Highlights on Spanish Astrophysics X, Proceedings of the XIII Scientific Meeting of the Spanish Astronomical Society held on July 16 – 20, 2018, in Salamanca, Spain. B. Montesinos, A. Asensio Ramos, F. Buitrago, R. Schödel, E. Villaver, S. Pérez-Hoyos, I. Ordóñez-Etxeberria (eds.), 2019

Gaia DR2: contents and properties of the second Gaia data release.

Xavier Luri[1], on behalf of Gaia DPAC

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

Abstract

The second data release of ESA's Gaia mission (Gaia DR2) was published in April 2018. It is just another step in the Gaia release scenario, to be followed by other more complete and precise releases in the coming years, but already represents a huge step ahead. Gaia DR2 contains data positions and G magnitudes for more than 1.7 billion objects, and provides five-parameter astrometry for more than 1.3 billion of them, the biggest astrometric catalogue ever by several orders of magnitude. The release contains also additional photometric, radial velocity and physical parameters data for different subsets and specialized datasets for variability and solar system objects.
In this paper we review the properties of Gaia DR2 and discuss some of the shortcomings and limitations that have to be taken into account to properly use its data. We also review the archive systems where the data can be accessed and the expected contents of future Gaia data releases.

1 Introduction: the Gaia mission

The European Space Agency's (ESA) astrometric mission Gaia was launched in September 2013. Its main aim was to generate an astrometric catalogue with highly accurate positions, parallaxes and proper motions for more than 1 billion sources brighter than magnitude 20. A description of the mission, its principles and scientific case can be found in [7], while a more extensive scientific motivation for the mission is presented in [6].

The data processing required to transform the spacecraft's telemetry into a scientifically useful dataset was entrusted by ESA to an european-wide consortium, the Gaia Data Processing and Analysis Consortium (DPAC, also described in [7]). This consortium is currently formed by more than 450 scientists and engineers from all around Europe, with a significant presence of Spanish teams. At the time of its creation, and as part of its agreements with ESA, a plan for a staged publication of the mission results was defined. This so-called "data release scenario" includes several intermediate data releases based on preliminary calibrations and the subset the measurements available at a given time, and a final release with the complete satellite telemetry. The current details of this data release scenario (which is periodically revised) can be found in ESA's Gaia web site.

As part of this release scenario ESA and DPAC have published in April 2018 the Gaia Second Data Release (Gaia DR2). The properties of this release are reviewed in the coming sections, and we encourage the reader to consult two key sources of information for it: the Astronomy & Astrophysics special edition devoted to Gaia DR2 and the Gaia DR2 on-line documentation at the Gaia archive.

The Gaia data can be accessed through the main Gaia archive or through one of the Gaia partner data centers. There are also replicas of the archive in other sites around the world and in addition the Gaia data can be downloaded for local use through the Gaia download site.

2 Gaia DR2 contents

A full review of the Gaia DR2 contents can be found in [2]. Here we present a brief summary of its characteristics.

First of all, the Gaia DR2 main content is the overall catalogue. This table, named GaiaSource in the main archive, contains 1,692,919,135 sources for which *at least* a sky position and a G magnitude (see [8] for a review of the Gaia photometry) are available. Besides this basic information, additional data is available for subsets of this main dataset as described in Table 1.

Subset	number of objects
Full catalogue	1,692,919,135
Objects with 5-parameter astrometry	1,331,909,727
Objects with 2-parameter astrometry	361,009,408
Objects with G_{BP} magnitude	1,381,964,755
Objects with G_{RP} magnitude	1,383,551,713
Objects with radial velocity	7,224,631
Objects with estimated T_{eff}	161,497,595
Objects with estimated A_G and $E(G_{BP} - G_{RP})$	87,733,672
Objects with estimated radius and luminosity	76,956,778

Table 1: Number of sources for which a given data product is available in Gaia DR2. Extracted from Table 1 in [2]

In addition to this main dataset, DR2 contains other specialised datasets:

- A dataset for 550,737 identified variable stars, including light curves and classification into variability type for most of them

- A dataset for 14,099 Solar System Objects, including epoch photometry and astrometry
- Pre-computed cross-match catalogues against DR2 for a dozen of external catalogues (2mass, HIP, AllWise, and others)
- Cross-Id table for AllWiseAgn and ICRF sources (QSOs) used to define the celestial reference frame of Gaia DR2 (Gaia-CRF2)

3 Gaia DR2 properties

The sky distribution of the global DR2 dataset is presented in Figure 1. Notice that one can still faintly appreciate the distribution of the different sets and subsets as a function of the G magnitude is presented in Figure 2. Notice that the main dataset contains sources well beyond the nominal limit of Gaia observations of $G = 20.7$, but as discussed in the next section it is complete only around magnitude $G \simeq 19$.

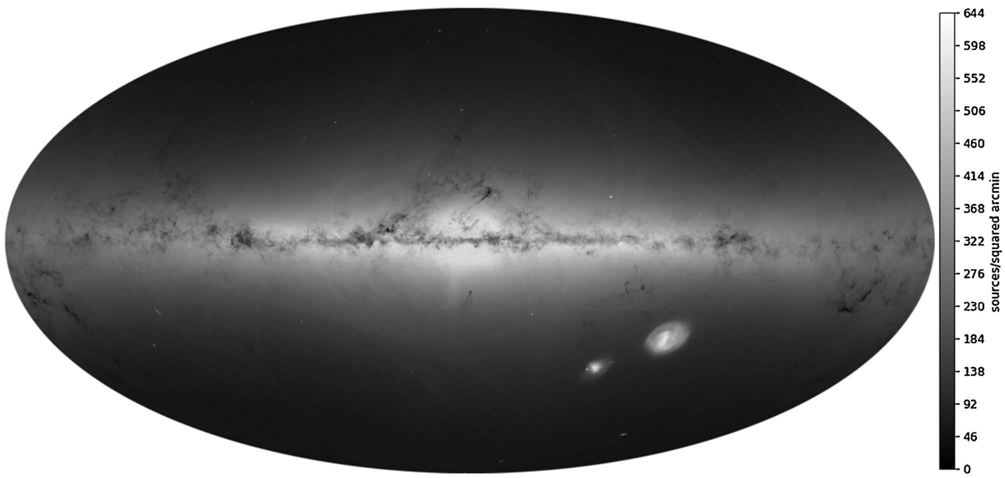

Figure 1: Sky distribution of all Gaia DR2 sources in Galactic coordinates. The figure has been extracted from [2].

3.1 Astrometric data

The astrometric processing and the properties of the Gaia DR2 astrometry are described in detail in [5]. The formal uncertainties for the Gaia astronomy are individually provided for each object in the catalogue but in general vary with the G magnitude, being more precise for bright objects and progressively less precise for fainter ones, as illustrated in Figure 3. A summary of these formal uncertainties is presented in Table 2 and a detailed list of values as a function of the magnitude can be found in Table B.1 of [5].

As seen in the figures these uncertainties are dominated by photon noise above $G \simeq 15$, but show a plateau at brighter magnitudes, where the limitation is the present knowledge

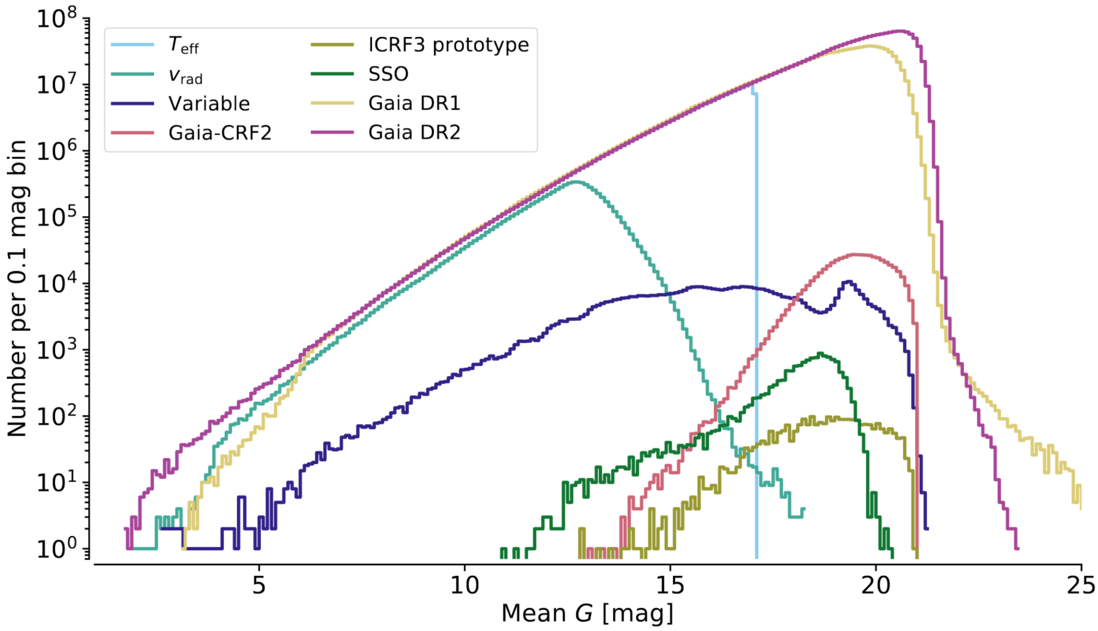

Figure 2: Distribution of the mean values of G magnitude for the different sets and subsets of Gaia DR2 sources shown as histograms with 0.1 mag wide bins. It also includes the distribution for Gaia DR1 sources. The figure has been extracted from [2].

of the instrument calibrations. We expect to improve these limitations in the instrument calibrations and, together with an increase in precision due to the accumulation of additional data, reach precisions in the bright end around $10 - 20\mu as$.

When using the astrometric data it is important to take into account that, due to the characteristics of the processing, the uncertainties of the five astrometric parameters are correlated. A correlation matrix is included as part of the Gaia data to allow a proper treatment of the joint uncertainity distribution.

The astrometric data processing has been designed to minimize as much as possible any systematic error present in the data. However, at this stage of the mission, still working with just a part of the total expected observations and with a partial knowledge of the instrument calibrations, there are still uncontrolled effects that can creep into the astrometry. Some of these are already known (and its magnitude estimated) after the validation process carried out by DPAC on the astrometric solution, including a global zero point on the parallaxes ($-30\mu as$ in the sense Gaia minus external data), a residual rotation on the proper motions and some regional effects. Also, 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).

G	position (mas)	parallax (mas)	proper motion (mas yr^{-1})
15	0.03	0.04	0.06
17	0.08	0.09	0.15
20	0.5	0.7	1.1
21	1.5	2.1	3.2

Table 2: Astrometric formal uncertainities as a function of the G magnitude. Extracted from Table B.1 in [5]

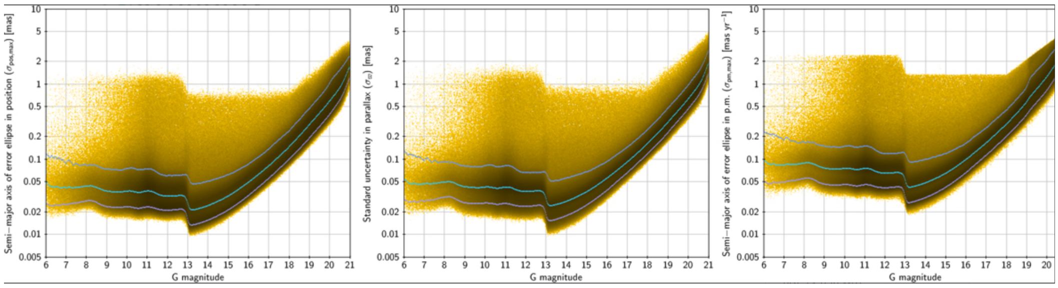

Figure 3: Formal uncertainties for parallax (left), proper motion in right ascension (middle) and declination (right) of the objects in Gaia DR2 as a function of the G magnitude. The cyan curve is the median uncertainty and the blue curves are the 10th and 90th percentiles. The yellow points correspond to a representative subset added as background. The figure has been extracted from [5].

A review of these and other known issues in the data are collected in Gaia's know issues page, and specifically for astrometry a document from Lennart Lindegren linked in this same page provides a useful summary of the known systematics. See also Section 6.3.1 in [2].

3.2 Photometric data

The photometric processing Gaia DR2 is described in detail in [8] and the properties of the DR2 photometry are discussed in [4]. As discussed in the previous section all the 1.7 billion published objects have at least a G magnitude, and about 1.3 billion of them also have a G_{BP} (blue) and/or G_{RP} (red) magnitude. Thus, in many cases the Gaia objects have three-band photometry; from these values a figure of the average color in each sky region has been composed (Figure 4).

The passbands definining the G, G_{BP} and G_{RP} filters are published as a numerical table in the Gaia archive. However, these passbands have been revised after the publication of DR2 using a larger library of standard stars, providing updated estimations. We recommend to use the passbands published in [1].

The photometric data processing considered three types of sources, "Gold", "Silver", and "Bronze", which represent decreasing quality levels of the photometric calibration achieved;

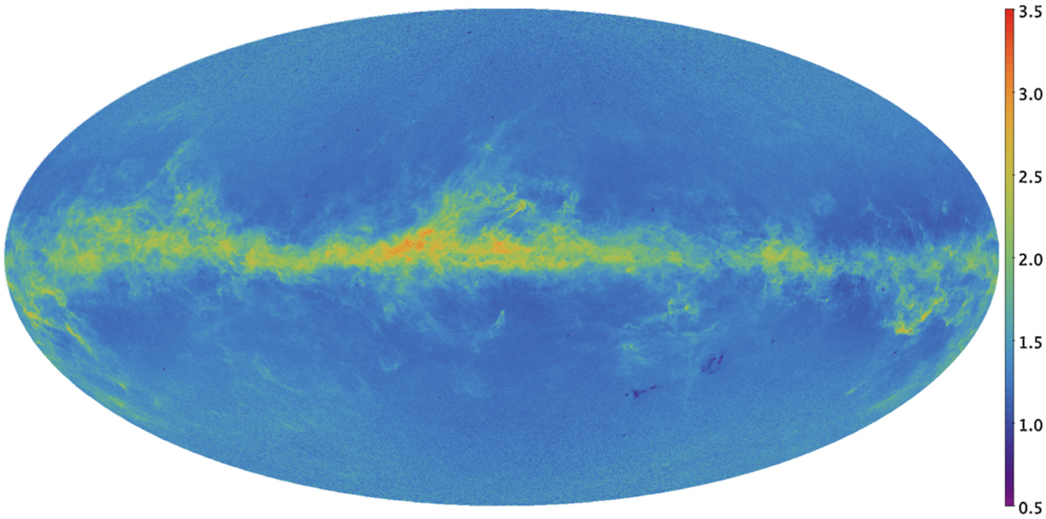

Figure 4: Colour distribution as a function of sky position in Galactic coordinates. Each pixel represents the median colour ($G_{BP} - G_{RP}$) of all sourceswith $G < 19$ in that pixel. The figure has been extracted from [4].

in particular, in the case of the Bronze sources no colour information is available. This is indicated in the released catalogue by a numeric field (`phot_proc_mode`) assuming values 0, 1 and 2 for gold, silver, and bronze sources respectively. Similarly to the astrometric data the uncertainties are strongly dependent on the G magnitude; at the bright end the photometric uncertainties are dominated by calibration effects and flatten around a few millimagnitudes, except for the very bright stars where they are substantially higher due to saturation effects. At the faint end the uncertainties are dominated by the photon noise and reach 0.1 mag around $G = 20$. The distribution of uncertainties as a function of G magnitude is depicted in Figure 5.

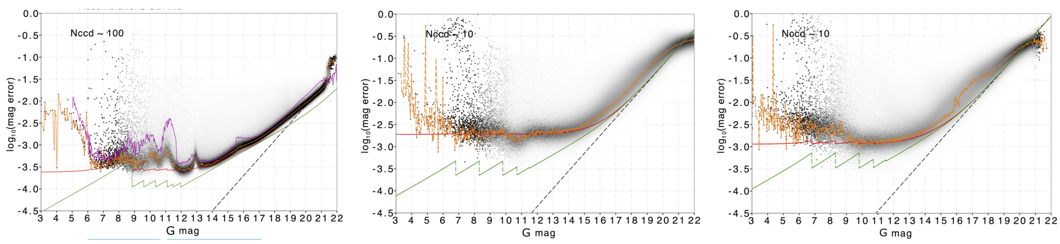

Figure 5: Distribution of uncertainty on the weighted mean magnitude for the three Gaia bands as a function of the G magnitude. The figures have been extracted from [4].

Please refer to section 6.3.2 of [2] for a discussion of the limitations of the photometric Gaia data at this stage of the mission, including a discussion about the so-called “flux excess”, the photometric zero points and the calibrations. See also [1] for a discussion of a magnitude trend in G for objects brigther than $G \simeq 16$.

3.3 Radial velocities

As described in previous sections, besides the astrometric and photometric data Gaia DR2 also provides radial velocity data for a subset of the objects. These data are acquired using the Radial Velocity Spectrometer (RVS) instrument described in [3]. The processing of the RVS data is described in detail in [9]. The processing pipeline delivered median radial velocities for Gaia stars with narrow-band near-IR magnitude $G_{RVS} \leq 12$ (approximately brighter than $V \simeq 13$). Stars identified as double-lined spectroscopic binaries were removed from the pipeline, while variable stars, single-lined, and non-detected double-lined spectroscopic binaries were treated as single stars. Furthermore, for the hottest ($T_{eff} \geq 7000$ K) and coolest ($T_{eff} \leq 3500$ K) starts the accuracy and precision of the stellar parameter estimates were not sufficient to allow selection of appropriate templates and the radial velocities obtained for these stars are not published in DR2. The overall accuracy of radial velocity measurements is around 200-300 m s^{-1}, and the overall precision is 1 km s^{-1}, reaching 200 m s^{-1} for the brightest stars. The overall distribution of uncertainties of radial velocities as a function of the G magnitude is depicted in Figure 6.

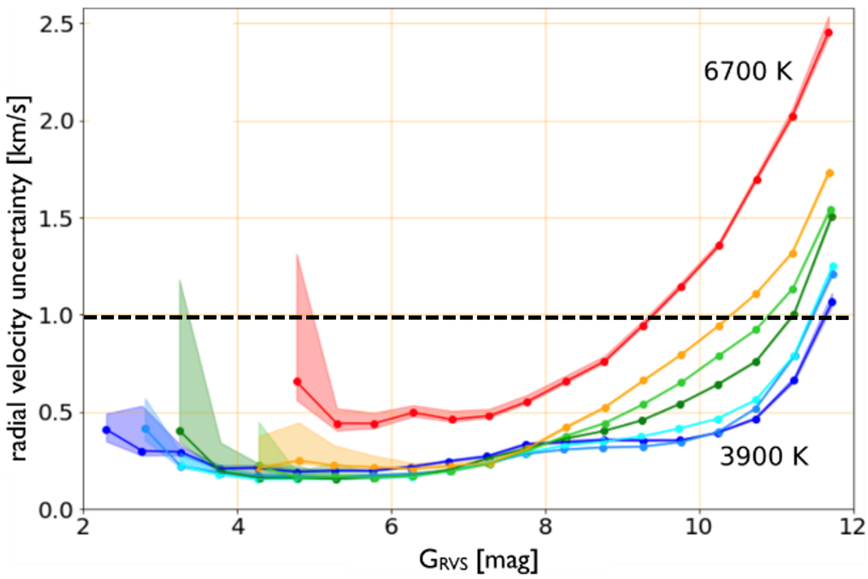

Figure 6: Distribution of uncertainty on the mean radial velocity as a function of the G magnitude.

Please refer to section 6.3.3 of [2] for a discussion of the limitations of the Gaia radial velocities at this stage of the mission, including a discussion about the treatment of spectroscopic binaries.

4 Conclusions

The publication of the Gaia DR2 has been an outstanding success. After a huge peak in publication of papers in the months immediately after the release (about 120 papers from non-DPAC authors in two months, about 17% of them including Spanish authors) we are now seeing about 2 papers per day based on Gaia data. The science with Gaia is becoming a reality and in many areas the Gaia data is allowing significant breakthroughs. We also direct the reader to the Astronomy & Astrophysics special issue on Gaia DR2 where a collection of science demonstration papers by DPAC can be found along the release papers.

Furthermore, in the not too far future, around mid-2021, the next Gaia data release (DR3) will arrive, with even more precise and more varied data. The expected contents of this new release include:

- Covering 34 months of data (22 months in DR2)
- Better precision and accuracy in astrometry and photometry
- Accelerated/orbital solutions in astrometry: binary systems
- More RVs: fainter magnitude, wider surface temperature range, more precise
- Mean spectra in the BP/RP and RVS bands
- Stellar parameters from spectra (BP/RP, RVS): more precise, fainter magnitude, surface gravity (log g), age, [Fe/H] and various element abundances
- Source classifications (e.g. quasars, binaries, galaxies)
- More solar system objects, reflectance spectra, colours
- More light curves, new variable source classifications

And more to come with other future releases. Updated information about the upcoming publications of Gaia data can be found in the Gaia data release scenario web page.

Acknowledgments

This work has made use of results from the European Space Agency (ESA) space mission *Gaia*, the data from which were processed by the *Gaia Data Processing and Analysis Consortium* (DPAC). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. The *Gaia* mission website is `http://www.cosmos.esa.int/gaia`. The authors are current or past members of the ESA *Gaia* mission team and of the *Gaia* DPAC.

This work was supported by: the MINECO (Spanish Ministry of Economy) through grant ESP2016-80079-C2-1-R (MINECO/FEDER, UE) and ESP2014-55996-C2-1-R (MINECO/FEDER, UE) and MDM-2014-0369 of ICCUB (Unidad de Excelencia 'María de Maeztu'); the DLR (German space agency) via grant 50 QG 1403.

References

[1] Apellaniz, J. & Weiler, M., accepted, http://cdsads.u-strasbg.fr/abs/2018arXiv180802820M

[2] Gaia Collaboration, Brown, A. et al., 2018, A&A , 616, A1

[3] Gaia Collaboration, Cropper, M. et al., 2018, A&A , 616, A5

[4] Gaia Collaboration, Evans, D. et al., 2018, A&A , 616, A4

[5] Gaia Collaboration, Lindegren, L. et al., 2018, A&A , 616, A2

[6] Perryman, M. A. C. et al., 2001, A&A , 369, 339

[7] Gaia Collaboration, Prusti, T. et al., 2016, A&A , 595, A1

[8] Gaia Collaboration, Riello, M. et al., 2018, A&A , 616, A3

[9] Gaia Collaboration, Sartoretti, P. et al., 2018, A&A , 616, A6