Astronomy and Astrophysics in the Gaia sky
Proceedings IAU Symposium No. 330, 2017
A. Recio-Blanco, P. de Laverny, A.G.A. Brown
& T. Prusti, eds.

doi:10.1017/S1743921317006664

The Gaia sky: version 1.0

Anthony G. A. Brown

Sterrewacht Leiden, Leiden University, Niels Bohrweg 2, 2333 CA, Leiden, Netherlands
email: `brown@strw.leidenuniv.nl`

Abstract. In this contribution I provide a brief summary of the contents of *Gaia* DR1. This is followed by a discussion of studies in the literature that attempt to characterize the quality of the *Tycho-Gaia* Astrometric Solution parallaxes in *Gaia* DR1, and I point out a misconception about the handling of the known systematic errors in the *Gaia* DR1 parallaxes. I highlight some of the more unexpected uses of the *Gaia* DR1 data and close with a look ahead at the next *Gaia* data releases, with *Gaia* DR2 coming up in April 2018.

Keywords. catalogs, surveys, astrometry

1. Overview of *Gaia* DR1

With the announcement on September 14 2016 of the first data release from the ESA *Gaia* mission (*Gaia* DR1) the astronomical community truly entered the *Gaia* era. This data release is the culmination of over 10 years of effort by ESA and the members of the *Gaia* Data Processing and Analysis Consortium (DPAC), the community of European astronomers responsible for the data processing for the *Gaia* mission. The *Gaia* satellite was launched in December 2013 to collect data that will allow the determination of highly accurate positions, parallaxes, and proper motions for over one billion sources brighter than magnitude $G = 20.7$ in the *Gaia* white-light photometric band. The astrometry is complemented by multi-colour photometry, measured for all sources observed by *Gaia*, and radial velocities which are collected for stars brighter than $G_{\mathrm{RVS}} \sim 16$ in the pass-band of Gaia's radial velocity spectrograph. The scientific goals of the mission and the scientific instruments on board *Gaia* are summarised in Gaia Collaboration, *et al.* (2016a). The raw data collected during the first 14 months of the mission were processed by the DPAC, involving some 450 astronomers and IT specialists, and turned into the first version of the *Gaia* catalogue of the sky (Gaia Collaboration, *et al.* 2016).

The bulk of *Gaia* DR1 consists of celestial positions (α, δ) and G-band magnitudes for about 1.1 billion sources. The distribution of the *Gaia* DR1 sources in magnitude is show in Fig. 1. With median positional accuracies of 2.3 milli-arcsec (mas) and a spatial resolution comparable to the Hubble Space Telescope, the *Gaia* DR1 catalogue represents the most accurate map of the sky to date, including the most precise and homogeneous all-sky photometry, ranging from milli-magnitude uncertainty at the bright end of the *Gaia* survey to 0.03 magnitude uncertainty at the faint end. In addition the combination of *Gaia* data and the positions from the *Hipparcos* and *Tycho-2* catalogues allowed the derivation of highly precise proper motions and parallaxes for the 2 million brightest sources in *Gaia* DR1 (the so-called *Tycho-Gaia* Astrometric Solution or TGAS; Michalik *et al.* 2015, Lindegren *et al.* 2016). The typical parallax uncertainty is 0.3 mas, while the proper motion uncertainties are about 1 mas yr^{-1} for the stars from the *Tycho-2* catalogue and as small as 0.06 mas yr^{-1} for the stars from the *Hipparcos* catalogue. *Gaia* DR1 in addition contains light curves and variable type classifications for a modest sample of some 2600 RR Lyrae and 600 Cepheid variables (Clementini *et al.* 2016) as

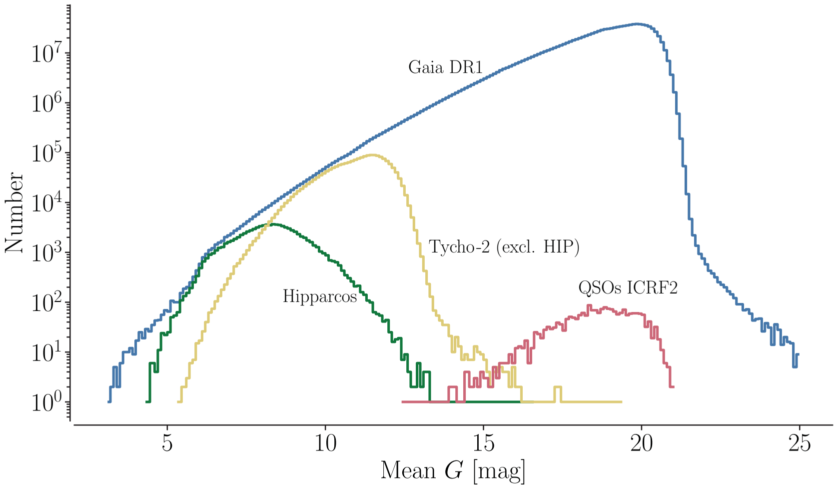

Figure 1. Magnitude distribution of the sources in the *Gaia* DR1 catalogue. The TGAS magnitude distribution is split into the *Hipparcos* stars and the stars from the *Tycho-2* catalogue (excluding *Hipparcos*). The distribution of the magnitude of the $\sim$ 2000 ICRF2 QSOs is also shown separately.

well as the optical positions of about 2000 ICRF2 sources (Mignard *et al.* 2016). More details can be found in Gaia Collaboration, *et al.* (2016).

2. On the quality of the TGAS parallaxes

Since the publication of *Gaia* DR1 various papers have treated the topic of the quality of the TGAS parallaxes, in particular focusing on the possibility of systematic offsets with respect to independent parallax measurements or distance estimates. The *Gaia* DR1 catalogue validation done by DPAC (Arenou *et al.* 2017) confirms the estimate by Lindegren *et al.* (2016) that the global parallax zero-point for TGAS is at the ± 0.1 mas level (an average of ~ -0.04 mas for various comparison samples was found, see table 2 in Arenou *et al.* 2017). The validation effort also confirmed the conclusion by Lindegren *et al.* (2016) that locally additional offsets at the ± 0.2 mas level can exist (see for example figure 24 in Arenou *et al.* 2017). The latter are spatially correlated and colour-dependent. Analyses of the period-luminosity relations of local Cepheids and RR Lyrae revealed no global offset in the TGAS parallaxes to $\sim$0.02 and $\sim$0.05 mas precision, respectively (Casertano *et al.* 2016, Sesar *et al.* 2017). In contrast Jao *et al.* (2016) and Stassun & Torres (2016) find rather large systematic offsets of 0.24 and 0.25 mas, where the Gaia parallaxes are too small by these amounts compared to independently measured or predicted parallaxes for a sample of nearby stars and a sample of eclipsing binaries, respectively.

De Ridder *et al.* (2016) analyzed a sample of nearby (parallaxes larger than $\sim$5 mas) dwarfs and sub-giants as well as a sample of more distant red giants (parallaxes less than $\sim$3 mas), for which asteroseismic distances are available based on data from the Kepler mission (Borucki *et al.* 2010). In both cases the asteroseismically derived distances were converted to parallaxes and a linear relationship, $\varpi_{\mathrm{predicted}} = \alpha + \beta\varpi_{\mathrm{TGAS}}$, was fit between these parallaxes and the TGAS values. For the dwarfs and sub-giants the one-to-one relation between the parallax values, with a zero offset, was found to be a plausible model, while for the red giants De Ridder *et al.* (2016) find $\alpha \sim 0.3$ mas and $\beta \sim 0.75$, implying that the TGAS parallaxes are too small. Davies *et al.* (2017) used a sample of Red Clump (RC) stars in the Kepler field to assess the TGAS parallaxes. They predicted the parallaxes from the apparent magnitudes of the RC stars and the assumed mean absolute magnitude for the RC, using the K_s band. Fitting a linear relation they find $\alpha \sim 0.24$ and $\beta \sim 1.64$, implying that the TGAS parallaxes are too large. Finally, Huber *et al.* (2017) present an analysis of 2200 stars in the Kepler field, from the main sequence to the giant branch. They conclude that the offsets between the TGAS parallaxes and distances derived from the properties of eclipsing binaries and asteroseismic samples have been overestimated for parallaxes in the 5–10 mas range and find a significantly smaller deviation than De Ridder *et al.* (2016) for smaller parallaxes. They also find that the remaining differences can be partially compensated by adopting a hotter T_eff scale, leaving differences between TGAS parallaxes and asteroseismic distances at the ∼2 per cent level.

These apparently contradictory conclusions on the differences between TGAS parallaxes and independent distance indicators merit a couple of remarks:

- Spurious differences between two sets of parallax measurements can occur due to the effects of truncating the sample on the value of the parallax or that of the relative parallax uncertainty. For example, when comparing parallax measurements that differ in precision and truncating the sample on the value of the lower precision parallaxes it can be shown (assuming no systematic errors in either sample and Gaussian errors) that for a case similar to the Jao *et al.* (2016) study (using only parallaxes larger than 40 mas in the comparison drawn from an underlying sample reaching to 25 mas in parallax) the mean difference in the parallaxes is expected to be ~ -0.1 mas, in the sense of high minus low precision parallax values. This is of the order of the offsets claimed and can be understood when considering that truncation on the low precision parallaxes combined with the steep increase in stars toward smaller true parallaxes, leads to an excess of stars in the low precision sample having overestimated parallaxes. It should be stressed that Jao *et al.* (2016) are aware of this issue and have made the comparison for various subsamples and also by truncating on the value of the high precision (TGAS) parallaxes. The differences they find cannot readily be explained away by the sample truncation effect only.
- Similarly a truncation of the sample on apparent magnitude can introduce a spurious difference between two sets of parallax estimates. This issue could play a role in the studies that use standard candles, such as RC stars, to estimate parallaxes from the apparent magnitude. In the study of Davies *et al.* (2017) many of the RC stars have apparent magnitudes around the completeness limit of TGAS (this is also pointed out in the work by Gontcharov & Mosenkov, 2017). This will lead to favouring the intrinsically brighter RC stars for which the parallaxes will be underestimated when using the mean RC absolute magnitude to calculate the parallax. This specific effect is probably not very important in the Davies *et al.* (2017) study if the intrinsic spread in RC star absolute magnitudes is as small as derived in Hawkins *et al.* (2017). In general the properties of a selected sample, the properties of the parent population it is drawn from, as well as the survey selection functions (in both TGAS and other surveys the samples are drawn from) need to be well understood in order to properly interpret any differences between different sets of parallaxes.
- Offsets between TGAS and other parallax estimates that increase with the size of the parallax could be indicative of a scaling error in the non-astrometric parallaxes. Silva Aguirre *et al.* (2017), Huber *et al.* (2017), and Yildiz *et al.* (2017) point out the possibility that the T_eff scale used in the calculation of asteroseismic distances (which scale as $T_\mathrm{eff}^{2.5}$) could partly explain the offsets between TGAS and their distances estimates. Likewise an underestimate of the absolute magnitude in the case standard candles are used could explain the trend seen in the study of Davies *et al.* (2017), where the ratio of estimated to

true parallax scales as $10^{0.2\Delta M}$, with ΔM the difference between the estimated and true absolute magnitude. However $\Delta M \sim 1$ is required to explain the effect found by Davies *et al.* (2017), and it is highly unlikely that the existing Red Clump absolute magnitude calibrations are in error by that amount. An overestimate of the extinction has a similar effect but again an implausibly large overestimate of the extinction toward the Kepler field would be required.

- Comparisons of TGAS parallaxes to independent parallax measurements or to parallaxes estimated by other means should always consider that systematic effects can occur in either set of parallaxes. Independently estimated distances that are converted to parallaxes (using $\varpi = 1/d$) will suffer from the same non-linear transformation problems as when calculating distances as $1/\varpi$. Depending on the relative distance error, the error on the predicted parallax will be highly non-Gaussian and the mean can be biased away from the true value.
- The actual properties of the parallax uncertainties in both TGAS and the independent parallax measurements/estimates play an important role. All studies above implicitly assume normally distributed errors for which the values are correct. However Lindegren *et al.* (2016) show that the distribution of the normalized TGAS-Hipparcos parallax differences contains exponential tails, while its width hints at uncertainties being underestimated in one or both data sets. The differences between TGAS and previously published trigonometric parallaxes are modelled as Lorentzians by Jao *et al.* (2016) in order to accommodate for extended wings, again hinting at partly non-Gaussian errors in the parallaxes of one or both samples. The parallaxes estimated from astrophysical properties of a particular sample of stars can likewise suffer from non-Gaussianity and/or under- or overestimation of the errors. Errors that deviate significantly from normal behaviour will amplify the effects discussed above (sample truncation, scale errors in distance estimators).
- A number of studies rely on stars from the Kepler field only and extrapolating the results to the entire TGAS catalogue is a dubious undertaking. The special validation solutions for TGAS discussed in appendix E of Lindegren *et al.* (2016), as well as the QSO analysis in Arenou *et al.* (2017) show that there are regional systematics on the sky. Stars distributed over the Kepler field may well suffer from similar parallax systematics as suggested by figure E.1 in Lindegren *et al.* (2016). However the offsets between TGAS and independent parallax estimates for the Kepler field should not blindly be extrapolated to other regions on the sky.

The studies of TGAS parallaxes carried out so far have mostly not or only partly addressed the above issues. This makes it difficult to come to clear interpretation of the offsets seen between TGAS parallaxes and alternative parallax measurements or estimates. I conclude that there is no strong reason to revise the estimates of the level of systematic errors in TGAS parallaxes as presented in Lindegren *et al.* (2016) and Arenou *et al.* (2017).

2.1. *TGAS parallax uncertainties, treatment of (spatially correlated) systematics*

As described in Lindegren *et al.* (2016) an inflation factor was applied to the formal uncertainties on the astrometric parameters (as determined in the astrometric data processing) to arrive at the uncertainties quoted in the *Gaia* DR1 catalogue. This inflation factor was derived from a comparison between the TGAS and the Hipparcos parallaxes. There are indications in various studies that the inflated parallax uncertainties may be overestimated at the 10–20% level (Casertano *et al.* 2016, Gould *et al.* 2016, Sesar *et al.* 2017). If desired the inflation factor applied in Lindegren *et al.* (2016) can be undone (see their section 4.1) but then this factor should be re-estimated as part of the data analysis (see Sesar *et al.* 2017, for an example).

Since the publication of *Gaia* DR1 there has been some confusion in the astronomical community (also reflected in the literature) as to how to deal with the systematic uncertainties known to be present in the TGAS astrometry, in particular for the parallaxes. This was partly caused by a misleading statement in the paper describing *Gaia* DR1 (Gaia Collaboration *et al.* 2016b) in which it is recommended to 'consider the quoted uncertainties on the parallaxes as $\varpi \pm \sigma_\varpi$ (random) ±0.3 mas (systematic)'. This creates the impression that the typical 0.3 mas systematic uncertainty should be added in quadrature to the uncertainty quoted in the *Gaia* DR1 catalogue. It should be stressed here that this *should not be done*. The reason is that the calibration of the TGAS parallax uncertainties by comparison to the Hipparcos parallaxes automatically leads to the inclusion of the local systematics in the quoted uncertainty. There is no simple recipe to account for the systematic uncertainties. The advice is to proceed with one's analysis of the *Gaia* DR1 data using the uncertainties quoted in the catalogue, but to keep the systematics in mind when interpreting the results of the data analysis.

As illustrated in Arenou *et al.* (2017) and Lindegren *et al.* (2016) the systematic uncertainties on the parallaxes vary over the sky and are spatially correlated in the sense that the systematics over small patches of the sky tend to be in the same direction. No attempt was made during the *Gaia* DR1 processing and validation to derive a correlation length scale. Zinn *et al.* (2017) made use of the precise asteroseismic distances for stars in the Kepler field to calibrate the spatial correlation length scale of systematic parallax uncertainties in *Gaia* DR1. They also provide a model of the spatial correlations that can be used to construct a covariance matrix for data analyses that involve TGAS parallaxes. It is not obvious that this finding for the Kepler field holds for the entire sky, but it does provide a useful estimate of the local correlations and perhaps the appropriate values of the model parameters for other sky regions can be estimated as part of the data analysis.

3. Science from *Gaia* DR1

Notwithstanding the complexity of dealing with its error characteristics, the scientific exploitation of the first *Gaia* data release has been taken up enthusiastically by the world-wide astronomical community, as evidenced by the numerous workshops organized to collectively work on the analysis of *Gaia* data†, and the over 300 papers that have appeared in the literature since September 14 2016 which are based on or make use of the *Gaia* DR1 data.

The Gaia Collaboration has published two performance verification papers that provide a new inventory of the nearby open clusters (Gaia Collaboration *et al.* 2017a), and a test of the TGAS parallaxes through a thorough study of the local Cepheid and RR Lyrae populations (Gaia Collaboration *et al.* 2017b), where the K-band period luminosity relations show a substantial improvement in the TGAS parallaxes compared to the *Hipparcos* values. I highlight below a few of the more creative and unexpected analyses of the *Gaia* DR1 data.

Mapping the structure of the Magellanic clouds. Belokurov *et al.* (2017) describe a very clever method for tracking down variable stars in *Gaia* DR1, even though their light curves have not been published and no explicit indication is included on the possible variability of catalogue sources (keep in mind that light curves and variable star charac-

---

† For example the 'Gaia Sprints' (`http://gaia.lol/`), and the Gaia 2016 Data Workshop (`https://www.cosmos.esa.int/web/gaia-2016-data-workshop/home`).

terizations were included in *Gaia* DR1 only for a very modest sample of 2600 RR Lyrae and 600 Cepheid variables, see Clementini *et al.* 2016). Belokurov *et al.* (2017) make use of the fact that the photometric uncertainties quoted in the *Gaia* DR1 catalogue reflect the scatter in the individual observations made for each source. This leads to overestimates of the uncertainty on the mean G band value for variable sources, making these stars stand out in a diagram of the uncertainty in G vs. the value of G. By calibrating against samples of known variable stars Belokurov *et al.* (2017) were able to identify candidate RR Lyrae stars in a field covering the Magellanic clouds. These candidate RR Lyrae beautifully outline the LMC and SMC and in particular reveal the bridge of old stars between the two Milky Way companions. A combination of *Gaia* DR1 with *GALEX* (Bianchi *et al.* 2014) data revealed the existence of a bridge of younger stars, offset from the bridge of old stars and coincident with the known HI bridge between the LMC and SMC. The technique to find variable stars in *Gaia* DR1 was also applied by Deason *et al.* (2017a) in order to map the structure of the Magellanic system through Mira variables.

A cluster hiding near Sirius. The power of a high spatial resolution, high dynamic range, all-sky star map was demonstrated nicely in the paper by Koposov *et al.* (2017). Although *Gaia* observations of bright sources suffer from CCD saturation effects, there is no need to avoid the vicinity of even the brightest stars on the sky and hence *Gaia* can observe sources very near such stars. Koposov *et al.* (2017) made use of this by creating an all-sky map of potential source over-densities and in that way discovered a hitherto unknown star cluster very near the brightest star in the sky, Sirius. The reality of the *Gaia* 1 cluster was confirmed by combining the *Gaia* DR1 information with the photometry form the *2MASS* (Skrutskie *et al.* 2006), *WISE* (Wright *et al.* 2010), and Pan-Starrs1 (Chambers *et al.* 2016) surveys. Simpson *et al.* (2017) carried out spectroscopic follow-up observations and concluded that *Gaia* 1 is an intermediate age ($\sim$3 Gyr) open cluster with a mass of roughly $10^4\ M_\odot$.

De-noising the TGAS colour magnitude diagram. A central goal of the *Gaia* mission is the establishment of a precise and accurate empirical description of the colour magnitude diagram, which opens the way to accurate luminosity calibrations of stars across the CMD and to an accurate calibration of the theoretical Hertzsprung-Russell diagram. The measurement of stellar distances plays a fundamental role in this endeavour but accurate distances cannot always be obtained through parallaxes alone. In particular for the more luminous and rarer stars near the bright end of the CMD, even the end of mission *Gaia* parallaxes may have relative errors above the level where one should not simply invert the parallax to obtain a distance (see also Bailer-Jones 2015). Hence it is imperative to combine multiple pieces of information to estimate accurate distances to stars. Two papers based on*Gaia* DR1 (Leistedt & Hogg 2017 and Anderson *et al.* 2017) present approaches in which the information contained in the photometry of stars (apparent brightness and colour) is combined with the TGAS parallax information to arrive at more precise representations of the CMD than can be obtained through TGAS parallaxes alone. In both cases a hierarchical Bayesian model is employed albeit with a different approach to constructing the prior on the distribution of stars in the CMD. Both studies successfully demonstrate how this type of modelling leads to shrinkage in the errors on the inferred distances (absolute magnitudes) of the stars, even if strictly speaking they only provide a more precise description of the contents of the TGAS CMD, rather than of the CMD per se (which would require folding in selection functions and considerations on the degree to which the solar neighbourhood is representative). The hope is that eventually this type of analysis of the *Gaia* data leads to a data-driven predictive models of stars which would very tightly constrain our physical models of stars.

The needle in the haystack. The work by Marchetti *et al.* (2017) shows how machine learning (in this case a neural network) can be applied to large and rich data sets such as *Gaia* DR1. The goal of this work was to find the very few hyper-velocity stars (which were ejected from the Galactic centre) expected to be present in the TGAS catalogue (a few hundred to a few thousand hyper-velocity stars are expected in the full billion star *Gaia* data set). This was done by training and artificial neural network on simulated data containing both the Milky Way and a population of hyper-velocity stars. The optimized neural network was then applied to the TGAS data which resulted in the isolation of 80 hyper-velocity star candidates purely on the basis of astrometric information. A careful follow-up of these candidates through the collection of radial velocity information and the assessment of whether the orbits of the stars imply that they come from the Galactic centre, resulted in one candidate hyper-velocity star that might be unbound from the Milky Way and 5 candidates that appear to be bound. The results of this study greatly strengthen the confidence that in future *Gaia* data releases (where the application of machine learning techniques will be more important) many more hyper-velocity stars can be uncovered.

Stellar occultations. Although this was not an unforeseen application of *Gaia* DR1 it is an excellent illustration of the benefits of an accurate star map. The accurate prediction of the path on the earth from where the occultation of a star by minor body in the solar system can be observed depends very much on the accuracy to which the orbit of the body is known and the accuracy to which the star's position at the observation epoch is known. The latter is greatly improved by the availability of *Gaia* DR1, with more improvements expected in future *Gaia* data releases when proper motions and parallaxes are available for all sources observed by *Gaia*. A taste of the possibilities was provided in the summer of 2016 through the exceptional early release of the Gaia position for a star that would be occulted by Pluto. The better knowledge of Pluto's ephemeris due to the New Horizons flyby was combined with the more accurate *Gaia* position for the star to enable a much more accurate prediction of the occultation path on earth†. The subsequent successful occultation campaign allowed to add a further observational point to the evolution of the atmospheric pressure on Pluto, showing a hint that the pressure increase seen since 1988 (despite Pluto's moving away from the Sun) is now coming to an end, perhaps indicating the start of Pluto's predicted atmospheric 'collapse' due to the lower solar flux.

The above examples are an illustration of the new and complementary ways in which astronomical science can be pursued in the era of large surveys. Creative 'playing' with the data can lead to significant discoveries and new understanding, while at the same time the hard work of developing statistical/numerical/data-driven methods that can efficiently deal with the large amount of information to be uncovered is indispensable.

Finally, it should be noted that *Gaia* DR1 has quickly become the standard against which other surveys are calibrated both astrometrically and photometrically. An example of *Gaia* DR1 serving as the astrometric standard for another large survey is provided by SMASH (Nidever *et al.* 2017) for which the astrometry was re-reduced to the *Gaia* DR1 reference frame. A number of proper motion catalogues have been constructed from the *Gaia* DR1 positions in combination with other surveys. The 'Hot Stuff for One Year' proper motion catalogue (Altmann *et al.* 2017) combines *Gaia* DR1 with the PPMXL (Röser *et al.* 2010) positions in order to derive proper motions for over 500 million stars. The combination of Pan-Starrs1 and *Gaia* DR1 led to the GPS1 proper motion catalogue, covering three quarters of the sky (Tian *et al.* 2017), and Deason *et al.* (2017b) make

† `https://www.cosmos.esa.int/web/gaia/iow_20160914`

use of a proper motion catalogue derived by combining *Gaia* DR1 and SDSS (York *et al.* 2000). Finally the UCAC series of proper motion catalogues was extended with the creation of UCAC5 by re-reducing the existing UCAC observations to the *Gaia* DR1 reference frame and then combining them with the *Gaia* DR1 positions to to derive new proper motions (Zacharias *et al.* 2017).

4. Looking ahead

Although the exploitation of the *Gaia* DR1 data is still in full swing the next data release will arrive soon, in April 2018. *Gaia* DR2 will be based on 22 months of input data and allow for a *Gaia* stand-alone astrometric solution (so the *Hipparcos/Tycho-2* positions will no longer be used), including parallaxes and proper motions for a much larger number (of order one billion) sources. The larger amount of data, the improvements in the various instrument calibrations, and the introduction of colour terms in the astrometric solution will lead to large reductions of the astrometric uncertainties. A major difference between *Gaia* DR2 and*Gaia* DR1 will be the presence of radial velocities for stars brighter than $G_{\mathrm{RVS}} = 12$ and the availability of a broad-band colour, $(G_{\mathrm{BP}} - G_{\mathrm{RP}})$, for all stars on an all-sky homogeneous photometric system. Perhaps these two elements represent the biggest advance from *Gaia* DR1 to *Gaia* DR2. In addition for stars brighter than $G = 17$ the effective temperature and extinction will be determined from the broad-band photometry in combination with the parallaxes, and a major extension of the variable star catalogue is foreseen, including an all-sky RR Lyrae survey. Finally, the epoch astrometry for a pre-selected list of about 10 000 asteroids will be released.

In connection with *Gaia* DR2 it is important to be aware of the following issue concerning the traceability of sources from *Gaia* DR1 to *Gaia* DR2. The data processing leading up to a data release starts with a process that groups individual Gaia observations and links them to sources on the sky. This leads to a working catalogue of sources ('the source list') and their corresponding observations, which forms the basis for the subsequent data processing. The algorithm that carries out the grouping and linking had been much improved before the start of the *Gaia* DR2 processing and this led to many changes in these groups.

When using the Gaia data one should thus be aware that the source list for *Gaia* DR2 should be treated as independent from *Gaia* DR1. Although the majority of sources in *Gaia* DR1 can be identified with the same source in *Gaia* DR2 through the Gaia source identifier, the improved source list will lead to the following changes in the linking of the observations to the source identifiers for a substantial fraction of entries in the source list:

- The merging of groups of observations previously linked to more than one source will lead to a new source associated to the merged observations (with a new source identifier) and the disappearance of the original sources (along with their source identifiers).
- The splitting of groups of observations previously linked to one source will lead to new sources associated to the split groups of observations (with new source identifiers) and the disappearance of the original source (along with its source identifier).
- The list of observations linked to a source may change (and hence the source characteristics may change), while the source identifier remains the same.

A means to trace sources from *Gaia* DR1 to *Gaia* DR2 will be provided, but a one-to-one relation will not exist for all sources. It will then be up to the catalogue user to judge which *Gaia* DR2 source (best) matches a given *Gaia* DR1 source.

Beyond *Gaia* DR2 we can look forward to *Gaia* DR3, targeted for mid to late 2020, and *Gaia* DR4, targeted for the end of 2022. Details on the contents of these releases can be found on the *Gaia* web pages†. Note that *Gaia* DR4 will be the final release for the nominal (5 year) *Gaia* mission. Should the *Gaia* mission be extended, at least one additional data release is foreseen at the end of the extended mission operations. There is much more to come!

Acknowledgements

The early release of Gaia data and the scientific success of this IAU symposium have only been possible due to the excellence and the tireless efforts of the DPAC and ESA teams.

References

Altmann, M., Roeser, S., Demleitner, M., Bastian, U., & Schilbach, E., 2017, *A&A* 600 L4

Anderson, L., Hogg, D. W., Leistedt, B., Price-Whelan, A. M., & Bovy, J., 2017, arXiv:1706.0505

Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., Helmi, A., *et al.*, 2017, *A&A* 599, A50

Bailer-Jones, C. A. L., 2015, *PASP* 127, 994

Belokurov, V., Erkal, D., Deason, A. J., Koposov, S. E., De Angeli, F., *et al.*, 2017, *MNRAS* 466, 4711

Bianchi, L., Conti, A. & Shiao, B., 2014, *Advances in Space Research* Vol. 53, 900

Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., *et al.*, 2010, *Science* 327, 977

Casertano, S., Riess, A. G., Bucciarelli, B., & Lattanzi, M., 2016, *A&A* 599, A67

Chambers, K. C., Magnier, E. A., Metcalfe, N., Flewelling, H. A., Huber, M. E., *et al.*, 2016, arXiv:1612.05560

Davies, G. R., Lund, M. N., Miglio, A., Elsworth, Y., Kuszlewicz, J. S., *et al.*, 2017, *A&A* 598, L4

Deason, A. J., Belokurov, V., Erkal, D., Deason, A. J., Koposov, S. E., & Mackey, D., 2017, *MNRAS* 467, 2636

Deason, Alis J., Belokurov, V., Koposov, S. E., Gómez, F. A., Grand, R. J., *et al.*, 2017, *MNRAS* 470, 1259

De Ridder, J., Molenberghs, G., Aerts, C., & Eyer, L., 2016, *A&A* 595, L3

Clementini, G., Ripepi, V., Leccia, S., Mowlavi, N., Lecoeur-Taibi, I., *et al.*, 2016, *A&A* 595, A133

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., Vallenari, A., & Babusiaux, C., *et al.* 2016a, *A&A* 595, A1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., & Mignard, F., *et al.* 2016b, *A&A* 595, A2

Gaia Collaborations, van Leeuwen, F., Vallenari, A., Jordi, C., Lindegren, L., Bastian, U., *et al.*, 2017a, *A&A* 601, A19

Gaia Collaborations, Clementini, G., Eyer, L., Ripepi, V., Marconi, M., Muraveva, T., *et al.*, 2017b, *A&A* in press, arXiv:1705.00688

Gontcharov, G., & Mosenkov, A., 2017, arXiv:1705.09063

Gould, A., Kollmeier, J. A., & Sesar, B., 2016, arXiv:1609.06315

Hawkins, K., Leistedt, B., Bovy, J., & Hogg, D. W., 2017, arXiv:1705.08988

Huber, D., Zinn, J., Bojsen-Hansen, M., Pinsonneault, M., Sahlholdt, C., *et al.*, 2017, arXiv:1705.04697

Jao, W.-C., Henry, T. J., Riedel, A. R., Winters, J. G., Slatten, K. J., & Gies, D. R., 2016, *ApJL* 832, L18

Kopsov, S. E., Belokurov, V., & Torrealba, G., 2017, *MNRAS* 470, 2702

Leistedt, B. & Hogg, D. W., 2017, arXiv:1703.08112

Lindegren, L., Lammers, U., Bastian, U., Hernández, J., Klioner, S., *et al.*, 2016, *A&A* 595, A4

---

† `https://www.cosmos.esa.int/web/gaia/release`

Marchetti, T., Rossi, E. M., Kordopatis, G., Brown, A. G. A., Rimoldi, A., *et al.*, 2017, *MNRAS* 470, 1388

Michalik, D., Lindegren, L., & Hobbs, D., 2015, *A&A* 574, A115

Mignard, F., Klioner, S., Lindegren, L., Bastian, U., Bombrun, A., *et al.*, 2016, *A&A* 595, A5

Nidever, D. L., Olsen, K., Walker, A. R., Vivas, A. K., Blum, R. D., *et al.*, 2017, arXiv:1701.00502

Röser, S., Demleitner, M., & Schilbach, E., 2010, *AJ* 139, 2440

Sesar, B., Fouesneau, M., Price-Whelan, A., Bailer-Jones, C. A. L., Gould, A., & Rix, H.-W., 2017, *ApJ* 838, 107

Silva Aguirre, V., Lund, M. N., Antia, H. M., Hall, W. B., Basu, S., *et al.*, 2017, *ApJ* 835, 173

Simpson, J. D., De Silva, G. M., Martell, S. L., Zucker, D. B., Ferguson, A. M. N., *et al.*, 2017, arXiv:1703.03823

Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., *et al.*, 2006, *AJ* 131, 1163

Stassun, K. G. & Torres, G., 2016, *ApJL* 831, L6

Tian, H.-J., Gupta, P., Sesar, B., Rix, H.-W., & Martin, N. F., 2017, arXiv:1703.06278

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., Ressler, M. E., Cutri, R. M., *et al.*, 2010, *AJ* 140, 1868

Yildiz, M., Çelik Orhan, Z., Örtel. S., & Roth, M., 2017, arXiv:1705.08313

York, D. G., Adelman, J., Anderson, J. E., Jr., Anderson, S. F., Annis, J., *et al.*, 2000, *AJ* 120, 1579

Zacharias, N., Finch, C., & Frouard, J., 2017, *AJ* 153, 166

Zinn, J. C., Huber, D., Pinsonneault, M. H., & Stello, D., 2017, arXiv:1706.09416