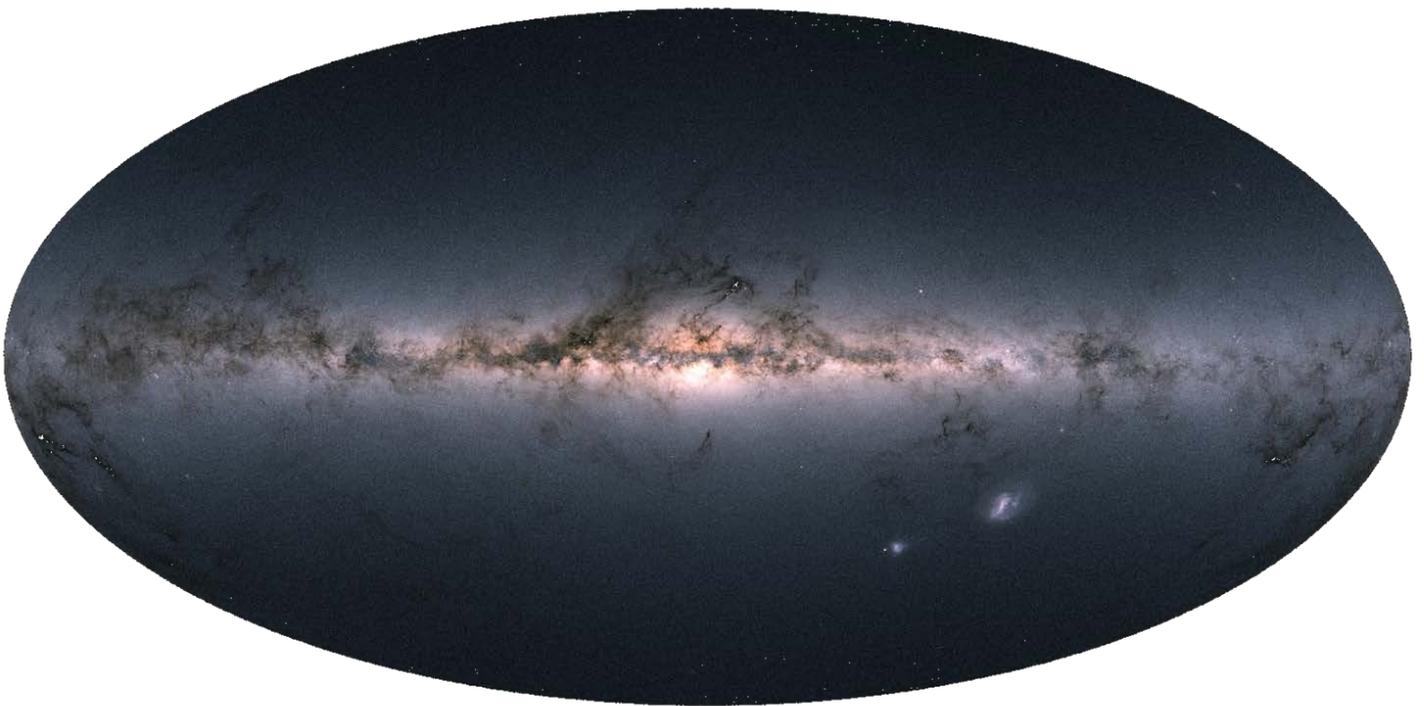


# Gaia DR2 primer



Everything you wish you had known before  
you started working with Gaia Data Release 2

# Contents

1. [Introduction](#)
2. [Gaia mission](#)
  - [Mission overview](#)
  - [Focal plane](#)
  - [Mission and publication statistics](#)
  - [Acronyms](#)
3. [Gaia scanning law](#)
  - [Scanning law](#)
  - [Predicting past and future transits](#)
4. [Gaia Data Release 2](#)
  - [Overview](#)
  - [Gaia DR2 data](#)
5. [Astrometry](#)
  - [Astrometric model and Gaia DR2 astrometric contents](#)
  - [Negative parallaxes](#)
  - [Spurious astrometric solutions](#)
  - [Astrometric quality – renormalised unit weight error \(RUWE\)](#)
  - [Astrometric quality – visibility periods](#)
  - [Astrometric quality – astrometric excess noise](#)
  - [Distance estimates](#)
  - [Parallax zero point](#)
  - [Formal uncertainties \('standard errors'\)](#)
  - [Correlations between the astrometric parameters](#)
  - [Systematic astrometric errors](#)
  - [Galactic coordinates](#)
  - [Cartesian coordinates and velocities](#)
  - [Avoiding biases](#)
6. [Photometry](#)
  - [Broad-band, white-light G magnitudes](#)
  - [Blue and Red Photometer BP and RP magnitudes](#)
  - [Photometric quality – photometric flux-excess factor](#)
  - [Variability information and light curves \(epoch photometry\)](#)
  - [Gaia DR2 extinction coefficients  \$k\_G\$ ,  \$k\_{BP}\$ , and  \$k\_{RP}\$](#)
  - [Gaia DR2 passbands](#)
  - [Relationships with other photometric systems](#)
  - [Gaia DR2 magnitude errors](#)
  - [Systematic photometric errors and corrections](#)
  - [Gaia DR2 synthetic photometry](#)
7. [Spectroscopy](#)
  - [Radial velocities](#)
  - [Radial-velocity zero point](#)
  - [Spurious radial velocities](#)
8. [Astrophysical parameters](#)
  - [Astrophysical parameters](#)
  - [Caveats and systematic errors](#)



9. [Reference frame](#)
  - [International Celestial Reference Frame \(ICRF\) and System \(ICRS\)](#)
  - [Gaia DR2 reference frame, reference epoch, and time coordinate](#)
  - [Epoch propagation](#)
  - [Orientation and rotation of the Gaia DR2 reference frame for bright objects](#)
10. [Solar-system objects](#)
11. [Limitations](#)
  - [Filtering](#)
  - [Survey completeness and selection function](#)
  - [Double and multiple stars](#)
12. [Data access](#)
  - [Gaia Archive](#)
  - [User registration](#)
  - [Query limits](#)
  - [Tips and tricks](#)
  - [Python](#)
  - [Partner data centres](#)
13. [Data model](#)
14. [Data queries](#)
  - [Astronomical Data Query Language \(ADQL\)](#)
  - [ADQL tutorials](#)
  - [ADQL resources](#)
15. [Cross-matches](#)
  - [Gaia source identifiers and designations](#)
  - [Cross-matching Gaia DR2 to Gaia DR1](#)
  - [Find Gaia data for your list of sources](#)
  - [Cross-matching Gaia DR2 to other catalogues](#)
16. [Visualisation and exploration](#)
  - [Gaia Archive Visualisation Service \(GAVS\)](#)
  - [Tool for OPERations on Catalogues And Tables \(TOPCAT\)](#)
  - [ESA Sky](#)
  - [Gaia Sky](#)
17. [Documentation](#)
  - [User documentation](#)
  - [Data release papers](#)
  - [Guide-to-Scientists videos](#)
  - [Issues discovered after Gaia DR2 was released \('known issues'\)](#)
  - [How to cite and acknowledge Gaia DR2](#)
18. [Gaia Data Release 3 and beyond](#)
  - [Gaia Data Release 3](#)
  - [Gaia Data Release 4 and beyond](#)
  - [Mission extension](#)
  - [Science performance](#)
19. [Keeping up to date](#)
  - [Websites](#)
  - [Newsletters](#)
  - [Social media](#)
20. [Get in touch](#)
  - [Gaia Helpdesk](#)
  - [Communicate your Gaia science results with support from ESA](#)
  - [Feedback](#)
  
  - [Change log](#)
  - [Credits front page](#)



# 1. Introduction

The aim of this primer, written by the [Gaia Helpdesk](#), is to collect all information, tips and tricks, pitfalls, caveats, and recommendations relevant to Gaia's second data release ([Gaia DR2](#)) in one place and to provide pointers to where more detailed information can be found. This primer is aimed to serve any astronomer interested in Gaia DR2 data, from starting undergraduates to emeriti and amateurs.

# 2. Gaia mission

## Mission overview

Gaia is the one-billion-star surveyor of the [European Space Agency](#). It has been continuously [scanning the sky](#) since July 2014 from its vantage point close to the second Lagrange point of the Sun-Earth/Moon system. Gaia's aim is to map the entire sky, down to magnitude  $\sim 20.7$ , and to collect micro-arcsecond-level [astrometry](#) and milli-magnitude-level [photometry](#) for the brightest 1,000+ million stars as well as medium-resolution [spectroscopy](#) for mainly radial-velocity determination of the brightest subset of  $\sim 150$  million objects. The resulting data set of 3D positions and 3D velocities of stars, combined with their detailed [astrophysical classification and characterisation](#) obtained from the photometry and spectroscopy, is bound to revolutionise our view of the Milky Way and its constituent stars. More information about the Gaia science case, mission, payload, and spacecraft can be found in the [Gaia mission paper](#) and on the [Gaia mission website](#).

## Focal plane

The [Gaia focal plane](#) is equipped with 106 CCD detectors and consists of several functional sections which are traversed in sequence by a given object as the spacecraft slowly rotates to scan the heavens:

- A Sky Mapper (SM) to [autonomously detect objects](#). Stars brighter than  $\sim 3$  mag are too bright to be detected automatically. The faint detection threshold is set at 20.7<sup>th</sup> magnitude in the [Gaia G band](#) but is not infinitely sharp due to magnitude estimation errors of the on-board software which amount to a few tenths of a magnitude.
- An Astrometric Field (AF) to collect the instantaneous positions and fluxes of all objects detected by the Sky Mapper. Astrometric measurements are made in a white-light bandpass, covering the range 330–1050 nm, which is referred to as the [Gaia G band](#).
- A Blue Photometer (BP) to collect low-resolution spectra (with dispersion varying between 3 and 27 nm pixel<sup>-1</sup>) of all objects over the wavelength range 330–680 nm. The integrated magnitude is referred to as the [G<sub>BP</sub> or BP magnitude](#).
- A Red Photometer (RP) to collect low-resolution spectra (with dispersion varying between 7 and 15 nm pixel<sup>-1</sup>) of all objects over the wavelength range 630–1050 nm. The integrated magnitude is referred to as the [G<sub>RP</sub> or RP magnitude](#).



Gaia focal plane. Credit: [Gaia Collaboration et al. \(2016\)](#).



Collectively, the BP and RP spectra are sometimes referred to as XP spectra.

- An integral-field [Radial Velocity Spectrometer](#) (RVS) to collect medium-resolution spectra ( $R \sim 11,700$ ) over the wavelength range 845–872 nm, around the ionised-Calcium triplet, for all objects brighter than  $\sim 16^{\text{th}}$  magnitude in this bandpass. The integrated magnitude in the RVS bandpass is referred to as the  $G_{\text{RVS}}$  magnitude. The RVS has a reduced field of view orthogonal to the scan direction such that fewer observations up to the RVS limiting magnitude are collected compared to SM/AF/BP/RP in a ratio of 4:7.

A passage of a star through the focal plane is called a field-of-view transit. Over the nominal, five-year mission, stars typically accumulate [~70 transits](#) (~40 for RVS). During each transit, Gaia collects instantaneous, so-called epoch data of each object. Except for photometric epoch data for selected [variable stars](#) and astrometric and photometric epoch data for selected [Solar-system objects](#), Gaia DR2 does not contain astrometric, photometric, or spectroscopic epoch data. Publication of all epoch data, however, is scheduled for the final release(s) except for the Gaia Andromeda Photometric Survey (GAPS) that is part of [Gaia DR3](#) and that will contain photometric time series for all  $\sim 1$  million stars in a selected pencil beam down to  $G \sim 20.7$  mag.

## Mission and publication statistics

The [Gaia Helpdesk](#) maintains an [overview of the mission statistics](#) (updated daily), such as the volume of science data collected on board and the number of object transits through the focal plane, as well as an [ADS library of peer-reviewed scientific papers](#) that use Gaia data (updated a few times per year) and a list of completed [PhD theses](#) that use Gaia data (updated on request).

## Acronyms

An actively maintained list of Gaia mission acronyms can be found on the Gaia mission [website](#).

# 3. Gaia scanning law

## Scanning law

The scanning law describes how Gaia's fields of view scan the sky as function of time. The scanning is composed of two independent, superimposed motions: a rotation around the spacecraft spin axis with a period of 6 hours plus a slow, 63-day-period precession of the spin axis around the Solar direction at a fixed Solar-aspect angle of 45 degrees. Over the nominal five-year mission, Gaia has completed 29 of these precession periods, leading to an optimally uniform sky coverage with, on average,  $\sim 70$  astrometric and photometric transits across the focal plane (and  $\sim 40$  for the spectroscopic instrument). In the [extended mission](#) phase that started in July 2019, a similar scanning law is being employed but with a reversed precession direction during the first year to aid astrometric calibrations. Limited to the time interval underlying Gaia DR2, the commanded scanning law, which the spacecraft materialises to within a few dozen arcseconds, is [available](#) such that, for instance, transit times and scan directions for each transit of an object in Gaia DR2 can be [reconstructed](#).

## Predicting past and future transits

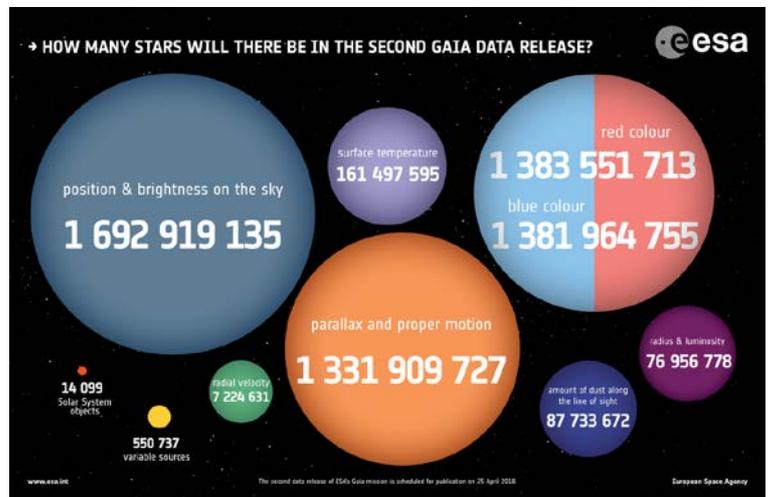
The [Gaia Observation Forecast Tool](#) allows finding out when specific targets have been or will be observed by Gaia (up to and including the [extended mission](#)) with the caveat that individual observations cannot be guaranteed (many reasons may lead to situations in which the data for an object at a predicted time will not be available).



## 4. Gaia Data Release 2

### Overview

The second Gaia data release, Gaia DR2, encompasses astrometry, photometry, radial velocities, astrophysical parameters (stellar effective temperature, extinction, reddening, radius, and luminosity), and variability information for stars plus astrometry and photometry for a sample of pre-selected bodies in the Solar system. Gaia DR2 is based on data collected by the spacecraft between 25 July 2014 and 23 May 2016, spanning a period of 22 months. An overview of Gaia DR2 can be found in the [Gaia Data Release 2: Summary of the contents and survey properties paper](#) and in the [Gaia DR2 media kit](#).



Gaia DR2 contents. Credit: ESA, [CC BY-SA 3.0 IGO](#).

### Gaia DR2 data

Gaia data is made available without limitations or proprietary rights albeit with a [licence](#). Gaia DR2 data can be [accessed](#) through the ESA [Gaia Archive](#) as well as through our [partner data centres](#). Of course, it has extensive [documentation](#) as well as instructions on [how to cite and acknowledge](#) it.

## 5. Astrometry

### Astrometric model and Gaia DR2 astrometric contents

The astrometric parameters contained in Gaia DR2 result from fitting a [standard, five-parameter model of uniform, linear motion](#) to the calibrated spacecraft measurements of the instantaneous positions of each object collected over time. These five parameters, which can only accurately describe the motion of single stars with respect to the barycentre of the Solar system, are the position on the sky (right ascension and declination), the parallax, and the proper-motion vector on the sky (one component in right ascension and one component in declination). For 1332 million objects, this fit behaved well and all five parameters, with [reference epoch](#) J2015.5, have been released. For 361 million, primarily faint objects, the [sub-optimum quality](#) of the fit resulted in only the 'mean positions' being considered reliable enough for publication; these two-parameter sources therefore have 'parallax', 'pmra', and 'pmdec' values of 'null'. Details are contained in the paper [Gaia Data Release 2: The astrometric solution](#). In cases in which the object is a [double or multiple star](#), the published Gaia DR2 astrometry may be severely biased as a result of the uniform application of a non-appropriate, single-star model.

### Negative parallaxes

A significant fraction of sources in Gaia DR2 has negative parallax measurements. Whereas this may seem surprising at first sight since stars have positive distances, negative (or zero) parallaxes are fully legitimate and the expected outcome of a measurement process in cases where the measurement error exceeds the magnitude of the measured quantity. Clearly, the physical interpretation of a negative parallax measurement is non-trivial but methods exist to [properly deal with the information](#) content of such data. Excluding stars with negative parallaxes (or with parallaxes with relatively large uncertainties) is not only removing scientific information but can also introduce serious [biases](#) and is hence not advisable.



## Spurious astrometric solutions

For closely aligned sources (separated by  $\sim 0.2\text{--}0.3$  arcseconds on the sky) that are not systematically resolved in all of the Gaia transit observations, Gaia DR2 occasionally contains spurious parallax values which are either very large or have a strongly negative value, with in either case small [formal uncertainties](#). These problematic sources tend to be faint and located in crowded regions such as the Galactic plane and are also often associated with unreliable (and large) proper motions. Guidance on how to clean samples from spurious astrometry is provided in Appendix C of the paper [Gaia Data Release 2: The astrometric solution](#); a particularly powerful quantity for filtering out spurious astrometry is the renormalised unit weight error (RUWE) filter, that supersedes the filter based on the unit weight error (UWE) in equation (C.1) in the aforementioned paper.

## Astrometric quality – renormalised unit weight error (RUWE)

The renormalised unit weight error, referred to as [RUWE](#), is a powerful and recommended astrometric-quality diagnostic. Statistically speaking, RUWE is expected to be around unity for sources where the single-star model provides an acceptable fit to the astrometric transit observations. A RUWE value significantly greater than unity normally indicates that the source is non-single or otherwise problematic. A RUWE filter hence increases the model-data consistency of the sample. The precise threshold to be used in this filter depends on the science application and on the sample in use and some considerations to define it can be found in the document [Re-normalising the astrometric chi-square in Gaia DR2](#). For historical reasons, RUWE values are published for five-parameters sources in the [Gaia Archive](#) in a [separate RUWE table](#) from the main [gaia\\_source table](#). In cases in which [Gaia partner data centres](#) served RUWE values before their official release through the Gaia Archive, small inconsistencies may appear for selected sources, with the Gaia Archive values being authoritative.

## Astrometric quality – visibility periods

A necessary (but not sufficient) condition for a source to have five-parameter astrometry published in Gaia DR2 is that the transit observations are well spread out in time over the associated 22 months. In practice, this has been achieved by applying the condition 'visibility\_periods\_used  $\geq 6$ ', where 'visibility\_periods\_used' is a published field in the data that describes the number of groups of observations that are separated from other groups by a gap of at least four days. In general, a higher number of visibility periods allows a better separation of parallactic and proper-motion effects which results in more reliable astrometry. As explained in the paper [Gaia DR2: Catalogue validation](#), applying a more stringent condition, for instance 'visibility\_periods\_used  $> 8$ ', can be beneficial for certain scientific applications by removing astrometric outliers at the expense of decreased completeness of the sample.

## Astrometric quality – astrometric excess noise

Although users may be tempted to clean samples using the '[astrometric\\_excess\\_noise](#)' field, this is generally not recommended as a result of the degrees of freedom bug described in Appendix A of the paper [Gaia Data Release 2: The astrometric solution](#).

## Distance estimates

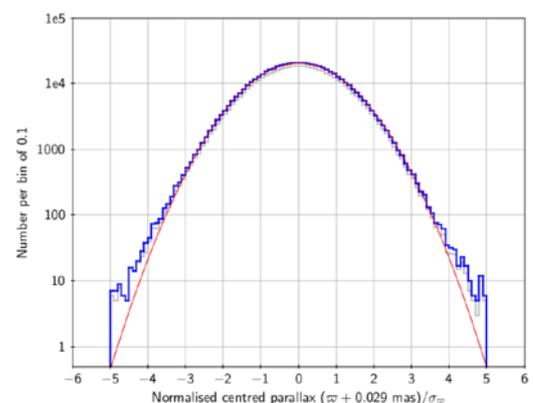
Often, science applications need stellar distances and not parallaxes. Since distance and parallax are related in a non-linear way, estimating distances from parallaxes by simple inversion is in general not advisable and in particular not when the relative parallax precision exceeds  $\sim 10\%\text{--}20\%$ . Not only does inversion lead to biased distance estimates, also the interpretation of distance confidence intervals from parallax uncertainties is not straightforward. In general, one should always treat the derivation of astrophysical parameters from astrometric data, in particular when parallaxes are involved, as an inference problem which should preferably be handled with a [Bayesian approach](#). Extensive guidelines are given in the paper [Gaia Data Release 2: Using Gaia parallaxes](#) that also has a series of ready-to-use Python and R [tutorials](#), for instance for inferring [the distance to a single source](#), [the distance to and size of a cluster](#), or [the distance and tangential velocity of a single source](#).

## Parallax zero point

Gaia DR2 has an overall parallax zero point offset which, from an examination of faint quasar parallaxes (see figure), is estimated to be around  [\$-0.03\$  mas](#) in the sense of the Gaia DR2 parallaxes being too small. This offset has not been corrected in the published data. The estimated parallax zero point [depends on the sample of sources](#) under scrutiny through dependencies (at levels below  $\lesssim 0.1$  mas) on celestial position, magnitude, and colour. It is therefore no surprise that community researchers, for stellar samples and using a variety of methods on redder and brighter objects, typically converge on an offset of around  [\$-0.05\$  mas](#).

## Formal uncertainties ('standard errors')

The astrometric uncertainties listed in Gaia DR2 are derived from the formal errors resulting from the astrometric data processing. As such, they represent the internal consistency of the measurements and do



**Fig. 8.** Distributions of the normalised centred parallaxes for the same samples as in Fig. 6. The red curve is a Gaussian distribution with the same standard deviation (1.081) as the normalised centred parallaxes for *Quasar parallax histogram*. Credit: [Lindgren et al. \(2018\)](#).



not represent the total (external) error. Unlike for Gaia DR1 (TGAS), the formal uncertainties in Gaia DR2 have not been externally calibrated. The Gaia DR2 uncertainties are known to be underestimated by ~8% for faint sources ( $G \gtrsim 16$  mag) outside the Galactic plane and by up to ~20% for bright stars ( $G \lesssim 13$  mag) and proposed corrections to obtain external errors have been [published](#).

## Correlations between the astrometric parameters

The five astrometric parameters published for each source come with a five-by-five covariance matrix in which the diagonal elements refer to the [formal uncertainties](#) or standard errors (in practice, Gaia DR2 contains standard errors  $\sigma_x$  and  $\sigma_y$  combined with correlation coefficients  $\rho_{xy}$  so that the covariance  $\sigma_{xy}$  needs to be reconstructed as  $\sigma_{xy} = \sigma_x \sigma_y \rho_{xy}$ ). The off-diagonal elements quantify correlations between the astrometric parameters of the source, for instance between right ascension and parallax. These covariances should be included, following the guidelines provided in the paper [Gaia Data Release 2: Using Gaia parallaxes](#), for instance when propagating uncertainties on subsets and/or linear combinations of astrometric parameters, when estimating model parameters using, for instance,  $\chi^2$ -fitting, maximum likelihood, Bayesian inference, etc., or when sampling the astrometric uncertainties in a Monte Carlo procedure.

## Systematic astrometric errors

Bright-star astrometry ( $G \lesssim 13$  mag) in Gaia DR2 is limited by calibration uncertainties. Special care is needed with the parallaxes of very bright stars ( $G \lesssim 5$  mag) as they have [additional systematic errors](#) due to calibration imperfections. As described on the Gaia [website](#) and in Section 5 of the paper [Gaia Data Release 2: The astrometric solution](#), significant spatial (source-to-source) correlations, of up to 0.04 mas in parallax and 0.07 mas per year in proper motion, are present on small ( $\lesssim 1$  deg) and intermediate (~20 deg) angular scales. Gaia DR2 also contains spatial (source-to-source) correlations in the uncertainties which also affect the astrometric parameters. These correlations have implications when, for instance, [estimating the \(systematic\) uncertainty of the mean parallax or proper motion of a cluster](#) or for the [measurement of the internal kinematics of star clusters](#).

## Galactic coordinates

For each source, the [Gaia Archive](#) provides pre-computed Galactic coordinates in the `gaia_source` table, named 'l' and 'b'. Their computation, as well as the transformation of the proper motion vector and the associated [covariance data](#), is described in the [on-line documentation](#).

## Cartesian coordinates and velocities

The [on-line documentation](#) describes a generic recipe for the transformation of astrometric data, including the propagation of the [covariance matrix](#). A special case of such a transformation is the derivation of Cartesian positions and velocities from the five measured astrometric parameters and the measured radial velocity. Although the formal expression for the true space velocity vector  $\mathbf{v}$  in terms of the true source parameters is simple, namely,

$$\mathbf{v} = (A_v / \varpi) \cdot (\mathbf{p} \mu_{\alpha^*} + \mathbf{q} \mu_{\delta} + \mathbf{r} \mu_r),$$

where  $\varpi$  is the true parallax in mas,  $\boldsymbol{\mu} = (\mu_{\alpha^*}, \mu_{\delta})$  is the true proper motion vector in  $\text{mas year}^{-1}$ ,  $\mu_r = v_r \cdot \varpi / A_u$  is the radial proper motion,  $v_r$  is the radial velocity in  $\text{km s}^{-1}$ ,  $[\mathbf{p}, \mathbf{q}, \mathbf{r}]$  constitutes the normal triad ( $\mathbf{p} = [-\sin \alpha, \cos \alpha, 0]'$ ,  $\mathbf{q} = [-\cos \alpha \sin \delta, -\sin \alpha \sin \delta, \cos \delta]'$ ,  $\mathbf{r} = [\cos \alpha \cos \delta, \sin \alpha \cos \delta, \sin \delta]'$ ), and  $A_u = 149,597,870,700$  m and  $A_v = 4.740,470,446$   $\text{km year s}^{-1}$  equal the astronomical unit, the presence of parallax in the denominator introduces the same risks when applying this equation with observed parallaxes as for the case of [distances estimated from observed parallaxes](#) by simple inversion. For Cartesian coordinates and velocities, the same recommendations therefore apply as for distances.

## Avoiding biases

Generic advice on how to treat and interpret astrometric data and optimally exploit their content is given in the paper [Gaia Data Release 2: Using Gaia parallaxes](#) and the associated [astrometry inference tutorials](#). In general, users should treat the derivation of quantities or model parameters from the astrometric data as a [Bayesian inference problem](#) using priors that represent the already available information. Wherever possible, one should formulate the inference problem in the space of observables, in which the measurement uncertainties are well understood and the handling of covariances is straightforward, and infer parameters by comparison of the measurements to predicted observations using a forward model using all relevant information that is available (for instance proper motions, magnitudes, and colours all contain distance information). Data should preferentially not be truncated or filtered on, for instance, relative precision since such operations distort the underlying sample and can introduce severe biases. In case filtering is applied, this shall be properly mimicked in the (forward) model.



# 6. Photometry

## Broad-band, white-light G magnitudes

Gaia DR2 contains G magnitudes, averaged over the 22 months of observations, for more than 1.6 billion sources. Details are provided in the paper [Gaia Data Release 2: The photometric content and validation](#).

## Blue and Red Photometer BP and RP magnitudes

Gaia DR2 also contains average, integrated BP and RP magnitudes for more than 1.38 billion sources. Details are provided in the paper [Gaia Data Release 2: The photometric content and validation](#).

## Photometric quality – photometric flux-excess factor

The photometric BP and RP colour information in Gaia DR2 suffers from systematics at the faint end ( $G \gtrsim 19$  mag) of the survey, in crowded regions, and in the vicinity of bright stars. In these cases, the lack of a specific treatment of blending and decontamination from nearby sources, combined with inaccurate removal of the astronomical and stray light background contributions from the integrated fluxes in the BP and RP photometers, has resulted in overestimated fluxes in primarily BP and to a lesser extent RP. This, in turn, leads to measured fluxes that are inconsistent between the G band on the one hand and the BP and RP bands on the other hand in the sense that the sum of the flux values in the latter two bands may be significantly larger than that in G whereas it is expected that, for “normal” spectral energy distributions, the sum of fluxes in BP and RP should be comparable to that in G. A quantitative indication of this effect is included in Gaia DR2 in the form of the so-called flux-excess factor (field 'phot\_bp\_rp\_excess\_factor' in the data). This indicator can be used to judge the extent to which the photometry of a given source is compromised. Following the paper [Gaia Data Release 2: The photometric content and validation](#), the standard recommendation is to filter out spurious photometry by using the (ADQL) conditions:

```
AND phot_bp_rp_excess_factor < 1.3 + 0.060*power(phot_bp_mean_mag - phot_rp_mean_mag, 2)
AND phot_bp_rp_excess_factor > 1.0 + 0.015*power(phot_bp_mean_mag - phot_rp_mean_mag, 2).
```

## Variability information and light curves (epoch photometry)

Gaia DR2 contains photometric times series for more than 550,000 variable sources consisting of [Cepheids](#), [RR Lyraes](#), [Miras](#), and semi-regular variables as well as [high-amplitude Delta Scuti](#), [BY Draconis](#), and [SX Phoenicis](#) stars and [short-timescale variables](#). Light curves, which are ancillary data to the main catalogue, are not provided as tables. Instead, they can be retrieved from the [Gaia Archive](#) using the Virtual Observatory [DataLink](#) interface that is in preparation for [Gaia DR3 and beyond](#), when large amounts of non-tabular data such as spectra will be available. The dedicated [light curves and DataLink tutorial](#) provides a full worked example of how to retrieve epoch photometry for variable stars.

Only a subset of the variable stars classified as a certain type have been characterised in detail. Hence not every variable star appears in the more detailed [vari\\_\\* tables](#) and a fraction of the classifications may be erroneous. As explained in the [on-line documentation](#), the subset that has been classified by variability type has been based on at least two transits; the overlapping subset that has been studied in detail has been based on either 12 or more or 20 or more transits. Of course, sources that are not explicitly listed as 'VARIABLE' in Gaia DR2 are not necessarily photometrically constant. Details about the variability processing can be found in the paper [Gaia Data Release 2: Summary of variability processing and analysis results](#).

## Gaia DR2 extinction coefficients $k_G$ , $k_{BP}$ , and $k_{RP}$

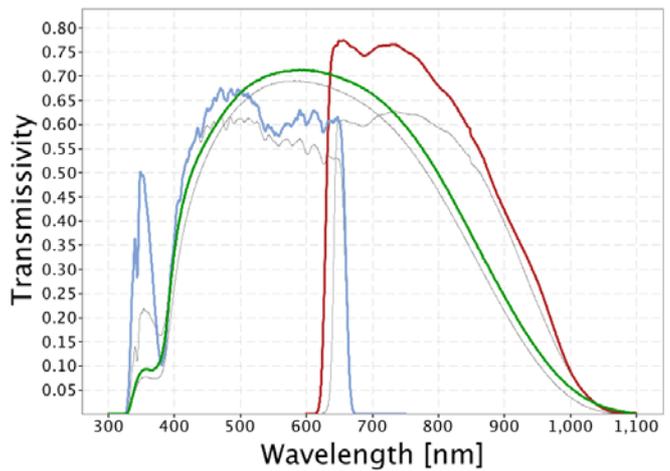
As explained in the paper [Gaia Data Release 2: First stellar parameters from Apsis](#), extinction  $A_G$  and reddening  $E(BP-RP)$  estimates for selected individual objects are included in the set of published Gaia DR2 [astrophysical parameters](#). These extinction/reddening values must be treated with care, with guidance provided in Section 6 of the aforementioned paper.

Recipes for computing extinction coefficients for the three [Gaia passbands](#) as a function of colour and extinction and as a function of effective magnitude and metallicity are described in Table 1 in the paper [Gaia Data Release 2: Observational Hertzsprung-Russell diagrams](#) and Table 2 of [this paper](#), respectively. Both studies assume a ratio of total to selective extinction of  $R = 3.1$ . Both for the extinction/reddening estimates and for the extinction coefficients, users should not blindly apply the Gaia DR2 values but should carefully verify, using the aforementioned papers, that the applicable caveats do not invalidate the intended usage for their particular science case.



## Gaia DR2 passbands

As explained on the [Gaia website](#), nominal, pre-launch passband curves were published in 2010. They should not be used for Gaia DR2 studies. A first, preliminary version for the three Gaia passbands defining the Gaia DR2 photometric system was released for internal use in the Gaia DR2 data processing in July 2017. These passbands underlie the [zero points](#) of the [G](#), [BP](#), and [RP](#) magnitudes as well as the astrophysical parameters published in Gaia DR2. A new set of so-called 'revised' passbands, with a significant change in the RP cutoff wavelength position, was subsequently derived in October 2017 and published alongside Gaia DR2. After the Gaia DR2 release, [Maíz Apellániz & Weiler \(2018\)](#) published a third set of response curves for the Gaia DR2 photometric system, together with corresponding [zero points](#) and with suggested modifications of the [G flux scale](#). The combination of the suggested corrections and their set of response curves, by mitigating inconsistencies in the published photometry, allows the user to obtain more accurate synthetic photometry and is a valuable attempt to obtain the best possible results from the Gaia DR2 data. Differences between the published Gaia DR2 magnitudes and those generated using the Maíz Apellániz & Weiler (2018) passbands and corrections are below  $\sim 0.01$  mag for the calibration sources used to derive them. However, for "peculiar" objects, for instance very red objects, strongly reddened early-type stars, or quasars, differences can be larger. When using stellar tracks and/or isochrones, users should check carefully which Gaia DR2 passbands have been implemented to predict the Gaia DR2 photometry.



Revised Gaia DR2 passbands. Credit: ESA/Gaia/DPAC, P. Montegriffo, F. De Angeli, C. Cacciari.

Differences between the published Gaia DR2 magnitudes and those generated using the Maíz Apellániz & Weiler (2018) passbands and corrections are below  $\sim 0.01$  mag for the calibration sources used to derive them. However, for "peculiar" objects, for instance very red objects, strongly reddened early-type stars, or quasars, differences can be larger. When using stellar tracks and/or isochrones, users should check carefully which Gaia DR2 passbands have been implemented to predict the Gaia DR2 photometry.

## Relationships with other photometric systems

Polynomial relations between [Gaia DR2 passbands](#) and common photometric systems, in particular the Hipparcos, Tycho-2, Sloan SDSS DR12, Johnson-Cousins, 2MASS, and GSC-2.3 systems, can be found in the [on-line documentation](#).

## Gaia DR2 magnitude errors

Gaia DR2 contains measured fluxes and associated uncertainties (referred to as errors in the [Gaia Archive](#)). Whereas the fluxes have been converted to magnitudes, the latter have been published without associated uncertainties. The fundamental reason for this is that the non-linear relation between magnitudes and fluxes causes the classical, first-order error propagation of low signal-to-noise flux measurements to break down. In particular, the asymmetric nature of the magnitude error cannot be captured in this way. Nonetheless, for most applications in the high signal-to-noise flux regime, users can safely assume that  $\text{magnitude\_error} \approx (2.5 / \log(10)) * (\text{flux\_error} / \text{flux}) = 1.086 / \text{flux\_over\_error}$ , where log is the natural logarithm.

## Systematic photometric errors and corrections

Gaia DR2 photometry is affected by systematic errors and the [following corrections](#) are recommended to be applied to the published Gaia DR2 magnitudes:

- $2 < G \leq 6$ : compute corrected magnitudes as  $G_{\text{cor}} = -0.047344 + 1.16405 * G - 0.046799 * G^2 + 0.0035015 * G^3$  (following Appendix B of the paper [Gaia Data Release 2: The photometric content and validation](#));
- $2 < G \leq 4$ : compute corrected magnitudes as  $BP_{\text{cor}} = BP - 2.0384 + 0.95282 * G - 0.11018 * G^2$  (ibidem);
- $2 < RP \leq 3.5$ : compute corrected magnitudes as  $RP_{\text{cor}} = -13.946 + 14.239 * RP - 4.230 * RP^2 + 0.4532 * RP^3$  (ibidem);
- $6 < G \leq 16$ : compute corrected magnitudes as  $G_{\text{cor}} = G - 0.0032 * (G-6)$  (following [Maíz Apellániz & Weiler 2018](#));
- $G > 16$ : compute corrected magnitudes as  $G_{\text{cor}} = G - 0.032$  (ibidem).

The quantities G, BP, and RP above are the magnitudes available via the [Gaia DR2 Archive](#).

## Gaia DR2 synthetic photometry

As explained on the [Gaia website](#), synthetic Gaia DR2 photometry in the [VEGAMAG system](#) can be derived as:

$$\text{synthetic\_mag} = -2.5 * \log_{10}(\text{synthetic\_flux}) + \text{synthetic\_ZP}.$$

If synthetic\_flux is expressed in units of photons per second per square metre, the values of the synthetic photometric zero points (synthetic\_ZP) are:

- 26.3914 mag for G;
- 26.1297 mag for BP ( $G \leq 10.87$  mag);
- 26.1698 mag for BP ( $G > 10.87$  mag), and;
- 25.3255 mag for RP.



These zero points supersede those published with Gaia DR2 and should only be used with synthetic fluxes derived using the [Maíz Apellániz & Weiler \(2018\) passbands](#). They should not be used with the Gaia DR2 fluxes available from the [Gaia Archive](#). Details on synthetic photometry, including zero points in the AB system, can be found in the [on-line documentation](#).

## 7. Spectroscopy

### Radial velocities

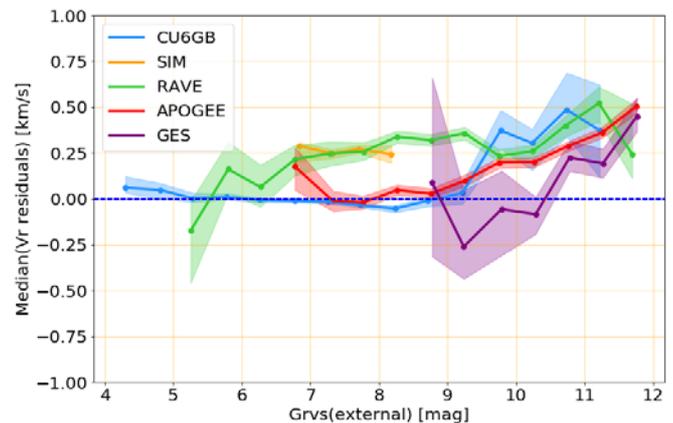
Gaia DR2 contains the median radial velocities for 7 million stars brighter than  $\sim 13^{\text{th}}$  magnitude limited to effective temperatures in the range of  $\sim 3550$  to  $\sim 6900$  K. The median refers to the (as yet unpublished) epoch radial velocities collected during the 22 months underlying Gaia DR2. In addition to the effective-temperature limitation, which is Gaia-DR2 specific and which will be relaxed in [forthcoming releases](#), specific object types have been [filtered out](#) before releasing the data, including emission-line stars and double-lined spectroscopic binaries. Single-lined spectroscopic binaries are present but have been treated as [single stars](#) in the data processing. Also for these objects, only a median radial velocity, together with the scatter in the underlying epoch velocities, is provided. Details can be found in the paper [Gaia Data Release 2: Properties and validation of the radial velocities](#).

### Radial-velocity zero point

Although the formal precision of the radial velocities at the bright end is of the order of  $\sim 0.2$ - $0.3$   $\text{km s}^{-1}$ , systematic errors of comparable size exist in the radial velocities. There is in particular a positive bias ([zero point offset](#)) with respect to other radial-velocity surveys which increases with magnitude up to about  $+0.5$   $\text{km s}^{-1}$  at the faint end of the  $G_{\text{RVS}}$  scale (which is the non-published, integrated magnitude in the RVS bandpass).

### Spurious radial velocities

Up to [~70,000 radial velocities in Gaia DR2](#) (<1% of the data), predominantly in [crowded regions](#), are potentially affected by unmodeled light from nearby objects. The handful of radial velocities in the tails of the distribution, with absolute values above  $\sim 550$   $\text{km s}^{-1}$ , should be treated with care.



Radial-velocity zero point. Credit: [Katz et al. \(2018\)](#).

## 8. Astrophysical parameters

### Astrophysical parameters

Gaia DR2 contains the effective temperatures for more than 161 million sources brighter than  $17^{\text{th}}$  magnitude with effective temperatures in the range  $3300$ – $8000$  K. For a subset of  $\sim 87$  million sources, also the line-of-sight extinction  $A_G$  in the [G band](#) and the reddening  $E(\text{BP}-\text{RP})$  are given while for a part of this subset containing  $\sim 76$  million sources, the luminosity and radius are available as well under the assumption of zero extinction. Details can be found in the paper [Gaia Data Release 2: First stellar parameters from Apsis](#).

### Caveats and systematic errors

The astrophysical parameters in Gaia DR2 were determined only from the three broad-band photometric measurements ([G](#), [BP](#), and [RP](#)) and the parallax on a star-by-star basis (parallax was not used for estimating effective temperature though). The



strong degeneracy between effective temperature and extinction / reddening that exists when using only broad-band photometry has necessitated strong assumptions in the data processing. This, in combination with limitations in the model grids that have been used to train the machine-learning algorithms employed in the data processing as well as specific choices made in the data processing such as a [positivity constraint for  \$A\_G\$](#) , has introduced severe systematics that are described in detail in the paper [Gaia Data Release 2: First stellar parameters from Apsis](#).

## 9. Reference frame

### International Celestial Reference Frame (ICRF) and System (ICRS)

As of [1 January 1998](#), the [International Celestial Reference System \(ICRS\)](#) is the standard celestial reference system adopted by the International Astronomical Union (IAU). The ICRS is the set of prescriptions and conventions together with the modelling required to define, at any given time, a triad of orthogonal axes. The ICRS has its origin at the barycentre of the Solar system, with axes that are space-fixed and kinematically non-rotating with respect to the most distant sources in the Universe. In practice, the ICRS is materialised by the [International Celestial Reference Frame \(ICRF\)](#) through the coordinates of a defining set of extra-galactic objects (quasars). The ICRS alleviates astronomers from worrying about precession, nutation, etc. In the ICRS, connecting different epochs only involves [propagating positions](#) based on proper motions.

For historical background, it is worth pointing out that pre-1998, prior to astronomers being able to define and use the ICRS and ICRF, dynamical reference systems were used based on observations of star positions tied in some way to moving objects in the Solar system. These old reference systems refer to a mean equator and equinox at a given reference epoch (typically B1950 or J2000), requiring precession / nutation models and corrections to deal with the time-variable fundamental plane. To within  $\sim 25$  mas, mean J2000 equatorial coordinates are the same as ICRS coordinates such that, for “ordinary” applications, they can in practice be considered to be the same. For high-accuracy applications, however, the appropriate frame conversion shall be used when converting old, mean J2000 equatorial coordinates to ICRS coordinates.

### Gaia DR2 reference frame, reference epoch, and time coordinate

Gaia DR2 astrometry provides stellar coordinates (and parallaxes and proper motions) valid for epoch J2015.5 (roughly mid-2015, where J stands for Julian year). The time coordinate for Gaia DR2 is [barycentric coordinate time](#) (TCB). TCB is equivalent to the proper time experienced by a clock at rest in a coordinate frame co-moving with the barycentre of the Solar system but outside its gravity well, therefore not influenced by the gravitational time dilation caused by the Sun and the other bodies in the Solar system. The [celestial reference frame](#) of Gaia DR2, described in the [Gaia Data Release 2: The celestial reference frame](#) paper and referred to as Gaia-CRF2, is based on the Gaia DR2 positions of a sample of 556,869 quasars obtained from a positional cross-match with the ICRF3-prototype and AllWISE quasar catalogues. This set defines a kinematically non-rotating reference frame in the optical domain. A subset of 2820 quasars used in Gaia-CRF2 have accurate VLBI positions that allow the axes of the Gaia-CRF2 reference frame to be aligned with the [International Celestial Reference System](#) (ICRF3-prototype) radio frame.

### Epoch propagation

For some applications, for instance [cross-matching](#) Gaia DR2 to other catalogues, it may be needed to propagate the Gaia DR2 astrometry from J2015.5 to a different epoch, for instance J2000.0. Often, a simple linear correction based on the proper motion and epoch difference will be sufficient. However, in general, and in particular close to the poles or when long time intervals ( $\sim 100$  year) and/or high proper motions ( $\sim 500$  mas year<sup>-1</sup>) are involved, it is recommended to apply the fully-reversible time transformation of the six-element vector (positions, parallax, proper motions, radial velocity) and their associated errors and correlation coefficients. This framework is described in [The Hipparcos and Tycho Catalogues](#) documentation and has been implemented in the [Gaia Archive](#) by means of ready-to-use [ADQL](#) functions, in particular:

- [EPOCH\\_PROP](#) to propagate the six observables (positions, parallax, proper motions, radial velocity);
- [EPOCH\\_PROP\\_ERROR](#) to propagate the errors and correlations coefficients of the propagated coordinates.

Further details can be found in the [epoch propagation tutorial](#).

### Orientation and rotation of the Gaia DR2 reference frame for bright objects

As described on the Gaia [website](#) and in Section 5.1 of the paper [Gaia Data Release 2: The astrometric solution](#), there are strong indications that the Gaia DR2 positions and proper motions for stars with  $G \lesssim 12$  mag have a significant orientation and



rotation bias, respectively. It is advisable to correct for these biases for high-accuracy applications and/or applications that cover long time intervals, for instance using the first-order Equations (6)—(9) in [this paper](#):

- $(\alpha_{cor}-\alpha) \cos\delta = +\epsilon_X \cos\alpha \sin\delta + \epsilon_Y \sin\alpha \sin\delta - \epsilon_Z \cos\delta;$
- $\delta_{cor} - \delta = -\epsilon_X \sin\alpha + \epsilon_Y \cos\alpha;$
- $\mu_{\alpha^*,cor} - \mu_{\alpha^*} = +\omega_X \cos\alpha \sin\delta + \omega_Y \sin\alpha \sin\delta - \omega_Z \cos\delta;$
- $\mu_{\delta,cor} - \mu_{\delta} = -\omega_X \sin\alpha + \omega_Y \cos\alpha,$

where  $\mu_{\alpha^*,cor} = (d\alpha_{cor}/dt) \cos\delta_{cor}$  and  $\mu_{\delta,cor} = d\delta_{cor}/dt$  are the corrected components of the proper motion in the ICRS and  $\mu_{\alpha^*} = (d\alpha/dt) \cos\delta$  and  $\mu_{\delta} = d\delta/dt$  are the components as published in Gaia DR2 (and similar for the coordinates). The corrective orientation and spin vectors at J2015.5 are, respectively,  $(\epsilon_X, \epsilon_Y, \epsilon_Z) = (-0.019 \pm 0.137, +1.304 \pm 0.245, +0.553 \pm 0.045)$  mas and  $(\omega_X, \omega_Y, \omega_Z) = (-0.068 \pm 0.051, -0.051 \pm 0.042, -0.014 \pm 0.036)$  mas year<sup>-1</sup>.

## 10. Solar-system objects

Gaia DR2 contains epoch astrometry for 14,099 known, pre-selected Solar-system objects based on ~1.5 million CCD observations. [G-band photometry](#) is available for about half of the transits. Further details can be found in the paper [Gaia Data Release 2: Observations of Solar-System Objects](#).

## 11. Limitations

### Filtering

Gaia DR2 is an early release based on a limited amount of data, simplifications in the data processing, and imperfect calibrations. The [validation of Gaia DR2](#) was done in various stages and led to the decision to filter out parts of the processed data prior to publication. The applied filters are summarised in the [Gaia DR2 release paper](#) and details can be found in the various data-processing papers.



## Survey completeness and selection function

The Gaia DR2 catalogue covers the magnitude range  $G \sim 3\text{--}21$  mag, with the bright limit set by a hard instrument constraint and the faint limit resulting from the commanded threshold at 20.7 mag combined with on-board magnitude estimation errors of a few tenths of a magnitude. The catalogue is essentially complete between  $G \sim 12$  and  $\sim 17\text{--}18$  mag. Away from the Galactic poles, Gaia DR2 has an ill-defined faint magnitude limit, resulting from the on-board management of observation window conflicts and resource limitations in crowded fields: above  $\sim 300,000\text{--}400,000$  stars per square degree, the faint magnitude limit of Gaia DR2 is as bright as  $\sim 17^{\text{th}}\text{--}18^{\text{th}}$  magnitude.

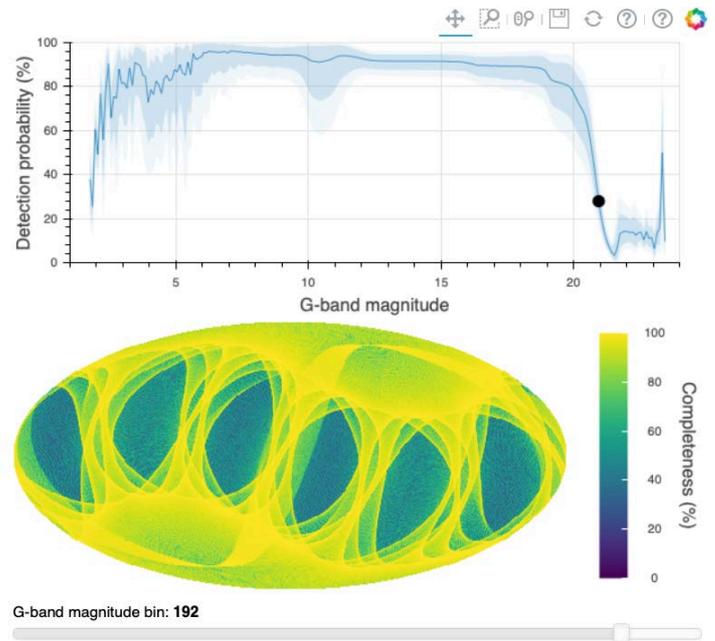
The source list is also incomplete for high proper-motion stars:  $\sim 20\%$  of stars with proper motions exceeding 0.6 arcseconds per year are missing in Gaia DR2. The [completeness near bright sources](#) is limited and the [effective angular resolution of Gaia DR2](#) is  $\sim 0.4$  arcseconds (compared to a PSF size of  $\sim 0.1$  arcseconds) and an [instrumental resolution of  \$\sim 0.23\$  arcseconds](#).

The combination of the non-uniform sky coverage over the limited, 22-month time span underlying the release combined with [gaps in the data](#) and the filtering on data quality that has been implemented prior to the publication of Gaia DR2 has resulted in regions of the sky showing spurious source-density fluctuations that reflect the [scanning-law](#) pattern, in particular for sources fainter than  $\sim 17^{\text{th}}$  magnitude.

Useful attempts to create a Gaia DR2 selection function have been made by community scientists, e.g., by Boubert, Everall, and Holl in their [Completeness of the Gaia-verse](#) and by [Rybizki et al. \(2018\)](#).

### Double and multiple stars

During the Gaia DR2 data processing, including the astrometric and radial-velocity processing, all sources were treated as single stars. This means that, for binary and multiple stellar systems, the astrometry and median radial velocity can be (much) less accurate than suggested by the quoted formal uncertainties as a result of unquantified model errors.



Gaia DR2 selection function tool. Credit: [Boubert & Everall \(2019\)](#).

## 12. Data access

### Gaia Archive

The Gaia DR2 data is available from the [ESA Gaia Archive](#). Four main access mechanisms exist:

1. "Basic" user interface for finding data for individual objects or [short object / coordinate lists](#). Data fields to be returned from the [gaia\\_source table](#) can be selected through "Display columns" and constraints can be added through "Extra conditions";
2. "Advanced (ADQL)" user interface accepting [ADQL queries](#) to select customised samples;
3. [ADQL](#) access using a Table Access Protocol (TAP) server, for instance from within Python or from the [command-line](#);
4. [Direct download](#) of the main [gaia\\_source](#) table for offline analysis (in total 550 GB of compressed CSV files).

For user convenience, selected external catalogues are available along with Gaia DR2 to enable joint queries ([AllWISE](#), [GSC-2.3](#), [Hipparcos new reduction](#), [Pan-STARRS DR1](#), [PPMXL](#), [SDSS DR9](#), [2MASS](#), [Tycho-2](#), [URAT-1](#)). It is also possible in the "Advanced (ADQL)" search tab to use an [external TAP search](#) to combine, for instance, any [VizieR](#) table from a refereed paper with Gaia DR2. As of February 2020, the ESA Gaia Archive only accepts secure communication (https).



## User registration

Gaia Archive users are advised to access the data using a self-registration user account (click the “SIGN IN” button in the top-right corner on the [entry page](#)). A registered account has the benefits of a dedicated, private user space, allowing to save and share uploaded tables, queries, and result tables with other registered users, accompanied with cross-match functionalities and [enhanced user quota](#). Upon demonstrated need, power users can request (temporary) changes to their quota by sending a motivated request to the [Gaia Helpdesk](#).

## Query limits

Authenticated, registered users benefit from:

- Asynchronous ADQL query time-out of 120 minutes (compared to 90 minutes for anonymous users);
- Unlimited number of rows for asynchronous ADQL queries (compared to 3 million rows for anonymous users);
- ADQL query output up to 20 GB;
- User table space of 1 GB (non-existing for anonymous users).

## Tips and tricks

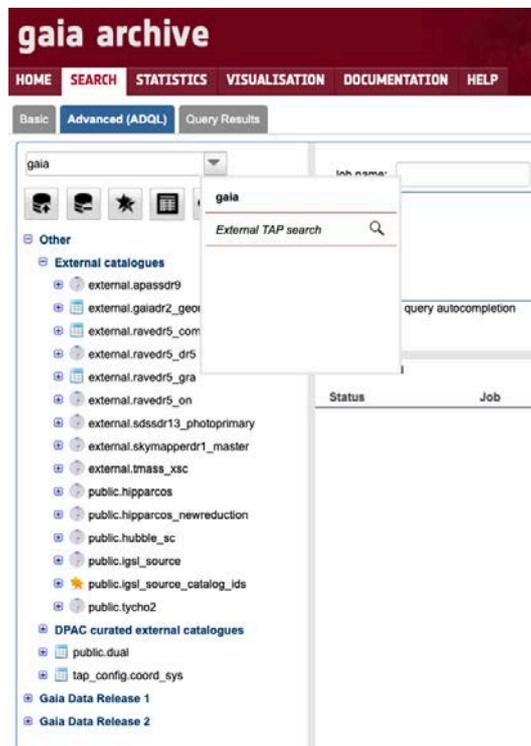
Due to the ease of writing [ADQL queries](#), it is straightforward to launch unwieldy requests resulting in run times of days and/or hundreds of millions of selected sources. We have collected a series of [tips and tricks](#) and a [dedicated tutorial](#) to avoid query time-outs. Most importantly, be careful with on-the-fly computations, use the COUNT(\*) function to first estimate the size of the output of your query, use the 'random\_index' field to create statistical samples of reduced size, and use indexed fields (indicated in **bold font** in the “Advanced (ADQL)” search tab in the [Archive](#)) as well as pre-computed fields (such as 'bp\_rp' for BP-RP colour index or 'parallax\_over\_error') to speed up queries.

## Python

ESA provides an [Astropy](#) package named [astroquery.gaia](#) to access the [Gaia Archive](#) using [Python](#) and/or [Jupyter Notebook](#). A generic package named [astroquery.utils.tap](#) allows accessing any [TAP and TAP+](#) compliant service from within Python.

## Partner data centres

Besides the main [ESA Gaia Archive](#), four partner data centres provide the same data, each using their own interface: the Institut für Astrophysik Potsdam ([AIP](#)), the Astronomisches Rechen-Institut ([ARI](#)), the ASI Space Science Data Centre ([ASI SSSDC](#)), and the Centre de Données astronomiques de Strasbourg ([CDS](#)). Note that some data centres (e.g., CDS) serve rounded numbers (e.g., parallax with four decimal places) so that the data sets are, strictly speaking, not identical.



*Gaia Archive interface. Credit: ESA/Gaia/DPAC.*

# 13. Data model

The Gaia DR2 catalogue is stored in and made available as a series of tables in a database, a key component of the [Gaia Archive](#). The [data model](#) defines the table names as well as the names and attributes such as the units of the various fields (“columns”) in the various tables. For stars (and quasars), the [main table](#) is called [gaiadr2.gaia\\_source](#) (for historical reasons, the renormalised unit weight error – [RUWE](#) – values for each object are published in a separate table called [gaiadr2.ruwe](#)). The [gaia\\_source](#) table contains all summary information related to the [astrometry](#), [photometry](#) (G, BP, and RP magnitudes for Gaia DR2), [spectroscopy](#) (radial velocities for Gaia DR2), and [astrophysical parameters](#) of the objects. Epoch photometry of all stars classified as variable is contained in [DataLink tables](#). Further tables contain [photometric variability information](#) and [Solar-system object information](#).



# 14. Data queries

## Astronomical Data Query Language (ADQL)

[Astronomical Data Query Language](#) – ADQL – is a simple and intuitive yet powerful means to retrieve information from tables stored in a database. ADQL is an SQL dialect specifically developed for astronomy by the Virtual Observatory. As an example, to retrieve the source identifier, right ascension, declination, and parallax for all sources in Gaia DR2 that have parallaxes in the range 15–50 mas as well as G-band magnitudes in the range 9–9.5 mag, one would simply execute, in the “Advanced (ADQL)” sub-tab of the “SEARCH” tab:

```
SELECT source_id, ra, dec, parallax
FROM gaiadr2.gaia_source
WHERE parallax between 15 AND 50 AND
phot_g_mean_mag >= 9 and phot_g_mean_mag <= 9.5
```

This example query results in a list of 1929 objects that can be downloaded from the Gaia Archive, for instance in [VOTable](#), CSV, or [FITS](#) format, for further offline analysis. Before being able to draft this query, it is of course necessary to inform oneself on the name and units of the various data fields ('source\_id', 'ra', 'dec', 'parallax', and 'phot\_g\_mean\_mag' in this example) through inspection of the [data model](#).

Be careful that ADQL discriminates between integers and floats: “WHERE parallax > 1000/150” is different from “WHERE parallax > 1000.0/150.0”.

## ADQL tutorials

The [Gaia Archive](#) contains a series of written as well as video [tutorials](#) to demonstrate how to select subsets of Gaia DR2 data that match your science case. Beginners are advised to start with the [white-dwarfs exploration](#) and [open-cluster analysis](#) tutorials. A worked example of getting access to [epoch photometry for variable stars](#) is contained in the [DataLink service](#) tutorial.

## ADQL resources

A variety of ADQL resources exists, among others [ADQL Gaia Archive examples](#), [ADQL Gaia Archive syntax overview](#) (including, for instance, special functions, mathematical functions, trigonometric functions, and data-type casting functions), the [ADQL Cookbook](#), an [ADQL Reference Card](#), and an [ADQL short course for Gaia](#).



## External TAP Query examples

ConeSearch sorted by distance	<pre>SELECT DISTANCE( POINT(ICRS', ra, dec), POINT(ICRS', 266.41683, -29.00781)) AS dist, * FROM gaiadr2.gaia_source WHERE 1=CONTAINS( POINT(ICRS', ra, dec), CIRCLE(ICRS', 266.41683, -29.00781, 0.08333333)) ORDER BY dist ASC</pre>
Cone search filtered by magnitude, ordered by magnitude Galactic centre, radius 5'	<pre>SELECT * FROM gaiadr2.gaia_source WHERE 1=CONTAINS( POINT(ICRS', ra, dec), CIRCLE(ICRS', 266.41683, -29.00781, 0.08333333)) AND phot_g_mean_mag &gt;= 10 AND phot_g_mean_mag &lt;= 15 ORDER BY phot_g_mean_mag ASC</pre>
TGAS catalogue filtered by magnitude and parallax	<pre>SELECT * FROM gaiadr1.tgas_source WHERE parallax &gt;= 15 AND parallax &lt;= 50 AND phot_g_mean_mag &gt;= 9 and phot_g_mean_mag &lt;= 9.5</pre>
ADQL positional cross-match: Hipparcos vs Gaia, 1" radius	<pre>SELECT TOP 10 *, distance( POINT(ICRS', hip.ra, hip.dec), POINT(ICRS', gaia.ra, gaia.dec)) AS dist FROM gaiadr2.gaia_source AS gaia, public.hipparcos AS hip</pre>

Cancel

*ADQL query examples. Credit: ESA/Gaia/DPAC.*

# 15. Cross-matches

## Gaia source identifiers and designations

Sources in Gaia DR2 and subsequent releases are unambiguously identified by their so-called designation, which is a composition of the release name plus the source identifier (e.g., “Gaia DR2 3212878018378138752”). For practical reasons, designations are only provided in the main [gaia\\_source](#) table in the [Gaia Archive](#), while source identifiers ([source\\_id](#)) are used in that table as well as in all other tables for efficient joins.



Due to improved calibrations and observation-matching recipes regularly becoming available in the course of the data processing, the source list of each data release differs from and is independent of that of previous releases. Each data release comes with a neighbourhood table to help users trace source identifiers to the [preceding release](#).

## Cross-matching Gaia DR2 to Gaia DR1

Around 25% of sources in Gaia DR1 have no Gaia DR2 counterpart with the same source\_id. However, a close object in Gaia DR2 exists for most of them (99.9% of the Gaia DR1 sources have a Gaia DR2 source closer than 1 arcsecond). A cone search, possibly supplemented by a magnitude-difference threshold, is typically sufficient to carry out a suitable cross-match. The auxiliary table [gaiadr2.dr1\\_neighbourhood](#) contains pre-computed angular separations and G-band magnitude differences for all stars [with small proper motions](#).

## Find Gaia data for your list of sources

For lists up to ~100 objects, it is convenient to use the "File" tab in the "Basic" search tab. The description of the format of the input file can be found by hovering the mouse over the "Select a file with Target Names" link and clicking once the question mark appears. The input file has no header and should contain one entry per line, either specifying object name or coordinates (name resolution is performed using the [Simbad and NED name resolvers](#)). In the background, an intermediate job with the source properties (name, coordinates, parallax, and proper motion) is created and [cross-matched](#) against Gaia DR2.

For lists exceeding ~100 objects, users may find it more convenient to upload their input file in the "Advanced (ADQL)" tab and perform the [cross-matching](#) in the [Gaia Archive](#), following the example given in the [white-dwarfs exploration tutorial](#).

## Cross-matching Gaia DR2 to other catalogues

There is no single recipe for cross-matching Gaia DR2 to other catalogues: there is no general rule to determine whether, for instance, a Gaia DR2 observation (visible, sub-arcsecond resolution, epoch 2015.5) and an IRAS measurement (far-infrared, 2 arcminute resolution, epoch 1983) are physically related to the same object. In general, the challenge is to treat the different spatial resolutions between Gaia operating from space and other (often ground-based) surveys, in addition to (possible) differences in the [epoch of the data](#), the underlying [celestial reference frames](#), the spectral response / [photometric passbands](#), etc. In other words: cross-matching catalogues is a scientific exercise, always involving a case-by-case, user-driven trade-off between completeness and contamination.

Nonetheless, the [Gaia Archive](#) provides a series of pre-computed cross-matches with popular catalogues ([AllWISE](#), [APASS DR9](#), [GSC-2.3](#), [Hipparcos new reduction](#), [Pan-STARRS DR1](#), [PPMXL](#), [RAVE DR5](#), [SDSS DR9](#), [2MASS](#), [Tycho-2](#), and [URAT-1](#)). In these cases, simple geometrical conditions have been used to select candidate match pairs based on a maximum distance on the celestial sphere and priority has been given to purity over completeness (for the case of [Hipparcos](#), there is also a [cross-match table](#) that prioritises completeness). The [Gaia DR2 cross-match paper](#) provides details on the two tables that materialise each cross-match: "best\_neighbour" and "neighbourhood". While the "best\_neighbour" table contains a single entry, namely the neighbour with the highest value of the figure of merit, the "neighbourhood" table contains all good neighbours, i.e., all objects whose position error ellipses overlap within a  $5\sigma$  confidence level with the target. Of course, users can impose supplemental restrictions besides proximity to refine candidate matches, e.g., maximum flux difference (e.g.,  $|G - K| < 1$  mag) and/or 3D location in space (e.g.,  $4 \text{ mas} < \text{parallax} < 5 \text{ mas}$ ).

For fast moving objects and/or when the time difference between the two catalogues subject to the cross-matching is significant, proper motions may need to be considered to bring the two catalogues to a common epoch. To put this into perspective, Gaia DR2 contains ~1 million stars moving faster than 50 mas per year, i.e., 5 arcseconds per century. A final word of warning: [VizieR](#) catalogues include the synthetic columns "RAJ2000" and "DEJ2000", which aim at providing [ICRS coordinates propagated to epoch J2000](#) and which are used in the [CDS cross-match service](#). Users should keep in mind that these computed coordinates only reflect true ICRS coordinates at epoch J2000 if the catalogue contains proper motions. For example, in a VizieR cross-match between [Gaia DR2](#) and [Pan-STARRS DR1](#), positions in the former can be properly propagated from [J2015.5](#) to J2000, but positions in the latter cannot be shifted from ~J2013 to J2000 because Pan-STARRS lacks proper motions. In this case, following the [cross-match tutorial](#), it is advisable to first [propagate](#) the Gaia DR2 positions from J2015.5 to the Pan-STARRS epoch and then do the cross-matching. The proper motion dispersion of the source population being matched can be used to estimate positional uncertainties at the other epoch.

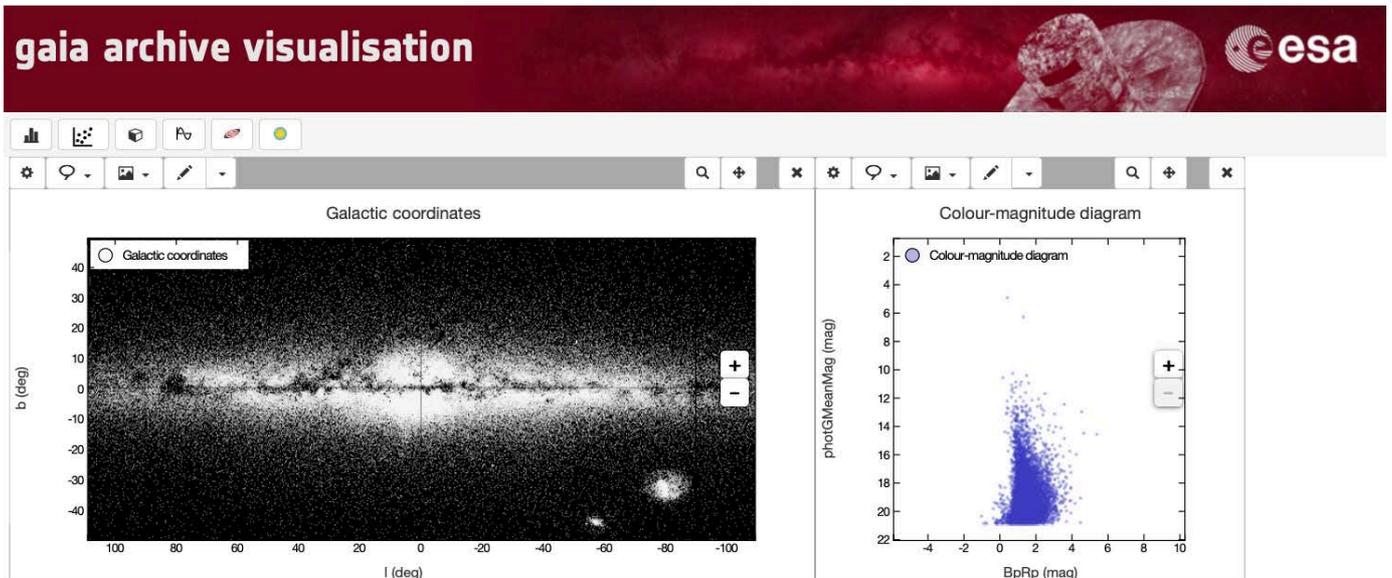
In general, the [cross-match tutorial](#) that comes with Gaia DR2 is highly-recommended reading for anything having to do with cross-matches. It shows, for instance, how to bulk download the pre-computed cross-match tables in the Gaia Archive using a salami technique of chopping the tables into pieces, downloading these, and then re-assembling the pieces into one table using the "SELECT TOP" and "OFFSET" ADQL commands. The tutorial also provides hints to improve the efficiency of cross-match queries (e.g., minimise the use of "OFFSET", and ensure that the coordinates of the smaller catalogue are used inside the POINT function in constructs like "WHERE CONTAINS(POINT('ICRS', u.ra, u.dec), CIRCLE('ICRS', g.ra, g.dec, 0.003))=1").



# 16. Visualisation and exploration

## Gaia Archive Visualisation Service (GAVS)

The [Gaia Archive](#) features a powerful, flexible, interactive, built-in [visualisation portal](#) that provides a multi-panel visualisation desktop in a browser tab supporting, for instance, the creation of 1D histograms, 2D scatter plots and binned plots, 3D scatter plots, time series, and all-sky source density and integrated luminosity maps.



[Gaia Archive Visualisation Service](#). Credit: ESA/Gaia/DPAC.

## Tool for OPERations on Catalogues And Tables (TOPCAT)

[TOPCAT](#) – developed by Mark Taylor (University of Bristol, United Kingdom) – is a stand-alone, interactive graphical viewer and editor for tabular data. It offers a variety of ways to view and analyse tables, including a browser for the cell data themselves, viewers for information about table and column metadata, and facilities for sophisticated interactive 1D, 2D, 3D, and higher-dimensional visualisation, calculating statistics and joining tables using flexible cross-matching algorithms. TOPCAT uses Virtual Observatory ([VO](#)) standards, has a built-in [Table Access Protocol](#) (TAP) client, supports various data formats (e.g., [FITS](#), [VOTable](#), and [CDF](#)), and is ideal for exploring (subsets of) the Gaia DR2 catalogue.

## ESA Sky

[ESA Sky](#) – developed by the [ESA Science Data Centre](#) – is a web-based, science-driven discovery portal providing full access to the entire sky as observed with a range of multi-wavelength space astronomy missions. It provides the possibility to explore the Gaia DR2 catalogue superimposed on [background sky views](#), ranging from the ultraviolet to the far-infrared. When small portions of the sky are selected, Gaia DR2 data can be downloaded and proper motions can be visually inspected.

## Gaia Sky

[Gaia Sky](#) – developed by the Gaia group at the Astronomisches Rechen-Institut (ZAH, Universität Heidelberg, Germany) – is a real-time, 3D astronomy visualisation software that runs on Windows, Linux, and Mac OS X. As shown by the examples contained in the [YouTube playlist](#), Gaia Sky is extremely versatile, for instance covering full-dome displays.

# 17. Documentation

## User documentation

Gaia DR2 comes with extensive user documentation which can be accessed [on-line](#). The separate chapters have also been indexed at [ADS](#) to allow citing them. This documentation covers an overview of the mission, detailed descriptions of the data processing and calibration models, and the [Gaia DR2 data model](#).

## Data release papers

Gaia DR2 comes with a [series of peer-reviewed papers](#) in Astronomy & Astrophysics that [present the mission, summarise the release](#), provide an overview of the [validation of the data](#), and provide overviews of the processing and detailed validation of the various data products. Six more performance verification Gaia collaboration papers describe the scientific quality of the data through examples focused on [the celestial reference frame](#), [the observational Hertzsprung-Russell diagram](#), [observations of Solar-system objects](#), [mapping the Milky-Way disk kinematics](#), [the kinematics of globular clusters and dwarf galaxies around the Milky Way](#), and [variable stars in the colour-magnitude diagram](#).

## Guide-to-scientists videos

Gaia DR2 is accompanied by a collection of YouTube videos entitled [Gaia Data Release 2: a guide for scientists](#). These videos, which are based on interviews with Gaia scientists who processed the data, provide an overview of Gaia DR2, a guide (to the use) of the data, and warnings and explanations on the limitations of the data.

## Issues discovered after Gaia DR2 was released ('known issues')

Of course, Gaia DR2 as published is not perfect and many caveats are described in the [accompanying papers](#). Any known issues that affect the data and that are not included in these papers or in the [on-line documentation](#), including recommendations on how to deal with these issues in your science research, are being collected on the [Gaia DR2 known issues](#) webpage. These issues, where relevant, have been included in this primer.

## How to cite and acknowledge Gaia DR2

When using Gaia DR2 data in your research, please cite both the [Gaia DR2 release](#) and the [Gaia mission](#) papers and, in addition, please acknowledge the people and institutes in the Gaia DPAC [Data Processing and Analysis Consortium](#) following [these guidelines](#).



# 18. Gaia Data Release 3 and beyond

## Gaia Data Release 3

Preparations for Gaia DR3 are currently in full swing. Gaia DR3 is based on more data (34 months compared to 22 months for Gaia DR2) and features improved calibration models and reduced systematic errors. The [foreseen contents](#) of this release include, among others, non-single star results as well as BP/RP and RVS spectra for selected sources. The third data release comes in two parts, with an early release (Gaia EDR3) of astrometry and integrated G/BP/RP photometry in the third quarter of 2020 and the remainder of the data products in the second half of 2021 (but see the latest news [here](#)).

## Gaia Data Release 4 and beyond

Gaia Data Release 4 (Gaia DR4) is the [full release of the nominal-mission data](#), including “everything”, for instance exo-planet lists and all epoch and transit data for all sources. A release schedule remains to be defined. Further data releases have not been defined but it is clear that a fifth release covering the [extended mission](#) is needed as a minimum.

## Mission extension

Gaia’s nominal, five-year lifetime ended on 16 July 2019, with the [Whitehead Eclipse Avoidance Manoeuvre](#). Since then, Gaia has been operating nominally in its extended mission phase. ESA funding for continued spacecraft and science operations has been confirmed until the [end of 2020](#) and has been provisionally approved until the end of 2022. The cold-gas supply used to control the spin rate of the spacecraft is in principle sufficient to continue spacecraft operations until at least the end of 2024.

## Science performance

The predicted astrometric, photometric, and spectroscopic uncertainties of the 4<sup>th</sup> data release (Gaia DR4), that will be based on all data collected during the nominal, five-year mission, is described on our [website](#) and in the paper [Science performance of Gaia, ESA’s space-astrometry mission](#) (see also the [PyGaia Python toolkit](#)). It has been assumed in these predictions that all calibrations in the data processing have been successfully applied and have converged and that fundamental calibration floors exist: ~0.01 mas per field-of-view transit for astrometry, ~3 milli-magnitude at CCD level for the G band and ~5 milli-magnitude at CCD level for the BP/RP bands, and ~0.5 km s<sup>-1</sup> for median radial velocities.

Predictions for Gaia DR3 are best achieved by scaling based on Gaia DR2 (while respecting the above-mentioned calibration floors, in particular at the bright end, for  $G \lesssim 13$  mag). As an example, position and parallax uncertainties scale with time<sup>0.5</sup> whereas proper-motion uncertainties scale with time<sup>1.5</sup>. For Gaia DR3, therefore, proper motion uncertainties for faint stars are expected to be a factor  $(34 \text{ months} / 22 \text{ months})^{1.5} \sim 1.9$  smaller compared to Gaia DR2.

# 19. Keeping up to date

## Websites

Extensive information and news on the mission, the science it produces, the people involved, image galleries, education resources, etc. is provided on the [Gaia website](#) for professional astronomers. Additional information, mainly targeted at broader audiences, may also be found on the [ESA science & technology website](#) and on the [ESA corporate web portal](#).

## Newsletters

Our e-mail newsletters deliver regular [Gaia mission updates](#) and [ESA science archives updates](#).

## Social media

We also maintain a Gaia [Twitter](#) feed, a Gaia [Facebook](#) page, and a Gaia [LinkedIn](#) account.



# 20. Get in touch

## Gaia Helpdesk

Do our [tutorials](#) or [documentation](#) fail? Are our [Archive tips & tricks](#) insufficient? Is your question not in our [FAQ](#)? Feel free to contact the [Gaia Helpdesk](#).

## Communicate your Gaia science results with support from ESA

We are always keen to promote new science based on Gaia data. Please [contact us](#) around the time of paper submission at the latest if you are willing to share your work with the Gaia mission and get more attention for your research by having a story or press release published through ESA. We have professional story writers and graphics designers to support you and we are experienced in dealing with embargoed material.

## Feedback

Any [feedback](#) (positive or negative), for instance on this primer or on our [website](#), is very welcome.

## Change log

- Issue 1.0, 9 December 2019: First release.
- Issue 1.1, 17 January 2020: Fixed typos. Added sentence on TCB. Added [guide-to-scientists YouTube videos](#). Added hyperlink to known issue on Gaia DR2 to DR1 cross-match.
- Issue 1.2, 21 January 2020: Updated numerical values of uncertainties of the bright-faint star reference frame offset.
- Issue 1.3, 14 February 2020: Added sentence that the ESA Gaia Archive exclusively accepts secure communication (https). Added sentence that some partner data centres serve rounded data. Minor text updates in photometric passband text.
- Issue 1.4, 6 March 2020: Updated numerical values of the bright-faint star reference frame offset following a [corrigendum](#).
- Issue 1.5, 12 June 2020: Updated section on selection function (added links to [Completeness of the Gaia-verse](#) and [Gaia DR2 gaps](#) webpages) and added paragraph in cross-match section promoting the [cross-match tutorial](#).

## Credits front page

- [Fairing logo](#): ESA.
- [The Gaia sky in colour](#): ESA/Gaia/DPAC, A. Moitinho / A.F. Silva / M. Barros / C. Barata, University of Lisbon, Portugal; H. Saviotto, Fork Research, Portugal.

