Gaia Data Release 2

Variable stars in the colour-magnitude diagram**


(Affiliations can be found after the references)

Received 25 April 2018 / Accepted 3 November 2018

* A movie associated to Fig. 11 is available at https://www.aanda.org
** Data are only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via https://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/623/A110

Article published by EDP Sciences
telescope; surveys using multiple sites and telescopes and instruments require cross-calibration, which unavoidably introduces systematics and reduces precision because of the increased scatter. Thus, Gaia will play an important role as a standard source in cross-calibrating heterogeneous surveys and instruments, much like the HIPPARCOS mission (Perryman et al. 1997; ESA 1997) did in the past. Of course, Gaia represents a quantum leap from HIPPARCOS in many regards, including an increase of four orders of magnitude in the number of objects observed, additional types of observations (spectrophotometry and spectroscopy), and significantly improved sensitivity and precision for all types of measurements.

Fourthly, there are unprecedented synergies for calibrating distance scales using the dual astrometric and time-domain capabilities of Gaia (e.g. Eyer et al. 2012). Specifically, Gaia will enable the discovery of unrecognised numbers of standard candles residing in the Milky Way, and anchor Leavitt laws (period-luminosity relations) to trigonometric parallaxes (see Gaia Collaboration 2017; Casertano et al. 2017, for two examples based on the first Gaia data release).

Variable stars have for a long time been recognised to offer crucial insights into stellar structure and evolution. Similarly, the Hertzsprung–Russell diagram (HRD) provides an overview of all stages of stellar evolution, and together with its empirical cousin, the colour-magnitude diagram (CMD), it has shaped stellar astrophysics like no other diagram. Henrietta Leavitt (1908) was one of the first to note the immense potential of studying variable stars in populations, where distance uncertainties did not introduce significant scatter. Soon thereafter, Leavitt & Pickering (1912) discovered the period-luminosity relation of Cepheid variables, which has become a cornerstone of stellar physics and cosmology. It appears that Eggen (1951, his Fig. 42) was the first to use (photoelectric) observations of variable stars (in this case, classical Cepheids) to constrain regions where Cepheids occur in the HRD; these regions are today referred to as instability strips. Eggen further illustrated how Cepheids change their loci in the colour-magnitude diagram (CaMD) during the course of their variability, thus developing a time-dependent CaMD for variable stars. Kholopov (1956) and Sandage (1958) later illustrated the varying locations of variable stars in the HRD using classical Cepheids located within star clusters. By combining the different types of Gaia time-series data with Gaia parallaxes, we are now in a position to construct time-dependent CaMD towards any direction in the Milky Way, building on previous work based on HIPPARCOS (Eyer et al. 1994; Eyer & Grenon 1997), but on a much larger scale.

Many variability (ground- and space-based) surveys have exploited the power of identifying variable stars in stellar populations at similar distances, for example, in star clusters or nearby galaxies such as the Magellanic Clouds. Ground-based microlensing surveys such as the Optical Gravitational Lensing Experiment (OGLE; e.g. Udalski et al. 2015), the Expérience pour la Recherche d’Objets Sombres (EROS Collaboration 1999), and the Massive Compact Halo Objects project (MACHO; Alcock et al. 1993) deserve a special mention in this regard. The data will continue to grow with the next large multi-epoch surveys such as the Zwicky Transient Facility (Bellm 2014) and the Large Synoptic Survey Telescope (LSST Science Collaboration 2009) from the ground, and the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015) and PLATO (Rauer et al. 2014) from space.

Another ground-breaking observational trend has been the long-term high-precision high-cadence uninterrupted space photometry with CoRoT/BRITE (Auvergne et al. 2009; Pablo et al. 2016, with time bases of up to five months) and Kepler/K2 (Gilliland et al. 2010; Howell et al. 2014, with time bases of up to four years and three months, respectively) provided entirely new insights into micro-magnitude level variability of stars, with periodicities ranging from minutes to years. These missions opened up stellar interiors from the detection of solar-like oscillations of thousands of sun-like stars and red giants (e.g. Bedding et al. 2011; Chaplin & Miglio 2013; Hekker & Christensen-Dalsgaard 2017, for reviews), as well as hundreds of intermediate-mass stars (e.g. Aerts et al. 2015; Bowman 2017) and compact pulsators (e.g. Hermes et al. 2017). The results we provide in Sects. 3 and 4 on the variability fractions and levels are representative of milli-mag level variability and not of micro-mag levels, as are found in space asteroseismic data.

Any of these astroseismic surveys can benefit from Gaia astrometry, however, so that distances and luminosities can be derived, as De Ridder et al. (2016) and Huber et al. (2017) reported with Gaia DR1 data. Gaia will also contribute to these surveys with its photometry, and some surveys will also benefit from the Gaia radial velocities (depending on their operating magnitude range).

Stellar variability comprises a great variety of observable features that are due to different physical origins. Figure 1 shows the updated variability tree (Eyer & Mowlavi 2008), which provides a useful overview of the various types of variability and their known causes. The variability tree has four levels: the distinction of intrinsic versus extrinsic variability, the separation into major types of objects (asteroid, stars, and AGN), the physical origin of the variability, and the class name. In this article, we follow the classical distinction of the different causes of the variability phenomena: variability induced by pulsation, rotation, eruption, eclipses, and cataclysmic events. A large number of variability types can be identified in the Gaia data even now, as described in the subsequent sections.

We here provide an overview of stellar variability in the CaMD, building on the astrometric and photometric data of the second Gaia data release (DR2). Future Gaia DRs will enable much more detailed investigations of this kind using longer temporal baselines, greater number of observations, and added classes of variable stars (such as eclipsing binaries, which will be published in DR3).

This paper is structured as follows. Section 2 shows the location of different variability types in the CaMD, making use of known objects from the literature that are published in Gaia DR2, but without any further analysis of the Gaia data. Section 3 presents the fraction of variables as a function of colour and absolute magnitude, obtained by processing the Gaia time series for the detection of variability (Eyer et al. 2018). Section 4 investigates the variability level in the CaMD by employing statistics and classification results (some of which are related to unpublished Gaia time series). Section 5 shows the motion of known variables stars in the CaMD, that is, a time-dependent CaMD, which also includes sources that are not available in the DR2 archive but are online material. Section 6 summarises our results and presents an outlook to future Gaia DRs. Further information on the literature cross-match and on the selection criteria applied to our data samples can be found in Appendices A and B, respectively.

2. Location of variability types in the CaMD

The precision of the location in the CaMD depends on the precision of the colour on the one hand and on the determination of the absolute magnitude on the other. The precision of the
absolute magnitude of variable stars depends on the photometric precision, the number of measurements, the amplitude of variability, and the relative parallax precision $\sigma_{\pi}/\pi$. The upper limits of $\sigma_{\pi}/\pi$ employed in this article vary between 5% and 20%, which means that the uncertainty of the absolute magnitude that is solely due to the parallax uncertainty can be as large as $(10)^{-2}\sigma_{\pi}/\pi \approx 0.43$ mag.

As we determined the colour as a function of integrated BP and integrated RP spectro-photometric measurements with tight constraints on the precision of these quantities (see Appendix B), there are parts of the CaMD that are not explored here. For example, the faint end of the main sequence presented in Fig. 9 of Gaia Collaboration (2018) does not fulfill the condition on the precision in BP, so our diagrams do not include L and T brown dwarfs (which are fainter than $M_G \sim 14$ mag). If we cross-match the Gaia data (conditional on the selection of Appendix B) with the catalogue of M dwarfs of Lépine & Gaidos (2011), only a few M6, M7, and M8 dwarf stars are found.

In Fig. 2 we introduce the Gaia CaMD, which is displayed as a background in subsequent figures. For clarity, we note basic astronomical features such as the main sequence, the red clump (and its long tail due to interstellar extinction), the horizontal branch, the extreme horizontal branch (see D’Cruz et al. 1996, for its physical origin), the red giant branch, the asymptotic giant branch, the white dwarf sequence, the subdwarfs, the supergiants, and the binary sequence. There are additional subtle features above and below the red clump that are described in Fig. 10 of Gaia Collaboration (2018) and are known as the asymptotic giant branch bump and the red giant branch bump, respectively. On the right-hand side of Fig. 2, we also note the typical limiting distance that can be reached because of the selection of $\sigma_{\pi}/\pi$, up to 1 kpc (which was the largest distance we considered for background stars).

Several effects can influence the average location of a star in the CaMD (in both axes), including interstellar extinction, stellar multiplicity, rotation, inclination of the rotation axis, and chemical composition. In this work, we do not correct for such phenomena and instead rely on the apparent magnitudes and colours measured by Gaia, computing “absolute” magnitudes using Gaia parallaxes. We note that interstellar extinction and reddening can be significant at the considered distances (up to 1 kpc), in particular for objects in the Galactic plane. This leads to distortions of certain observed features, such as the long tail in the red clump, which extends to redder and fainter magnitudes.

The stellar variability aspects covered in the second Data Release of Gaia include a limited number of variability classes (Holl et al. 2018): long-period variables, Cepheids, RR Lyrae stars, SX Phoenicis/Scuti stars, and rotation-modulated solar-like variability (i.e. all late-type BY Draconis stars). Short-timescale variability (within one day) was explored regardless of the physical origin of the variability (Roelens et al. 2018), although stars that are classified as eclipsing binaries were removed as planned to first appear in the third Data Release of Gaia. The stars presented in this section are solely based on the cross-match with known objects in the literature. The list of variability types presented here is not meant to be comprehensive.

Figures 3–7 illustrate the locations of known variable stars from catalogues in the literature that are cross-matched with the Gaia data. We indicate these targets according to their known variability type published in the literature (the references are listed in Table A.1), and only the stars that satisfy the selection criteria described in Appendix B are kept. Each of these figures includes as reference the location and density (in grey scale) of all stars, regardless of stellar variability, that satisfy the astrometric and photometric criteria of Appendix B with the additional constraint of a minimum parallax of 1 mas (i.e. within 1 kpc.
Fig. 2. CaMD with its most striking known features (see text). The points in grey denote objects with parallax greater than 1 mas, with relative parallax precision better than 20% and other criteria described in Appendix B.

to the Sun). This radius seems a good compromise between a large number of stars and a limited effect of interstellar matter. Variable stars whose variability type was previously known are represented by combinations of symbols and colours. Following the structure of the variability tree in Fig. 1, we show in separate figures the CaMDs of stars whose variability is induced by different causes, such as pulsations, rotation, eruptions, eclipses, and cataclysmic events.

Several caveats apply to Figs. 3–7 and should be kept in mind for their interpretation. (a) The quality of catalogues published in the literature can be rather different, in part because variability is often classified without knowledge of a parallax. To reduce the impact of misclassified objects on these figures, we selected subsets of all available catalogues as reference for specific variable star classes, depending on their agreement with the expected locations in the CaMD. In certain cases, we have excluded sources from the literature by choice of specific catalogues (Table A.1) and by using the Gaia astrometry and the multi-band photometric time-series data for occasional cuts in magnitude or colour. Future Gaia data releases will provide a more homogeneous variability classification that will rely primarily on the results of the variability processing (Holl et al. 2018). (b) The CaMDs are not corrected for extinction, which leads to increased scatter, in particular for objects that primarily reside in heavily attenuated areas such as the Galactic disc and the Galactic bulge. (c) The cross-match of sources can be erroneous when stars are located in crowded regions or have high proper motion, especially if the positions of stars in the published catalogues are not sufficiently precise or if proper motion information is not available. (d) Some variability types like magnetically active stars (e.g. RS CVn stars) exhibit different observational phenomena, such as rotational modulation variations as well as flares. To avoid overcrowding the CaMD diagrams, these types are represented in only one of the relevant diagrams (e.g. with rotational or eruptive variables). Furthermore, we note that the time sampling and the waveband coverage of a given survey might favour the detection of only some of these aspects. (e) Gaia represents a milestone for space astrometry and photometry. Nevertheless, some sources can be affected by problems such as corrupt measurements so that their location in the CaMD may be incorrect (Arenou et al. 2018). However, we stress that these problems are limited to a small fraction of sources so that most known variable classes are recovered as expected. The cyclic approach of the Gaia data processing and analysis will allow us to correct for these unexpected features in the future data releases.

2.1. Pulsating variable stars

Figure 3 shows the positions of different classes of pulsating variable stars based on the Gaia data and can be compared to its theoretical counterpart in recent textbooks on asteroseismology (Fig. 1.12 in Aerts et al. 2010) and on pulsating stars (Catelan & Smith 2015). We refer to these books for further details of specific variability classes. Here, we only consider the following types of pulsating variable stars:

1. Long-period variables, which are red giant stars that populate the reddest and brightest regions of the CaMD. They include Miras, semi-regular variables, slow irregular variables, and small-amplitude red giants.
Fig. 3. Known pulsating variable stars retrieved from published catalogues are placed in the observational CaMD, with symbols and colours representing types as shown in the legend (see Table A.1 for the references from the literature per type). All stars satisfy the selection criteria described in Appendix B. The background points in grey denote a reference subset of objects with a stricter constraint on parallax ($\varpi > 1$ mas), which limits the sample size, extinction, and reddening. The effects of interstellar matter and other phenomena (see text) are not corrected for. The condition on the relative precision of $G_{BP}$ measurements introduces artificial cuts in the distributions of low-mass main-sequence stars and red (super)giants.

2. $\alpha$ Cygni stars, which are luminous supergiant stars that pulsate in non-radial modes. They are particularly affected by interstellar extinction as they are young massive stars that reside in the Galactic disc, so that their position in Fig. 3 must be treated with caution.

3. $\delta$ Scuti stars, which are Population-I stars of spectral types A and F with short periods ($\leq 0.3$ d) that dominantly pulsate in pressure modes, but may also reveal low-order gravity modes of low amplitude.

4. SX Phoenicis stars, which are Population-II high-amplitude $\delta$ Scuti stars.

5. $\gamma$ Doradus stars, which are stars with spectral types A and F with periods from 0.3 to 3 d that dominantly pulsate in high-order gravity modes, but may also reveal low-amplitude pressure modes.

6. RR Lyrae stars (Bailey’s type ab and c), which are Population-II horizontal branch stars with periods from 0.2 to 1 d that pulsate in pressure mode. C-type RR Lyrae stars are bluer than ab-type stars.

7. Slowly pulsating B (SPB) stars, which are non-radial multi-periodic gravity-mode pulsators of spectral type B with periods typically in the range from 0.4 to 5 d.

8. $\beta$ Cephei stars, which are late-O to early-B spectral type non-supergiant stars with dominant low-order pressure and gravity modes that feature periods in the range from 0.1 to 0.6 d. Several of them have been found to also exhibit low-amplitude high-order gravity modes as in the SPB stars (e.g. Stankov & Handler 2005). The $\beta$ Cephei stars are located in the Galactic disc so that their CaMD position is easily affected by interstellar extinction.

9. Classical Cepheids (prototype $\delta$ Cephei), which are evolved Population-I (young intermediate-mass) stars featuring radial $p$-mode pulsations with periods of approximately 1–100 d. Cepheids can be strongly affected by interstellar extinction as they reside in the Galactic disc and can be observed at great distances.

10. Type-II Cepheids, which are Population-II stars pulsating in $p$-mode that were historically thought to be identical to classical Cepheids. Type-II Cepheids consist of three different sub-classes (separated by period) that are commonly referred to as BL Herculis, W Virginis, and RV Tauri stars. Their evolutionary scenarios differ significantly, although the three sub-classes together define a tight period-luminosity relation.

11. PV Telescopii stars, which include the sub-classes V652 Her, V2076 Oph, and FQ Aqr (Jeffery 2008). They are rare hydrogen-deficient supergiant stars that cover a wide range of spectral types and exhibit complex light and radial velocity variations.
Fig. 4. Same as Fig. 3, but for rotational-induced variability types.

Fig. 5. Same as Fig. 3, but for eclipsing binaries (of types EA, EB, and EW) and known host-stars that show exoplanet transits. As expected, eclipsing binaries can be anywhere in the CaMD, which explains why they are the main source of contamination of pulsating stars, for instance.
Fig. 6. Same as Fig. 3, but for eruptive variability types.

Fig. 7. Same as Fig. 3, but for cataclysmic variables and some sub-types.
12. Rapidly oscillating Am and Ap stars, which are chemically peculiar A stars that exhibit multiperiodic non-radial pressure modes in the period range of about 5–20 min.

13. V361 Hydæ (or EC 14026) stars, which are subdwarf B stars on the extreme horizontal branch that pulsate in pressure modes with very short periods of ~1–10 min.

14. V1093 Her (or PG 1716) stars, which are subdwarf B stars on the extreme horizontal branch that pulsate in gravity modes with periods of about 1–3 h.

15. ZZ Ceti stars, which are white dwarfs featuring fast non-radial gravity-mode pulsations with periods of 0.5–25 min. The CaMD of pulsating stars carries a great deal of information, much of which has shaped the understanding of stellar structure and evolution and can be found in textbooks. 

16. Interstellar reddening blurs the boundaries between variability classes. Correcting for interstellar extinction will be crucial to delineate the borders of the instability strips in the CaMD, as well as to deduce their purity in terms of the fraction of stars that exhibit pulsations while residing in such regions.

17. Practical difficulties involved in separating variable star classes in the way required to construct Fig. 3 include (a) that variable stars are often subject to multiple types of variability (e.g. γ Doradus/δ Scuti, β Cephei/SPB hybrid pulsators, pulsating stars in eclipsing binary systems, or pulsating white dwarfs that exhibit eruptions), and (b) that naming conventions are often historical or purely based on light-curve morphology, so that they do not account for different evolutionary scenarios (e.g. type-II Cepheids). With additional data and a fully homogeneous variable star classification based on Gaia alone, such ambiguities will be resolved in the future unless they are intrinsically connected to the nature of the variability.

We note multiple groups of ZZ Ceti stars along the white dwarf sequence. The most prominent of these is located at $G_B - G_R \approx 0$ and $M_G \approx 12$, as reported in Fontaine & Brassard (2008).

2.2. Variability due to rotation and eclipses

Figure 4 shows stars whose variability is induced by rotation. There are three primary categories: spotted stars, stars deformed by tidal interactions, and objects whose variability is due to light reflected by a companion. Following the nomenclature in the literature (Table A.1), we list the following variability classes separately, although we note occasional overlaps among the definitions of these variability classes. The following types are included in Fig. 4:

1. $\alpha^2$ Canum Venaticorum stars, which are highly magnetic variable Bp and Ap MS stars.

2. Spotted stars, which show rotational modulation variability from spots.

3. BY Draconis stars, which are main-sequence stars with late spectral types (K and M) that exhibit quasi-periodic light curves due to spots and chromospheric activity.

4. RS Canum Venaticorum stars, which are spotted stars whose rotation-induced variability is frequently accompanied by other phenomena, such as eclipses and flares.

5. Ellipsoidal variables, which show variability (without eclipses) due to orbital motion of a star that is distorted by a stellar companion; solar-like stars with magnetic activity. 

6. Stars known to exhibit exoplanetary transits from the Kepler spacecraft, and only very few of them have detectable transits in the Gaia data because of the different photometric precision and time sampling.
1. S Doradus stars, also known as luminous blue variables, are massive evolved stars that feature major and irregular photometric variations that are due to heavy mass loss by a radiation-driven wind.

2. R Coronae Borealis stars, which are carbon-rich supergiants that emit obscuring material and as a consequence have drastic rapid dimming phases.

3. Wolf-Rayet (WR) stars, which are the almost naked helium core that is left over from originally very high mass evolved stars. They feature strong emission lines of hydrogen, nitrogen, carbon, or oxygen. WR stars undergo very fast mass loss and can be significantly dust-attenuated.

4. γ Cassiopeiae stars and stars with B spectral types that exhibit hydrogen emission lines, that is, Be stars. These are emitting shell stars. During their “eruptive” phenomena, they become brighter.

5. Flare stars, which are magnetically active stars that display flares. This category includes many subtypes of magnetically active stars, such as UV Ceti-type, RS CVn-type, and T Tauri stars.

6. UV Ceti stars, which usually are K-M dwarfs that show flares.

7. T Tauri stars (classical and weak lined), which are young pre-main sequence stars that either accrete strongly (classical) or show little sign of accretion (weak lined). These stars show variability that is due either to magnetic activity (e.g. rotational modulation and flares) or accretion (quasi-periodic, episodic, or stochastic variations), in addition to pulsations that may also occur in some of them.

In Fig. 6, we notice the absence of eruptive variables among hot main-sequence stars (non-supergiants). This region is populated by pulsating stars, such as γ Doradus and δ Scuti stars, cf. Fig. 3. Moreover, WR stars, R Coronae Borealis stars, and S Doradus stars are among the most luminous stars in this diagram.

Figure 7 illustrates three types of cataclysmic variables.

1. Cataclysmic variables (generic class), typically novae and dwarf novae involving a white dwarf. Many of these stars are situated between the main and white dwarf sequences.

2. U Geminorum stars, which are dwarf novae that in principle consist of a white dwarf with a red dwarf companion that experiences mass transfer.

3. Z Andromedae stars, which are symbiotic binary stars composed of a giant and a white dwarf.

Further information on cataclysmic variables can be found for example in Warner (2003) and Hellier (2001).

In Fig. 7, we note a clump of cataclysmic variables located in the ZZ Ceti variability strip near $G \sim 12$ and $G_{BP} - G_{RP} \sim 0.1$. The most significant clump of cataclysmic variables is near $G \sim 4$ and $G_{BP} - G_{RP} \sim 0.1$ mag. These are probably binary systems with stars from the extreme horizontal branch and the main sequence.

3. Variable object fractions in the CaMD

The different types of brightness variations as presented in the CaMD may strongly depend on the colour and absolute magnitude, as seen in Sect. 2, because they are driven by different physical mechanisms. Similarly, the variable object fraction, which is defined as the number of variable objects per colour-magnitude bin divided by the total number of objects in the same bin, is expected to depend on the location in the CaMD. The variable object fraction was previously determined based on variable objects detected in the HIPPARCOS time series (ESA 1997), for example. Here we significantly expand this investigation using 13.5 million stars with heliocentric distances of up to 1 kpc that satisfy the astrometric and photometric selection criteria listed in Appendix B as well as (a) at least 20 observations in the $G$, $G_{BP}$, and $G_{RP}$ bands, and (b) a relative parallax uncertainty of <5%. In order to reduce the number of objects that are affected by significant extinction, stars at low Galactic latitudes (from $-5$ to $5^\circ$) are excluded. This effectively reduces the number of disc variables such as classical Cepheids and βCephei stars.

Figure 8 illustrates this Gaia-based high-resolution map of the variable object fraction in the CaMD at the precision level of approximately 5–10 mmag. Variability is identified in about 9% of the stars based on a supervised classification of Gaia sources. This method depends heavily on the selection of the training set of constant and variable objects. Minor colour-coded features can be due to training-set related biases. The detection of variability further depends upon the amplitude of the variables, their apparent magnitude distribution, and the instrumental precision. The accuracy of the fraction of variables is also affected by the number of sources per bin of absolute magnitude and colour, which can be as low as one in the tails of the two-dimensional source number density distribution.

Figure 8 contains many informative features, despite possible biases. Future data releases will significantly improve upon Fig. 8 by correcting for reddening and extinction and using a larger number of objects with more accurate source classifications. For the time being, we remark the following.

– The classical instability strip is clearly visible with a variability in about 50–60% of the stars (although extinction limits the precision of this estimate).

– For evolved stars, red giants, and asymptotic giant branch stars, we find that higher luminosity and redder colour implies a higher probability of variability.

– The red clump has a very low fraction of variable stars in the Gaia data. Kepler photometry of red clump stars has revealed complex variability at the micro-mag level that has been used extensively for asteroseismology, cf. Sect. 1 and references therein.

– The classical ZZ Ceti location is extremely concentrated in colour and magnitude, with variability in about half of the stars. The concentration is due to the partial ionisation of hydrogen in the outer envelope of white dwarfs, which is developed only in extremely narrow ranges of effective temperatures (see Fontaine & Brassard 2008).

– Extreme horizontal branch stars show a high probability of variability.

– The hottest and most luminous main-sequence stars are very frequently variable.

– There is a clear gradient towards larger fractions of variables above the low-mass main-sequence stars.

4. Variability amplitudes in the CaMD

Figure 9 shows variability amplitudes as a function of position in the CaMD. Here, we quantify variability amplitudes using the $G$-band inter-quartile range (IQR). Objects are selected according to the general criteria described in Appendix B, with stricter conditions on the parallax (greater than 1 mas) and its relative precision (better than 5%). To prevent the false impression that faint (and very bright) sources have intrinsically higher amplitudes, we corrected for the instrumental spread of the IQR as a function of the median $G$ magnitude. This correction was determined using sources that are classified as constant in the all-sky classification (Rimoldini et al. 2018) and subtracted in.
Fig. 8. Variable object fraction in the CaMD shown as a colour scale as labelled. This figure is not based on variable objects from the literature. Instead, variability is detected directly using Gaia data and employing supervised classification for sources with at least 20 observations in the G, G<sub>BP</sub>, and G<sub>Rp</sub> bands. All objects satisfy the selection criteria described in Appendix B, but with more restrictive constraints on the parallax precision (parallax<sub>over_error</sub> > 20) and on the parallax value (ϖ > 1 mas), which limits the sample (size, extinction, and reddening). In order to reduce the extinction effect, objects at low Galactic latitudes (from −5° to 5°) are excluded. About 9% of the 13.5 million stars that satisfy these criteria are variable. Some of the bins (especially the outlying ones) can contain only a few or even single sources. The condition on the relative precision of G<sub>BP</sub> measurements introduces artificial cuts in the distributions of low-mass main-sequence stars and red (super)giants.

Figure 9 contains the following stellar variability types based on the all-sky classification (Rimoldini et al. 2018): α<sup>2</sup> Canum Venaticorum, α Cygni, β Cephei, cataclysmic, classical Cepheids, δ Scuti, γ Cassiopeiae, γ Doradus, Mira, ellipsoidal, RR Lyrae of Bailey’s type ab and c, semiregular, slowly pulsating B stars, solar-like variability due to magnetic activity (flares, spots, and rotational modulation), SX Arietis, and SX Phoenicis. We did not include other classes (listed in Eyer et al. 2018) for clarity or because there were too few objects. We note that any specific selection criteria applied to the objects shown in Fig. 9 introduce biases that can highlight or diminish the prominence of certain phenomena. Nevertheless, Fig. 9 provides a first detailed illustration of some of the most important amplitude-related variability features in the CaMD. A number of clumps and instability regions are visible in Fig. 9, which are related to the variability classes described in Sects. 2 and 3. We notice the following features.

– The classical instability strip that contains classical Cepheids and RR Lyrae stars is not very prominent, although some clumps (in red or cyan) are apparent.
– The instability regions linked to SPB stars and β Cephei stars are broad and uniform.
– Higher amplitude variations are clearly correlated with redder colours for long-period variables.
– Variables with the highest amplitude (IQR > 0.1 mag) occur in several regions in the CaMD, including the classical instability strip, long-period variables, below the red clump, above the main sequence of low-mass stars (in correspondence of the observed gradient in the fraction of variables), and between the white dwarf sequence and the main sequence.
– Significant amplitudes of >0.04 mag are found very frequently among the coolest white dwarfs.
– The stars between the main sequence and the white dwarfs sequence feature large variability amplitudes and extend into the clump of ZZ Ceti stars in the white dwarf sequence. This intermediate region is populated in particular by the high-amplitude cataclysmic variables, cf. Fig. 7. A closeup view of the white dwarf sequence is shown in Fig. 10, which represents all classified variables within 200 pc. Each object is plotted without binning to emphasise the variability of the ZZ Ceti stars.
Fig. 9. Amplitude of variability in the CaMD based on a selection of classified variables within 1 kpc and with a relative uncertainty for the parallax of 5%. The colour scale shows the corrected $G$-band IQR (see text) with a cut-off at 0.1 mag to emphasise the low- and mid-level variability. The background points in grey represent classified constant stars. All objects satisfy the selection criteria described in Appendix B, in addition to the stricter conditions on parallax and its precision, as mentioned above. The effects of interstellar extinction are not corrected for.

Fig. 10. Same as Fig. 9, but focusing on the white dwarf sequence and plotting all classified variables within 200 pc with a relative uncertainty for the parallax better than 5%. A close inspection of this sequence reveals amplitudes at the level of 40 mmag in various regions.

5. Variability-induced motion in the CaMD

In this section, we visualise the variability-induced motion of stars in the time-dependent CaMD using all-sky measurements made in the $G$, $G_{BP}$, and $G_{RP}$ passbands. Gaia data are uniquely suited to create this time-dependent CaMD, since the different data types (astrometric and photometric in three bands) are acquired in a quasi-simultaneous fashion at many epochs that are distributed over a multi-year time span. The first of such representations, although much less detailed, was presented for individual classical Cepheids in the Milky Way (Eggen 1951) and in Galactic star clusters (Kholopov 1956; Sandage 1958). Similarly minded representations in the literature were based on data from the SDSS (Ivezić et al. 2003, mostly Galactic objects), EROS (Spano et al. 2009, LMC objects), and, very recently, the HST observations of M51 (Conroy et al. 2018).

Figure 11 illustrates the variability-induced motion of stars in the CaMD. As elsewhere in this paper, no correction for interstellar extinction is applied. Individual stars are shown by differently (arbitrarily) coloured lines that connect successive absolute $G$ magnitudes and $G_{BP} - G_{RP}$ measurements, that is, the observations are ordered in time as opposed to variability phase. This choice was made to avoid uncertainties related to phase-folding the relatively sparsely sampled light curves based on 22 or fewer months of observations and to include both periodic and non-periodic variable objects.

Figure 11 is limited to a subset of all available variable stars in order to avoid overcrowding the diagram. As a preview for future data releases, we include here the variability-induced motions of some stars whose time series and variability types are not published in DR2 (but which are available as online material). Figure 11 includes the following variability types as defined in Sect. 2: $\alpha^2$ Canum Venaticorum variables, B-type emission
Fig. 11. Motions of selected variable stars in the CaMD, highlighted by segments connecting successive absolute $G$ magnitudes and $G_{BP} - G_{RP}$ measurements in time with the same colour for the same source. Preferential directions and amplitudes of magnitude and colour variations can be inferred as a function of variability type ($\alpha^2$ Canum Venaticorum, Be-type and $\gamma$ Cassiopeiae, cataclysmic, classical and type-II Cepheid, $\delta$ Scuti and SX Phoenicis, eclipsing binary, long period, and RR Lyrae), as labelled in the figure. For clarity of visualisation, the selection of eclipsing binaries (and partially other types) was adjusted to minimise the overlap with other types. Selection criteria of all sources represented in colour or grey are the same as in Fig. 3. Additional conditions are described in the text. An animated version of this figure is available online.
line /γ Cassiopeiae stars, cataclysmic variables, classical and type-II Cepheids, δ Scuti stars, eclipsing binaries, RR Lyrae stars, long-period variables, and SX Phoenicis stars. All sources shown satisfy the general criteria described in Appendix B and typically have at least ten available observations 1. We further prioritised the selection of objects featuring wider ranges of variations in the G band (with a minimum of about 0.1 mag) 2. The number of sources shown for each variability type ranges from a few to several tens and was selected to ensure clarity in case of high source density or overlapping variability classes in certain regions of the CaMD. In order to limit the effect of outlying values, time-series data are filtered by operators as described in Holl et al. (2018), and the 10 % of the brightest and faintest observations in the $g_{BP}$ band are excluded for sources with $g_{BP} - g_{RP}$ lower than 1.5 mag. Non-variable objects are shown as a grey background to provide a visual reference for the variable object locations in the CaMD. These stars satisfy the criteria described in Appendix B as well as the stricter condition of $r > 1$ mas. Stars whose variability is caused by different physical effects exhibit different motions within the time-dependent CaMD. We briefly summarise the different motions seen in Fig. 11 as follows.

1. Pulsating stars, including long-period variables, Cepheids, RR Lyrae, and δ Scuti/SX Phoenicis stars, exhibit a similar behaviour. These stars are bluer when brighter in G, which illustrates that brightness variations of pulsating stars are dominated by the effect of change in temperature rather than radius. For the longest-period variables, the 22-month time span of the Gaia data is similar to the pulsation cycle, so that in some cases, loop-like shapes are apparent. For variable stars with shorter periods (e.g. RR Lyrae stars or classical Cepheids), successive measurements in time are not generally ordered in phase, so that an overall “envelope” of many cycles is revealed.

2. The motions of eclipsing binary systems in the CaMD depend on the colour difference between the two stars. The components of eclipsing systems of the EW type have similar mass (and colour), leading to a rather vertically aligned motion (parallel to the absolute magnitude axis). For eclipsing binary systems with stars of different mass (and thus colour) close to the main sequence, the deepest eclipse is usually slightly redder, since the secondary (less massive and redder) component eclipses part of the light of the primary star. The slope of the motion of eclipsing binaries in the CaMD is very different (much steeper) than the one of pulsating stars.

3. Rotationally induced variables (shown here: $\alpha^2$ Canum Venaticorum stars) feature small amplitudes in absolute G and are rather horizontal in the CaMD. This is as expected from starspots, which have a lower temperature than the surroundings and hence absorb the light at bluer frequencies to re-emit it by back-warming effect at redder frequencies. Therefore, the magnitude change in a broad band like G is smaller than it would be if measured in narrow bands.

4. Eruptive stars (shown here: γ Cassiopeiae and Be-type stars) become redder when brighter because of additional extinction during their eruptive phase. The slopes of their motions in the CaMD therefore have the opposite sign with respect to the sign of pulsating stars.

5. The variability of cataclysmic variables (shown here: novae) primarily features strong outbursts in the ultraviolet and blue part of the spectrum that are understood to be caused by mass transfer from donor stars in binary systems. These outbursts very significantly change the colour of the system towards bluer values.

The current version of Fig. 11 represents a first step towards a more global description of stellar variability. The motions described by the variable stars in the time-dependent CaMD provide new perspectives on the data that can be exploited as variable star classification attributes to appreciate improve the classification results. Gaia data will definitively help identify misclassifications and problems in published catalogues, thanks to its astrometry and the quasi-simultaneous measurements.

In future Gaia data releases, there will be more data points per source, which will enable us to refine Fig. 11. In particular, periods can be determined with an accuracy inversely proportional to the total Gaia time base for periodic objects. In this way, the motion in the CaMD can be represented more precisely by connecting points that are sorted in phase (rather than in time). This leads to Lissajous-type configurations for pulsators. For sufficiently bright stars, radial velocity time series will add a third and unprecedented dimension to Fig. 11.

An animated version of Fig. 11 is provided online and at https://www.cosmos.esa.int/web/gaia/gaiadr2_cu7. We provide material at the CDS that includes the time series in the G, BP, and RP bands of the selected field-of-view transits for 224 sources that are not published in Gaia DR2, but are plotted in Fig. 11.

6. Conclusions

The Gaia mission enables a comprehensive description of phenomena related to stellar variability. We here focused on stellar variability in the CaMD and showed locations that are occupied by different variability types as well as variable object fractions, variability amplitudes, and variability-induced motions that are described by different variability classes in the CaMD.

The wealth of information related to variable stars that is contained in Gaia DR2 is unprecedented for the Milky Way. The CaMD can provide guidance for further detailed studies, which can focus on individual regions or clumps, for instance, to investigate the purity of instability strips and how sharply such regions are truly defined or how they depend on chemical composition. Of course, additional work is required to this end, and accurately correcting for reddening and extinction will be crucial. The (time-dependent) CaMD will play an important role for improving the variable star classification by providing additional attributes, such as the expected direction of variability for specific variable classes, and for illustrating stellar variability to non-expert audiences.

The CaMD of variable stars can further point out interrelations between variability phenomena that are otherwise not easily recognised, and it might be able to identify new types of variability. Detailed follow-up observations from the ground will help correct previous misclassifications and enable in-depth studies of peculiar and particularly interesting objects. Based on the variable stars that reside in the Milky Way, as presented here, it will be possible to obtain data with particularly high S/N, for example through high-resolution spectroscopy. Finally, the observed properties of variable stars in the CaMD, such as instability strip boundaries or period-luminosity relations, provide

---

1 The minimum number of observations per source is increased to 20 in the case of long-period variables, but the condition on the number of observations is removed for cataclysmic variables.

2 A minimum range in the G band is not required for $\alpha^2$ Canum Venaticorum stars and cataclysmic variables as their variability may be small in the “white” G band.
crucial input and constraints for models describing pulsational instability, convection, and stellar structure in general. Future Gaia data releases will further surpass the variability content of this second data release 3. By the end of mission, Gaia data are expected to comprise many tens of millions of variable celestial objects, including many additional variability types, as well as time series BP and RP spectra. Eventually, time series of radial velocities and spectra from the radial velocity spectrometer will be published for subsets of variables. Finally, the variability classification of future Gaia data will also make use of unsupervised clustering techniques aimed at discovering entirely new (sub-)clusters and classes of variable phenomena.

Acknowledgements. We would like to thank Laurent Roudraissy for tests done on the representation of time series. This work has made use of data from the ESA space mission Gaia, processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC has been provided by national institutions, some of which participate in the Gaia Multilateral Agreement, which include, for Switzerland, the Swiss State Secretariat for Education, Research and Innovation through the ESA Proex program, the “Mesures d’accompagnement”, the “Activités Nationales Complémentaires”, the Swiss National Science Foundation, and the Early Postdoc.Mobility fellowship; for Belgium, the Fonds de la Recherche Scientifique (F.R.S.-FNRS), the “Activités Nationales Complémentaires”, the Swiss National Science Foundation, the Early PostdocMobility fellowship, and the BELgian Fonds pour la Recherche Scientifique (F.R.S.-FNRS); for France, the centre national d’Etudes Spatiales (CNES); Part of this research has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Advanced Grant agreements N° 670519: MAMSIE (“Mixing and Angular Momentum transport in Massive stars”)). This research has made use of NASA's Astrophysics Data System, the VizieR catalogue access tool, CDS, Strasbourg, France, and the International Variable Star Index (VXSI) database, operated at AAVSO, Cambridge, Massachusetts, USA. We gratefully acknowledge Mark Taylor for creating the astronomy-oriented data handling and visualization software TOPCAT (Taylor 2005).

References

Acknowledgements. We would like to thank Laurent Roudraissy for tests done on the representation of time series. This work has made use of data from the ESA space mission Gaia, processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC has been provided by national institutions, some of which participate in the Gaia Multilateral Agreement, which include, for Switzerland, the Swiss State Secretariat for Education, Research and Innovation through the ESA Proex program, the “Mesures d’accompagnement”, the “Activités Nationales Complémentaires”, the Swiss National Science Foundation, and the Early Postdoc.Mobility fellowship; for Belgium, the Fonds de la Recherche Scientifique (F.R.S.-FNRS), the “Activités Nationales Complémentaires”, the Swiss National Science Foundation, the Early PostdocMobility fellowship, and the BELgian Fonds pour la Recherche Scientifique (F.R.S.-FNRS); for France, the centre national d’Etudes Spatiales (CNES); Part of this research has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Advanced Grant agreements N° 670519: MAMSIE (“Mixing and Angular Momentum transport in Massive stars”)). This research has made use of NASA's Astrophysics Data System, the VizieR catalogue access tool, CDS, Strasbourg, France, and the International Variable Star Index (VXSI) database, operated at AAVSO, Cambridge, Massachusetts, USA. We gratefully acknowledge Mark Taylor for creating the astronomy-oriented data handling and visualization software TOPCAT (Taylor 2005).

References

122 QUASAR Science Resources for ESA/ESAC, Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
123 Fork Research, Rua do Cruzado Osberno, Lt. 1, 9 esq., Lisboa, Portugal
124 APAVE SUDEUROPE SAS for CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
125 Nordic Optical Telescope, Rambla José Ana Fernández Pérez 7, 38711 Breña Baja, Spain
126 Spanish Virtual Observatory, Spain
127 Fundación Galileo Galilei – INAF, Rambla José Ana Fernández Pérez 7, 38712 Breña Baja, Santa Cruz de Tenerife, Spain
128 INSA for ESA/ESAC, Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
129 Departamento Arquitectura de Computadores y Automática, Facultad de Informática, Universidad Complutense de Madrid, C/ Prof. José García Santandemases s/n, 28040 Madrid, Spain
130 H H Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK
131 Institut d’Estudis Espacials de Catalunya (IEEC), Gran Capita 2-4, 08034 Barcelona, Spain
132 Applied Physics Department, Universidade de Vigo, 36310 Vigo, Spain
133 Stellar Astrophysics Centre, Aarhus University, Department of Physics and Astronomy, 120 Ny Munkegade, Building 1520, 8000 Aarhus C, Denmark
134 Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
135 Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611 Australia
136 Sorbonne Universités, UPMC Université Paris 6 et CNRS, UMR 7095, Institut d’Astrophysique de Paris, 98 bis bd. Arago, 75014 Paris, France
137 Department of Geosciences, Tel Aviv University, Tel Aviv 6997801, Israel
## Appendix A: Literature for each variability type

See Table A.1 for details on the references from the literature regarding the objects included in Figs. 3–7 and 11.

### Table A.1. Literature references of stars as a function of variability type and the corresponding number of sources depicted in Figs. 3–7, after selections based on reliability, photometric accuracy, and astrometric parameters (Appendix B).

<table>
<thead>
<tr>
<th>Variability Type</th>
<th>Reference</th>
<th># Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pulsating</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α Cygni</td>
<td>Hip97, VSX16</td>
<td>17</td>
</tr>
<tr>
<td>β Cephei</td>
<td>PDC05</td>
<td>20</td>
</tr>
<tr>
<td>Cepheid</td>
<td>ASA09, Hip97, INT12</td>
<td>155</td>
</tr>
<tr>
<td>δ Scuti</td>
<td>ASA09, Hip97, JD07, Kep11b, Kep11c, SDS12</td>
<td>724</td>
</tr>
<tr>
<td>γ Doradus</td>
<td>FKA16, Kep11b, Kep11c, VSX16</td>
<td>561</td>
</tr>
<tr>
<td>Long Period Variable</td>
<td>ASA12, Hip97, INT12, Kep11b, NSV04</td>
<td>5221</td>
</tr>
<tr>
<td>PV Telescopii</td>
<td>VSX16</td>
<td>3</td>
</tr>
<tr>
<td>Rapidly Oscillating Am star</td>
<td>VSX16</td>
<td>8</td>
</tr>
<tr>
<td>Rapidly Oscillating Ap star</td>
<td>VSX16</td>
<td>25</td>
</tr>
<tr>
<td>RR Lyrae, fundamental mode (RRab)</td>
<td>ASA09, ASA12, Cat13a, Cat13b, Cat14b, Cat15, Hip97, INT12, LIN13, NSV06, VFB16, VSX16</td>
<td>1676</td>
</tr>
<tr>
<td>RR Lyrae, first overtone (RRc)</td>
<td>ASA09, ASA12, Cat13b, Cat14b, Hip97, INT12, Kep11b, LIN13, MA14, VFB16, VSX16</td>
<td>611</td>
</tr>
<tr>
<td>RV Tauri</td>
<td>ASA12, Hip97, VSX16</td>
<td>48</td>
</tr>
<tr>
<td>Slowly Pulsating B star</td>
<td>IUE03, Hip97, PDC05</td>
<td>78</td>
</tr>
<tr>
<td>SX Phoenixis</td>
<td>ASA12, Hip97, VSX16</td>
<td>41</td>
</tr>
<tr>
<td>Type-II Cepheid</td>
<td>ASA12, Cat14b, Hip97, VSX16</td>
<td>21</td>
</tr>
<tr>
<td>V361 Hya (also EC 14026)</td>
<td>VSX16</td>
<td>41</td>
</tr>
<tr>
<td>V1093 Her (also PG 1716)</td>
<td>VSX16</td>
<td>1</td>
</tr>
<tr>
<td>ZZ Ceti</td>
<td>VSX16</td>
<td>61</td>
</tr>
<tr>
<td><strong>Rotational</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α² Canum Venaticorum</td>
<td>Hip97, VSX16</td>
<td>598</td>
</tr>
<tr>
<td>Binary with Reflection</td>
<td>VSX16</td>
<td>27</td>
</tr>
<tr>
<td>BY Draconis</td>
<td>VSX16</td>
<td>713</td>
</tr>
<tr>
<td>Ellipsoidal</td>
<td>ASA12, Cat14b, Hip97, Kep11b, VSX16</td>
<td>398</td>
</tr>
<tr>
<td>FK Comae Berenices</td>
<td>Hip97</td>
<td>3</td>
</tr>
<tr>
<td>Rotating Spotted</td>
<td>Kep15b</td>
<td>16</td>
</tr>
<tr>
<td>RS Canum Venaticorum</td>
<td>ASA12, Cat14b, Hip97, VSX16</td>
<td>1381</td>
</tr>
<tr>
<td>Solar-Like Variations</td>
<td>HAT10</td>
<td>176</td>
</tr>
<tr>
<td>SX Arietis</td>
<td>Hip97, VSX16</td>
<td>14</td>
</tr>
<tr>
<td><strong>Eclipsing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA, β Persei (Algol)</td>
<td>ASA09, Cat14b, Hip97, LIN13, VSX16</td>
<td>8123</td>
</tr>
<tr>
<td>EB, β Lyrae</td>
<td>ASA09, ASA12, Cat14b, Hip97, LIN13, VSX16</td>
<td>3096</td>
</tr>
<tr>
<td>EW, W Ursae Majoris</td>
<td>ASA09, Hip97, VSX16</td>
<td>3248</td>
</tr>
<tr>
<td>Exoplanet</td>
<td>JS15</td>
<td>278</td>
</tr>
<tr>
<td><strong>Eruptive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B-type emission-line star</td>
<td>ASA12, VSX16</td>
<td>86</td>
</tr>
<tr>
<td>Classical T Tauri Star</td>
<td>VSX16</td>
<td>75</td>
</tr>
<tr>
<td>Flares (UV, BY, TTS)</td>
<td>Kep11a, Kep13, Kep15a, MMT15</td>
<td>478</td>
</tr>
<tr>
<td>γ Cassiopeiae</td>
<td>Hip97, VSX16</td>
<td>84</td>
</tr>
<tr>
<td>R Coronae Borealis</td>
<td>VSX16</td>
<td>4</td>
</tr>
<tr>
<td>S Doradus</td>
<td>ASA12, INT12</td>
<td>2</td>
</tr>
<tr>
<td>T Tauri Star (TTS)</td>
<td>VSX16</td>
<td>173</td>
</tr>
<tr>
<td>UV Ceti</td>
<td>INT12, VSX16</td>
<td>425</td>
</tr>
<tr>
<td>Weak-lined T Tauri Star</td>
<td>VSX16</td>
<td>119</td>
</tr>
<tr>
<td>Wolf-Rayet</td>
<td>INT12, VSX16</td>
<td>15</td>
</tr>
<tr>
<td><strong>Cataclysmic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cataclysmic Variable (generic)</td>
<td>Cat14a, OGL15, VSX16</td>
<td>132</td>
</tr>
<tr>
<td>U Geminorum</td>
<td>INT12, VSX16</td>
<td>4</td>
</tr>
<tr>
<td>Z Andromedae</td>
<td>INT12, VSX16</td>
<td>5</td>
</tr>
</tbody>
</table>

Notes. Figure 11 includes only subsets of variability types and of sources per type.

References. ASA09: Pigulski et al. (2009); ASA12: Richards et al. (2012); Cat13a: Drake et al. (2013a); Cat13b: Drake et al. (2013b); Cat14a: Drake et al. (2014a); Cat14b: Drake et al. (2014b); Cat15: Torrealba et al. (2015); FKA16: Kahraman Aliçavuş et al. (2016); HAT10: Hartman et al. (2010); Hip97: ESA (1997); INT12: Alfonso-Garzón et al. (2012); IUE03: Niemczura (2003); JD07: Debosscher et al. (2007); JS15: J. Southworth, http://www.astro.keele.ac.uk/jkt/tepcat/observables.html (as of Aug. 2015); Kep11a: Walkowicz et al. (2011); Kep11b: Debosscher et al. (2011); Kep11c: Uytterhoeven et al. (2011); Kep13: Shibayama et al. (2013); Kep15a: Wu et al. (2015); Kep15b: Reinhold & Gizon (2015); LIN13: Palaversa et al. (2013); MA14: Abbas et al. (2014); MMT15: Chang et al. (2015); NSV04: Wozniak et al. (2004); NSV06: Kinemuchi et al. (2006); OGL15: Mróz et al. (2015); PDC05: P. De Cat, http://www.ster.kuleuven.ac.be/~peter/Bstars/ (as of Jan. 2005); SDS12: Süveges et al. (2012); VFB16: Braga et al. (2016); VSX16: Watson et al. (2016, 2015, 2006).
Appendix B: Selection criteria

Astrometric and photometric conditions are applied to all CaMDs for improved accuracy of the star locations in such diagrams. Astrometric constraints include limits on the number of visibility periods (observation groups separated from other groups by at least four days) per source used in the secondary astrometric solution (Gaia Collaboration 2018), the excess astrometric noise of the source postulated to explain the scatter of residuals in the astrometric solution for that source (Gaia Collaboration 2018), and the relative parallax precision (herein set to 20% although in other applications it was set to 5 or 10%):

1. visibility_periods_used > 5;
2. astrometric_excess_noise < 0.5 mas;
3. parallax > 0 mas;
4. parallax_over_error > 5.

Photometric conditions set limits for each source on the relative precisions of the mean fluxes in the \(G_{BP}\), \(G_{RP}\), and \(G\) bands, as well as on the mean flux excess in the \(G_{BP}\) and \(G_{RP}\) bands with respect to the \(G\) band as a function of colour (Evans et al. 2018):

5. \(\frac{\text{phot bp mean flux error}}{\text{phot bp mean flux}} < 0.05\);
6. \(\frac{\text{phot rp mean flux error}}{\text{phot rp mean flux}} < 0.05\);
7. \(\frac{\text{phot g mean flux error}}{\text{phot g mean flux}} < 0.02\);
8. \(\frac{(\text{phot bp mean flux} + \text{phot rp mean flux})}{\text{phot g mean flux} * [1.2 + 0.03 * (\text{phot bp mean mag} - \text{phot rp mean mag})^2]} < 1.2\).

The ADQL query to select a sample of sources that satisfy all of the above listed criteria follows.

```
SELECT TOP 10 source_id
FROM gaiadr2.gaia_source
WHERE visibility_periods_used > 5
  AND astrometric_excess_noise < 0.5
  AND parallax > 0
  AND parallax_over_error > 5
  AND phot_bp_mean_flux_over_error > 20
  AND phot_rp_mean_flux_over_error > 20
  AND phot_g_mean_flux_over_error > 50
  AND phot_bp_rp_excess_factor < 1.2 * (1.2 + 0.03 * power(phot_bp_mean_mag - phot_rp_mean_mag, 2))
```