

THE VARIABLE SOURCES IN THE GAIA ARCHIVE

L. EYER

Département d'Astronomie de l'Université de Genève, Chemin Pegasi 51, 1290 Versoix, Switzerland
e-mail: laurent.eyer@unige.ch

Abstract. At the core of the Gaia mission is a multi-epoch survey consisting of astrometric, photometric, spectrophotometric, and spectroscopic measurements. The astrometric time series provides parallaxes and proper motions, along with information on astrometric binary systems. The photometric time series offers a means to investigate the variability of the sources. Due to their whole-sky, multi-epoch nature, multiple instruments, their magnitude range covering 21 magnitudes, and their remarkable photometric precision, these data allow us to describe the variability of celestial phenomena in an unprecedented manner. For the third Gaia Data Release (DR3), the data collection spanned 34 months, with a median number of field-of-view measurements in the G band of about 44, reaching up to 270. At publication time, DR3 delivered the largest collection of variable sources with an associated classification across the entire sky. All these sources have their G , G_{BP} , G_{RP} epoch data published and accessible in the Gaia ESA archive. In summary, there are 10.5 million variable sources, including 9.5 million variable stars and 1 million QSOs. Additionally, 2.5 million galaxies were identified thanks to spurious variability caused by the non-axisymmetric nature of galaxies and the way Gaia collects data. Moreover, all the epoch data and time series of nearly 1.3 million sources in a pencil beam around the Andromeda galaxy are published, regardless of their status (constant or variable); This dataset is known as the Gaia Andromeda Photometric Survey. We also introduce the citizen science project, GaiaVari to classify variable stars, the Focused Product Release delivered on October 10, 2023. In the future, DR4 will cover 66 months, and we hope DR5 will have accumulated 10.5 years of data.

Keywords: stars: variables: general - Galaxy: stellar content - catalogues - stars: oscillations - binaries: eclipsing - starspots

1 Introduction

In astronomy, from its beginnings to now, the primary means of learning about the Universe and its constituents has been through the photons. We can find the following text¹ by Evry Schatzman in his book "Astrophysique" Schatzman (1963):

The fundamental basis of astrophysics lies in the radiation that reaches us from celestial objects. Whether in the visible or invisible spectrum, it is always a form of radiation that we measure.

Only from the late 20th century were alternative methods used to learn about our Universe, first as exceptions to the sentence above, such as the cosmic rays (Hess 1912), then with important implications, such as the neutrinos (from the Sun, Davis et al. 1968), and the gravitational waves in 2015 (Abbott et al. 2016).

The classical subdivision in observational astronomy about the three main observation branches are:

- Astrometry: Knowledge of celestial objects obtained from their positions, motions, and shapes. Probably this discipline is as old as humanity.
- Photometry: Knowledge of celestial objects obtained from the measurement of their integrated light over a range of wavelengths. Historically, photometry went hand in hand with astrometry as they are the most basic perceptions of the human eye. The magnitude scale, as it is well known, is inherited from the ancient greeks (and formalized by Norman Pogson, cf. Pogson 1856).

¹Translated from French.

- Spectroscopy: Knowledge of celestial objects obtained from their spectra. This technique emerged in the 19th century, with Joseph von Fraunhofer discovering dark lines in the spectra of the Sun and other stars, and Gustav Kirchhoff and Robert Bunsen correctly interpreting these dark lines as being caused by specific atomic elements (for a brief historical account see [Appenzeller 2013](#)).

It should be remarked that we gain knowledge not only from the source but also from what is between the source and the observer, e.g. the interstellar medium. Observations from the ground can also allow us to determine the properties of the atmosphere and its evolution, such as in [Burki et al. \(1995\)](#).

Each of these three main observation branches benefits from multi-epoch observations. The most knowledge is obtained when these different multi-epoch domains are combined. Such merging is fruitful for finding and calibrating standard candles, identifying stellar populations, probing the invisible with microlensing events (e.g. merging the photometry and the astrometry as in [Sahu et al. 2022](#)) and determining astrophysical parameters thanks to binaries ([Eyer et al. 2015](#)), to pulsating stars (Baade-Wesselink method, [Carney et al. 1992](#)), etc.

It should be noted that on the side of photometric surveys, there has been a booming period for these past 30 years with OGLE ([Udalski 2003](#)), HAT ([Bakos et al. 2004](#)), Kepler ([Borucki et al. 2010](#)), and ZTF ([Bellm et al. 2019](#)), to name just a few. The future is also bright with the LSST ([Ivezić et al. 2019](#)) and PLATO ([Rauer et al. 2014](#)), again to name just a few. Spectroscopy is entering into the era of systematic multi-epoch global surveys, obviously with Gaia, but also with 4MOST ([de Jong et al. 2019](#)), and in the future with WST ([Mainieri et al. 2024](#)).

Now, if we turn to the Gaia cornerstone mission, the "tour de force" of Gaia is to assemble multi-epoch data for the entire sky on these three main observation branches: astrometry, photometry, and spectroscopy.

In this short text, we will mostly restrict the presentation to what has been put in the Gaia archive for the variable sources.

2 The Gaia mission: An introduction

Let us review key aspects of the Gaia mission: Gaia is a cornerstone mission of the European Space Agency science program ([Gaia Collaboration, Prusti et al. 2016](#)), tasked with performing a systematic survey of all objects brighter than $G=20.7$, attaining a survey of more than 2 billion celestial sources. The measurements collected by Gaia consist of position (astrometry), brightness, and colours (photometry and spectrophotometry), along with radial velocity (spectroscopy). It should be noted that the survey for radial velocities is restricted to brighter sources (about $G = 15$). Gaia was launched on December 19, 2013 by a Soyuz rocket. The CCD camera aboard Gaia is the largest ever deployed in space, reaching nearly one billion pixels. Over its ten-year mission, each source, on average, will be observed 140 times in each of the 9 CCDs² in the G band, as well as in each of the BP and RP CCDs. The number of measurements for the radial velocity spectrometer instrument will be comparatively fewer, averaging approximately 80 times. The results of the Gaia mission, that the Data Processing and Analysis Consortium processes, are made accessible to the public through Data Releases. DR4 and DR5 are anticipated to be delivered in 2026 and no later than 2032, respectively.

We will focus on this article on the DR3 results, which compiles 34 months of observations and nearly one trillion CCD measurements over the whole celestial sphere.

Several factors contribute to Gaia's exceptional status as a mission:

- Unprecedented Astrometry: Gaia provides unparalleled astrometric data, i.e. positions, parallaxes, and proper motions for more than a billion stars over the whole celestial sphere. The astrometric precision is available on the Gaia webpage (<https://www.cosmos.esa.int/web/gaia>³): For Gaia DR4, that is, for the nominal mission, the parallax uncertainty is $22 \mu\text{as}$ at magnitude 15. This precision can be compared to the 20-25 μas estimates from pre-launch assessments in 2006 for a G2 star at magnitude 15 (see e.g. [Eyer 2006](#)). It's important to note that the errors in position and proper motion need to be multiplied by 0.8 and 0.5, respectively. The astrometry is also able to detect astrometric binaries, exoplanets, microlensing effects, etc...
- Three Instruments on a Single Platform: astrometry, [spectro-]photometry, spectroscopy with the derivation of radial velocities. This results in a unprecedented uniformity of data across the entire celestial sphere.

²for each of the 7 rows at the exception of row 5, for which the number is 8.

³Note: This page has undergone changes in URL and error estimates following the better understanding of the performance.

- **Multi-Epoch Measurements:** The mission conducts multi-epoch measurements of the entire sky, providing a comprehensive view of variable celestial objects, see also Section 3.
- **10-Year Time Baseline:** With a 10-year mission duration (if all goes as planned), Gaia offers high frequency precision, particularly beneficial for periodic objects.
- **Quasi-Simultaneous Measurements:** Gaia achieves quasi-simultaneous measurements. Other multi-band photometric surveys do not have this simultaneity and this complicates somewhat the analysis. For more details, see Section 3.
- **Time-Domain Selection Function:** Because of its predefined scanning law, Gaia allows us to determine the time-domain selection function, enabling the understanding of observational biases.
- **Extensive Number of Measurements:** Gaia’s data volume is significant, with nearly 1 trillion CCD measurements for DR3 (i.e. 34 months).
- **High Dynamical Range:** Gaia covers a high dynamical range, capturing data from the brightest sources ($G \approx 1.7$ in DR3) to magnitudes as faint as $G \approx 21/22$.
- **Space-based Operation:** Gaia’s position in space provides stability and allows us to access the entire celestial sphere from a single platform, which is not achievable by any single ground-based optical telescope. However, operating in space presents its own challenges, such as the high cost of space missions⁴, the impossibility of repairs, micrometeoroid hits, and solar flares/the radiation damages. Indeed Gaia’s location at L2, 1.5 million kilometers from Earth in the direction opposite to the Sun, leaves it unshielded by Earth’s magnetic field.
- **Cyclic Improvements:** The Gaia consortium implements cyclic improvements, performing systematic data analyses, with not only more data at each cycle, but with improved calibrations, and enhanced outlier detection.

Each of these elements makes Gaia extremely fruitful for variable and binary stars. These are among the objects that benefit the most from Gaia. Due to the semi-regular Gaia sampling, strictly periodic objects can be studied in detail despite the gaps in the scanning law. Also transient sources are and will be classified in the data releases (Eyer et al. 2023; Rimoldini et al. 2023) and are also detected in real time by the Gaia Science Alert System (Hodgkin et al. 2021). Indeed, since the beginning of its science operations, Gaia has continuously (with few interruptions) provided alerts, identifying potentially time-sensitive events. If not addressed promptly, these events could lead to significant scientific loss. To date (June 2024), the Science Alerts team has released over 25,000 of these alerts. About 27% of the alerts are classified, with approximately 60% of these classified alerts identified as supernovae.

3 The Gaia time sampling

The time sampling⁵ of Gaia is very particular and semi-regular. It is optimized to achieve astrometric precision as uniform as possible across the entire celestial sphere. This sampling is known as the Nominal Scanning Law (NSL): Gaia continuously sweeps the sky, with the spacecraft’s rotation axis precessing on a cone with a 45-degree opening angle to the Sun. A full precession cycle on the cone is completed every 63 days. The spacecraft rotates on its axis every 6 hours, and the 106.5-degree angle between Gaia’s two fields-of-view (preceding and following fields-of-view) results in a time interval of 1 hour and 46 minutes between them (the exemplary angle of 253.5-degree between the following and preceding fields-of-view takes thus 4 hours 14 minutes).

Figure 1 shows the sky plot for the number of observations (the field-of-view transit). We took the field `matched_transits`, i.e. the total number of field-of-view transits matched to this source. This number is higher or equal to the `num_selected_g_fov` of `gaiadr3.vari_summary` table, because it does not take into account the number of observations that are removed.

The second plot, Fig. 2, presents the number of CCD observations as a function of ecliptic latitude. On a field crossing the CCD observations are separated by 4.85 seconds. The Annex A of Eyer et al. (2017) provides

⁴As Prof. Bohdan Paczynski humorously noted, "One dollar in space is worth less than one dollar on Earth".

⁵referred to as the cadence in LSST/Vera C. Rubin observatory.

detailed properties of the scanning law for 5 years, including various representations, histograms of the lags, and spectral windows for different regions of the sky.

Because the scanning law is well-defined, it is possible to study the selection function. [Eyer & Mignard \(2005\)](#) conducted a study on periodic signals based on an earlier design of the spacecraft and a 5-year scanning law. They concluded that the ability to detect the correct period of a periodic signal strongly depends on the ecliptic latitude. For DR3, significant differences in detection exist even at fixed ecliptic latitudes due to the scanning law creating regions with a low number of observations; see [Fig. 1](#).

There are two exceptions to the Nominal Scanning Law:

1. At the beginning of the mission, a 28-day period was planned to assist with photometric calibrations. During this time, the same stars and regions of the sky were scanned nearly continuously. The axis of rotation of the spacecraft was kept in the ecliptic plane (at a 45-degree angle to the Sun) without precession, allowing the ecliptic poles to be scanned regularly during these 28 days, with repeated sequences of intervals 1h46m, 4h14m, 1h46m, etc. As the spacecraft orbits around the Sun, stars at lower ecliptic latitudes (in absolute value) are observed for shorter sequences.
2. The second exception occurred when, for 12 months⁶, a modified NSL was introduced to improve the astrometric solution. As a positive side effect it reduces aliasing peaks. In this mode, the precession of the spacecraft's rotation axis was reversed. However, the spacecraft's rotation direction remained unchanged due to the constraints of the Time Delay Integration (TDI) mode. Note that the reverse motion of the rotation axis on the cone increased the consumption of cold gas.

Although the scanning law determines which stars are observed, understanding the temporary absence of an observed source is challenging. Several factors could explain a missing observation: the star might have truly disappeared (the most interesting case) by becoming very faint due to an eclipse or occultation, it could be just below the detection threshold of the less precise sky mapper and thus not observed, the star can be perturbed by the other field-of-view (e.g. bright star/extended object) which is superposed on the same focal plane, the data might have been deleted due to excessive data volume onboard, or technical problems might have occurred.

4 The Gaia G-band photometric precision

[Fig. 3](#) shows the uncertainty in the mean G-band photometry as a function of the G magnitude. There are two plots: the results from DR1 and DR3. DR2 is not displayed to avoid too many figures and it falls between the two. The plot is based on a selection of stars with absolute ecliptic latitudes below 60 degrees to avoid the perturbation of ecliptic pole scanning. The uncertainty is estimated from the flux standard deviation divided by the square root of the number of measurements (after some clipping). Naturally, this uncertainty decreases as the number of measurements increases. In case of variable sources, the standard deviation is obviously larger. Although more measurements may reveal more instrumental calibration features, this is not the case here. On the contrary, some features visible in DR1 are less prominent or have disappeared in DR3 demonstrating the improvements of the calibrations. It should be noted that in the Gaia archive, uncertainties are provided in flux rather than in magnitude. Therefore, these uncertainties need to be converted into magnitude uncertainties for [Fig. 3](#). We used the formula $\text{Uncertainty}(\text{magnitude}) = 1.086 \times \text{Uncertainty}(\text{flux})/\text{flux}$, where the term 1.086 is coming from $2.5/\ln(10)$.

We will not enter into the debate on using flux versus magnitudes to evaluate the brightness of a celestial source here, opening Pandora's box. The main point is to avoid being too dogmatic about these issues and to use the approach that makes the most sense for the problem at hand. As a side note, negative fluxes were unfortunately excluded in the mean calculation process ([Riello et al. 2021](#); [Evans et al. 2023](#)). Negative fluxes can occur due to statistical variations or overestimation of the background for example. This exclusion introduces a bias in the magnitude estimation at the faint end: the sources are brighter than they are in reality. Starting from DR4, negative fluxes will be included.

5 The Gaia data releases and the variability

The data release process is iterative and advances with each step in both quantity and quality, this is particularly evident for the variability processing and analysis. So let us have a quick look at the DR1 and DR2 data releases.

⁶6 months will be included in DR4, the rest in DR5.

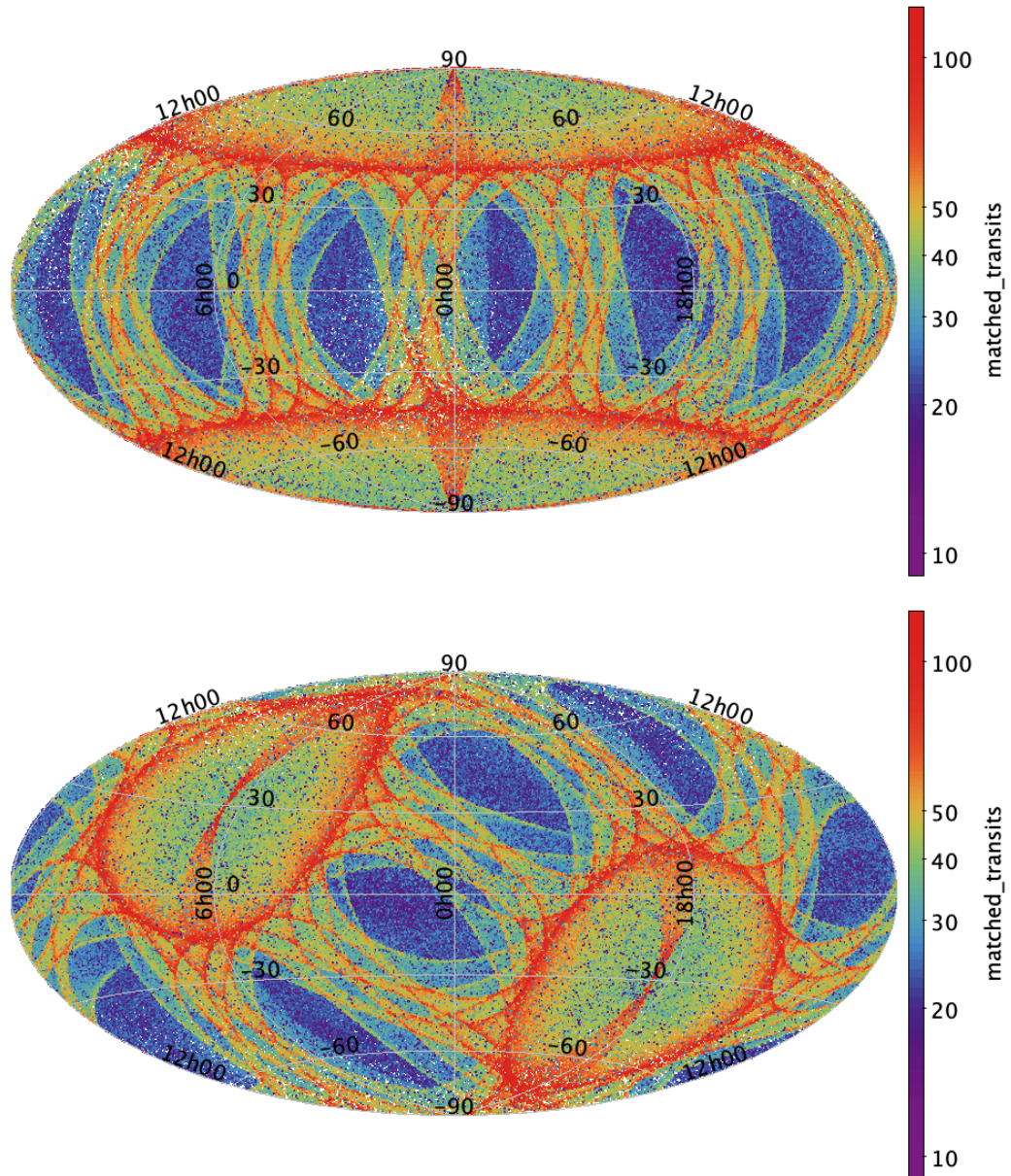


Fig. 1. Number of fields-of-view data per source in ecliptic coordinates (top) and Galactic coordinates (bottom) for DR3 in Aitoff projections. We see the complex sampling which is clearly not completed yet, with blue regions, i.e. low number of transits, around the ecliptic plane. We see that one of these low number of measurements is in the direction of the lower bulge at positive Galactic longitudes.

A global pipeline was set up for the variability processing and analysis of Gaia data. Initially, it was applied to the cycle 1 data, the details of this processing can be found in [Eyer et al. \(2017\)](#). The DR1 publication also covers the data properties of cycle 1, along with certain aspects of the 5-year sampling obtained through simulations. Based on 14 months of data, the DR1 variable star catalogue showcased a selection of 3,194 variable stars. This data release consisted of two types of variability, namely Cepheids and RR Lyrae stars in the direction of the Large Magellanic Cloud ([Clementini et al. 2016](#)). The Large Magellanic Cloud, situated near the South Ecliptic pole, benefited from the regular sampling during the 28 days of the Ecliptic Pole Scanning Law. In addition, a performance verification paper ([Gaia Collaboration, Clementini et al. 2017](#)) presented the TGAS parallaxes for Cepheids and RR Lyrae stars, demonstrating that already the improvements with respect to Hipparcos were

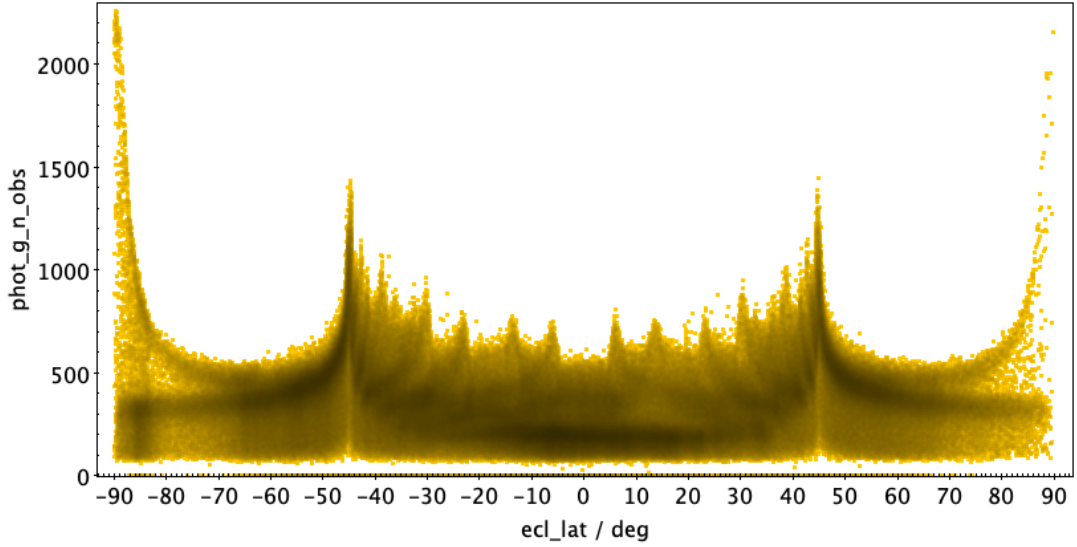


Fig. 2. Number of CCD observations per source as a function of ecliptic latitude (β) for DR3. A large number of observations are visible at $\pm 45^\circ$ due to the Nominal Scanning Law. Additionally, stars at the ecliptic poles (both ends of the plot) have an exceptionally high number of observations, exceeding 2000, due to the Ecliptic Pole Scanning Law, which was in effect for the first 28 days of the mission.

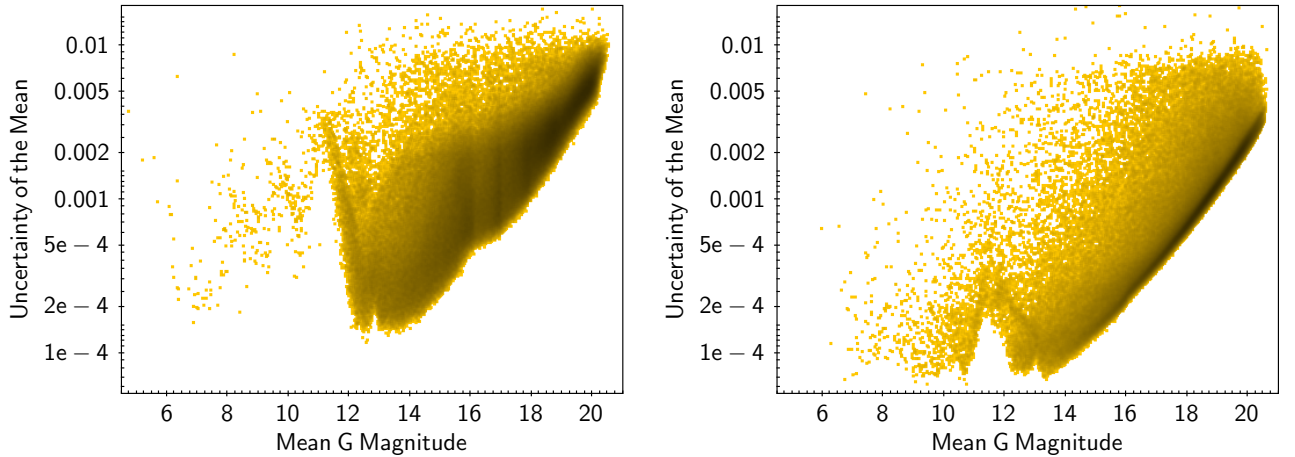


Fig. 3. Uncertainty on the mean versus the mean G magnitude. Two plots are displayed side by side, for the comparison of DR1 (left) and DR3 (right) photometry. We see that not only has the noise gone down thanks to the square root of the number of measurements but also that the calibration has significantly improved. The various changes in windowing schemes or gates are better taken into account in DR3 than in DR1.

impressive.

The DR2 based on 22 months of operation data reached 550,737 variable stars (Holl et al. 2018). The DR2 contained the following variability types: RR Lyrae stars, Cepheids, long period variables, δ Scuti/SX Phoenicis stars, stars with rotation modulation and stars with short time scale variability. A performance verification paper (Gaia Collaboration, Eyer et al. 2019) presented the properties of variable sources across the Hertzsprung-Russel Diagram. As a remarkable result, we can follow the path of stars, their motion in the Hertzsprung-Russel Diagram as their magnitude and colour change⁷.

More generally, information and details for all data releases can be found in various articles (many are listed in this section for details on variability processing and analysis) or in the Gaia documentation associated with

⁷https://www.cosmos.esa.int/web/gaia/gaiadr2_cu7

the data releases on the archive website, <https://gea.esac.esa.int/archive/documentation/GDR3/> for DR3. This document, available online and also in PDF format, contains more than 1,500 pages.

It is also important to consult the "known issues" page. For each data release, there is a section on the Data tab of <https://www.cosmos.esa.int/web/gaia/data>. Under these sections, you can find the known issues for each data release. For example, the known issues for DR3 can be found at the following address: <https://www.cosmos.esa.int/web/gaia/dr3-known-issues>.

A note on the archive names: you have in the ESA Gaia archive (<https://gea.esac.esa.int/archive/>) the different data sets associated with the different data releases, `gaiadr1.table_name`, `gaiadr2.table_name`, `gaiadr3.table_name`, `gaiadr3.table_name`, `gaiadr3.table_name`, `gaiadr3.table_name` which relates the Gaia Data Release 1 (DR1), Gaia Data Release 2 (DR2), Gaia Early Data Release 3 (EDR3), Gaia Data Release 3 (DR3), Focused Product Release (FPR) respectively. For a given source, the `source_id` can change. A last remark on the nomenclature related to variability: for the tables names associated to the variability processing and analysis, starting from the second data release, these table names start with "vari_table_name".

In the following sections, we will focus more on the DR3 as it has the most numerous outputs and diversity in variability types.

5.1 The DR3 results

For the third data release of Gaia, the variability analysis is described in Eyer et al. (2023). In summary, as of the data release in June 2021, Gaia provided the largest catalogue ever of classified variable sources across the entire sky. In DR3, we have 10.5 million variable sources, comprising 9.5 million variable stars and 1 million AGN. Additionally, thanks to a spurious perturbation in the photometry, we identified 2.5 million galaxies.

A general overview of the DR3 result, as mentioned above, can be found in Eyer et al. (2023). In the Gaia archive <https://gea.esac.esa.int/archive/>, the different associated tables related to the variability processing and analysis can be seen in Fig. 4.

Variability

- `gaiadr3.vari_agm`
- `gaiadr3.vari_cepheid`
- `gaiadr3.vari_classifier_class_definition`
- `gaiadr3.vari_classifier_definition`
- `gaiadr3.vari_classifier_result`
- `gaiadr3.vari_compact_companion`
- `gaiadr3.vari_eclipsing_binary`
- `gaiadr3.vari_epoch_radial_velocity`
- `gaiadr3.vari_long_period_variable`
- `gaiadr3.vari_microlensing`
- `gaiadr3.vari_ms_oscillator`
- `gaiadr3.vari_planetary_transit`
- `gaiadr3.vari_planetary_transit_13june2022`
- `gaiadr3.vari_rad_vel_statistics`
- `gaiadr3.vari_rotation_modulation`
- `gaiadr3.vari_rrlyrae`
- `gaiadr3.vari_short_timescale`
- `gaiadr3.vari_summary`
 - `solution_id`
 - `source_id`**
 - `num_selected_g_fov`

Fig. 4. Tables available related to the variability processing and analysis in the Gaia archive. We opened the `vari_summary` table with just the first three lines.

Variability types	Variability types
(1) ACV CP MCP ROAM ROAP SXARI: α 2 CVn, (Magnetic) Chemical Peculiar, Rapidly Oscillating Am/Ap ⁸ SX Arietis stars	(13) MICROLENSING: Star with microlensing event
(2) ACYG: Alpha Cygni-type variable	(14) RCB: R Coronae Borealis stars
(3) AGN: Active Galactic Nuclei (including Quasars)	(15) RR: RR Lyrae stars
(4) BCEP: Beta Cephei type variable	(16) RS: RS Canum Venaticorum type variable
(5) BE GCAS SDOR WR: Subset of eruptive variable types: B-type emission line star, Gamma Casiopeiae, S Doradus, and Wolf-Rayet	(17) S: Set of stars with short timescale variability
(6) CEP: Cepheid variable types: delta Cepheid, anomalous Cepheid, and type-II Cepheid	(18) SDB: Subdwarf B stars
(7) CV: Cataclysmic variable	(19) SPB: Slowly Pulsating B-star variable
(8) DSCT GDOR SXPHE: delta Scuti, gamma Doradus, and SX Phoenicis	(20) SN: Supernovae
(9) ECL: Eclipsing Binaries of types: Beta Persei (Algol), Beta Lyrae, and W Ursae Majoris	(21) SOLAR_LIKE: Stars with solar-like variability due to flares, spots, and rotational modulation
(10) ELL: Ellipsoidal variable	(22) SYST: Symbiotic variable star
(11) EP: exoplanetary transits	(23) WD: Variable White Dwarf of types: ZZ Ceti (DAV, ZZA), V777 Her (DBV, ZZB), and GW Vir (DOV, ZZO)
(12) LPV: Long Period Variable stars of types: omicron Ceti (Mira), OGLE Small Amplitude Red Giants, and semiregular	(24) YSO: Young Stellar Object

Table 1. The 24 variability types considered in DR3. Please note that in some cases, a denomination encompasses several different variability types.

A vast and systematic compilation of the literature was done in [Gavras et al. \(2023\)](#); it combines the sources of 152 catalogues crossmatched with the Gaia data and contains 4.9 million variable sources. The catalogue is available at the CDS (Centre de Données astronomiques de Strasbourg, <https://cds.unistra.fr>, [Gavras et al. \(2022\)](#)).

Part of this compilation was used as a training set for the supervised classification ([Rimoldini et al. 2023](#)). For the classification, the methods employed are Random Forest and eXtreme Gradient Boosting models in a mixture of multi-class and binary classifiers. There are 24 variability (grouped) classes. Some of these are combined types. Therefore, when counting the subclasses, we reach a total of 35 variability types. The variability classes list can be found in Table 1.

The table that contains the variability type class from the Machine Learning is `gaiadr3.vari_classifier_result`.

A validation of the classification for the young stellar object class was done by [Marton et al. \(2023\)](#).

After the classification, Specific Object Studies are done for various topics. They act as a validation of the classification, and also these activities compute further parameters that describe the class. It is important to note that there are two sources for a given variability type: (1) The Machine Learning classification and (2) the classification of Specific Object Studies. There may be differences between these two classifications.

The Specific Object Studies covers the following topics (listed in the alphabetical order, as in the Gaia archive). Here we just give the number of sources and also some selected highlights:

- **AGN:** `gaiadr3.vari_agn` [Carnerero et al. \(2023\)](#). There are 872,228 sources in this table, this number is obtained with the ADQL query (just shown for this example):

```
SELECT COUNT(*)
FROM gaiadr3.vari_agn
```

We can compare this number from the number of AGN from classification, we have 1,035,207 sources, see following ADQL command:

⁸here probably the small oscillations are not detected.


```
SELECT COUNT(*)
FROM gaiadr3.vari_classifier_result
WHERE best_class_name = 'AGN'
```

A surprising result is that even with the gaps of the Gaia time series, the time delay for a source, the lens system DESJ0501-4118, was derived.

- **Cepheids:** `gaiadr3.vari_cepheid` [Ripepi et al. \(2023\)](#). There are 15,021 sources. A Fourier series model of the light curve is fitted and the parameters are published in this table. The estimation of the mean should be taken for this table and not from `gaia_source` table. A subclassification is also given. Radial velocities time series are published for 798 sources. It is the largest catalogue of Cepheids having radial velocities.
- **Compact companion:** `gaiadr3.vari_compact_companion` [Gomel et al. \(2023\)](#). There are 6,306 sources. If the variability is caused by ellipsoidal deformation, a mass ratio can be determined from the Fourier parameters of the light curve. This mass ratio helps identify potential candidates for white dwarfs, neutron stars, and black holes. Other methods to detect black holes are based on astrometry and radial velocities ([El-Badry et al. 2023](#); [Chakrabarti et al. 2023](#); [El-Badry et al. 2023](#); [Gaia Collaboration, Panuzzo et al. 2024](#)).
- **Eclipsing binaries:** `gaiadr3.vari_eclipsing_binary` [Mowlavi et al. \(2023\)](#). There are 2,184,477 sources. The largest catalogue of eclipsing binaries ever published on the whole sky. In this table, we adopt a geometrical modelling of the light curve and extract some relevant statistics. A sub-sample is then studied with a physical modelling, see [Pourbaix et al. \(2022\)](#). The results are published in the non-single star table `gaiadr3.nss_two_body_orbit`. It contains 86,918 systems with the flag `nss_solution_type` `EclipsingBinary` and 155 for `EclipsingSpectro`.
- **Long Period Variables:** `gaiadr3.vari_long_period_variable` [Lebzelter et al. \(2023\)](#). There are 1,720,588 sources. Thanks to the RP spectrophotometry, 546,468 stars were classified as carbon rich candidates.
- **Microlensing:** `gaiadr3.vari_microlensing` [Wyrzykowski et al. \(2023\)](#). This is the first whole sky microlensing event list, which contains 363 sources.
- **Main Sequence Oscillators:** `gaiadr3.vari_ms_oscillator` [Gaia Collaboration, De Ridder et al. \(2023\)](#). There are 54,476 sources. See Section 5.1.1.
- **Planetary transits:** `gaiadr3.vari_planetary_transit` [Panahi et al. \(2022b\)](#). The table contains 214 sources. Please note that the table `gaiadr3.vari_planetary_transit_13june2022` is incorrect. For details, please see Section 5.1.3.
- **Rotation modulation:** `gaiadr3.vari_rotation_modulation` [Distefano et al. \(2023\)](#). There are 474,026 sources. The Period Amplitude diagram shows different regimes.
- **RR Lyrae stars:** `gaiadr3.vari_rrlyrae` is taking care of the RR Lyrae variables and is described in [Clementini et al. \(2023\)](#). There are 271,779 sources. RR Lyrae stars are well known standard candles. There are several subtypes of them: the Bailey's ab and c types corresponding to fundamental mode pulsation and first overtone, respectively. The ab type light curves have a very recognisable shape with a sharp rise and slower declined with a period typically of half a day, the c types have a more sinusoidal light curve. An example of the quite astonishing results is the metallicity map derived from the RR Lyrae light curves in Figure 5. There are 1,100 sources with radial velocities time series which are in the table of RR Lyrae stars.
- **Short Time scales:** `gaiadr3.vari_short_timescale` is described in [Eyer et al. \(2023\)](#). There are 471,679 sources. This procedure did not go through removal of sources and reveals also instrumental and/or calibration artefacts.

For the completeness of the archive description, three additional tables are listed under the variability folder: one is a summary of the variability analysis, and the other two are associated with the radial velocities of Cepheids and RR Lyrae stars:

- `gaiadr3.vari_summary` table provides a global overview. It helps determine which table contains a given source (Eyer et al. 2023), and provides statistical attributes of the time series.
- `gaiadr3.vari_epoch_radial_velocity`: the table contains the time series of the radial velocities.
- `gaiadr3.vari_rad_vel_statistics`: the table contains statistics of the radial velocities.

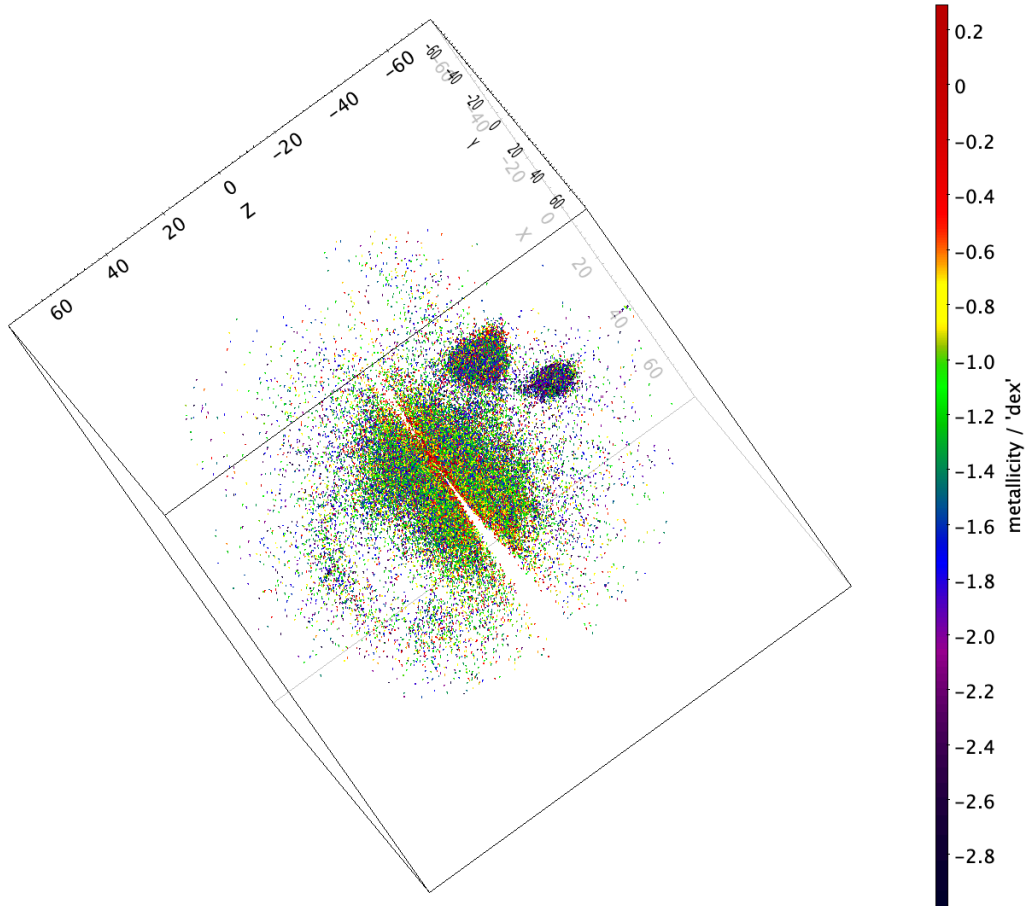


Fig. 5. Distribution of RR Lyrae stars in a 3D projection using the distance modulus. The LMC and SMC can be seen behind the Milky Way RR Lyrae stars. The metallicity near the Galactic plane is higher, lower in the halo, then lower in the LMC and then even lower in the SMC, as expected. We also see the Sagittarius stream. We can guess the presence of other structures in the halo.

5.1.1 The Performance Verification Paper on variability

A performance verification paper (PVP) authored by Gaia Collaboration, De Ridder et al. (2023) investigate the pulsation properties in main sequence OBAF-type stars. Among its noteworthy findings are: (1) the establishment of a period-luminosity relation for δ Scuti stars. (2) It is further shown that stellar rotation seems to reduce the amplitude of variability. (3) Discrepancies arise when comparing the observed locations of these variables with the predictions of theoretical models. Such inconsistencies may be attributed to inaccuracies in determining astrophysical parameters using Gaia instruments, and/or it is conceivable that rotational effects play a significant role, potentially altering the instability regions within the Hertzsprung-Russell diagram. (4) A quite remarkable result (see Appendix B of the work by Gaia Collaboration, De Ridder et al. (2023)) which demonstrates Gaia performance: with as few as 40 measurements, Gaia detects two accurate frequencies in certain cases when compared with the reference observations from the Kepler mission. As Gaia extends its

measurement duration to a decade, the precision of frequency determination, which scales inversely with the observation time basis, is anticipated to reach exceptional levels of accuracy.

There is also a publication on the spurious effect of spurious photometric variability introduced by the angle by which the satellite scans a (non axi-symmetric) source (Holl et al. 2023a).

5.1.2 The Gaia Andromeda Photometric Survey

To offer a teaser of DR4 and DR5, which will include epoch photometry/time series for **all sources**, we provide in the Gaia archive a pencil beam with a 5.5-degree opening angle centred toward the Andromeda galaxy, ensuring photometry availability for all sources within this region (Evans et al. 2023). This selection was made to provide a sufficiently broad sample of celestial sources reaching nearly 1.3 million sources; in addition to the Andromeda galaxy stars, the foreground stars, with significant parallaxes, allow us to provide a well-populated Hertzsprung-Russell Diagram. This dataset contains all the G , G_{BP} , G_{RP} time series in the Gaia archive. The Evans et al. (2023) article also presents a renormalisation of the photometric uncertainties (as these uncertainties are usually underestimated). We also show a way to combine the three photometric bands with a Principal Component Analysis method to enhance the detection and the characterisation of the variable sources, an approach similar to what was done with SDSS data (Süveges et al. 2012).

5.1.3 The exoplanets from Gaia with the transit method

Although the time span of DR3 was limited to 34 months, the quality of the photometry is sufficient to discover exoplanets by the transit method (Panahi et al. 2022b). Currently, there are two exoplanets detected by the transit method and confirmed by radial velocity measurements, Gaia-1b and Gaia-2b. The list of candidates is provided in the Gaia table `vari_planetary_transit` and contains 214 sources: 173 systems are known, and 41 are new (among them the two above already cited); the list of these 41 candidates can be found in Eyer et al. (2023). There is further work with the exoplanets with the transit methods, there has been a special agreement with the TESS mission (Ricker et al. 2015) to verify the TESS candidates by Gaia. Again, with only 34 months, we are able (1) to show that about 5% of the TESS candidates are background eclipsing binaries and also (2) to confirm about 5% of the cases where the transit can be identified (Panahi et al. 2022a).

As a side note, it is noteworthy to mention that: (1) astrometric exoplanet candidates were proposed by Holl et al. (2023b). At the time of writing, one of these candidates, named Gaia-3b, was confirmed thanks to ground-based radial velocity measurements (Sozzetti et al. 2023). (2) Exoplanets have also been discovered through microlensing. Gaia21dnc (Ban et al., in Prep) and Gaia22dkv (Wu et al. 2023) are the first two exoplanets detected through follow-ups of microlensing events alerted by the Gaia Science alert system (Hodgkin et al. 2021).

5.2 The focused product release: Time domain radial velocities for Long Period Variable stars

After the third data release, on October 10, 2023, there was a focused product release by the Gaia consortium, consisting of several products:

- Astrometry and photometry from engineering images taken in the ω Centauri region.
- The first results of quasars' environment analysis for gravitational lenses search.
- Extended radial velocity epoch data for Long Period Variables.
- Diffuse Interstellar Bands from aggregated RVS spectra.
- Updated astrometry for Solar System objects.

Here we present the Long Period Variables (LPV) (Gaia Collaboration, Trabucchi et al. 2023). The results of this focused product release are based on the DR3 processing both in photometry and radial velocities. A reminder of DR3 on Long Period Variables: It contains 1.7 million LPV. A selection was performed of stars fainter than $G = 6$ and brighter than $G = 14$. The total number of sources is 9,614 stars with radial velocities times series. Though this number may look small, it is a record holder: it is the largest sample ever published of Long Period Variable stars with radial velocity time series. It should be noted that the knowledge of radial velocities allows us to distinguish between pulsation (so true LPVs) and ellipsoidal variability.

The tables in the Gaia archive are located in the directory Gaia Focused Product Release and the subdirectory Variability. There are three tables:

- `gaiafpr.vari_epoch_radial_velocity`: Contains the radial velocity time series.
- `gaiafpr.vari_long_period_variable`: Similar to the DR3 table of the same name, but with added radial velocity data (frequencies and amplitudes are recomputed in both the radial velocity and G photometric band time series).
- `gaiafpr.vari_rad_vel_statistics`: Contains statistics on the radial velocities.

5.3 Citizen Science Project: GaiaVari

GaiaVari is a citizen science project available on the Zooniverse platform that invites the community to classify variable stars from the public Gaia DR3 results based on the following various visual elements:

- The light curve, representing the G band magnitude measurements as a function of time (colour-coded with time).
- The folded curve, displaying the G band magnitude measurements as a function of phase computed from the observed time and an estimated period (the colour-coded according to the light curve, so according to the time, not the phase).
- The position in the Hertzsprung-Russell Diagram (the absolute G magnitude as a function of $G_{BP} - G_{RP}$) and the motion of the variable star within this diagram. There are no corrections from the extinction/reddening.
- The position in the Milky Way (the distance is taken as the inverse of the parallax).

The project is limited to stars with sufficiently good parallax measurements, allowing for the determination of a distance so that it can be placed within the Milky Way. The sample to classify is not exclusively focused on periodic objects, although the majority of objects chosen fall into this category. An example of the representation is given in Figure 6 for an RR Lyrae star.

Up to now, there have been one beta campaign and two official campaigns of GaiaVari. The first one classified the source into the possible following variability types: Eclipsing binaries, Cepheids, RR Lyrae stars and long-period variable stars. In the second campaign, the following variability types were added to the previous list: δ Scuti stars and stars with ellipsoidal variability. In both campaigns, a "None of the above" option was also introduced.

After making a classification choice, and only then, there is the possibility to leave comments. By choosing this option, one can also discover the Gaia sourceID and the classification from the automated supervised machine learning method (Rimoldini et al. 2023). There is also a link towards ESASky (Merín et al. 2015), and then one can obtain information from databases like Simbad at "Centre de Données astronomiques de Strasbourg" (CDS) (Wenger et al. 2000), etc.

Some citizen scientists systematically classify RR Lyrae stars and Cepheids into subgroups, such as RRab, RRc types, or Type-I, Type-II Cepheids. Such sub-classifications can be compared with the Gaia classification (Clementini et al. 2023; Ripepi et al. 2023) or the literature.

The community easily identifies the Blazhko effect, although, at times, it is challenging to distinguish Blazhko from an RRd type (double-mode RR Lyrae stars). Additionally, systematic differences in classification between what is published in DR3 and the classifications by some highly educated citizen scientists, particularly in RR Lyrae and Cepheids and their sub-classification, are noteworthy. It will be very interesting to compare these differences.

Due to demand we also created a permanent educational campaign for teachers and schools that can use it as an introduction resource to variable stars and HR diagram understanding.

We plan to conduct additional campaigns. At a point, the plan is to include radial velocity information for a more comprehensive dataset offered by Gaia.

Furthermore, there are plans to write articles on the results of GaiaVari from the first two campaigns,

GaiaVari is funded by ESA and driven by Sednai Sarl with a collaboration with the University of Geneva, ESA and ScienceNow, see <https://www.zooniverse.org/projects/gaia-zooniverse/gaia-vari>.

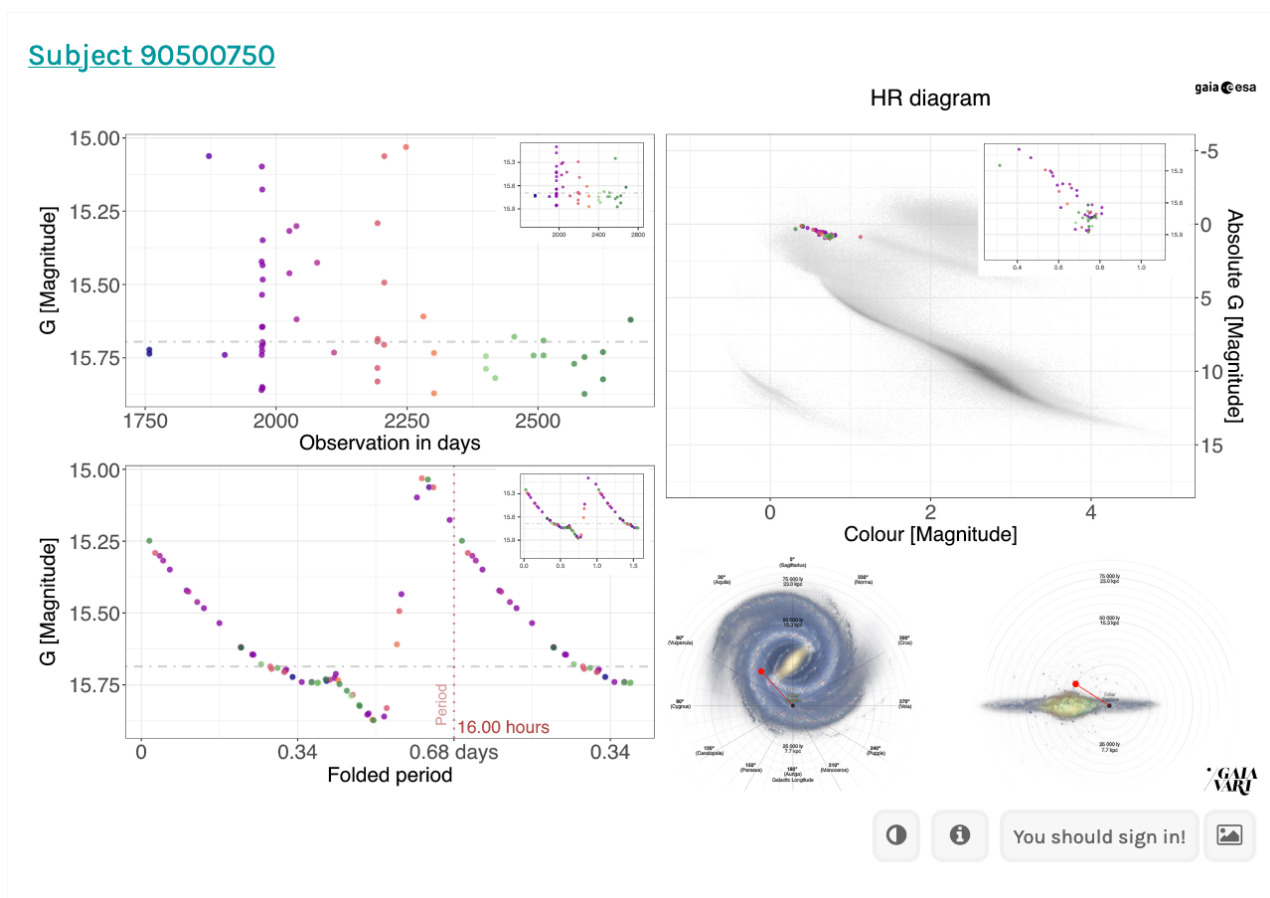


Fig. 6. RR Lyrae star of Bailey’s type ab in GaiaVari. Upper left panel: the time series: Lower left panel: the folded curve, with the period fo 16 hours. Upper right panel: the motion of the star in the HR diagram. Lower right panel: the position in the Milky Way.

6 Conclusions

The quantity and quality of Gaia’s data are impressive. Among the many scientific subjects that Gaia is touching, one is particularly positively impacted: time domain astronomy. In this context, we mostly presented the analysis of photometrically variable sources of DR3, covering 34 months of data with its multitude of results. Even with the relatively short DR3 time span, numerous records were broken in the number of variable sources, and the unprecedented description of them sets a new standard. Looking ahead, Gaia’s data will nearly double its number of measurements in DR4 and, again, nearly double it for DR5, reaching about 10.5 years of science operational data (if all goes as planned). The Gaia variability pipeline for processing and analysing data in DR4 and DR5 will receive additional variability classes, features and improvements to the current ones, ensuring not only an increase in quantity but also an enhancement in quality. This holds the promise of even more interesting results for both the fourth and fifth Gaia data releases. Undoubtedly, Gaia’s data alone provide remarkable insights into variable sources. However, we can anticipate exciting developments, such as combining Gaia data with other photometric (e.g. Vera Rubin observatory/LSST (Ivezić et al. 2019), ZTF (Bellm et al. 2019), PLATO (Rauer et al. 2014), etc.) or spectroscopic surveys (e.g. WST, Maineri et al. (2024)).

I am grateful to the organisers of this School, in particular to Carine Babusiaux. I also appreciated the discussions with PhD students and postdocs. I would like to thank Mathieu Van der Swaelmen, whose question on the zero point parallax correction prompted a change in the Gaia consolidated catalogue for DR4. I thank Nicholas Chornay and Krzysztof Nienartowicz for reading and commenting on this manuscript.

I am thankful to the Geneva Gaia team, the Gaia Variability Processing and Analysis international team (CU7), and DPAC consortium. The success of such a mission lies in the competence and dedication of these Consortium members. Such an article would not exist without their contributions.

During the writing of this document, I had to face the declining health and the passing of my mother, Andrée Eyer Pétermann. Throughout my life, she was a constant source of support for my high aspirations and my work. I therefore dedicate this text to her.

I met Evry Schatzman only once when I was a young Ph.D. student. I was deeply moved by his diligence and interest in the young researcher I was, working on Hipparcos variable stars at the time. This is why I am quite certain that Evry Schatzman would have been most enthusiastic about the Gaia mission and its results.

The texts and figures are original; several software were used to improve the readability and the quality of the text, such as ChatGPT, Grammarly and Google Translate. This work made use of TOPCAT software (Taylor 2005). This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular, the institutions participating in the *Gaia* Multilateral Agreement.

References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *Phys. Rev. Lett.*, 116, 061102
- Appenzeller, I. 2013, *Introduction to Astronomical Spectroscopy*
- Bakos, G., Noyes, R. W., Kovács, G., et al. 2004, *PASP*, 116, 266
- Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, *PASP*, 131, 018002
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, 327, 977
- Burki, G., Rufener, F., Burnet, M., et al. 1995, *A&AS*, 112, 383
- Carnerero, M. I., Raiteri, C. M., Rimoldini, L., et al. 2023, *A&A*, 674, A24
- Carney, B. W., Storm, J., & Jones, R. V. 1992, *ApJ*, 386, 663
- Chakrabarti, S., Simon, J. D., Craig, P. A., et al. 2023, *AJ*, 166, 6
- Clementini, G., Ripepi, V., Garofalo, A., et al. 2023, *A&A*, 674, A18
- Clementini, G., Ripepi, V., Leccia, S., et al. 2016, *A&A*, 595, A133
- Davis, R., Harmer, D. S., & Hoffman, K. C. 1968, *Phys. Rev. Lett.*, 20, 1205
- de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, *The Messenger*, 175, 3
- Distefano, E., Lanzafame, A. C., Brugaletta, E., et al. 2023, *A&A*, 674, A20
- El-Badry, K., Rix, H.-W., Quataert, E., et al. 2023, *MNRAS*, 518, 1057
- Evans, D. W., Eyer, L., Busso, G., et al. 2023, *A&A*, 674, A4
- Eyer, L. 2006, *Mem. Soc. Astron. Italiana*, 77, 549
- Eyer, L., Audard, M., Holl, B., et al. 2023, *A&A*, 674, A13
- Eyer, L. & Mignard, F. 2005, *MNRAS*, 361, 1136
- Eyer, L., Mowlavi, N., Evans, D. W., et al. 2017, arXiv e-prints, arXiv:1702.03295
- Eyer, L., Rimoldini, L., Holl, B., et al. 2015, in *Astronomical Society of the Pacific Conference Series*, Vol. 496, *Living Together: Planets, Host Stars and Binaries*, ed. S. M. Rucinski, G. Torres, & M. Zejda, 121
- Gaia Collaboration, Clementini, G., Eyer, L., Ripepi, V., et al. 2017, *A&A*, 605, A79
- Gaia Collaboration, De Ridder, J., Ripepi, V., Aerts, C., et al. 2023, *A&A*, 674, A36
- Gaia Collaboration, Eyer, L., Rimoldini, L., Audard, M., et al. 2019, *A&A*, 623, A110
- Gaia Collaboration, Panuzzo, P., Mazeh, T., Arenou, F., et al. 2024, *A&A*, 686, L2
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., et al. 2016, *A&A*, 595, A1
- Gaia Collaboration, Trabucchi, M., Mowlavi, N., Lebzelter, T., et al. 2023, *A&A*, 680, A36
- Gavras, P., Rimoldini, L., Nienartowicz, K., et al. 2023, *A&A*, 674, A22
- Gavras, P., Rimoldini, L., Nienartowicz, K., et al. 2022, *VizieR Online Data Catalog*, J/A+A/674/A22
- Gomel, R., Mazeh, T., Faigler, S., et al. 2023, *A&A*, 674, A19
- Hess, V. F. 1912, *Phys. Z.*, 13, 1084
- Hodgkin, S. T., Harrison, D. L., Breedt, E., et al. 2021, *A&A*, 652, A76
- Holl, B., Audard, M., Nienartowicz, K., et al. 2018, *A&A*, 618, A30
- Holl, B., Fabricius, C., Portell, J., et al. 2023a, *A&A*, 674, A25
- Holl, B., Sozzetti, A., Sahlmann, J., et al. 2023b, *A&A*, 674, A10
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, 873, 111
- Lebzelter, T., Mowlavi, N., Lecoœur-Taibi, I., et al. 2023, *A&A*, 674, A15
- Mainieri, V., Anderson, R. I., Brinchmann, J., et al. 2024, arXiv e-prints, arXiv:2403.05398
- Marton, G., Ábrahám, P., Rimoldini, L., et al. 2023, *A&A*, 674, A21
- Merín, B., Salgado, J., Giordano, F., et al. 2015, arXiv e-prints, arXiv:1512.00842

- Mowlavi, N., Holl, B., Lecoœur-Taïbi, I., et al. 2023, *A&A*, 674, A16
- Panahi, A., Mazeh, T., Zucker, S., et al. 2022a, *A&A*, 667, A14
- Panahi, A., Zucker, S., Clementini, G., et al. 2022b, *A&A*, 663, A101
- Pogson, N. 1856, *MNRAS*, 17, 12
- Pourbaix, D., Arenou, F., Gavras, P., et al. 2022, Gaia DR3 documentation Chapter 7: Non-single stars, Gaia DR3 documentation, European Space Agency; Gaia Data Processing and Analysis Consortium. Online at <https://gea.esac.esa.int/archive/documentation/GDR3/index.htm>, id. 7
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *Experimental Astronomy*, 38, 249
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 014003
- Riello, M., De Angeli, F., Evans, D. W., et al. 2021, *A&A*, 649, A3
- Rimoldini, L., Holl, B., Gavras, P., et al. 2023, *A&A*, 674, A14
- Ripepi, V., Clementini, G., Molinaro, R., et al. 2023, *A&A*, 674, A17
- Sahu, K. C., Anderson, J., Casertano, S., et al. 2022, *The Astrophysical Journal*, 933, 83
- Schatzman, E. 1963, *Astrophysique (Masson et Cie)*
- Sozzetti, A., Pinamonti, M., Damasso, M., et al. 2023, *A&A*, 677, L15
- Süveges, M., Sesar, B., Váradi, M., et al. 2012, *MNRAS*, 424, 2528
- Taylor, M. B. 2005, in *Astronomical Society of the Pacific Conference Series*, Vol. 347, *Astronomical Data Analysis Software and Systems XIV*, ed. P. Shobbell, M. Britton, & R. Ebert, 29
- Udalski, A. 2003, *Acta Astron.*, 53, 291
- Wenger, M., Ochsenein, F., Egret, D., et al. 2000, *A&AS*, 143, 9
- Wu, Z., Dong, S., Yi, T., et al. 2023, arXiv e-prints, arXiv:2309.03944
- Wyrzykowski, Ł., Kruszyńska, K., Rybicki, K. A., et al. 2023, *A&A*, 674, A23