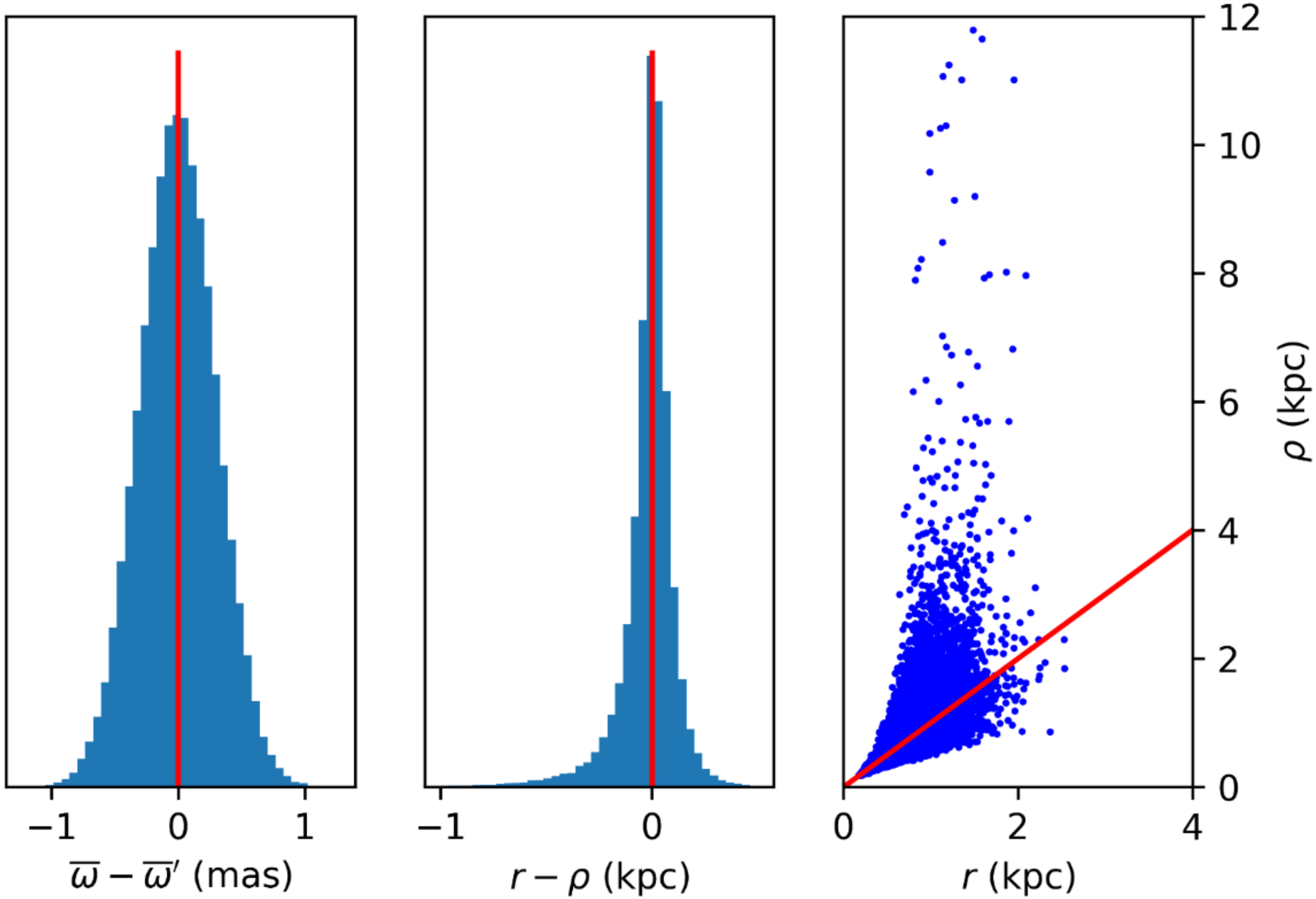
Note: ρ can be negative if the parallax is negative. Or can be infinite if the parallax is zero → <u>unphysical distance vàlues</u>

This is a first problem of ρ as an estimator of distance

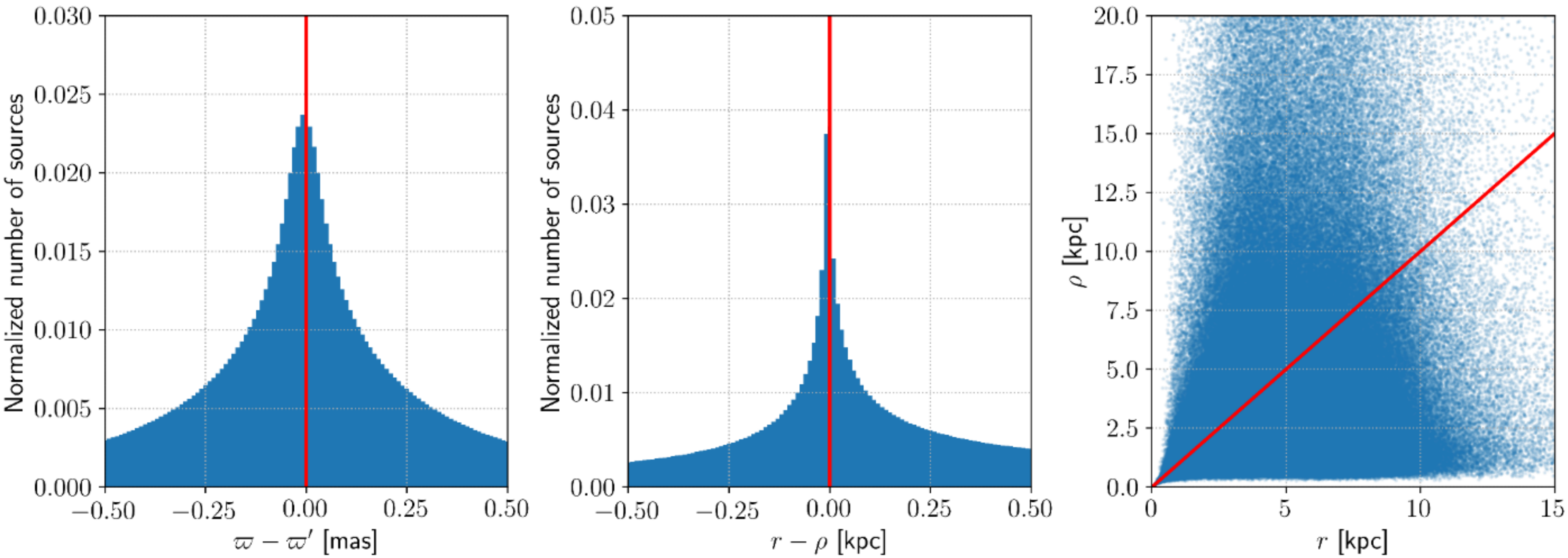
Properties of ρ

- It is a biased estimator of the true distance r.
- It is a high-variance estimator (large errors); the variance can quickly become pathological.
- The distribution is highly asymmetrical, leading to asymmetrical errors. There is a long tail for positive values.

Simple simulated sample with σ_ϖ= 0.3mas



DR2 simulation



Summary

$\rho \equiv \frac{1}{\varpi}$ is just a distance estimator (one among many), not the true distance. And it is not a specially good one.

It is handy and can be used for preliminary exploratory work, but it is not recommended for detailed analysis. <u>If used, always quote and take into account the asymmetrical error.</u>

End of this talk
(but not of the advice)

A paper discussing in detail the use of Gaia parallaxes

Astronomy & Astrophysics manuscript no. main ©ESO 2018
February 25, 2018

On the use of Gaia parallaxes

X. Luri[1], A. Brown[4], L. Sarro[6], F. Arenou[2], C.A.L. Bailer-Jones[3], A. Castro-Ginard[1], J. de Bruijne[5], T. Prusti[5], C. Babusiaux[2], and H.E. Delgado[6]

[1] Dept. Fisica Quantica i Astrofisica, Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí Franquès 1, E08028 Barcelona, Spain

[2] GEPI, Observatoire de Paris, Université PSL, CNRS, 5 Place Jules Janssen, 92190 Meudon, France

[3] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany

[4] Sterrewacht Leiden, Leiden University, P.O. Box 9513, 2300 RA, Leiden, The Netherlands

[5] Scientific Support Office, Directorate of Science, European Space Research and Technology Centre (ESA/ESTEC), Keplerlaan 1, 2 201AZ, Noordwijk, The Netherlands

[6] Dpto. Inteligencia Artificial, UNED, Juan del Rosal, 16, 28040 Madrid

Received date / Accepted date