Gaia Focused Product Release:
Radial velocity time series of long-period variables

Gaia Collaboration: M. Trabucchi[1,2,⋆], N. Mowlavi[2], T. Lebzelter[3], I. Lecoeur-Taibi[4], M. Audard[2,4],
L. Eyer[2], P. García-Lario[5], P. Gavras[6], B. Holl[2,4], G. Jevardat de Fombelle[2], K. Nienartowicz[7,4],
L. Rimoldini[4], P. Sartoretti[8], R. Blomme[9], Y. Frémat[9], O. Marchal[10], Y. Damerdji[11,12], A. G. A. Brown[13],
A. Guerrier[14], P. Panuzzo[8], D. Katz[8], G. M. Seabroke[15], K. Benson[15], R. Haigron[8], M. Smith[15], A. Lobel[9],
A. Vallenari[16], T. Prusti[17], J. H. J. de Bruijne[17], F. Arenou[8], C. Babusiaux[18], A. Barbier[14],
M. Biermann[19], O. L. Creevey[20], C. Ducourant[21], D. W. Evans[22], R. Guerra[5], A. Hutton[23], C. Jordi[24,25,26],
S. A. Klioner[27], U. Lammers[5], L. Lindegren[28], X. Luri[24,25,26], F. Mignard[20], S. Randich[29], R. Smiljanic[30],
P. Tanga[20], N. A. Walton[22], C. A. L. Bailer-Jones[31], U. Bastian[19], M. Cropper[15], R. Drimmel[32],
M. G. Lattanzi[32,33], C. Soubiran[21], F. van Leeuwen[22], J. Bakker[5], J. Castañeda[34,24,26], F. De Angeli[22],
C. Fabricius[26,24,25], M. Fouesneau[31], L. Galluccio[20], E. Masana[26,24,25], R. Messineo[35], C. Nicolas[14],
F. Pailler[14], F. Riclet[14], W. Roux[14], R. Sordo[16], F. Thévenin[20], G. Gracia-Abril[36,19], J. Portell[24,25,26],
D. Teyssier[37], M. Altmann[19,38], J. Berthier[39], P. W. Burgess[22], D. Busonero[32], G. Busso[22], H. Cánovas[37],
B. Carry[20], N. Cheek[40], G. Clementini[41], M. Davidson[42], P. de Teodoro[5], L. Delchambre[11], A. Dell'Oro[29],
E. Fraile Garcia[6], D. Garabato[43], N. Garralda Torres[44], N. C. Hambly[42], D. L. Harrison[22,45],
D. Hatzidimitriou[46], J. Hernández[5], S. T. Hodgkin[22], S. Jamal[31], S. Jordan[19], A. Krone-Martins[47,48],
A. C. Lanzafame[49,50], W. Löffler[19], A. Lorca[23], P. M. Marrese[51,52], A. Moitinho[48], K. Muinonen[53,54],
M. Nuñez Campos[23], I. Oreshina-Slezak[20], P. Osborne[22], E. Pancino[29,52], T. Pauwels[9], A. Recio-Blanco[20],
M. Riello[22], A. C. Robin[55], T. Roegiers[56], L. M. Sarro[57], M. Schultheis[20], C. Siopis[58], A. Sozzetti[32],
E. Utrilla[23], M. van Leeuwen[22], K. Weingrill[59], U. Abbas[32], P. Ábrahám[60,61], A. Abreu Aramburu[44],
C. Aerts[62,63,31], G. Altavilla[51,52], M. A. Álvarez[43], J. Alves[3], F. Anders[24,25,26], R. I. Anderson[64],
T. Antoja[24,25,26], D. Baines[65], S. G. Baker[15], Z. Balog[19,31], C. Barache[38], D. Barbato[2,32], M. Barros[66],
M. A. Barstow[67], S. Bartolomé[26,24,25], D. Bashi[68,69], N. Bauchet[8], N. Baudeau[70], U. Becciani[49],
L. R. Bedin[16], I. Bellas-Velidis[71], M. Bellazzini[41], W. Beordo[32,33], A. Berihuete[72], M. Bernet[24,25,26],
C. Bertolotto[35], S. Bertone[32], L. Bianchi[73], A. Binnenfeld[74], A. Blazere[75], T. Boch[10], A. Bombrun[76],
S. Bouquillon[38,77], A. Bragaglia[41], J. Braine[21], L. Bramante[35], E. Breedt[22], A. Bressan[78], N. Brouillet[21],
E. Brugaletta[49], B. Bucciarelli[32,33], A. G. Butkevich[32], R. Buzzi[32], E. Caffau[8], R. Cancelliere[79],
S. Cannizzo[80], R. Carballo[81], T. Carlucci[38], M. I. Carnerero[32], J. M. Carrasco[26,24,25],
J. Carretero[82,83], S. Carton[80], L. Casamiquela[21,8], M. Castellani[51], A. Castro-Ginard[13], V. Cesare[49],
P. Charlot[21], L. Chemin[84], V. Chiaramida[35], A. Chiavassa[20], N. Chornay[22,4], R. Collins[42], G. Contursi[20],
W. J. Cooper[85,32], T. Cornez[80], M. Crosta[32,86], C. Crowley[76], C. Dafonte[43], M. David[87],
P. de Laverny[20], F. De Luise[88], R. De March[35], J. De Ridder[62], R. de Souza[89], A. de Torres[76],
E. F. del Peloso[19], M. Delbo[20], A. Delgado[6], T. E. Dharmawardena[31], S. Diakite[90], C. Diener[22], E. Distefano[49],
C. Dolding[15], K. Dsilva[58], J. Durán[6], H. Enke[59], P. Esquej[6], C. Fabre[75], M. Fabrizio[51,52], S. Faigler[68],
M. Fatović[91], G. Fedorets[53,92], J. Fernández-Hernández[6], P. Fernique[10], F. Figueras[24,25,26], Y. Fournier[59],
C. Fouron[70], M. Gai[32], M. Galinier[20], A. Garcia-Gutierrez[26,24,25], M. García-Torres[93], A. Garofalo[41],
E. Gerlach[27], R. Geyer[27], P. Giacobbe[32], G. Gilmore[22,94], S. Girona[95], G. Giuffrida[51], R. Gomel[68],
A. Gomez[43], J. González-Núñez[96], I. González-Santamaría[43], E. Gosset[11,97], M. Granvik[53,98],
V. Gregori Barrera[26,24,25], R. Gutiérrez-Sánchez[37], M. Haywood[8], A. Helmer[80], A. Helmi[99], K. Henares[65],
S. L. Hidalgo[100,101], T. Hilger[27], D. Hobbs[28], C. Hottier[8], H. E. Huckle[15], M. Jabłońska[102,103], F. Jansen[104],
Ó. Jiménez-Arranz[24,25,26], J. Juaristi Campillo[19], S. Khanna[32,99], G. Kordopatis[20], Á. Kóspál[60,31,61],
Z. Kostrzewa-Rutkowska[13], M. Kun[60], S. Lambert[38], A. F. Lanza[49], J.-F. Le Campion[21], Y. Lebreton[105,106],
S. Leccia[107], G. Lecoutre[55], S. Liao[108,32,109], L. Liberato[20,110], E. Licata[32], H. E. P. Lindstrøm[32,111,112],
T. A. Lister[113], E. Livanou[46], C. Loup[10], L. Mahy[9], R. G. Mann[42], M. Manteiga[114], J. M. Marchant[115],
M. Marconi[107], D. Marín Pina[24,25,26], S. Marinoni[51,52], D. J. Marshall[116], J. Martín Lozano[40],

⋆ Corresponding author: M. Trabucchi, e-mail: michele.trabucchi@unige.ch

J. M. Martín-Fleitas[23], G. Marton[60], N. Mary[80], A. Masip[26,24,25], D. Massari[41], A. Mastrobuono-Battisti[8], T. Mazeh[68], P. J. McMillan[28], J. Meichsner[27], S. Messina[49], D. Michalik[17], N. R. Millar[22], A. Mints[59], D. Molina[25,24,26], R. Molinaro[107], L. Molnár[60,117,61], G. Monari[10], M. Monguió[24,25,26], P. Montegriffo[41], A. Montero[40], R. Mor[118,25,26], A. Mora[23], R. Morbidelli[32], T. Morel[11], D. Morris[42], D. Munoz[80], T. Muraveva[41], C. P. Murphy[5], I. Musella[107], Z. Nagy[60], S. Nieto[6], L. Noval[80], A. Ogden[22], C. Ordenovic[20], C. Pagani[119], I. Pagano[49], L. Palaversa[91], P. A. Palicio[20], L. Pallas-Quintela[43], A. Panahi[68], C. Panem[14], S. Payne-Wardenaar[19], L. Pegoraro[14], A. Penttilä[53], P. Pesciullesi[6], A. M. Piersimoni[88], M. Pinamonti[32], F.-X. Pineau[10], E. Plachy[60,117,61], G. Plum[8], E. Poggio[20,32], D. Pourbaix[†,58,97], A. Prša[120], L. Pulone[51], E. Racero[40,121], M. Rainer[29,122], C. M. Raiteri[32], P. Ramos[123,24,26], M. Ramos-Lerate[37], M. Ratajczak[102], P. Re Fiorentin[32], S. Regibo[62], C. Reylé[55], V. Ripepi[107], A. Riva[32], H.-W. Rix[31], G. Rixon[22], N. Robichon[8], C. Robin[80], M. Romero-Gómez[24,25,26], N. Rowell[42], F. Royer[8], D. Ruz Mieres[22], K. A. Rybicki[124], G. Sadowski[58], A. Sáez Núñez[26,24,25], A. Sagristà Sellés[19], J. Sahlmann[6], V. Sanchez Gimenez[26,24,25], N. Sanna[29], R. Santoveña[43], M. Sarasso[32], C. Sarrate Riera[34,24,26], E. Sciacca[49], J. C. Segovia[40], D. Ségransan[2], S. Shahaf[124], A. Siebert[10,125], L. Siltala[53], E. Slezak[20], R. L. Smart[32,85], O. N. Snaith[8,126], E. Solano[127], F. Solitro[35], D. Souami[105,128], J. Souchay[38], L. Spina[16], E. Spitoni[20,129], F. Spoto[130], L. A. Squillante[35], I. A. Steele[115], H. Steidelmüller[27], J. Surdej[11], L. Szabados[60], F. Taris[38], M. B. Taylor[131], R. Teixeira[89], K. Tisanić[91], L. Tolomei[35], F. Torra[34,24,26], G. Torralba Elipe[43,132,133], M. Tsantaki[29], A. Ulla[134,135], N. Unger[2], O. Vanel[8], A. Vecchiato[32], D. Vicente[95], S. Voutsinas[42], M. Weiler[26,24,25], Ł. Wyrzykowski[102], H. Zhao[20,136], J. Zorec[137], T. Zwitter[138], L. Balaguer-Nunez[26,24,25], N. Leclerc[8], S. Morgenthaler[139], G. Robert[80], and S. Zucker[74]

(Affiliations can be found after the references)

ABSTRACT

Context. The third *Gaia* Data Release (DR3) provided photometric time series of more than 2 million long-period variable (LPV) candidates. Anticipating the publication of full radial-velocity data planned with Data Release 4, this Focused Product Release (FPR) provides radial-velocity time series for a selection of LPV candidates with high-quality observations.

Aims. We describe the production and content of the *Gaia* catalog of LPV radial-velocity time series, and the methods used to compute the variability parameters published as part of the *Gaia* FPR.

Methods. Starting from the DR3 catalog of LPV candidates, we applied several filters to construct a sample of sources with high-quality radial-velocity measurements. We modeled their radial-velocity and photometric time series to derive their periods and amplitudes, and further refined the sample by requiring compatibility between the radial-velocity period and at least one of the G, G_{BP}, or G_{RP} photometric periods.

Results. The catalog includes radial-velocity time series and variability parameters for 9614 sources in the magnitude range $6 \lesssim G/\text{mag} \lesssim 14$, including a flagged top-quality subsample of 6093 stars whose radial-velocity periods are fully compatible with the values derived from the G, G_{BP}, and G_{RP} photometric time series. The radial-velocity time series contain a mean of 24 measurements per source taken unevenly over a duration of about three years. We identify the great majority of the sources (88%) as genuine LPV candidates, with about half of them showing a pulsation period and the other half displaying a long secondary period. The remaining 12% of the catalog consists of candidate ellipsoidal binaries. Quality checks against radial velocities available in the literature show excellent agreement. We provide some illustrative examples and cautionary remarks.

Conclusions. The publication of radial-velocity time series for almost ten thousand LPV candidates constitutes, by far, the largest such database available to date in the literature. The availability of simultaneous photometric measurements gives a unique added value to the *Gaia* catalog.

Key words. stars: AGB and post-AGB – stars: variables: general – stars: carbon – methods: data analysis – catalogs – techniques: radial velocities

1. Introduction

Evolved stars of low and intermediate mass show various kinds of light variability summarized in the class of long period variables (LPVs). Within this class, there are radially pulsating objects showing small to large light amplitudes in various pulsation modes and with various degrees of periodicity, but also stars whose light variability is due to the presence of a binary or due to eclipses of orbiting dust clouds. For disentangling the various causes for variability in these stars, sometimes even occurring in combinations, contemporaneous monitoring of radial-velocity (RV) variations has proven to be a useful approach.

Early measurements of RV variations in LPVs date back to the 1920s (Joy 1926). It was noted already then that emission and absorption lines in Miras show different kinds of velocity variations. This was supported by several further studies, all using lines in the blue part of the spectrum, but the observed variability pattern did not allow for a conclusive description of the pulsation in these stars (Joy 1954; Reid & Dickinson 1976). A major step forward was achieved by the first monitoring of RV changes in the near-infrared. The landmark paper by Hinkle et al. (1982) revealed the photospheric kinematics for the Mira χ Cyg, allowing for components related to stellar pulsation and to mass outflow to be identified, respectively. Line doubling of high-excitation CO lines was found near light maximum and, together with the appearance of hydrogen emission lines at those phases, interpreted as a trace of shock fronts. Combining velocity data from the violet to the radio regime allowed for a stratigraphy of

a Mira's atmosphere to be constructed out to its circumstellar layers (Wallerstein 1985).

Measurements of velocity amplitudes in Mira variables have played a key role in the discussion on the pulsation mode of these stars (see Wood & Sebo 1996, for a summary). In addition, these observations constrained dynamical models of LPV atmospheres and led to today's understanding that the levitation of the outer layers of the stellar atmosphere driven by pulsation is essential for driving an efficient mass loss during this evolutionary phase (e.g., Höfner & Olofsson 2018).

Since the periods of LPVs can reach values of a few hundred days, obtaining velocity curves at high resolution with a good phase coverage remained challenging. The total number of Miras with such datasets available in the literature is still limited to a few tens (Hinkle et al. 1984; Hinkle & Barnbaum 1996; Lebzelter et al. 1999, 2005a,b; Alvarez et al. 2001). However, this sample covers a wide range in period, metallicity, and chemistry, revealing a consistent pattern in the velocity variations with s-shaped velocity curves in the near-infrared and peak-to-peak velocity amplitudes, depending on the lines used to trace the variation, between 20 and 30 km s^{-1} (Lebzelter & Hinkle 2002; Nowotny et al. 2010).

For physical and observational reasons, most of these studies were done in the 1.6 or 2.2 μm range relying on the first and second overtone lines of CO. These lines trace parts of the stellar atmosphere close to the pulsation driving zone (Nowotny et al. 2010). Within the spectrum, they are located close to the maximum of the spectral energy distribution of Miras and in an area with comparably low line blending and telluric absorption. Atomic lines in the same spectral range show a behavior very similar to the molecular lines (Hinkle & Barnes 1979). Velocity time series from the 4000 Å region show a much less expressed pattern with an amplitude of only 8 km s^{-1}. In the 4000 to 6800 Å range covered in the study of Alvarez et al. (2001), amplitudes around 20 km s^{-1} were measured, and thus the lines in this range compare well with the near-infrared range.

The semiregular variables (SRVs) show significantly smaller light amplitudes and most of them are pulsating in an overtone mode (Wood & Sebo 1996). Consequently, RV amplitudes are expected to be smaller for these stars, which has been confirmed observationally for SRVs with light amplitudes ranging from 0.1 to more than 2.5 mag (Lebzelter 1999; Lebzelter et al. 2005a). For the small amplitude and short period end, velocities of 1–5 km s^{-1} have been reported. Some stars have characteristics somehow between SRVs and Miras, such as W Hya (with a period of 390 days and an amplitude of more than 2 mag in V), and reach velocity amplitudes around 10 km s^{-1}. Semiregular light variability is typically reflected in the velocity variations.

From the point of view of RV variations, the ellipsoidal variables form a group of special interest among the LPVs (Soszyński et al. 2004). From their location in the period-luminosity diagram of LPVs, these stars are also known as sequence E stars. They are close binaries with one object being a red giant and the other one typically being a main sequence star. While there is no visible eclipse, regardless of it being due to an angle of orbital inclination that is too steep or due to the red giant being orders of magnitude brighter than the companion, there is a gravitational distortion of the red giant, which fills the Roche lobe. This produces an elongated shape of the object, and as the star rotates, brightness variations are observed due to this asymmetry.

As a consequence, the light and RV curves of these stars show two light cycles, but only one velocity cycle within one orbital period (Nicholls et al. 2010). Nie & Wood (2014) presented an extensive database of RV curves for 81 ellipsoidal variables. About 20% of these systems show eccentric orbits, a fraction twice as high as derived from light-curve analysis alone (Nie et al. 2017), which stresses the importance of RV data for the understanding of these variables. During their further evolution, the unseen companion will gain mass from the red giant leading to a common envelope system at some point. Ellipsoidal variables are assumed to be precursors of close binary planetary nebulae (Nicholls et al. 2010).

Another group of binaries among the LPVs are the symbiotic stars consisting of a red giant and a degenerated star such as a white dwarf or a neutron star. In the case of D-type symbiotics, the evolved star is a Mira (Hinkle et al. 2013). Radial-velocity changes thus combine pulsation and orbital motion. However, orbital periods of these systems are typically longer than decades (Seaquist & Taylor 1990) and they are therefore difficult to detect even in long velocity time series.

Finally, RV curves play a critical role in the explanation of the mysterious sequence-D stars. These LPVs show radial pulsation in some overtone modes combined with a secondary period that is typically ten times longer. Fundamental mode pulsation has been excluded as the cause of this secondary period (Wood 2000). Binarity and strange modes were suggested as alternative solutions. Interestingly, these long periods seem to form a period-luminosity sequence by themselves.

The origin of this kind of variability remains a matter of debate. Nicholls et al. (2010) showed that sequence-D stars are not ellipsoidals. From an attempt to model the velocity curves of a small sample of sequence-D stars, Hinkle et al. (2002) concluded that binarity is unlikely the cause of the variation because almost all of the objects analyzed show extremely similar values for the orbital parameters K, e, and ω. Soszyński & Udalski (2014) and Soszyński et al. (2021) suggest from a careful analysis of light curves and infrared data that sequence-D variability can be explained by an orbiting dust cloud in combination with a low-mass companion in a close circular orbit. On the other hand, Saio et al. (2015) show that the sequence-D period-luminosity relation agrees with expectations from oscillatory convective modes.

The observation of reliable RV curves of LPVs plays an important role for interpreting various aspects of these stars and their evolution. Observational challenges have limited the collection of large datasets up to now. Considering the variety of objects found among LPVs, the small existing dataset remains insufficient.

Since its second data release, *Gaia* has provided high-quality data for the study of the variability of LPVs, with the publication of photometric time series in the G, G_{BP}, and G_{RP} bands of ~150 000 candidate LPVs in the second data release and over 2 million candidate LPVs in the *Gaia* Data Release 3 (DR3), respectively (Mowlavi et al. 2018; Lebzelter et al. 2023). Moreover, *Gaia* has the unique capability of simultaneously obtaining photometric and spectroscopic measurements owing to its Radial Velocity Spectrometer (RVS), thereby substantially boosting the possibility to investigate stellar variability. This feature was first exploited with the publication of RV time series of Cepheids and RR Lyrae as part of *Gaia* DR3 (Ripepi et al. 2023; Clementini et al. 2023). Here we extend this dataset to an additional 9614 sources that are part of the *Gaia* DR3 catalog of LPV candidates. In Sect. 2, we describe the procedures involved in the construction of this FPR catalog, while we present its content in Sect. 3 and discuss its quality in Sect. 4. In Sect. 5, we give an overview of the catalog, while Sect. 6 is dedicated to a summary and to conclusions.

Table 1. Summary of the steps involved in the construction of the catalog.

Steps	Sub-steps	Selection	# sources
Starting sample	`DR3-LPV`	2nd *Gaia* LPV catalog	1 720 588
	`DR3-RV`	radial velocity published in DR3	33 812 183
	`DR3-LPV-RV`	DR3-LPV and DR3-RV	501 308
Pre-filter	`bright`	DR3-LPV-RV and $G_{RVS} < 12$ mag	249 600
	`vis-periods`	DR3-LPV-RV and `rv_visibility_periods_used` ≥ 12	180 850
	`small-err-rv`	DR3-LPV-RV and $\varepsilon_{V_R} < 0.175 \times$ `rv_amplitude_robust`	224 696
	`filter-0`	bright and vis-periods and small-err-rv	110 654
`processing`		Outliers removal, computation of best-fit model and period of the RV time series.	
Post-filter	`num-outliers`	`filter-0` and number of outliers in the RV time series $N_{V_R}^{out} \leq 1$	105 715
	`no-trend`	`filter-0` and RV model's polynomial degree $D_p = 0$	90 942
	`high-snrv`	`filter-0` and $S/N_{V_R} > 1.5$	58 725
	`filter-1`	no-trend and num-outliers and high-snrv	44 216
	`Prv-low-limit`	`filter-1` and $P_{V_R} > 35$ days	24 118
	`Prv-up-limit`	`filter-1` and $P_{V_R} < \Delta t_{V_R}$	43 621
	`filter-2`	Prv-low-limit and Prv-up-limit	23 523
	`Prv-sim-Pg`	`filter-2` and $P_{V_R} \simeq P_G$	6 392
	`Prv-sim-Pbp`	`filter-2` and $P_{V_R} \simeq P_{G_{BP}}$	5 768
	`Prv-sim-Prp`	`filter-2` and $P_{V_R} \simeq P_{G_{RP}}$	6 534
	`Prv-sim-any-Pph`	`filter-2` and ($P_{V_R} \simeq P_G$ or $P_{V_R} \simeq P_{G_{BP}}$ or $P_{V_R} \simeq P_{G_{RP}}$)	7 551
	`Prv-sim-2Pg`	`filter-2` and $P_{V_R} \simeq 2P_G$	1 701
	`Prv-sim-2Pbp`	`filter-2` and $P_{V_R} \simeq 2P_{G_{BP}}$	1 646
	`Prv-sim-2Prp`	`filter-2` and $P_{V_R} \simeq 2P_{G_{RP}}$	1 703
	`Prv-sim-any-2Pph`	`filter-2` and ($P_{V_R} \simeq 2P_G$ or $P_{V_R} \simeq 2P_{G_{BP}}$ or $P_{V_R} \simeq 2P_{G_{RP}}$)	2 242
	`filter-3`[a]	(Prv-sim-any-Pph or Prv-sim-any-2Pph)	9 614
Additional statistics	`Prv-sim-Pph+2Pph`	`filter-3` and (Prv-sim-any-Pph and Prv-sim-any-2Pph)	179
	`Prv-sim-Pph-only`	`filter-3` and (Prv-sim-any-Pph and (not Prv-sim-any-2Pph))	7 372
	`Prv-sim-2Pph-only`	`filter-3` and ((not Prv-sim-any-Pph) and Prv-sim-any-2Pph)	2 063
	`Prv-sim-all-Pph`	`filter-3` and ($P_{V_R} \simeq P_G$ and $P_{V_R} \simeq P_{G_{BP}}$ and $P_{V_R} \simeq P_{G_{RP}}$)	4 899
	`Prv-sim-all-2Pph`	`filter-3` and ($P_{V_R} \simeq 2P_G$ and $P_{V_R} \simeq 2P_{G_{BP}}$ and $P_{V_R} \simeq 2P_{G_{RP}}$)	1 194
	`top-quality`[b]	(Prv-sim-all-Pph or Prv-sim-all-2Pph)	6 093

Notes. The exact criterion by which two periods are considered "similar" (indicated by $P_1 \simeq P_2$) is described in Sect. 2.3.2. [a]All sources belonging to the `filter-3` subset are published as part of the FPR. [b]The sources belonging to the subset `top-quality` are identified by the flag `flag_rv`.

Several Appendices complete the main body of the text. Appendix A gives additional details on the classification of LPV candidates presented in Sect. 3. Appendix B illustrates cases where the median RV differs significantly from the systemic RV. Appendix C analyzes the impact of the Java bug mentioned in Sect. 2.2 on the LPV results published in DR3. Finally, Appendix D gives some example queries to retrieve the data of the present catalog from the *Gaia* archive.

2. Catalog construction

Our starting dataset is the 2nd *Gaia* catalog of LPV candidates (Lebzelter et al. 2023) published as part of the *Gaia* DR3 (Gaia Collaboration 2023). More precisely, we consider the sources that appear in the table `gaiadr3.vari_long_period_variable` of the *Gaia* Archive. All these sources have their photometric time series already published in DR3. More than 70% of them do not have median RV published in *Gaia* DR3, most likely because they are too faint (see Katz et al. 2023). Therefore, we discard these sources, and focus on the remaining 501 308 LPV candidates having RV data in *Gaia* DR3. Hereinafter we adopt the notation V_R^{DR3} to indicate the median RV published as part of *Gaia*

DR3 (it corresponds to the field `radial_velocity` of the `gaiadr3.gaia_source` table in the *Gaia* archive[1]).

For the construction of the catalog, we proceed in three main steps. To begin with, we employ the quantities derived from the *Gaia* RVS, and published as part of *Gaia* DR3, to refine the input source list to be fed to the processing pipeline. We refer to this first step as "pre-filtering," and describe it in Sect. 2.1.

We then analyze the time series of the selected sources with an updated version of the pipeline used for variability processing in *Gaia* DR3 (Eyer et al. 2023; Lebzelter et al. 2023), as we describe in Sect. 2.2. Both the RV time series and the three photometric time series (in the *Gaia* G, G_{BP}, and G_{RP} bands) undergo this "processing" step, that involves the detection and removal of outlier epochs, the computation of time series statistics, and the determination of the best-fit model.

Lastly, we employ the resulting quantities to further refine the sample of sources for publication. This final step is referred to as "post-filtering," and is described in Sect. 2.3. The filtering conditions and number of selected sources of each step and the corresponding sub-steps are summarized in Table 1.
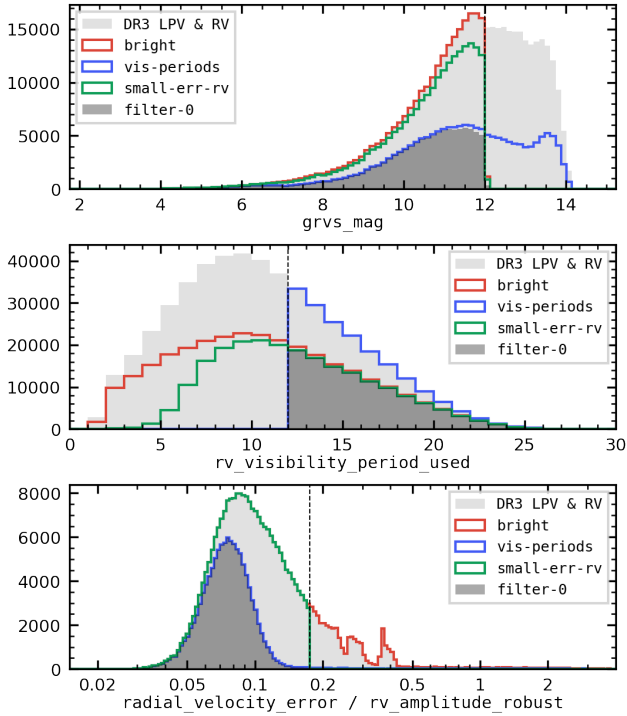
---

[1] https://gea.esac.esa.int/archive/

Fig. 1. Distribution of the pre-filtering parameters G_{RVS} (top panel), `rv_visibility_periods_used` (middle panel), and ε_{V_R}/`rv_amplitude_robust` (bottom panel), with vertical dashed lines indicating the filter limits. Different colors indicate the starting set (DR3-LPV-RV, gray filled histogram) and the individual pre-filters, labeled `bright` (red curve), `vis-periods` (blue curve), and `small-err-rv` (green curve) as in Table 1. The black filled histogram corresponds to the application of all three pre-filters (`filter-0`). We note that 50 349 sources in the starting sample of 501 308 sources lack a published value of G_{RVS} as it would be fainter than 14.1 mag, and that the quantity `rv_amplitude_robust` is not provided for sources with $G_{RVS} > 12$ mag (Sartoretti et al. 2023).

2.1. Pre-filtering

At this stage we aim to limit the sample to the objects with the highest-quality RV measurements by taking advantage of the information available from *Gaia* DR3 (namely in the `gaiadr3.gaia_source` table of the *Gaia* Archive). This is achieved by retaining only sources with large enough RVS flux, a sufficient number of RV measurements, and relatively small RV uncertainty ε_{V_R}. The relevant quantities and corresponding cuts involved in this process are illustrated in Fig. 1.

We begin by applying a filter that excludes the faintest objects in our dataset, using the median value G_{RVS} of the epoch G^t_{RVS} magnitudes (`grvs_mag` in the *Gaia* Archive, see Sartoretti et al. 2023), which are obtained by integration of the RVS epoch spectra. By requiring that $G_{RVS} < 12$ mag we limit our sample to "bright" stars (top panel of Fig. 1) following the distinction made for the DR3 RVS processing (Katz et al. 2023). Almost 250 000 sources meet this criterion. It is worth pointing out that several RV-related quantities published in DR3, such as the median RV and its uncertainty, are computed with different methods depending on whether the sources are brighter or fainter than $G_{RVS} = 12$ mag. Having required that $G_{RVS} < 12$ mag, these quantities are defined unequivocally for all sources in our sample. Namely, the RV is obtained as the median of the single-transit RVs, while the RV error is the uncertainty on the median of the epoch RVs, with a constant offset accounting for a calibration floor contribution (Sartoretti et al. 2022).

Then, we apply a condition to the number of data points in each RV time series. We note that the actual number of RV observations is not necessarily appropriate for this filtering step, as they often come in groups that span a relatively short period of time (often shorter than several days) because of the *Gaia* scanning law (see Eyer et al. 2017). This issue is often overcome through the concept of visibility period, that is a group of transits separated from other such groups by a gap of at least 4 days. The number of visibility periods used in the derivation of radial velocities is a parameter available from DR3, and we employ it to set the condition `rv_visibility_periods_used` ≥ 12. As will be explained in Sect. 2.2, a minimum number of 9 data points is necessary to obtain a time series model, but that may still not be enough for the model to be well-constrained. At the same time, raising too much the threshold would lead to the exclusion of too many sources, as can be appreciated from the middle panel of Fig. 1. We found that a good compromise could be attained by setting the threshold at 12. The condition on visibility periods is fulfilled by about 180 000 sources in our starting dataset.

The uncertainty ε_{V_R} on the median RV is provided in the `radial_velocity_error` data field of the `gaiadr3.gaia_source` table. Instead of setting an absolute upper limit to ε_{V_R}, we rather compare it with the amplitude of the RV curve (`rv_amplitude_robust`) estimated during DR3 processing after outlier removal[2]. We inspected visually the distribution of said parameters for the sources in our sample, before and after the application of the conditions on G_{RVS} and number of visibility periods (bottom panel of Fig. 1), and decided to construct the filtering condition in the form

$$\varepsilon_{V_R} < 0.175 \times \text{rv\_amplitude\_robust}, \qquad (1)$$

which retains almost 225 000 objects from the starting sample. The combination of the three conditions described above results in a pre-filtered sample of 110 654 RV time series, that we input into the variability pipeline.

2.2. Time series processing

Overall, the processing of the RV and photometric time series is performed in a very similar manner as it was done for the photometric time series of LPVs in DR3. Therefore, we briefly summarize the procedure, focusing on the specific parameters for RV analysis and the few differences due to updates to the pipeline, and refer the reader to Lebzelter et al. (2023, and references therein) for more details. The processing operations involve the detection and removal of outliers, followed by the calculation of time series statistics, and the derivation of the best-fit model.

The setup for detecting outliers in the photometric time series are unchanged with respect to DR3. For the RV time series, we exclude epochs with RV that:
- have an uncertainty larger than $5 \, \text{km s}^{-1}$;
- deviate from the median of the time series by more than $100 \, \text{km s}^{-1}$;
- deviate from the median of the time series by more than 10 times the median absolute deviation of the time series.

The choice of these parameters was guided by physical considerations concerning the typical RV amplitude for pulsation in LPVs, that is not expected to exceed several tens of km s^{-1}.

---

[2] We remark that the outlier removal procedure employed during DR3 processing is different from the one adopted in the pipeline used in the present work (Sect. 2.2), therefore the results are not necessarily the same (see Sect. 4.1.1).

However, we quickly realized that the sample contains a non-negligible fraction of high-quality RV curves likely originating from binarity, that we did not want to reject. Therefore, we have launched a few runs of the pipeline using rather permissive values, and tuned them by visual inspection of the distribution of the resulting time series statistics.

To describe the RV and photometric time series we adopt the same kind of mono-periodic model with frequency f_X (where the subscript $X \in \{V_R, G, G_{BP}, G_{RP}\}$ indicates the type of time series), that consists of the sum of a polynomial trend of degree $D_{p,X} \leq 1$ (i.e., no trend or a linear trend) and a Fourier series with up to $N_{h,X} = 3$ components (i.e., up to the second harmonic). Using a notation similar to that of Eyer et al. (2017), the model is defined as

$$y = \sum_{k=1}^{N_{h,X}} A_{k,X} \cos\left[2\pi k f_X(t - t_{0,X}) + \psi_{k,X}\right] + \sum_{i=0}^{D_{p,X}} c_{i,X}(t - t_{0,X})^i \quad (2)$$

where $A_{k,X}$ and $\psi_{k,X}$ are the amplitude and phase of the kth Fourier component, respectively, and $t_{0,X}$ is a reference epoch. To avoid overfitting, the number of Fourier components is limited by the condition $N_{h,X} < \pi/\Delta\phi_{max,X}$ on the maximum phase gap $\Delta\phi_{max,X}$ of the folded time series (cf. Eyer et al. 2017). While this approach is effective in most cases, it may fail for the few time series that end up having large and repeated gaps, and hence lack coverage of specific phase intervals, which makes them especially exposed to overfitting. A similar effect may result if the best-fit period is longer than the duration of the time series (see Sect. 2.3.2).

For each source, the RV time series and the three photometric time series are processed independently of each other. For each time series, we begin by computing the periodogram. It is computed over the frequency range $[7 \times 10^{-4}, 0.1]$ day^{-1}, with an even spacing in frequency of 0.33×10^{-4} day^{-1}. We take the period of the time series to be equal to the value corresponding to the highest peak of the periodogram. After a first determination of the best model in the form given by Eq. (2), we employ a non-linear Levenberg-Marquardt optimization algorithm to improve the result.

It should be clear that the main peak of the periodogram identifies the strongest periodic signal in a time series, which is not necessarily the same as the period of the underlying physical process. In particular, in the case of ellipsoidal red giants, the light curve shows two minima per cycle of possibly different depths, that mimic a variation with a period half as long as the true orbital period. This effect is not present in the RV time series, so that the occurrence is not uncommon of sources whose RV period is twice the photometric period as determined from the periodogram. This will be further discussed in Sect. 3.

We performed safety checks by comparing the newly derived photometric periods with the ones derived from G-band light curves and published in Lebzelter et al. (2023). To our surprise, despite the pipeline setup and sequence of operations being identical, we found that in several cases the results are not exactly the same. We traced this discrepancy to a bug of Java version 8 affecting the nonlinear modeling of the time series, that disappeared during the upgrade to Java 17 performed after the conclusion of *Gaia* DR3 operations. We remark that the deviations are small, and affect a minimal fraction of the sources. We provide a deeper analysis of this issue in Appendix C.

2.3. Post-filtering

We tackle the post-filtering in two successive sub-steps. The first one involves the properties of the cleaned RV curves (i.e., after
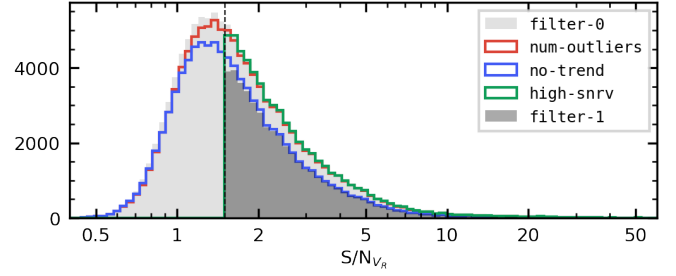
Fig. 2. Distribution of the RV signal-to-noise ratio, S/N_{V_R}, for the pre-filtered sample (gray histogram) and with the conditions involved in the `filter-1` post-filtering step, labeled `num-outliers` (red curve), `no-trend` (blue curve), and `high-snrv` (green curve) and described in Table 1. The dark gray histogram shows the combination of the three conditions. The vertical dashed line indicates the $SN_{V_R} = 1.5$ threshold.

outlier removal) revealed by the time series statistics as well as the parameters of the best-fit model, with the exception of the frequency. The latter quantity is considered in the second sub-step, aimed at excluding the objects whose best RV period (P_{V_R}) is uncertain. We construct a filtering criterion by comparing with each other the periods derived from the RV and photometric time series.

2.3.1. Selection on RV time series properties

To begin with, we assess the impact of outlier removal on the RV time series of the pre-filtered sample, and examine the number $N_{V_R}^{out} = N_{V_R}^{raw} - N_{V_R}$ of rejected epochs as a parameter for constructing an additional filter, where N_{V_R} is the number of epochs in the cleaned RV time series and $N_{V_R}^{raw}$ is the number of valid measurements in the original RV time series (i.e., excluding NaN values). The majority of the time series (about 88%) are unaffected, while $N_{V_R}^{out} = 1$ for about 7.5% of the sources, and the remaining 4.5% of RV curves had at least two rejected epochs. Visual inspection of time series and folded RV curves with varying number of outliers reveals satisfying results for $N_{V_R}^{out} \leq 1$, as well as a rapid degradation with increasing $N_{V_R}^{out}$. We therefore restrict our sample by requiring that no more than one RV epoch is excluded during outliers removal, a condition that selects 105 715 sources from the pre-filtered sample.

One of the properties that we found to be often associated with low-quality fits is the adoption of a first-degree polynomial in the RV curve model. About 18% of the RV time series in the pre-filtered sample are modeled this way. This is done automatically by the variability pipeline when the inclusion of a linear trend results in a better fit with the underlying data. However, the combination of a relatively small number of epochs and characteristic time scales comparable with the duration of the time series make this approach poorly suited for LPVs. In contrast with variable objects with shorter periods, for which the inclusion of a linear trend can significantly improve the characterization of the time series, in the case of LPVs it tends to pick up the signal associated with long periods, while in some cases erroneously detrends RV curves with poor phase coverage. We therefore retain only the time series modeled without a linear trend, which correspond to 90 952 sources in the pre-filtered dataset.

Finally, we apply a threshold to the signal-to-noise ratio of the cleaned RV time series at $S/N_{V_R} = 1.5$ (see Fig. 2). In the pre-filtered sample there are 58 725 sources above that limit. By combining the three conditions described above, we reduce the pre-filtered sample down to 44 216 sources.
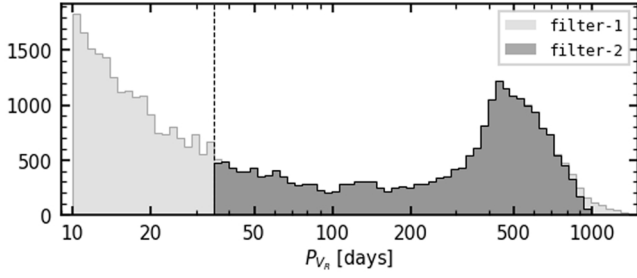
Fig. 3. Distribution of the RV periods of the sample after applying the post-filtering conditions `filter-1` (gray histogram, see Table 1) and after applying the conditions on the periods themselves (`filter-2`, black histogram). The vertical dashed line indicates the lower period limit at 35 days.

2.3.2. Selection on the periods

We follow the approach described in Lebzelter et al. (2023) to bound the range of RV periods, setting a fixed lower limit to 35 days, and excluding the cases in which P_{V_R} is longer than the duration Δt_{V_R} of the RV time series. We recall that the adoption of such a lower limit by Lebzelter et al. (2023) for the photometric time series was aimed at minimizing the contamination from spurious signals. These conditions further reduce our dataset to 23 523 sources. The distribution of RV periods before and after the application of these conditions is shown in Fig. 3.

Finally, we examine how close the RV period is to the periods obtained from modeling the photometric time series. To do so, we follow the method described by Mowlavi et al. (2023, their Sect. 4.1). We quantify the similarity between the RV period P_{V_R} and a photometric period P_{ph} (either P_G, $P_{G_{BP}}$, or $P_{G_{RP}}$) by the quantity

$$r_{V_R,\text{ph}} = \frac{|P_{V_R} - P_{\text{ph}}|}{P_{V_R}} \frac{\Delta t_{V_R}}{P_{V_R}}, \qquad (3)$$

which represents the maximum phase deviation a signal with period P_{ph} can accumulate with respect to P_{V_R} during the observation duration Δt_{V_R}. In order to better understand the meaning of Eq. (3), we note that $\delta\phi_{V_R,\text{ph}} = |P_{V_R} - P_{\text{ph}}|/P_{V_R}$ is the difference between the two periods normalized to P_{V_R}, whereas $\Delta t_{V_R}/P_{V_R} = n_{V_R}^{\text{cyc}}$ is the number of cycles with period P_{V_R} covered by the RV time series. Let us assume that the RV and photometric curves are in phase at the very beginning of the RV time series. Unless the two periods are identical, after one RV cycle P_{V_R} the two curves show a phase offset that is exactly equal to $\delta\phi_{V_R,\text{ph}}$. After two RV cycles the phase deviation is twice as large, and so on. At the very end of the RV time series, the phase offset is $\delta\phi_{V_R,\text{ph}} \times n_{V_R}^{\text{cyc}} = r_{V_R,\text{ph}}$. It is easy to see that this is also the maximum possible phase deviation for given P_{V_R}, P_{ph}, and Δt_{V_R}.

From Eq. (3) it is clear that $r_{V_R,\text{ph}}$ is defined asymmetrically, and that $r_{V_R,\text{ph}} \neq r_{\text{ph},V_R}$. However, the closer the values of the period being compared, and the smaller the asymmetry is. The distribution of the values of $r_{V_R,\text{ph}}$ and r_{ph,V_R} for all three photometric periods show that the two quantities rapidly converge when they are smaller than unity. We thus construct our "period similarity" condition in the form

$$r_{V_R,\text{ph}}^{\max} = r_{\text{ph},V_R}^{\max} = \max(r_{V_R,\text{ph}}, r_{\text{ph},V_R}) < 1. \qquad (4)$$

Figure 4 shows the distribution of $r_{V_R,\text{ph}}$ versus the number of cycles $n_{V_R}^{\text{cyc}}$ covered by the RV time series. A large number of sources accumulate along two slanted stripes in the diagram, that
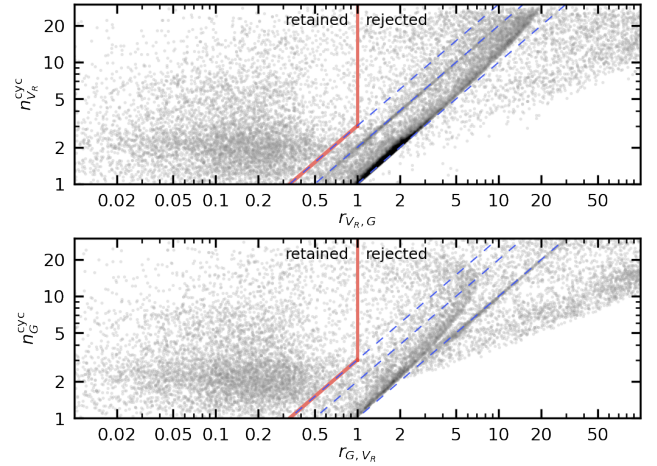
Fig. 4. Number of observed cycles versus phase deviation at the last cycle, comparing the RV and G-band curve models for the sample post-filtered down to `filter-2` (see Table 1). More precisely, the top panel shows the phase deviation $r_{V_R,G}$ with respect to the last RV cycle and the number $n_{V_R}^{\text{cyc}}$ of RV cycles, while the same quantities are referred to the G-band time series in the bottom panel. In each panel, the thick red line marks the upper limit to the phase difference employed in post-filtering (Eq. (4)), while the dashed lines indicate $n^{\text{cyc}}/r = 3$, 2, and 1. Data points to the right of the thick red line are rejected. A similar picture emerges when the G_{BP} or G_{RP} time series are considered in place of G.

are only partially rejected by the condition defined by Eq. (4). These stripes correspond to $r_{V_R,\text{ph}} \simeq n_{V_R}^{\text{cyc}}$ and $r_{V_R,\text{ph}} \simeq 2n_{V_R}^{\text{cyc}}$, respectively. It is easy to show that the former case corresponds to $P_{\text{ph}} \simeq 2P_{V_R}$ or $P_{\text{ph}} \ll P_{V_R}$, and the latter to $P_{\text{ph}} \simeq 3P_{V_R}$ (or $P_{\text{ph}} \simeq -P_{V_R}$, which is not possible as periods have positive values). They clearly indicate situations of incompatibility between pairs of periods, and should be excluded. To do so, we require that

$$\delta P_{V_R,\text{ph}} = \delta P_{\text{ph},V_R} = \frac{|P_{V_R} - P_{\text{ph}}|}{\min(P_{V_R}, P_{\text{ph}})} < \frac{1}{3}, \qquad (5)$$

which completes our condition on period similarity (Eq. (4)). The combination of Eqs. (4) and (5) for the G-band and RV periods is equivalent to taking only the data points that are on the left of the red lines in both panels of Fig. 4. Therefore, our definition of period similarity is given by

$$P_{V_R} \simeq P_{\text{ph}} \iff \left[\left(r_{V_R,\text{ph}}^{\max} < 1\right) \text{ and } \left(\delta P_{V_R,\text{ph}} < \frac{1}{3}\right)\right]. \qquad (6)$$

It should be noted that, if a source displays variability due to binarity, the main peak in the periodogram of any one of its photometric time series can be half of the true orbital period, and hence of the RV period. Therefore, the requirement that $P_{V_R} \simeq P_{\text{ph}}$ could lead to exclude these sources. In order to avoid this, we are also interested in using a requirement in the form $P_{V_R} \simeq 2P_{\text{ph}}$, which means that any occurrence of P_{ph} in Eq. (6) is replaced by $2P_{\text{ph}}$. Our final requirement is therefore

$$(P_{V_R} \simeq P_{\text{ph}} \text{ or } P_{V_R} \simeq 2P_{\text{ph}}) \text{ for any } P_{\text{ph}} \in \{P_G, P_{G_{BP}}, P_{G_{RP}}\}. \quad (7)$$

Figure 5 displays the period distributions for the sources displaying compatibility between the RV period and one or more photometric periods according to Eq. (6). Comparing with the distribution in Fig. 3 we note the effectiveness of this type of
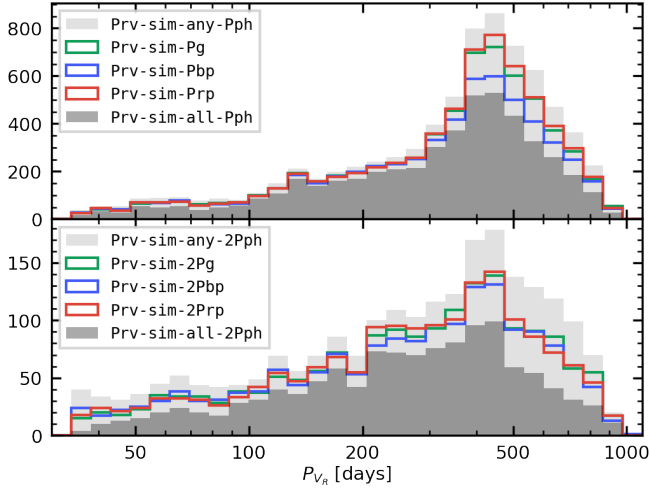
Fig. 5. Period distribution, after applying `filter-2`, for the subsets with RV period similar to one or more photometric periods (top panel), or twice of it (bottom panel). The light gray and dark gray histograms represent the sets in which the RV period is similar to at least one photometric period or to all of them, respectively. The colored curves represent the sets in which the RV period is similar to P_G (green), $P_{G_{BP}}$ (blue), or $P_{G_{RP}}$ (red).

selection at rejecting periods shorter than ∼200 days, where a higher rate of occurrence of spurious frequencies is expected (see Holl et al. 2023). This is also true when the RV period is compared with twice one of the photometric periods.

We note that the number of sources with $P_{V_R} \simeq P_{G_{RP}}$ is slightly higher than that with $P_{V_R} \simeq P_G$, which in turn are more numerous than the objects with $P_{V_R} \simeq P_{G_{BP}}$. The same trend is present when the comparison is made against twice the photometric period, but is less pronounced. This could be indicative of a color-dependence of the photometric variability features, typical of pulsating stars. The fact that this feature becomes less conspicuous when $P_{V_R} \simeq 2P_{ph}$ would support the interpretation that the variability of these sources is extrinsic and associated with binarity.

We construct the final filter by applying the condition defined by Eq. (7) inclusively to the three photometric periods, that is, we require that the RV period is similar to at least one of them. The final dataset consists of 9614 sources.

2.4. Top-quality sample

By the criteria defined above, we identified a subset of the FPR sample consisting of sources displaying a high degree of compatibility between the RV and photometric variability. Namely, there are 6093 sources whose RV period is consistent with each one of the three periods derived from the photometric time series. This means that these sources fullfil the condition

$$(P_{V_R} \simeq P_{ph} \text{ or } P_{V_R} \simeq 2\,P_{ph}) \text{ for all } P_{ph} \in \{P_G, P_{G_{BP}}, P_{G_{RP}}\}. \quad (8)$$

These sources are identified by the field `flag_rv =True` in the *Gaia* Archive (see Sect. 2.5), and form a subset that we dubbed the "top-quality sample" (TQS).

Such a high consistency between the RV and photometric periods is a strong indication that a signal originating from the same physical process is being detected in all four time series, with two important consequences. On the one hand, these sources can be used to investigate a given type of variability in its different aspects (physical motion, changes in brightness and

color) with a good degree of confidence that they all trace the same phenomenon. Given the multi-periodic nature of LPVs, this is far from trivial. On the other hand, there is a comparatively small probability that the periodic signal picked up by the variability processing pipeline is spurious.

Other than this self-consistency, the sources in the TQS have on the average the same properties as the remaining FPR sources, with the only exception that they include a larger fraction of sources identified as binary variables (see Sect. 3). Indeed, binary candidates are assigned to the TQS with a higher frequency (∼80%) than other sources in the FPR (∼60%). These trends are likely to be attributed to the fact that, owing to its geometric nature, binary-induced variability shows smooth variations compared with the pulsation of LPVs, known to display irregularities.

We compared the TQS and the other FPR sources in terms of the distributions of several quantities from *Gaia* DR3. The sources in the former set display slightly better astrometry (smaller errors in sky coordinates and proper motions), but the two sets are equivalent in terms of relative parallax uncertainty. These properties are to be attributed to a slightly higher number of visibility periods used in the astrometric solution. The uncertainty associated with both photometric and RV measurements as reported in the *Gaia* DR3 source table is slightly higher for the TQS sources, which simply reflects the fact that they tend to exhibit larger variability amplitudes. The TQS sources follow essentially the same brightness distributions of all other FPR sources in all three *Gaia* bands, except they are slightly brighter in G_{BP} and fainter in G_{RP}, and as a result they appear to have a slightly bluer color which reflects the larger fraction of binaries in the TQS compared to other FPR sources, see Sect. 5. The two sets do not show any particular difference in their RV distributions.

Some more significant differences between the two sets are found in terms of the variability parameters (Fig. 6). The requirement of period consistency effectively excludes from the TQS sources with short G-band periods (due to the lower limit at 35 days on P_{V_R}). Moreover, the amplitude distribution of TQS sources tends to be skewed toward slightly larger values compared with the full FPR sample, as they are associated with a higher signal-to-noise ratio (and hence a higher chance of picking the same periodicity in RV and in photometry). Other amplitude-related parameters (such as standard deviation, interquartile range, or Stetson variability index) show similar trends. For similar reasons the TQS sources display smaller Abbe values (Mowlavi 2014; Mowlavi et al. 2017) than other sources in all *Gaia* bands, indicating smoother light curves. However, such a difference is not present for the Abbe value computed for the RV time series.

Finally, we inspected the mean value of the uncertainties associated with single epochs (either of the RV or photometric time series) and the mean of the absolute residuals of the time series models, and found differences between the distribution associated with the TQS and with other sources that are consistent with the different amplitude distributions. Therefore, we remark that we consider this subsample to be of superior quality within the FPR because of its content of coherent physical information rather than in terms of actual quality of measurements.

2.5. Data fields

The present catalog follows the same scheme as the 2nd *Gaia* catalog of LPV candidates (Lebzelter et al. 2023), and has
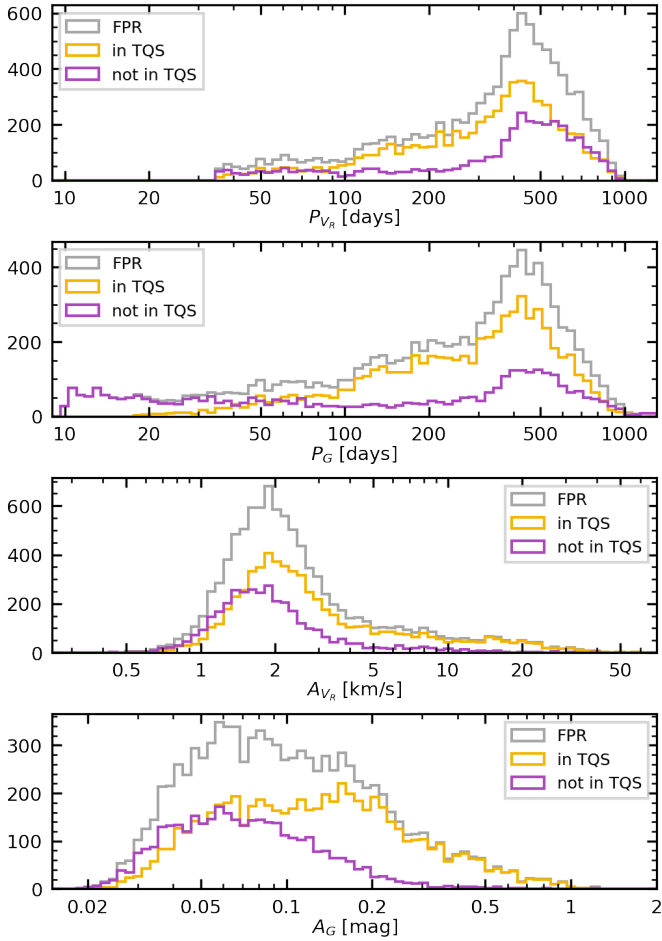
Fig. 6. Distribution of the RV and *G*-band variability parameters for the all the FPR sources (gray curves), and distinguishing between whether they are in the top-quality sample (TQS, orange curves) or not (purple curves). Panels from top to bottom show the distributions of RV periods (P_{V_R}), *G*-band periods (P_G), RV semi-amplitudes (A_{V_R}) and *G*-band semi-amplitudes (A_G).

therefore the same data fields, with the addition of the corresponding fields for the RV variability. More precisely, the fields `solution_id`, `source_id`, `median_delta_wl_rp`, and `isCstar` are left unchanged, while the fields `frequency`, `frequency_error`, and `amplitude` have their values replaced with the newly derived parameters of the best-fit model for the *G*-band time series (see Sect. 2.2 for the reason of the updated values). Finally, the following four data fields are added.

FREQUENCY_RV: *Frequency of the RV curve (double, Frequency [day⁻¹]).* This field provides the frequency determined from the RV time series.

FREQUENCY_ERROR_RV: *Uncertainty on the RV frequency (float, Frequency [day⁻¹]).* This field provides the uncertainty on the frequency of the RV time series.

AMPLITUDE_RV: *Amplitude of the RV curve (float, Velocity [km s⁻¹]).* This field gives the half peak-to-peak amplitude (semi-amplitude in km s⁻¹, based on the best-fit model; see Sect. 2.2).

FLAG_RV: *Flag identifying the top-quality subsample (boolean).* This field identifies the sources whose RV period
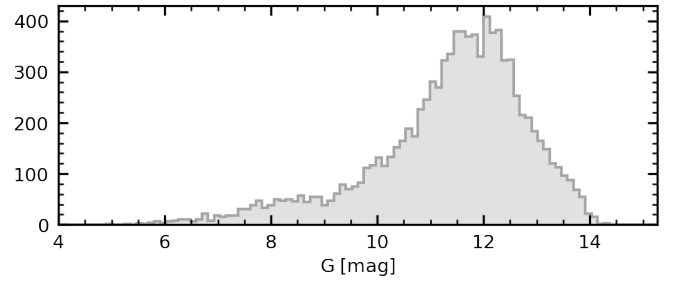
Fig. 7. *G*-band brightness distribution of the FPR sample.

is fully compatible with all three photometric periods (see Sect. 2.4).

The full RV time series for all sources in this FPR are available for download from the table `gaiafpr.vari_epoch_radial_velocity` in the *Gaia* archive, while the statistics for the cleaned RV time series are provided in the table `gaiafpr.vari_rad_vel_statistics`, following the same scheme adopted in *Gaia* DR3 for the RV time series of Cepheids and RR Lyrae (cf. Ripepi et al. 2023; Clementini et al. 2023). In Appendix D we provide some instructions on how to retrieve the FPR data.

Hereinafter we adopt the notation A_G and A_{V_R} to indicate the quantities `amplitude` and `amplitude_rv` published in this FPR, corresponding to the semi-amplitude of the fundamental component of the best-fit Fourier model of the *G*-band and RV time series, respectively. We note that roughly half of the *G*-band time series and more than 80% of the RV time series have been modeled with a single-component Fourier series, so the published value is exactly the semi-amplitude of the model. The remaining time series have harmonic components whose amplitude is typically much smaller than that of the fundamental component, so that the semi-amplitude of the latter is still representative to the semi-amplitude of the full Fourier model. Therefore, for simplicity, we often refer to A_G (not to be confused with the *G*-band extinction) and A_{V_R} as "semi-amplitude of the time series models", whereas their formal meaning should be clear.

3. Catalog content

This FPR provides epoch RVs for 9614 candidate LPVs, of which the *G*, G_{BP} and G_{RP} time series are available in DR3 as part of the second *Gaia* catalog of LPV candidates. Figure 7 shows the *G*-band distribution of these sources, which cover the range $6 \lesssim G/\mathrm{mag} \lesssim 14$. The RV time series have between 12 and 90 measurements, with an average of 24 epochs, unevenly sampling a time interval of about 3 yr. More precisely, the RV time series have a mean duration of 905 days, spanning a range between about 500 and 1000 days, but with a distribution skewed toward longer durations (typically $\gtrsim$800 days). The number of epochs in the RV and *G*-band time series, as well as the number of visibility periods adopted for deriving median RVs in DR3 (see Sect. 2.1), are illustrated in Fig. 8. We note that these numbers correspond to the cleaned time series, that is after outlier removal, and are the same values given in the *Gaia* archive summary tables, whereas the published *Gaia* light curves include the outlier epochs as well (flagged to indicate whether they have been rejected by the variability pipeline).
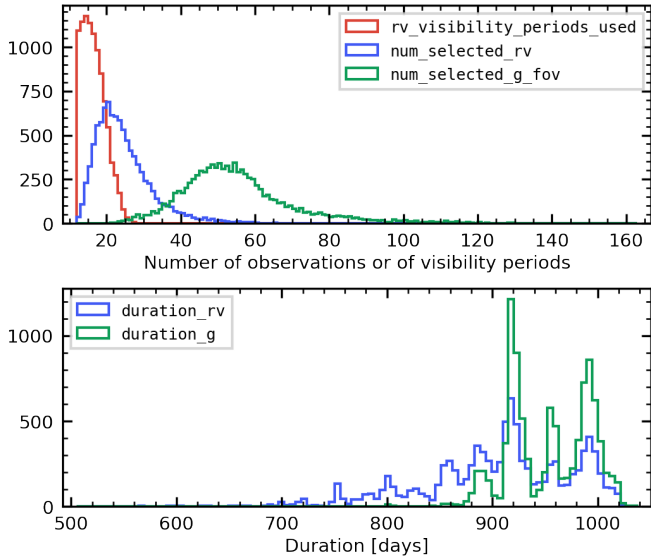
Fig. 8. Distribution of the number of observations (top) and duration (bottom) of the RV and G-band time series of the FPR sources. The red line indicates the number visibility periods used to derive the median RV published in DR3 (a single visibility period may contain multiple epochs, see Sect. 2.1). The blue and green lines indicate the number of measurements in the cleaned RV and G-band time series (top) or their duration (bottom).

Besides the filtering steps, in Table 1 we provide a summary of a few interesting subsets of the final sample, obtained by further comparing the periods derived from the RV and photometric time series. Following the criteria defined in Sect. 2.3.1, the period comparison for a given source has three possible outcomes: (1) $P_{V_R} \simeq P_{ph}$, (2) $P_{V_R} \simeq 2 P_{ph}$, or (3) $P_{V_R} \not\simeq P_{ph}$ and $P_{V_R} \not\simeq 2 P_{ph}$. These conditions are not necessarily the same for each of the three photometric periods. For instance, a source might be such that $P_{V_R} \simeq P_G$, $P_{V_R} \simeq 2 P_{G_{RP}}$, and at the same time $P_{V_R} \not\simeq P_{G_{BP}}$ and $P_{V_R} \not\simeq 2 P_{G_{BP}}$. However, due to the filters we applied, either conditions (1) or (2) must be verified for at least one of $P_{ph} \in \{P_G, P_{G_{BP}}, P_{G_{RP}}\}$ for all sources published in the FPR. These conditions allow us to distinguish between three types of sources:

– only 1:1 compatibility: $P_{V_R} \simeq P_{ph}$ for at least one photometric period, but none of the other periods meets the condition $P_{V_R} \simeq 2 P_{ph}$;
– only 2:1 compatibility: $P_{V_R} \simeq 2 P_{ph}$ for at least one photometric period, but none of the other periods meets the condition $P_{V_R} \simeq P_{ph}$;
– "mixed" compatibility: $P_{V_R} \simeq P_{ph}$ for at least one photometric period, and $P_{V_R} \simeq 2 P_{ph}$ for at least one of the other photometric periods (as in the example above).

The majority of the FPR sources fall in the first category, consisting of 7372 sources (about 77%). There is no direct indication that the variability of these sources results from binarity, as none of the photometric period is close to a 2:1 ratio with respect to the RV period. Of course, this does not prove that they are not binary variables. However, it is reasonable to assume that most of these sources are probably pulsating stars, at least for the purpose of assessing the relative fractions of these types of variables in the FPR. Similarly, the 2063 sources (about 21%) belonging to the second category in the list above are probably binary variables. More precisely, as they are selected among bright red giants, these sources are most likely ellipsoidal

variables (ELL), and will be referred as such hereinafter. Further evidence supporting this statement will be provided in Sect. 5.

Finally, there exist 179 sources such that their RV period is consistent with one or two of the photometric periods, and twice the value of the remaining ones. We examined visually the time series of a random sample of these sources, and found that the cleaned light curves often show large phase gaps when folded with the RV period. Figure 9 shows a clear example with a lack of data near minimum light. All time series show a similar trend, and the best-fit model to the RV, G, and G_{BP} time series is visually convincing, yet the G_{RP} has a best-fit model with half the period found in the other time series. A similar situation can arise when the time series covers a small number of RV cycles, so it becomes difficult to constrain the period precisely. It is clear that this kind of mixed consistency between photometric and RV periods has artificial causes and, in principle, none of the two periods can be confidently taken to be the correct one. It is not possible to make any inference on the nature of these sources based only on their periods. However, the fact that they represent less than 2% of the FPR is encouraging.

These three categories give us a general idea of the fractions of ellipsoidal and pulsating variables in the FPR based on weak conditions on period consistency. Stronger conditions can be imposed by restricting the analysis to the TQS, which includes 4899 probable pulsators ($P_{V_R} \simeq P_{ph}$ for each P_{ph}) and 1194 probable ELL ($P_{V_R} \simeq 2 P_{ph}$ for each P_{ph}). The two kinds of sources make up for about 80% and 20% of the TQS, respectively. These percentages are fully compatible with the values found in the previous paragraph.

3.1. Candidate ellipsoidal variables

Classifying the types of variables discussed above based only on the ratio between the RV period and photometric periods is not necessarily a good approach. In particular, it might be inappropriate if one or more of the cleaned time series end up having a small number of measurements, so that the corresponding period is poorly constrained. Therefore, we use the semi-amplitude A_{V_R} of the RV time series model, and the corresponding value A_G for the G-band model, to perform a deeper analysis. In doing so we consider only the TQS in the rest of this section, so to obtain as clean a picture as possible.

The G-band and RV semi-amplitudes derived for the sources in this sample are displayed in Fig. 10. Two groups are clearly separated in this diagram (in either panel). The first group shows G-band variations over a wide range ($0.02 \lesssim A_G/\mathrm{mag} < 2$), but is limited to relatively small RV amplitudes ($A_{V_R} \lesssim 10\,\mathrm{km\,s^{-1}}$, with only a few exceptions). The second group is characterized by large RV variations ($A_{V_R} \gtrsim 5\,\mathrm{km\,s^{-1}}$) and relatively small light amplitudes ($A_G \lesssim 0.2$ mag). We can readily interpret the former group as consisting of pulsating stars, whose brightness changes can become very large (owing to strong absorption by molecules that form efficiently in the expanding phase of the cycle, Reid & Goldston 2002) while they can hardly attain pulsation velocities larger than ~20 km s⁻¹ (Nowotny et al. 2010). In contrast, orbital velocities can easily exceed that value in binaries, but their G-band variations do not exceed a few tenths of magnitude. This interpretation is supported by the fact that the vast majority of sources with $P_{V_R} \simeq P_{ph}$ are found in the former group (black points in the top panel of Fig. 10), whereas most sources in the latter group have $P_{V_R} \simeq 2 P_{ph}$ (black points in the bottom panel), although some contamination is present in both.
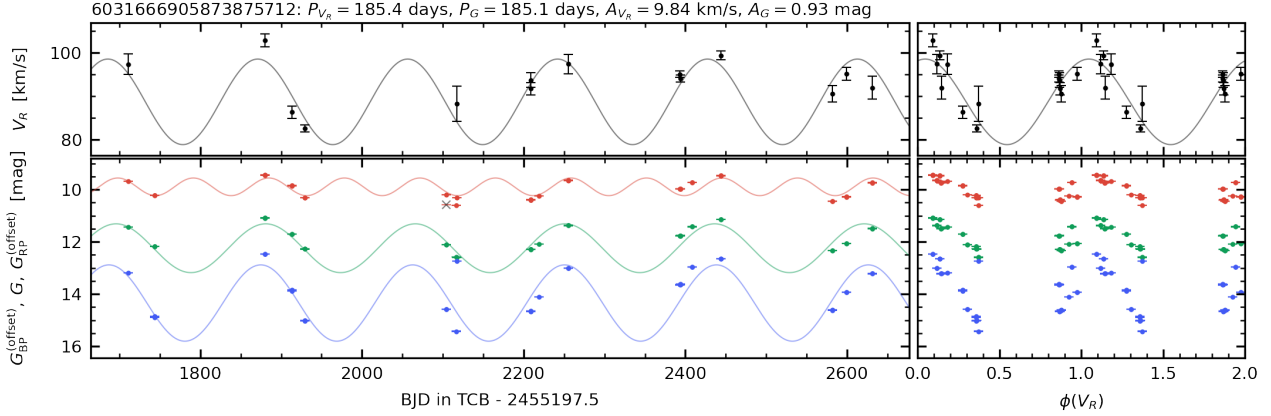
Fig. 9. Example time series for a source with mixed consistency between the photometric and RV time series. This source has $P_G \simeq P_{G_{BP}} \simeq P_{V_R}$, while $P_{G_{RP}} \simeq 0.5 P_{V_R}$. The panels in the top row show the RV data and model, while the photometric data and corresponding models are shown in the panels in the bottom row (in red, green, and blue for the G_{RP}, G, and G_{BP} bands, respectively). For visualization purposes, an arbitrary offset is applied to the G_{RP} and G_{BP} time series. The *Gaia* DR3 source ID of this object is indicated in the title, together with the period and semi-amplitude of the best-fit G-band and RV time series models. The panels on the right show the four time series folded by the RV period.

more confidently. For instance, we note that there are several sources in Fig. 10 (top panel) having $20 \lesssim A_{V_R}/\mathrm{km\,s^{-1}} \lesssim 50$ that are unlikely to be pulsators, but would be classified as such based only on the ratio between their RV period and photometric periods. Further evidence in support of this approach is given in Sect. 5.

At the same time, there are sources with $P_{V_R} \simeq 2\,P_{ph}$ that end up outside of the region associated with ELLs in Fig. 10 (bottom panel). While there is no a-priori reason why they should not be binaries, their distribution in this diagram is consistent with that of the sources with $P_{V_R} \simeq P_{ph}$ (top panel of Fig. 10), suggesting that they display the same kind of variability. Visual inspection of their RV and light curves indicates that the 2:1 period ratio is probably artificial. This most likely results from the fact that many stars in this part of the diagram are semi-regular variables with multi-periodic variability, not necessarily well-described by a single-period model.

3.2. Candidate LPVs: Pulsation and long secondary periods

For pulsating LPV stars the photometric amplitude of variability increases with the pulsation period, a trend that can be identified in the left and top sides of the diagram in Fig. 11 ($20 \lesssim P_G/\mathrm{days} \lesssim 500$). A second group of stars can be seen in the bottom-right corner of the diagram, characterized by long periods and comparatively small G-band amplitudes. While it is likely that these sources are also pulsating LPVs, the dominant period picked up by the variability processing pipeline is probably a long secondary period (LSP; see, e.g., Fig. 16 of Lebzelter et al. 2023).

We tentatively identify LSPs in the period amplitude diagram by the condition

$$A_G < 0.35\,\mathrm{mag} \quad \text{and} \quad \frac{A_G}{\mathrm{mag}} < 10^{-7} \cdot \left(\frac{P_G}{\mathrm{days}}\right)^{2.5}. \qquad (10)$$

We remark that this criterion and the resulting classification is necessarily approximate, and is adopted only for the purpose of characterizing the content of the FPR. In principle, a knowledge of the absolute brightness is required in order to accurately differentiate between pulsation periods and LSPs so that one can construct a period-luminosity diagram. This cannot be done for the entire FPR sample because of the relatively large

Fig. 10. Comparison between the semi-amplitudes A_G, A_{V_R} of the best-fit models of the G-band and RV time series for the TQS (light gray symbols in the background). The darker symbols indicate sources whose RV period is consistent with the photometric periods in a 1:1 ratio (top panel) or in a 2:1 ratio (bottom panel). The dashed red line corresponds to Eq. (9), and the size of each sample is indicated in the legend.

Based on the distributions displayed in Fig. 10, we identify ELL candidates by the condition

$$A_G < 0.3\,\mathrm{mag} \quad \text{and} \quad \frac{A_G}{\mathrm{mag}} < 2.5 \times 10^{-3} \cdot \left(\frac{A_{V_R}}{\mathrm{km\,s^{-1}}}\right)^2 \qquad (9)$$

which corresponds to the region in Fig. 10 below and to the right of the dashed red line. We prefer Eq. (9) to a condition based on the RV-to-photometric period ratios as it is based on physical arguments, and allows us to identify ellipsoidal variables

Fig. 11. Similar to Fig. 10, but comparing the period P_G and semi-amplitude A_G derived from G-band time series for TQS sources that are probable LPVs (not identified as ellipsoidal variable candidates). The dashed red line corresponds to Eq. (10), and the size of each sample is indicated in the legend.

uncertainties affecting the parallaxes of a number of sources, even though more than one third of them have relative parallax errors better than 10% (they are examined in more detail in Sect. 5.2).

We note that certain LPVs, such as some relatively massive AGB stars or red supergiants, have long pulsation periods and relatively small photometric amplitude, and their variability could thus mimic the LSP variation. These stars overlap with LSPs in the period-amplitude diagram (see e.g., Fig. 18 of Lebzelter et al. 2019), and might be misclassified by our criteria. However, given the rarity of these stars, this has a negligible impact in our quantification of the relative fraction of variability types in the FPR. They are further examined in Sect. 5.2.

We also note that pulsation and LSP usually coexist in LPVs that exhibit the LSP phenomenon. The knowledge of multiperiodicity and the amplitude associated with each period can then improve the classification derived from the period-amplitude diagram. However, as only one period is extracted from each time series in this FPR, we consider our selection appropriate enough for our purposes. Additional evidence to support this is provided in Appendix A.

3.3. Summary

Table 2 provides a summary of the number of sources identified as LPVs (either showing pulsation or LSP variability) or as ellipsoidal variables. The TQS consists by about 14% of ELLs, by about 38% of pulsating LPVs, and by about 48% of LPVs for which we detect LSP-like variability. If the conditions defined by Eqs. (9) and (10) are extended to the entire FPR sample, these percentages become about 12%, 42%, and 46%, respectively.

The numbers in Table 2 also show that, overall, the TQS includes roughly 60% of the LPVs (regardless of whether they

show pulsation or LSPs), and 80% of the ELL candidates, indicating that the latter enter more easily in the TQS. This is consistent with the fact that the geometric origin of the variability of the latter results in much smoother and regular variations than those presented by the former, increasing the chances of consistency between photometric and RV periods.

It is worth noting that the classification we adopted is generally consistent with ratios between RV and photometric periods. Among the sources in the TQS, 89% of the ones identified as ELLs show a 2:1 ratio between P_{V_R} and P_{ph}, whereas about 86% and 96% of the LPV candidates showing pulsation or LSP variability, respectively, are consistent with a 1:1 ratio. The corresponding percentages concerning the full FPR are roughly consistent with these values, although some differences arise due to the occurrence of sources showing mixed consistency. Some example time series of ELL, pulsating LPVs, and LSPs from the TQS are displayed in Figs. 12–14, respectively.

4. Catalog quality

In the present section we check the quality of the FPR catalog. We examine the average RVs in Sect. 4.1, then we consider the RV variability in Sect. 4.2, and finally we compare with literature data in Sect. 4.3.

4.1. Radial velocity estimates

4.1.1. Median radial velocity: Consistency with *Gaia* DR3

The values V_R^{DR3} of RV published in *Gaia* DR3 are derived with two different methods depending on the source brightness. In particular, if a source has G_{RVS} brighter than 12 mag (as is the case for all FPR sources), V_R^{DR3} is computed as median values over the RV time series (Katz et al. 2023). It is worth comparing these values with the median values $\langle V_R \rangle$ resulting from the variability processing pipeline employed for the FPR sample (the field median_rv in the table gaiafpr.vari_rad_vel_statistics). Some small deviations are expected between the two values because of slight differences in terms of data and methods. Indeed, the definition of median value adopted by the *Gaia* variability processing unit and by the *Gaia* spectroscopic data processing unit are slightly different. To compute the median value of a dataset consisting of an odd number of values, both methods sort the data and take the middle value. In contrast, if the number of values is even, after sorting the variability processing pipeline takes the smaller of the two middle values, whereas the spectroscopic processing pipeline takes the mean of the two middle values. This means that the former method systematically results in smaller median values than the latter. Moreover, compared to the RV time series used to compute V_R^{DR3}, the time series published in this FPR can have one fewer epoch as a result of outlier removal during the variability processing (Sect. 2.2).

In summary, we find differences between $\langle V_R \rangle$ and V_R^{DR3} for 305 sources whose RV time series had one outlier removed, and for 4672 sources whose time series had no outlier removed, but have an even number of epochs. These sources are displayed in Fig. 15, where the absolute value of the difference between $\langle V_R \rangle$ and V_R^{DR3}, normalized to the uncertainty ε_{V_R} on V_R^{DR3}, is shown against the number of epochs in the FPR time series. For the majority of the sources we find $|\langle V_R \rangle - V_R^{\mathrm{DR3}}| \simeq 0.2\,\varepsilon_{V_R}$, whereas only 104 sources have a difference between $\langle V_R \rangle$ and V_R^{DR3} that exceeds ε_{V_R}. These sources are usually characterized

Table 2. Number of sources assigned to different types (LPVs showing pulsation or LSP variability, or ellipsoidal variables) in the TQS and in the full FPR, distinguishing between the stars showing compatibility between the RV and photometric periods in a 1:1 or 2:1 ratio.

		TQS			FPR		
		All $P_{V_R} \simeq P_{\rm ph}$	All $P_{V_R} \simeq 2P_{\rm ph}$	Total	Any $P_{V_R} \simeq P_{\rm ph}$	Any $P_{V_R} \simeq 2P_{\rm ph}$	Total [a]
LPV	Pulsation	1992	304	2296	3266	906	4084
	LSP	2807	109	2916	4125	370	4421
ELL		100	781	881	160	966	1109
Total		4899	1194	6093	7551	2242	9614

Notes. [a] Sources that do not belong to the TQS may show photometric periods that are simultaneously compatible with P_{V_R} in both the 2:1 and 1:1 ratios, so the corresponding "total" column is not necessarily equal to the sum of the two previous columns.

by a relatively large RV amplitude or a small number of RV measurements. It is easy to see how such features can enhance the impact of outlier removal and methods differences on the calculation of the median RV, especially given the irregular time sampling of *Gaia* observations.

4.1.2. Mean and systemic radial velocities

The uneven sampling of *Gaia* time series also means that the median RV is not necessarily a good indicator of the systemic RV, that is the center-of-mass velocity, an accurate estimate of which generally requires the RV variability to be modeled. In principle, the zero-point V_R^0 of the RV time series model (i.e., the constant term c_{0,V_R} in Eq. (2), unpublished but computable with the published time series) is more representative of the systemic RV than an average over the time series. In order to assess this, we examine the fractional deviations between $\langle V_R \rangle$ and V_R^0 by taking the absolute value of their difference and scaling it to $|V_R^0|$. The distribution of this quantity is displayed in Fig. 16, together with the fractional difference with respect to V_R^0 of $V_R^{\rm DR3}$ and of the mean value $\overline{V_R}$ derived from variability processing. This diagram shows that in most cases the median values are compatible with V_R^0 to within few percents, whereas the mean value performs slightly better, typically within less than a 1%. Both indicators are consistent with V_R^0 within about 10% for the bulk of the sample.

As is discussed in Sect. 4.2, a subset of the FPR sources are also present in the *Gaia* DR3 non-single stars table, which provides center-of-mass velocities based on orbital solutions. We include in Fig. 16 a comparison between these values and V_R^0, that are found to be in excellent agreement, being usually compatible within 0.03%. We leave further comparisons with the non-single stars results for the next section.

While such fractional difference is helpful to have a quick glance at the general degree of agreement between various indicators, it fails to characterize the stars having a very small systemic RVs (comparable with or smaller than the RV uncertainty) which lead to the tail of the distribution at large values, as can also be appreciated from Fig. 17. Therefore, we further examine the sample by scaling the difference $|\langle V_R \rangle - V_R^0|$ by the mean value $\overline{\varepsilon}_{V_R}$ of the uncertainties on individual RV epochs on a given time series. The resulting quantity is displayed in Fig. 18 as a function of the semi-amplitude of the RV time series model. Its distribution peaks around one, indicating that for the majority of the time series the value $\langle V_R \rangle$ is compatible with V_R^0 within the mean RV uncertainty. However, there is a clear tendency for this compatibility to degrade with increasing RV amplitude. This is primarily due to the uneven and irregular time sampling of *Gaia*

observations. While this is rarely an issue for sources that display little or no variability, it is something to be kept in mind when dealing with LPVs and ELLs, because:

– they have periods comparable with the observational baseline, so the *Gaia* time series may cover only a small number of variability cycles;

– they have large variability amplitudes, and so observations taken at a random phase of the variability cycle may be very far from the central value;

– the *Gaia* scanning law is such that a large number of observations may be concentrated within an interval of time much shorter than the typical time scales of LPVs and ELLs; these "clusters" of epochs have a very high statistical weight, but carry little more information than a single measurement taken in the middle of that time interval.

The combination of these effects increases the chance of selectively over-sampling or under-sampling a specific phase of the variability cycle, thereby skewing the median away from the mid-point of the RV curve. For similar reasons, the amplitude of variability may be underestimated by statistical indicators such as the standard deviation or the interquartile range. Some examples of RV curves showing these effects are provided in Appendix B.

We note that the top-quality sample shows a slightly larger difference, on the mean, between $\langle V_R \rangle$ and V_R^0 than the full FPR sample. This is due to the presence of a larger fraction of ellipsoidal variables in the former sample, which have large RV ranges (see Sect. 3 and Fig. 19).

Finally, we recall that the epoch RVs are measured from RVS data by comparing observed spectra with synthetic stellar spectra used as templates (see Sect. 6.4.8 of the *Gaia* DR3 documentation; Sartoretti et al. 2022). In this FPR the templates have $T_{\rm eff}$ ranging from 3100 K to 7500 K (typically between 3300 K and 3800 K) and $\log(g)$ in the range –0.5 to 5.0, and are restricted to O-rich stars (Katz et al. 2023). As LPVs can have stellar parameters outside these ranges and include C-stars, the matched template is not necessarily the most appropriate in terms of atmospheric parameters $T_{\rm eff}$ and $\log(g)$, which might impact the derived zero-point and median RVs. Yet, the analysis of the overall features of the FPR sample does not highlight any clear systematic discrepancy with respect to what would be reasonably expected. For instance, the distribution of RV as a function of galactic longitude (Fig. 20) does not show a wider spread for C stars than for O-rich LPVs. Likewise, the sky map presented in Sect. 5 gives essentially the same picture as Fig. 5 of Katz et al. (2023), despite being limited to a selection of stars (i.e., cool red giants and including C-stars) that are substantially more exposed to the risk of template mismatch than the bulk of
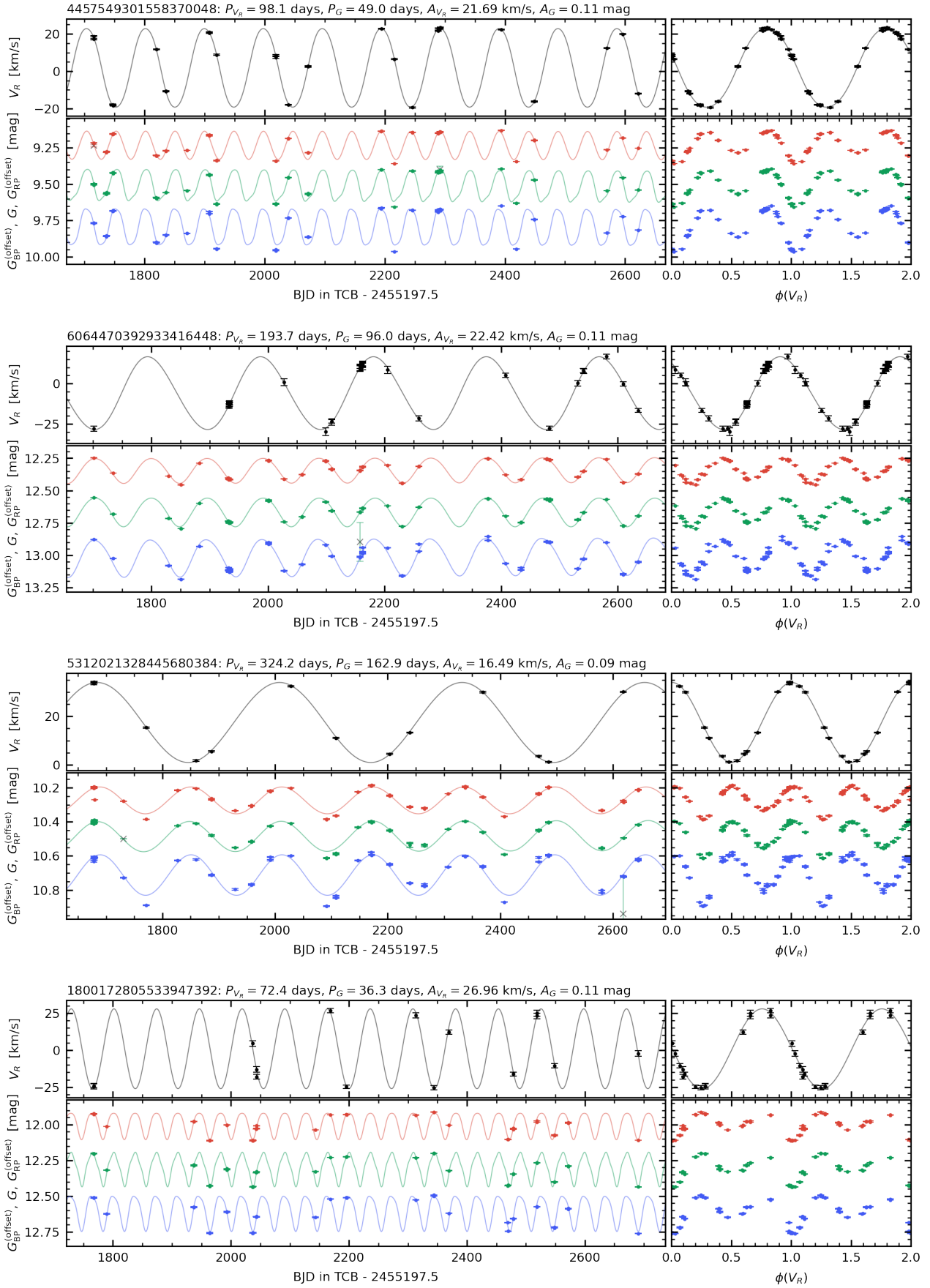
Fig. 12. Similar to Fig. 9, but showing some example time series of ELL candidates. All data in the panels on the right-hand column are folded by the FPR RV period.
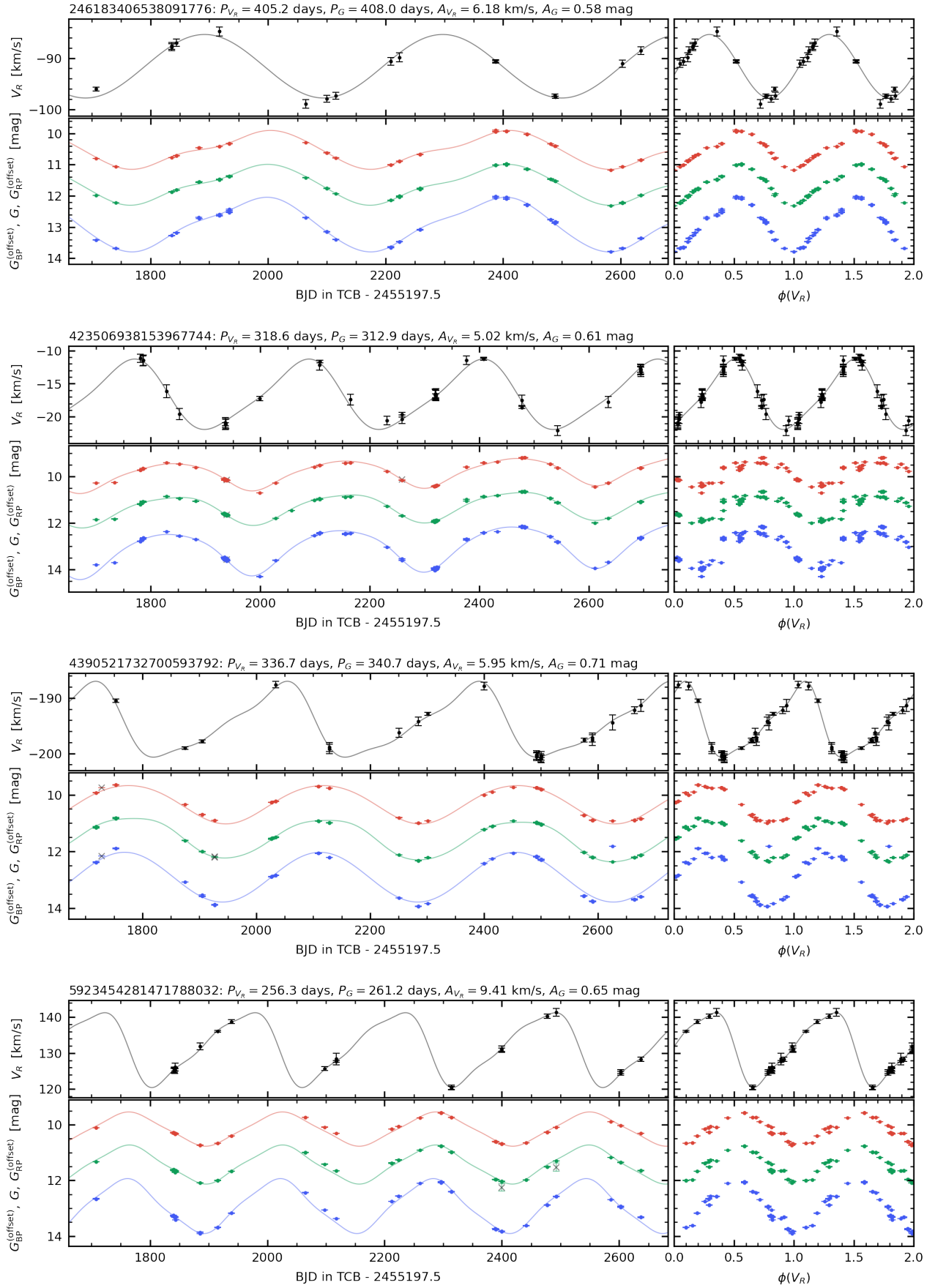
Fig. 13. Similar to Fig. 9, but for pulsating LPV candidates. All data in the panels on the right-hand column are folded by the FPR RV period.
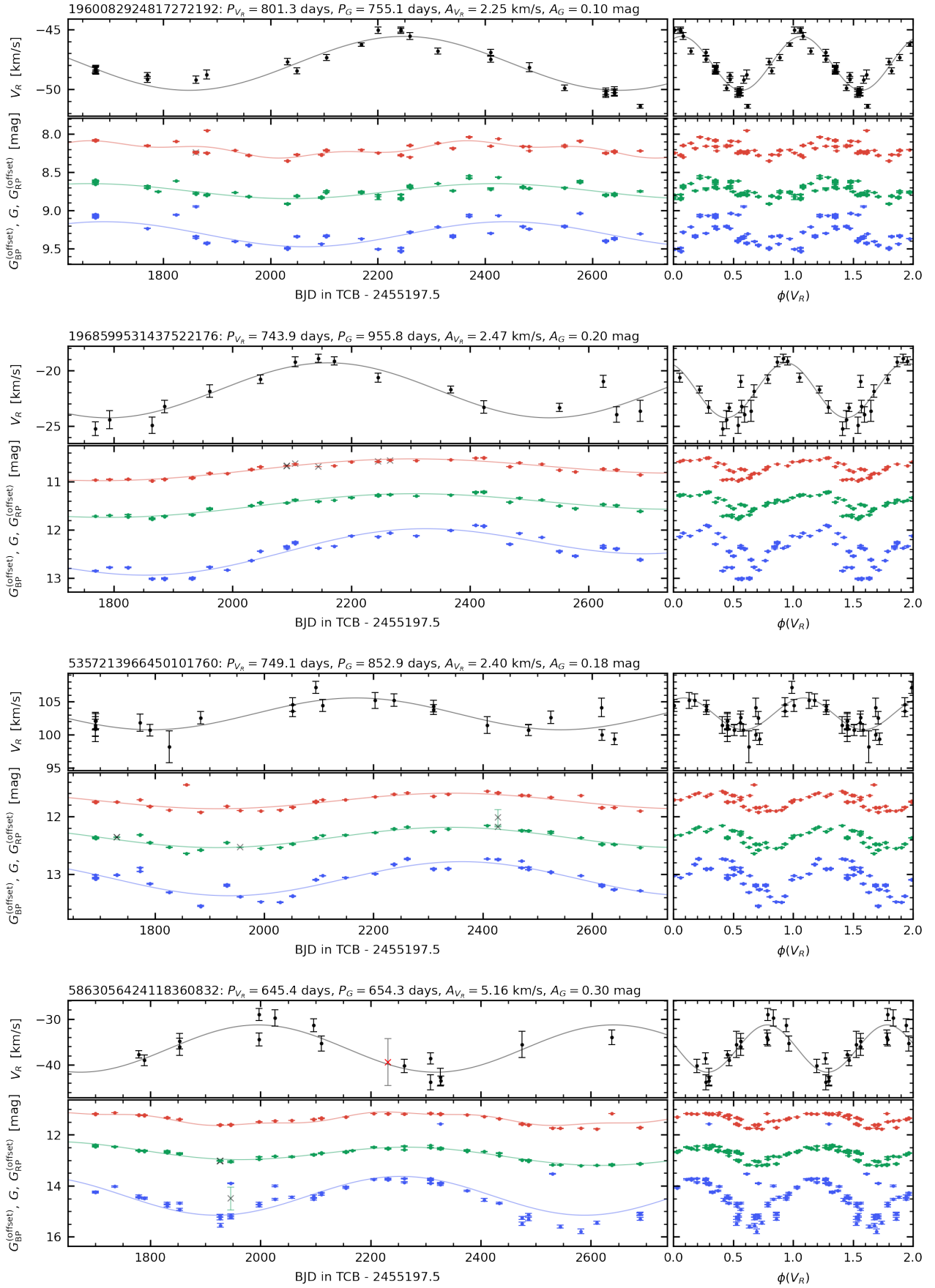
Fig. 14. Similar to Fig. 9, but for LPV candidates for which we likely detect a LSP. All data in the panels on the right-hand column are folded by the FPR RV period.
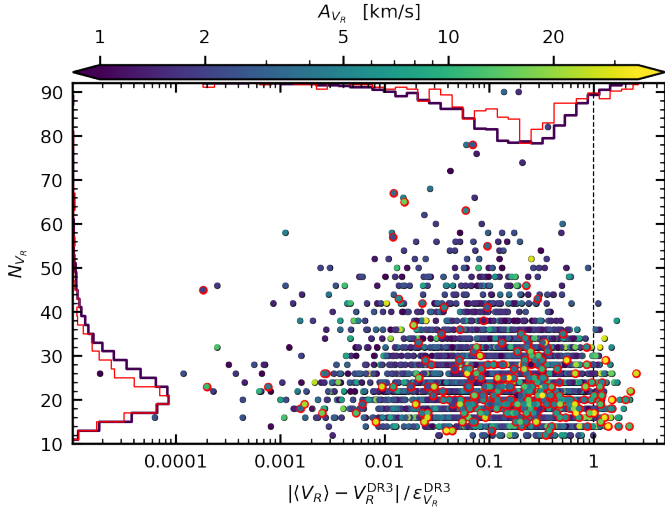
Fig. 15. Number N_{V_R} of epochs retained for RV variability processing against the absolute difference between the median RV derived by variability processing ($\langle V_R \rangle$) and published in *Gaia* DR3 (V_R^{DR3}), scaled to the RV uncertainty ε_{V_R} published in DR3. Data points are color-coded by the semi-amplitude of the RV time series model. The time series that had one RV epoch excluded by outlier removal during variability processing are circled in red. The value $\langle V_R \rangle$ for these sources is computed from one fewer epoch compared to V_R^{DR3}.
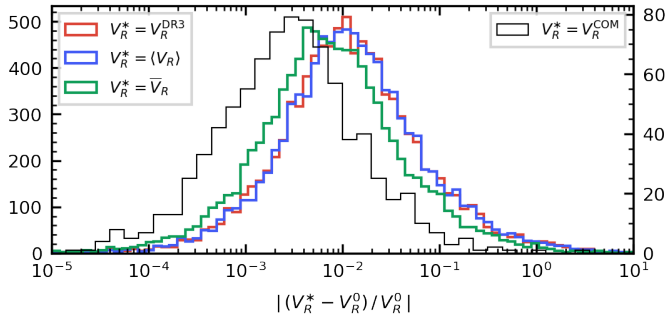
Fig. 16. Comparison of several average RV indicators with the zero point of the RV time series models. Different indicators are displayed in different colors (red: median value V_R^{DR3} published in *Gaia* DR3; blue: median value $\langle V_R \rangle$ computed by variability processing; green: mean value $\overline{V_R}$ computed by variability processing). The histograms show the distribution of absolute difference between each of the average values and V_R^0, normalized to the latter. The thin black histogram, limited to a subset of the FPR sample, compares V_R^0 with the center-of-mass velocity V_R^{COM} derived by the non-single stars processing pipeline for *Gaia* DR3 (see Sect. 4.1.2 for more details), and refers to the scale on the right-hand side axis.

sources used by these authors in their figure. More importantly, the variability properties of the RV curves, that are the object of this FPR, are not expected to be significantly affected by the template mismatches, as all epoch RV measurements should be equally impacted by the mismatch. The LPVs undergoing large changes in T_{eff} throughout the pulsation cycle may represent an exception as a different spectral template should be adopted for different epochs, but it is unclear how this might impact the RV variability data. Visual inspection of the RV curves and comparison with their photometric light curves give further support to these considerations.
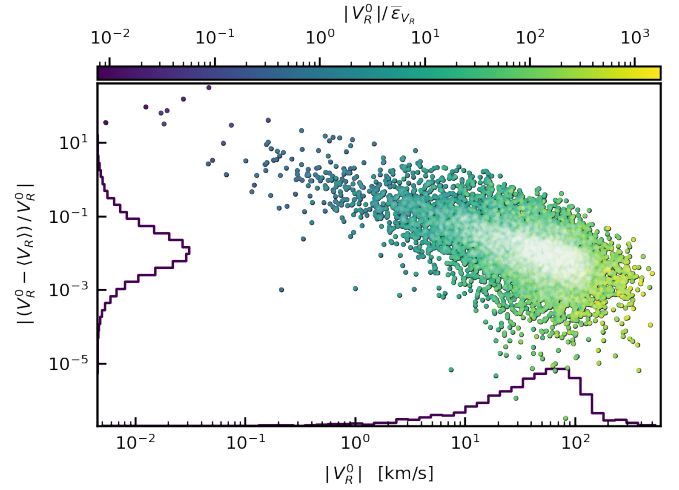
Fig. 17. Absolute difference between the median value $\langle V_R \rangle$ and the zero point V_R^0 RV estimates, scaled to the latter and shown against the absolute value $|V_R^0|$ of the latter. Data points are color-coded by the ratio $|V_R^0|/\overline{\varepsilon}_{V_R}$, showing that large discrepancies (top portion of the diagram) are associated with absolute values of the systemic RV comparable with or smaller than the RV uncertainty. A white shading indicates a more densely populated area of the diagram.
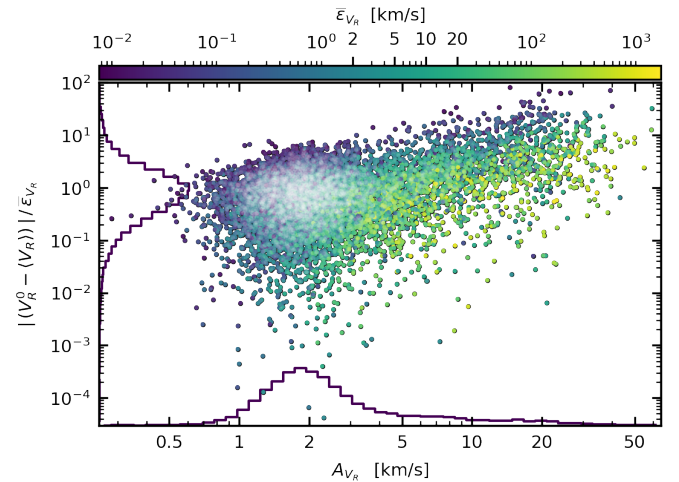
Fig. 18. Absolute difference between the median value $\langle V_R \rangle$ and the zero point V_R^0 RV estimates, scaled to the mean value of the uncertainties on individual RV epochs, and shown against the semi-amplitude A_{V_R} of the RV time series model. A white shading indicates a more densely populated area of the diagram.

4.2. Radial velocity variability

For the purpose of assessing the quality of the periods and amplitudes derived from RV time series of sources that are identified as ellipsoidal binary stars, we compare the FPR data with the results of non-single star (NSS) processing from *Gaia* DR3. In particular, we consider the data from the `nss_two_body_orbit` table from the *Gaia* archive (Gosset et al., in prep.). We find that 855 of the FPR sources are also found in that table. Based on the classification outlined in Sect. 3, we identify 353 of them as ELL candidates, 296 as pulsating LPVs, and 205 as LPVs with a LSP.

We compare our RV periods with the NSS values in Fig. 21, which displays the absolute difference between the two periods scaled to the latter. In most cases we find a good degree of period compatibility (typically within a few percents), even for
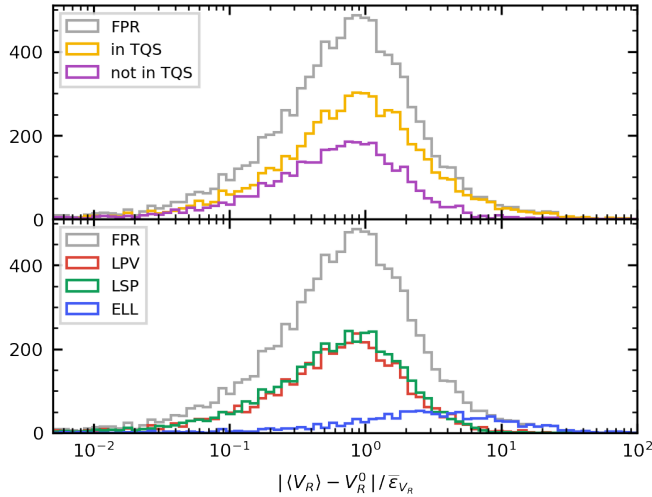
Fig. 19. Absolute difference between $\langle V_R \rangle$ and V_R^0, scaled by the mean of epoch RV uncertainties, for various subsets of the FPR sample. In the top panel, the orange curve corresponds to sources flagged for high consistency between RV and photometric periods, whereas all other sources are represented by the purple curve. In the bottom panel, the red, green and blue curve correspond to sources tentatively identified as pulsating LPVs, LPVs showing LSP variability, or ellipsoidal variables, respectively (see Sect. 3). The gray curves in both panels represent the whole FPR sample.
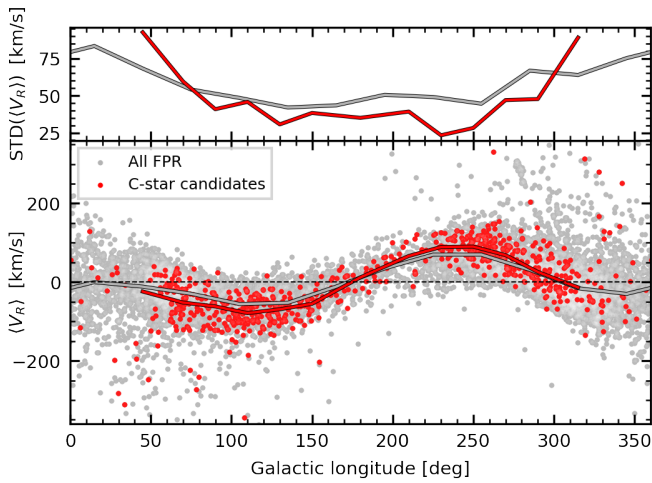
Fig. 20. Median RV ($\langle V_R \rangle$) of FPR sources as a function of their galactic latitude. Red sources are C-star candidates. The solid lines indicate median values over bins of galactic latitude. The curves show the median value and standard deviation of $\langle V_R \rangle$ in bins of galactic latitude (for C-stars these statistics are limited to galactic longitudes between 30° and 330°, excluding regions where they are scarce).

the sources that we do not classify as ELL candidates, although the agreement is better for the latter. In general, the comparison is slightly better for objects identified as pulsating LPVs compared with the LSP candidates. For sources that we identify as ellipsoidal variables, the periods are always within 10% of each other, and typically compatible within 0.1%. There is only one exception (whose time series are displayed in Fig. 22) for which we derive a 428.7 days period in stark difference with the NSS period of 0.35 days. As the RV time series only covers 12 visibility periods, it is hard to conclude which period is more realistic. Both values correspond to a strong signal in the periodogram, but we do not detect the latter as our processing is limited to periods
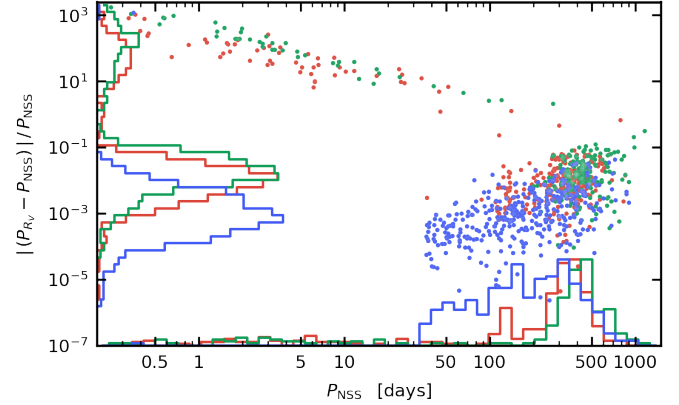
Fig. 21. Relative difference between the period P_{V_R} estimated by variability processing and the value derived from the non-single star pipeline, relative to the latter, for the FPR sources that have a counterpart in the *Gaia* DR3 table of orbital parameters of non-single stars. Data points are color-coded by variability type according to the classification presented in Sect. 3 (blue: ellipsoidal variables; red: pulsating LPVs; green: LPVs with LSP).

longer than 10 days. In any case, periods as short as 0.35 days are not expected for ellipsoidal red giants. A similar situation occurs for a number of other sources for which NSS results in relatively short periods, and hence there exists a large difference with respect to the FPR period (sequence of points in the upper part of Fig. 21 extending to low NSS periods). However, none of these sources is identified as a ELL candidate according to our classification criteria, which casts some doubt on the validity of the orbital model adopted for modeling their RV time series for the NSS processing.

We perform a similar analysis in terms of the RV semi-amplitude, comparing the value derived by the variability pipeline with the NSS results (Fig. 23). We find a qualitatively similar picture as for the periods, the compatibility between the NSS and our semi-amplitudes being typically within 2–3% for ELL candidates, 10% for pulsating LPV candidates, and 20% for LSP candidates. It is interesting to observe that the RV curve asymmetries frequently displayed by pulsating LPVs are often accounted for by a large eccentricity in the binary models adopted by NSS.

Finally, we note that the NSS data for the matched sources result from the assumption of a single-lined spectroscopic binary model, with the only exception of the source 2567779977831471232 for which an orbital astrometric binary model was adopted, obtaining a 928.7±85.3 days period. According to the classification presented in Sect. 3, we also identify that source as a binary (ELL) candidate, but we find a RV period significantly shorter ($P_{V_R} = 695.3$ days), although we note that the RV observation covers just about 800 days, hence the period cannot be accurately constrained.

4.3. Comparison with RV data from literature

We searched the literature for RV data to compare with this FPR, and found them to fall into two main categories. The first one concerns large-scale spectroscopic surveys providing extensive catalogs of RV data. The chances of finding matching objects against these source lists are relatively high, but at the same time they only allow for comparing average RVs, as they are based on single-epoch observations or a few epochs at most, and rarely
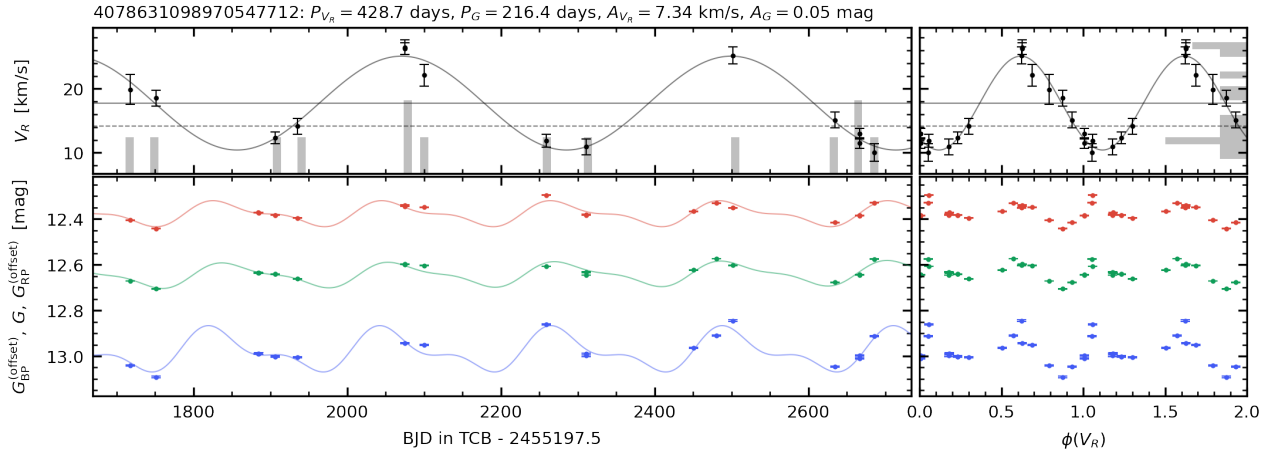
Fig. 22. Similar to Fig. 9, showing the *Gaia* RV and photometric curves of the ELL candidate showing a large mismatch between the FPR RV period (428.7 days) and the value published in the *Gaia* DR3 NSS table (0.35 days). For the purpose of visualization, the G_{BP} and G_{RP} time series are offset by an arbitrary amount. All data in the panels on the right-hand column are folded by the FPR RV period.
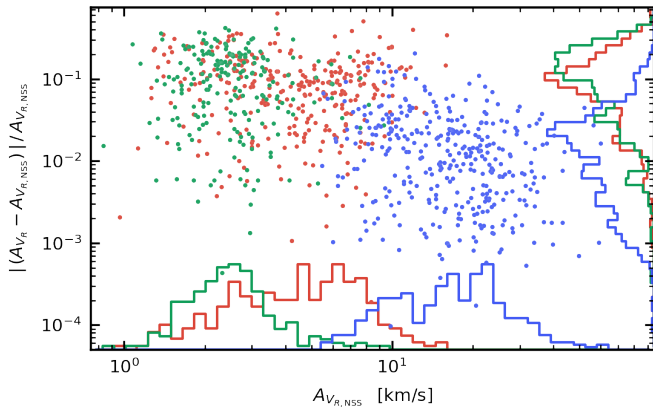
Fig. 23. Similar to Fig. 21, but comparing the RV semi-amplitudes rather than the periods.

provide RV time series. In Sect. 4.3.1 we present a comparison with a few such catalogs.

Besides these surveys there exist a number of smaller-scale observational programs targeting specific types of stars or fields of the sky. These studies involve a more focused analysis and validation of the RV data of the targets, and often result in the publication of the time series. We attempted to cross-match the FPR catalog with the source lists from various such literature works, but found only a few matches, that are examined in Sect. 4.3.2.

4.3.1. Comparison with RV surveys

We compare the FPR results with the data published by three spectroscopic surveys: the Apache Point Observatory Galactic Evolution Experiment survey (APOGEE, data release 16, Jönsson et al. 2020), the GALactic Archaeology with HERMES survey (GALAH, data release 3, Buder et al. 2021), and the RAdial Velocity Experiment survey (RAVE, data release 6, Steinmetz et al. 2020). For the former two, we rely on the cross-match with *Gaia* DR3 performed as part of their data release process (they provide best-match *Gaia* DR3 source IDs), whereas for the latter we adopt the results of the *Gaia* cross-match with external catalogs (`gaiadr3.ravedr6_best_neighbour` table in the *Gaia* archive).

Table 3. Number of FPR sources cross-matched with external spectroscopic surveys.

	FPR	APOGEE	GALAH [a]	RAVE
LPV	4084	34	58 (65)	503
LSP	4421	41	89 (95)	643
ELL	1109	14	20 (21)	80
Total	9614	89	167 (181)	1226
in TQS	6093	60	108 (119)	784

Notes. [a] Numbers indicate the sources with a published value of RV in GALAH, while the total number of matches is given in parentheses.
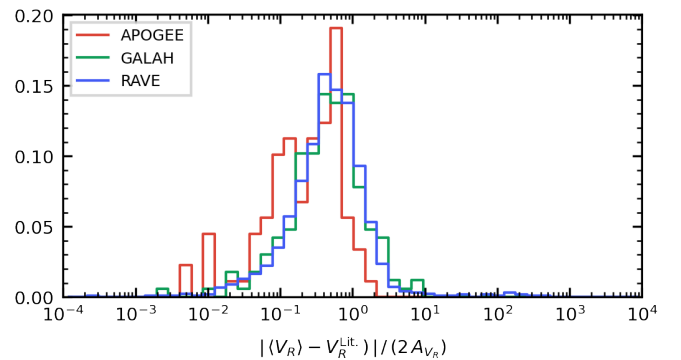
Fig. 24. Comparison between the median RV from this FPR with the average values provided by external catalogs (red: APOGEE; green: GALAH; blue: RAVE). We consider the absolute difference between the two values, normalized to the peak-to-peak model amplitude. We note that the histograms are normalized to their area.

A summary of the numbers of matched sources is provided in Table 3. We find 89 matches with APOGEE, 167 with GALAH, and 1226 with RAVE (of which 60, 119, and 784, respectively, are in the TQS). In Fig. 24 we provide an overview of the comparison between the median RV $\langle V_R \rangle$ from the FPR and the averages given by each of these surveys. Namely, that figure shows the distribution of the absolute difference between $\langle V_R \rangle$ and the literature value, divided by the peak-to-peak amplitude of the *Gaia* RV time series model. We employ the median RV rather

Table 4. Properties of the sources whose RV time series are compared with literature multi-epoch RV studies.

Name	*Gaia* DR3 source ID	Var. type [a]	Sp. type [a]	P_G(FPR) (days)	P_{V_R}(FPR) (days)	$P_{\text{Lit.}}$ (days)	Reference
AR Cep	2300884800185315712	SRb	M4III	575.2	550.4		Alvarez et al. (2001)
RS CrB	1372316959598411520	SRa	M7	330.9	305.9	328.3 ± 2.6	Hinkle et al. (2002)
R Nor	5985676640941632384	Mira	M3e-M6II	505.5	496.7	507	Lebzelter et al. (2005a)

Notes. [a] The variability and spectral types are taken from the corresponding reference paper.

than the zero-point RV for the purpose of comparison as the values published in each of the three surveys we compare with are also derived as averages (unless coming from single-epoch observations).

We find comparable results for all three surveys, with a slightly higher degree of compatibility with APOGEE. The differences in RV are usually comparable with or smaller than the amplitude of RV variability. As expected, larger deviations occur for sources having large RV amplitudes, and therefore typically for ellipsoidal variables, as well as for sources with few observations in the examined external surveys. Indeed, the GALAH RVs for the matched sources are all based on single-epoch measurements. Of the sources matched with APOGEE, 29 are RVs based on single-epoch data, 52 have between 2 and 4 epoch spectra, and the remaining 8 sources have at most 8 spectral observations. Similarly, the vast majority of the sources matched with RAVE (1141) has a single epoch, while 70 have been observed during two epochs, and only 15 have more than two epochs (at most 7).

While the poor time coverage in the external surveys we compare with is the most likely cause of discrepancy with our RV values, another possible cause could be related with spectral mismatch, whether in the *Gaia* pipeline (as mentioned in Sect. 4.1.2) or in the literature survey. This seems to be the case, for example, for a few sources for which the RAVE catalog indicates effective temperatures in excess of 8000 K and that we identify as LPVs showing the largest RV differences compared to the FPR. We found no evident correlation between the RV differences and the differences between the spectral parameters adopted by the *Gaia* and external survey pipelines. A mismatch between the values of T_{eff} or $\log(g)$ does not necessarily lead to large RV differences, at least as long as these values are not unreasonably far from the range of stellar parameters typical of LPVs.

4.3.2. Comparison with literature multi-epoch RV data

There exist a limited number of literature studies providing multi-epoch RV data of LPVs and ellipsoidal variables, hence we found only three FPR sources that we could compare with published RV time series. They are summarized in Table 4, where the literature period is compared with the values P_G and P_{V_R} given in the FPR.

The first source for which we found a match is the SRb star AR Cep observed by Alvarez et al. (2001). They obtained spectral observations at optical wavelengths and derived RVs by cross-correlation with template spectra of types K0-III or M4-V. For the matched source they provide two RV epochs separated by about 2 months (roughly 10% of the period found in the *Gaia* time series). They found RVs between $-16.0\,\text{km}\,\text{s}^{-1}$ and $-16.5\,\text{km}\,\text{s}^{-1}$ (with little differences between the two spectral templates) that are in good agreement with our results, but they do not provide a period to compare with. Their RVs are shown

on top of the *Gaia* RV curve folded with the best-fit RV model in the top section of Fig. 25 (an arbitrary phase offset is applied for the purpose of visual comparison). This source is not part of the TQS, and according to the criteria defined in Sect. 3 we identify it as a pulsating LPV.

Another source for which we found a match is the SRa star RS CrB, whose RV curve has been examined by Hinkle et al. (2002) based on near-IR spectral observations. They obtained measurements for 23 epochs spanning 5 yr, and derived a period of 328.3 ± 1.6 days that agrees with the value $P_G = 330.9$ days we derived from the *G*-band time series, but is less consistent with our RV period, $P_{V_R} = 305.9$ days. The time series of this source are compared in the middle section of Fig. 25, and in Fig. 26 we limit the comparison to the RV time series, folding them with both the FPR and literature RV periods. We note that both these periods are consistent with the longest of the three periods of RS CrB reported in literature, interpreted as resulting from binarity rather than pulsation. Indeed, the value given by Hinkle et al. (2002) is based on an orbital solution. Interestingly, we classify this source as a LPV but it lies very close to the boundary line defined by Eq. (10), so that its period is identified by pulsation. As clearly seen in Fig. 25, the photometric and RV time series of this source are consistent with each other, as a result this source is part of the TQS.

Finally, we found that the O-rich Mira R Nor is present both in the FPR and in the list of sources investigated by Lebzelter et al. (2005a). They derived RV measurements from nine near-IR spectra covering about an entire pulsation period of the star, for which they report a value of 507 days. This is in good agreement with both the periods we derived from the *G*-band and RV time series ($P_G = 505.5$ days, $P_{V_R} = 496.7$ days). The time series of this star are compared in the bottom section of Fig. 25. We identify this source as a pulsating LPV, but it is not part of the TQS.

It is worth noting that the RVs we are comparing for these three sources have been derived from different spectral ranges. Indeed, the observations by Alvarez et al. (2001) are taken at short wavelengths, between 390.6 nm and 681.1 nm, whereas the *Gaia* RVS covers the range between 846 nm and 870 nm. The spectral observations by Hinkle et al. (2002) and by Lebzelter et al. (2005a) cover an even more different range, being centered around 1.6 μm. This means that, if the observed variability results from pulsation, the RV measurements concern layers at different depths in the stellar atmosphere, and therefore the RV curves derived from different spectral ranges differ in terms of amplitude or show an offset between each other, which seems to be the case for R Nor (top-right panel in the bottom section of Fig. 25). While it should be kept in mind that the offset might be at least partially caused by a template mismatch in the processing of *Gaia* RVs (see Sect. 4.1.2), this figure clearly illustrates how such effects do not impact the quality of the RV variability parameters provided as part of this FPR. In the case of AR Cep
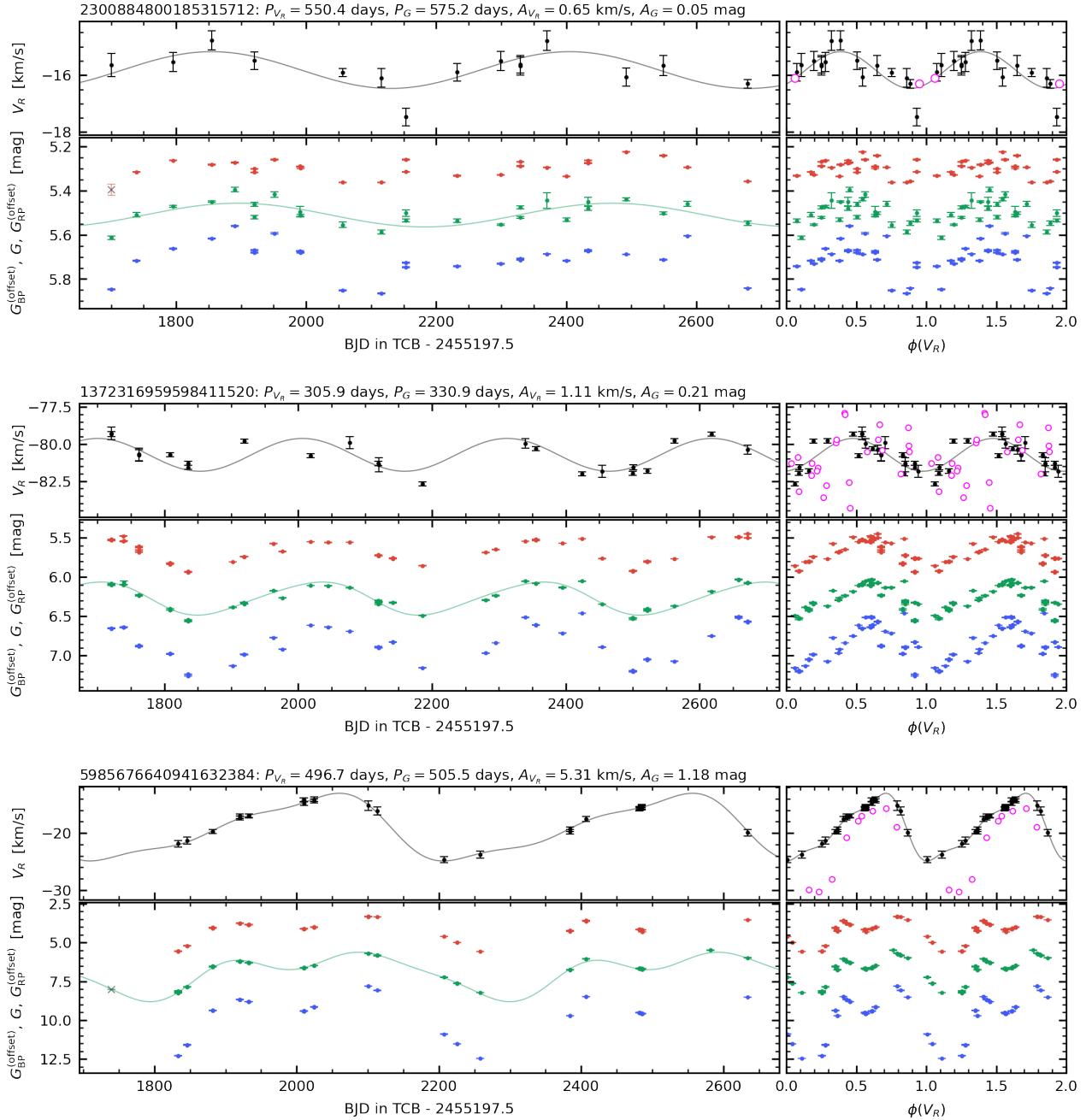
Fig. 25. Similar to Fig. 9, but showing the *Gaia* time series for the sources compared with literature as discussed in Sect. 4.3.2 (see also Table 4). The sources displayed from top to bottom are the SRb star AR Cep (compared with Alvarez et al. 2001), the binary SRa star RS CrB (compared with Hinkle et al. 2002), and the O-rich Mira R Nor (compared with Lebzelter et al. 2005a). Literature RV time series are displayed as magenta circles (with arbitrary phase offset) in the panels showing the folded RV curve. In each case, the *Gaia* RV period (indicated in the header of each panel) is used for folding.

(top section of Fig. 25) the small number of literature epochs to compare with makes it difficult to draw conclusions in this respect. In contrast, we note that the RV curves of RS CrB (middle section of Fig. 25) are well compatible in terms of average RV as well as amplitude (Fig. 26). This tends to confirm the orbital nature of the variability we detect for this source.

5. Catalog overview

In this section we provide an overview of the FPR catalog, with the purpose of showcasing its content of physical information.

We first present the sky distribution of the catalog in Sect. 5.1, then investigate in Sect. 5.2 the distributions in the color – absolute magnitude diagram of subsets of sources with good relative parallax precisions, and finally analyze in Sect. 5.3 the period-luminosity relations of various subsets of the top-quality sample.

5.1. Sky distribution

Figure 27 illustrates the sky distribution of the sources in this FPR, each being color-coded by the median RV $\langle V_R \rangle$ resulting from the *Gaia* variability processing pipeline. As discussed in
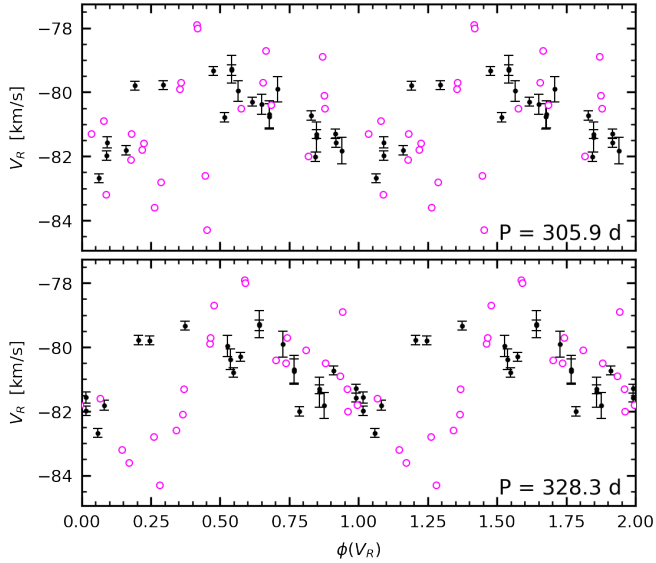
Fig. 26. Phased RV curve of RS CrB folded with the FPR RV period (top panel) and with the period derived by Hinkle et al. (2002; bottom panel). Symbols have the same meaning as in Fig. 25.

Sect. 4.1.1, for the vast majority of the FPR sources the value of $\langle V_R \rangle$ is compatible with the value of median RV derived from the *Gaia* spectroscopic data processing pipeline and published as part of *Gaia* DR3, even though the processing details are slightly different. Indeed, the physical picture emerging from Fig. 27 is entirely consistent with the expectations in terms of the Galactic rotation curve.

This also provides support to the expectation that the occurrence of spectral template mismatches has a minor impact on the accuracy of the median RV as an indicator of the center-of-mass velocity of the observed stars, even for the long-period, large-amplitude variables. We note that the RV difference between opposite parts of the Galaxy with respect to the Sun is $\gtrsim 200 \, \mathrm{km \, s^{-1}}$, which is much larger than any difference we encountered between $\langle V_R \rangle$ and the zero-point RV.

The sky distribution of the FPR shows some clear structures. Some of them are physical, such as the overdensity around the Galactic plane (but not on the plane itself, affected by strong interstellar extinction). The FPR contains a few sources located in the Small Magellanic Cloud (SMC) and Large Magellanic Cloud (LMC), with $V_R \sim 150 \, \mathrm{km \, s^{-1}}$ and $250 \lesssim V_R/\mathrm{km \, s^{-1}} \lesssim 300$, respectively. Most (possibly all) of them are red supergiant stars, following the period-luminosity relation of fundamental-mode pulsators (see Sect. 5.3). However, some structures are artificial and result from the *Gaia* scanning law and the selection filters applied to construct the catalog. The most evident such structures are the lack of sources in the Galactic center, as well as the largely empty regions on the bottom-left and top-right parts of the map. These areas are aligned on the ecliptic, and are characterized by a relatively small number of *Gaia* transits. As a result, they are easily removed by our condition on the number of RV visibility periods (Sect. 2.1). Conversely, the "stripes" around the central hole correspond to fields of the sky often observed by *Gaia*.

5.2. Color–absolute magnitude diagram

We restrict the presentation in this section to the FPR sources having precise parallaxes, so that we can confidently derive

their distance moduli μ and thus their absolute brightnesses. The distributions of the relative parallax errors for the FPR sample and various subsets thereof are displayed in Fig. 28 (excluding 92 sources with negative parallaxes). The results are largely independent on whether the sources are part of the TQS or not, whereas distinct distributions are observed for the different variability types present in the FPR. Indeed, ellipsoidal variables tend to have better parallax measurements than LPVs, and, within the latter category, LPVs showing an LSP have, on the average, slightly better parallax measurements than pulsating LPVs.

We set the upper limit on the relative parallax error at 15% (and require the parallax to be positive), thereby obtaining a sample of 5977 "good-parallax" sources (3740 in the TQS), that includes 794 ELL, 2136 pulsating LPVs, and 3047 LPVs showing a LSP. We examine this sample in the *Gaia* color – absolute magnitude diagram (CAMD, Fig. 29) constructed with the median magnitudes derived from the *Gaia* variability pipeline.

As expected, the majority of the LPVs are found on the asymptotic giant branch (AGB) of the CAMD, regardless of whether they show a LSP or a pulsation period. In contrast, ELL candidates are on the average fainter and bluer, and extend to the region of the red giant branch (RGB) in the CAMD. Most of the few stars with absolute G brighter than about 3 mag are identified as red supergiants (RSGs) in the SIMBAD astronomical database (Wenger et al. 2000), while some other ones could include massive AGB stars as well. Their periods tend to be identified as LSPs rather than pulsation by our classification method due to their relatively low variability amplitudes (see Sect. 3.2). Their true nature can clearly be revealed once their intrinsic brightness is known.

5.3. Period–luminosity diagrams

The sample can be further investigated in the period – luminosity diagram (PLD). To reveal the period – luminosity (PL) relations we adopt the *Gaia* Wesenheit index $W_{\mathrm{BP,RP}}$, which is an approximately reddening-free luminosity indicator (see Lebzelter et al. 2018, 2019, for the definition and details). We limit this analysis to the TQS in order to ensure higher period reliability, and note that we can construct two distinct PLDs depending on whether we adopt the period derived from photometric (taking here the G band) or RV time series. The two diagrams, both shown in Fig. 30, are entirely consistent with each other, except for the lower period limit, which is at 35 days for P_{V_R}, whereas it reaches ~17.5 days for P_G due to the presence of sources with $P_{V_R} \simeq 2P_G$ in the sample.

The main features of the PLDs are the following. Regardless of the adopted period (P_G or P_{V_R}), pulsating LPVs are found primarily on the period-luminosity sequences C′ and C associated with pulsation in the first-overtone mode and fundamental mode, respectively (top panels in Fig. 30). Some of them, that we further examine below, have a period in the area of sequence D, which is associated with LSP variability (see e.g., Pawlak 2021, and references therein). A few of them have $P_G \lesssim 40$ days, possibly on the bright part of sequence B. These might be identified with the type of variables that, in the context of the Optical Gravitational Lensing Experiment (OGLE, Udalski et al. 1992), are known as OGLE Small-Amplitude Red Giants (OSARGs, Wray et al. 2004). The LPVs whose variability we identify as LSP, on the other hand, are mostly found on sequence D (middle panels in Fig. 30), except for the RSGs that are located on the bright end of sequence C. The RV periods of ellipsoidal red giants align on
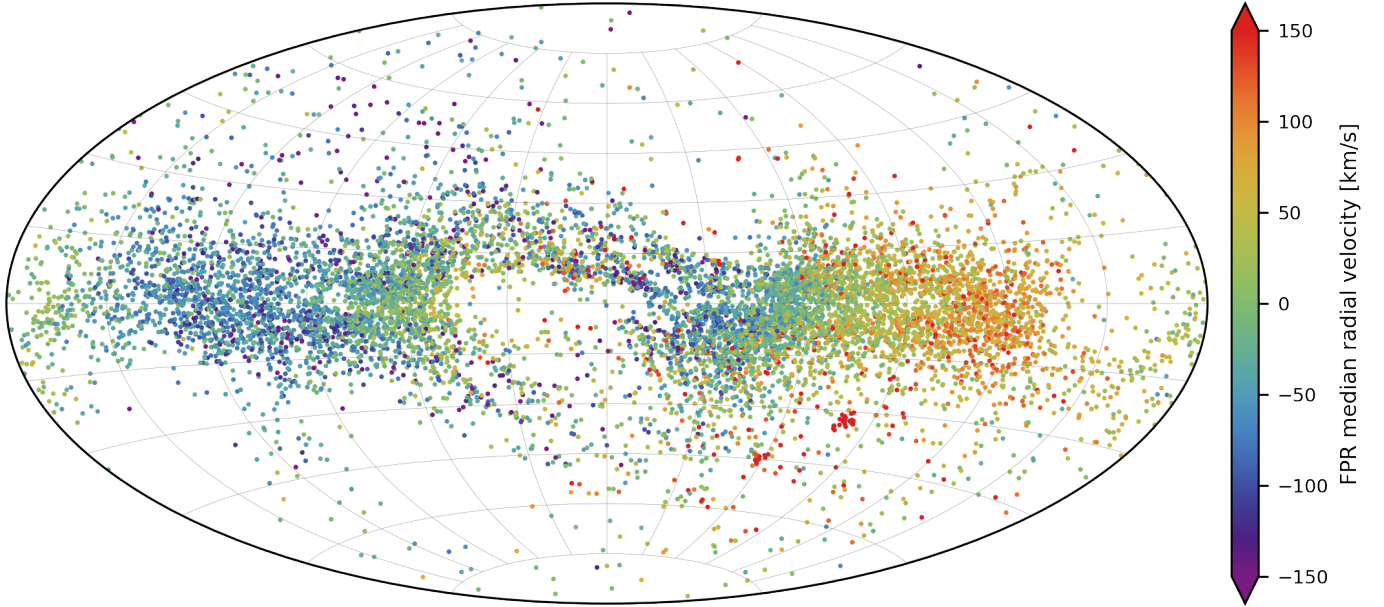
Fig. 27. Sky distribution in galactic coordinates of the sources in the FPR sample, color-coded by their median RV (the RV range is limited to $150 \, \mathrm{km \, s^{-1}}$ in absolute value for visibility). The velocity pattern resulting from the rotation of the Milky Way Galaxy is clearly visible, as well as several bright sources in the Magellanic Clouds having $V_R \gtrsim 150 \, \mathrm{km \, s^{-1}}$.
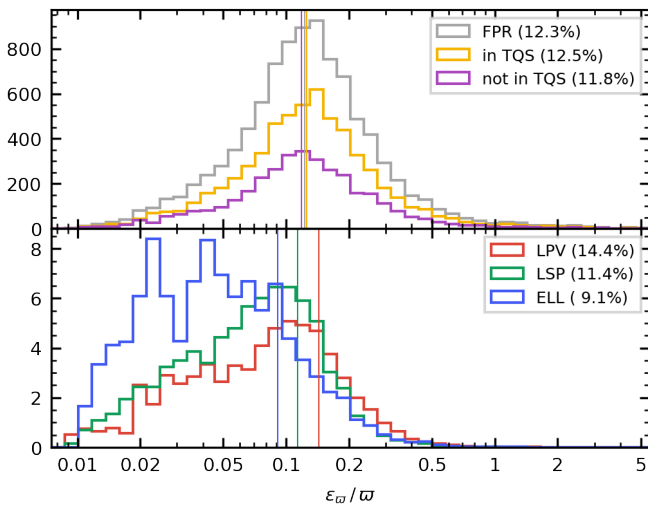
Fig. 28. Distribution of relative parallax errors in the FPR, excluding 92 sources with negative parallaxes. In the top panel, the histograms represent the full FPR (gray curve), the TQS (orange curve), and the sources that are not part of the TQS (purple). In the bottom panel, they indicate the sources identified as pulsating LPVs (red), LPVs showing a LSP (green), or ELL (blue). Note that the histograms in the bottom panel are normalized to their area. For each subset, the vertical line indicates the median relative error, which is reported as a percentage in the legends.

the PL sequence E (bottom panels in the figure), that forms a continuity with sequence D (see Fig. 1 of Soszyński et al. 2007).

In order to understand the presence of pulsation periods of LPVs on the PLD sequence D we have to distinguish between stars having O-rich and C-rich surface chemistry. Indeed, the PL relations corresponding to these two chemical types are different when examined through an optical-band Wesenheit index such as $W_{\mathrm{BP,RP}}$. A similar effect can be seen, for instance, in Fig. 1 of Soszyński et al. (2007; see also Lebzelter et al. 2018, 2019). To identify potential C-stars among the FPR sources we take advantage of the `is_cstar` flag provided with the *Gaia* DR3 catalog of LPV candidates (Lebzelter et al. 2023). About 7% of the FPR sources are classified as C-stars by this method, whereas 86% are identified as O-rich (another 7% are unclassified). After limiting the sample to pulsating LPV stars from the TQS and having good parallaxes, we find that 836 of them are O-rich, and 143 are C-rich. They are displayed in the PLD in Fig. 31, constructed using the photometric period P_G. In order to provide a visual reference for the positions of the PL relations, in the same diagrams we also show the sources from the OGLE-III catalog of LPVs in the LMC published by Soszyński et al. (2009). These authors also provide a photometry-based spectral-type classification which we use to discriminate between the O-rich and C-rich OGLE-III LPVs in Fig. 31.

As can be seen from the top panel of Fig. 31, there are some O-rich stars that we identified as pulsating LPVs and whose periods lie on the short-period, faint end of sequence D. These are LSPs mistakenly identified as pulsation periods by applying Eq. (10) because they lie close to the dividing line (between about 200 and 400 days, see Fig. 11). Let us now consider the C-rich sources in the bottom panel of Fig. 31. For the purpose of clear visualization, we do not show the C-rich OGLE sources whose primary period is flagged as a LSP. Indeed, the C-rich LPVs pulsating in the fundamental mode often lie below the corresponding PL relation (sequence C), overlapping with sequence D. This is a consequence of self-extinction by circumstellar dust which causes these sources to appear fainter than less dusty stars having similar periods (see e.g., Lebzelter et al. 2019, and references therein). This seems to be the case for the C-stars in our selection. Therefore, the presence of pulsating stars on sequence D can be correct if it results from circumstellar extinction.

Finally, we consider the TQS sources with good parallaxes and with $P_{V_R} \simeq 2 \, P_G$. These are displayed in the PLDs in Fig. 32, constructed with P_G (top panel) and P_{V_R} (bottom panel). It is clear that the majority of these sources obey the PL relation E of
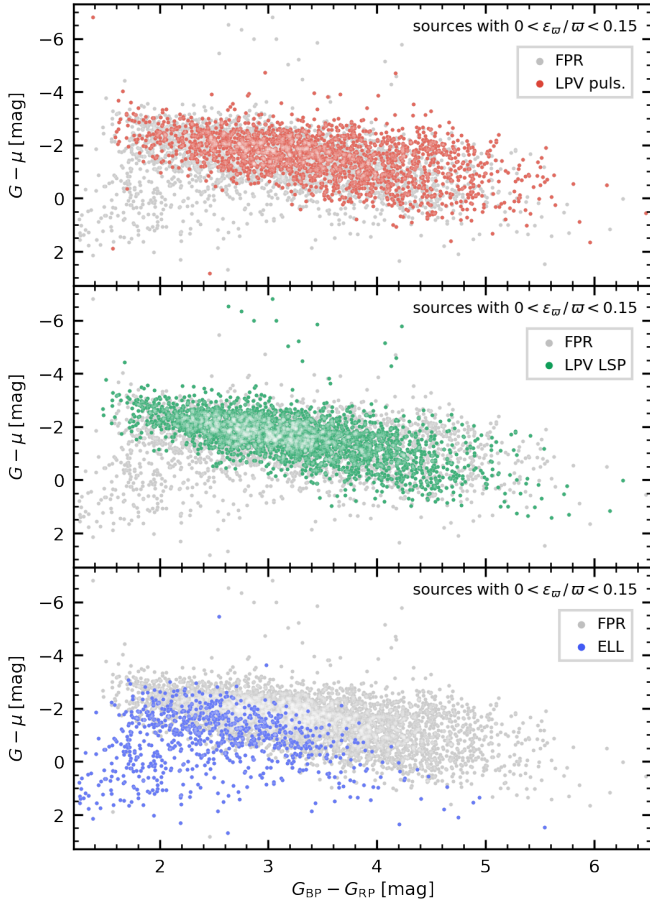
Fig. 29. *Gaia* color–absolute magnitude diagram of the FPR sources with positive parallaxes and parallax errors smaller than 15% (gray symbols), with μ being the distance modulus. The subsets of sources with different variability types are highlighted in the top (pulsating LPVs, red symbols), middle (LPVs with a LSP, green symbols) and bottom panels (ELL, blue symbols).

ellipsoidal red giants, leaving little doubt on the nature of their variability. Nonetheless, some of them populate other regions of the PLD, in particular the pulsation sequences C and C′ (and possibly B), as well as the LSP sequence D. This contradicts the naive expectation that the occurrence of P_{V_R} and P_G in a 2:1 ratio must indicate binary-induced variability. Let us then focus on the sources showing this feature and classified as LPVs (both pulsating and showing a LSP), which are highlighted in color in Fig. 32.

To begin with, we consider the G-band periods shorter than about 35 days, that are possibly compatible with the bright end of sequence B and that correspond to RV periods in the middle of sequence C′. If these periods are correct, it could mean that the variability pipeline identified the signature of first-overtone mode pulsation in the RV time series, and of second-overtone mode pulsation in the G-band time series. This would be consistent with the fact that, upon visual inspection, their photometric time series appear poorly regular, which can be taken as an indication of multi-periodicity. However, the first-overtone and second-overtone modes do not occur in a 2:1 ratio in LPVs (e.g., Wood 2015, and references therein). Furthermore, both sequences B and C′ are actually thought to result from pulsation in the first overtone mode (Trabucchi et al. 2017; Yu et al. 2020). Another explanation could be that these short G-band

periods are associated with spurious frequencies resulting from scan-angle dependent signals (Holl et al. 2023). However, the origin of these spurious frequencies as described in that paper should affect only the G band and not the G_{BP} and G_{RP} measurements. Yet another explanation could be that the periods derived from the RV time series are not correct. We note that this kind of stars often display irregular variations that may result in the detection of spurious periods.

Entirely similar arguments can be put forward for the G-band periods along sequence C′ in the top panel of Fig. 32 (between about 40 and 200 days), which correspond to RV periods on sequence C, in the range 80–500 days. However, in this case it is more likely that they actually correspond, respectively, to the first-overtone mode and fundamental mode periods, which can occur in a 2:1 ratio Soszyński et al. (see e.g., Fig. 6 of 2007). The photometric and RV periods would then both be correct and consistent with a pulsational origin.

Finally, we consider the LPVs that appear on sequence D in either panel of Fig. 32. The vast majority of them are probably LSPs, even though some are identified as pulsating LPVs (red points in the figure). Indeed, as discussed above, they could be pulsating C-rich stars suffering from circumstellar extinction, but only a few of the sources shown in Fig. 32 are classified as probable C-stars. Another explanation for these stars showing $P_{V_R} \simeq 2\,P_G$ is that their LSPs are indeed caused by binarity, which would be consistent with the scenario outlined by Soszyński et al. (2021).

6. Summary and conclusions

The *Gaia* DR3 saw the publication of average RVs for over 33 million stars based on 34 months of observations with the *Gaia* RVS, whereas epoch radial velocities were published only for a very restricted list of variable sources. Anticipating the publication of the full RV data with the fourth *Gaia* data release, we present RV time series for a selected sample of long-period variables as part of this *Gaia* FPR.

We describe the construction of the catalog, starting with the set of *Gaia* LPV candidates with a median RV published in *Gaia* DR3, and applying several filtering steps to ensure the highest quality of the final sample, leading to 9614 sources. In addition to the RV time series, for each source we provide the model-derived frequency and amplitude of RV variability, as well as the RV statistics, determined by the *Gaia* variability processing pipeline. In addition, we publish a flag allowing for the identification of a subset of 6093 sources that show a high degree of compatibility between the periods derived from the RV and photometric time series. We consider them to be of superior quality as all four of their *Gaia* time series (three photometric bands and the RV channel) are likely to carry a strong signature from the same physical process, enabling detailed studies of their variability.

We show how the catalog includes three groups of sources exhibiting different types of variability, namely ellipsoidal red giants, pulsating LPVs, and LPVs displaying a long secondary period. Stars from the first group are characterized by comparatively large RV amplitudes (that cannot be attained by pulsating stars) and small photometric amplitudes. They also frequently show RV periods that are twice as long as the periods derived from the photometric time series. They represent between 10 to 15% of the catalog. The remainder of the FPR consists in roughly equal parts of LPVs showing pulsation and long secondary periods, which we tentatively distinguish by their different positions in the period-amplitude diagram.
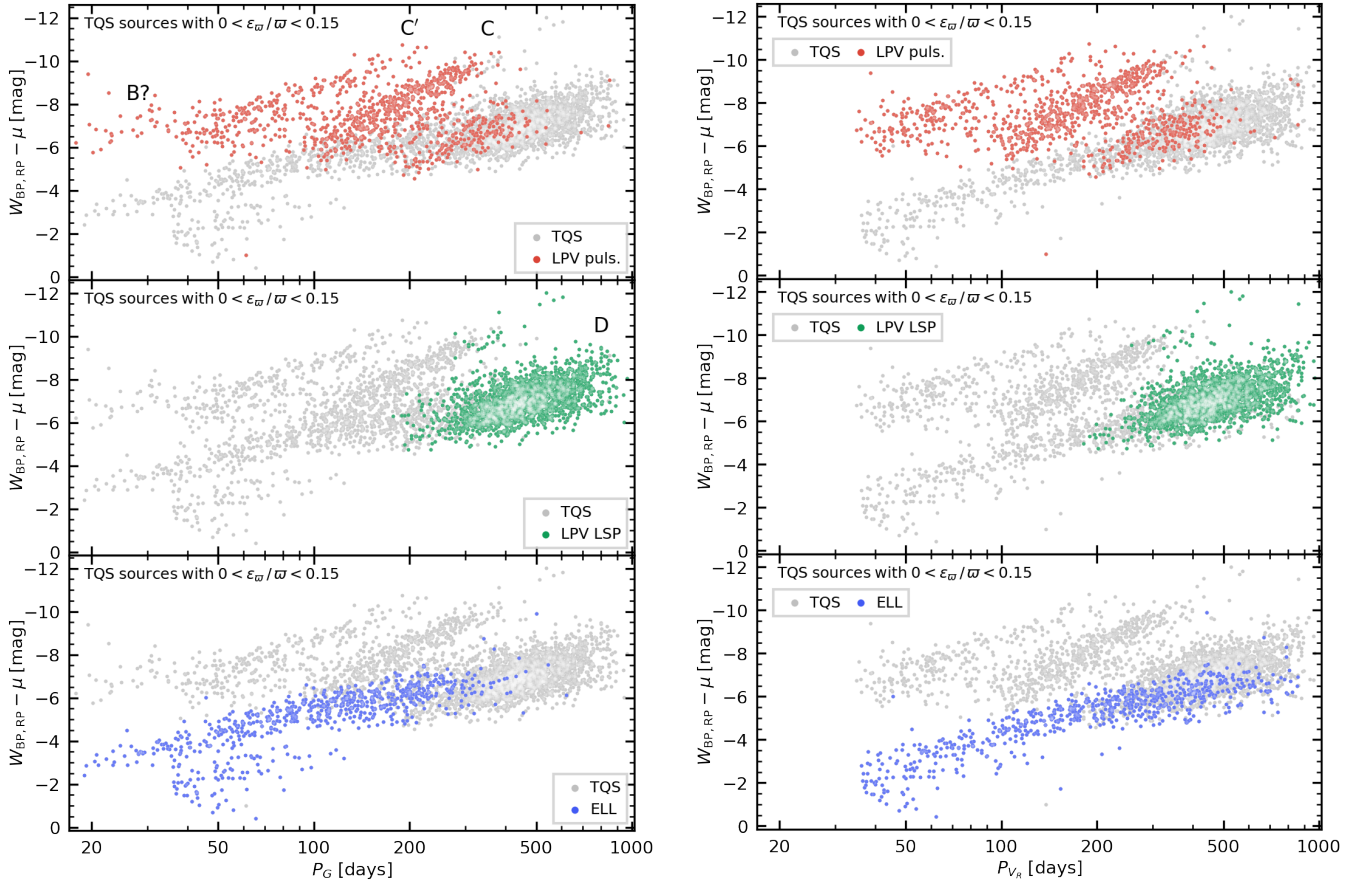
Fig. 30. Similar to Fig. 29, but showing the period – luminosity diagram. The panels in the left- and right-hand side columns employ the periods derived from the G-band and V_R time series, respectively, and the *Gaia* Wesenheit index $W_{BP,RP}$ is used as brightness indicator. The extent of the horizontal axis is intentionally kept the same in all panels for the purpose of comparison.

We further verify the quality of the FPR by comparison with *Gaia* DR3 products as well as other literature data. We show that, despite the use of a different pipeline, the median RV derived by our variability processing is entirely consistent with previous *Gaia* data.

When using the median RV of LPVs as a measure of the systemic velocity, one has to consider the following limitations. The first one is connected to the uneven time sampling of *Gaia* observations. Combined with the long periods and large amplitudes of the sources we examined, it can lead to a substantial degree of undersampling of specific phases of the variability cycle, skewing the RV distribution away from the true systemic velocity. In this case, the maximum uncertainty is given by half the peak-to-peak (true) RV variability amplitude. The occurrence of multi-periodicity might affect the median in a similar way.

The second limitation is connected with the possibly asymmetric shape of the RV curves of pulsating LPVs, in which case the median does not necessarily trace the systemic velocity, regardless of the sampling. Once again, the maximum deviation cannot be larger than half the peak-to-peak RV amplitude. Finally, a limitation could arise from the possible occurrence of template mismatches, that is the adoption, for the purpose of deriving RVs, of a spectral template whose atmospheric parameters are not suited for the target star. LPVs with C-rich chemistry, whose spectra show very distinctive molecular absorption features, are especially exposed to this risk, as all the adopted templates are for O-rich composition. In this case, it is not easy to assess the maximum deviation. However, our analysis

suggests that none of these three aspects significantly affect the RV variability parameters published here.

We then present an overview of the catalog. We first show that the distribution of the RVs on the sky is physically consistent with the Galactic rotation curve. We then analyze, in the color – absolute magnitude diagram and in the period – luminosity diagram, the distribution of a subsample with good parallaxes. The sources identified as LPVs showing a LSP and as ellipsoidal variables by our classification scheme are seen to follow the period-luminosity relations D and E, respectively, as expected. The periods of pulsating LPVs, on the other hand, are found mainly on sequences C' and C, corresponding to the first-overtone mode and the fundamental mode, respectively. Some of these sources show a 2:1 ratio between the RV period and the photometric period, which is consistent with simultaneous pulsation in those two modes. These results indicate the rich content of physical information available in both the RV and photometric time series.

This FPR includes the largest dataset of RV time series of LPVs to date. Moreover, it covers sources over a largely unexplored (in terms of RV time-series analysis) range of distances, intermediate between the extragalactic investigations of variable stars in the Magellanic Clouds and the studies of nearby LPVs. The RV time series, together with the photometric time series published in *Gaia* DR3 and spanning the same time baseline, offer an unprecedented opportunity to investigate, from different perspectives, the behavior of three different types of red giant variables, including the LSPs whose nature is still a matter of
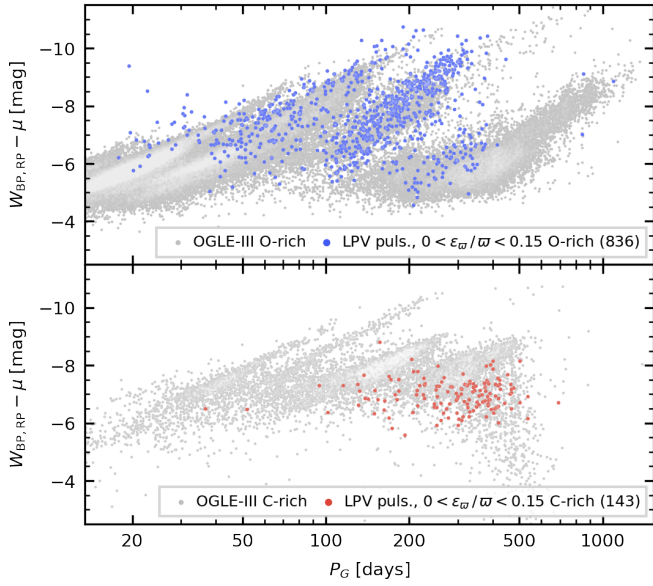
Fig. 31. Similar to the Fig. 30, but limited to pulsating LPVs and distinguishing between sources identified as C-stars (red symbols, bottom panel) or not (blue symbols, top panel) according to *Gaia* low-resolution spectra (Lebzelter et al. 2023). The numbers of these sources are indicated in the legends. Gray points in the background are O-rich (top panel) or C-rich (bottom panel) LPVs in the LMC from the OGLE-III catalog (sources whose OGLE primary period is flagged as LSP are excluded from the bottom panel). For the LMC, we adopt an average distance modulus $\mu_{LMC} = 18.49$ mag following de Grijs et al. (2017).
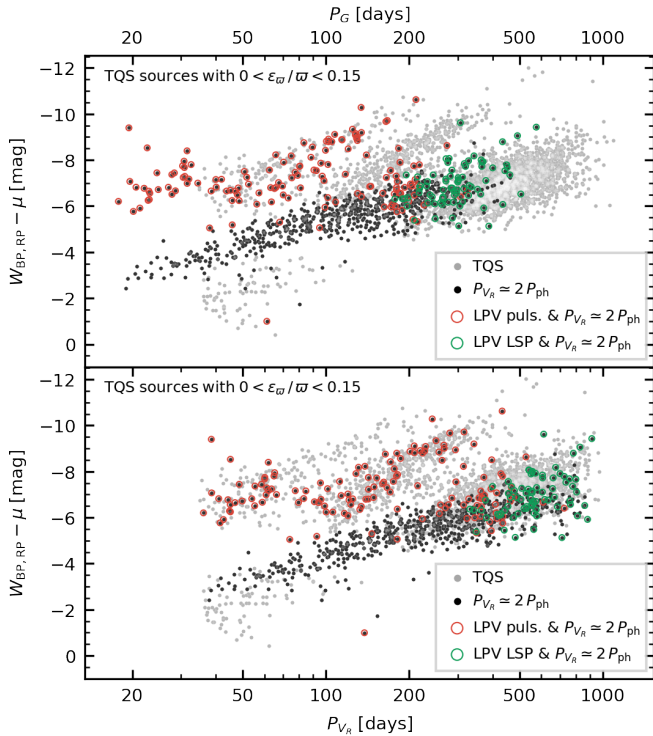
Fig. 32. Similar to the right-hand side panels in Fig. 30, but with the sources having $P_{V_R} \simeq 2P_{ph}$ highlighted in black. The PLD constructed with P_G is shown in the top panel and the one constructed with P_{V_R} in the bottom panel for the TQS sources with parallax errors better than 15% (gray symbols), and highlighting the sources with $P_{V_R} \simeq 2P_{ph}$. Red and green circles indicate LPVs showing pulsation and a LSP, respectively. Only five of them are identified as C-rich.

debate. Finally, the high-quality epoch RV data of this FPR will provide the astrophysical community with a means to prepare for the *Gaia* data release 4.

References

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, ApJS, 203, 21
Alvarez, R., Jorissen, A., Plez, B., et al. 2001, A&A, 379, 305
Albareti, F. D., Allende Prieto, C., Almeida, A., et al. 2017, ApJS, 233, 25
Astropy Collaboration (Price-Whelan, A., et al.) 2018, AJ, 156, 123
Boch, T., & Fernique, P. 2014, in Astronomical Society of the Pacific Conference Series, 485, Astronomical Data Analysis Software and Systems XXIII, eds. N. Manset, & P. Forshay, 277
Bonnarel, F., Fernique, P., Bienaymé, O., et al. 2000, A&AS, 143, 33
Breddels, M. A., & Veljanoski, J. 2018, A&A, 618, A13
Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, 150
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints
Clementini, G., Ripepi, V., Garofalo, A., et al. 2023, A&A, 674, A18
de Grijs, R., Courbin, F., Martínez-Vázquez, C. E., et al. 2017, Space Sci. Rev., 212, 1743
Eyer, L., Mowlavi, N., Evans, D. W., et al. 2017, arXiv e-prints [arXiv:1702.03295]
Eyer, L., Audard, M., Holl, B., et al. 2023, A&A, 674, A13
Fabricius, C., Høg, E., Makarov, V. V., et al. 2002, A&A, 384, 180
Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2020, ApJS, 251, 7
Gaia Collaboration (Vallenari, A., et al.) 2023, A&A, 674, A1
Gilmore, G., Randich, S., Worley, C. C., et al. 2022, A&A, in press
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Henden, A. A., Templeton, M., Terrell, D., et al. 2016, VizieR Online Data Catalogue: 2336
Hinkle, K. H., & Barnbaum, C. 1996, AJ, 111, 913
Hinkle, K. H., & Barnes, T. G. 1979, ApJ, 234, 548
Hinkle, K. H., Hall, D. N. B., & Ridgway, S. T. 1982, ApJ, 252, 697
Hinkle, K. H., Scharlach, W. W. G., & Hall, D. N. B. 1984, ApJS, 56, 1
Hinkle, K. H., Lebzelter, T., Joyce, R. R., & Fekel, F. C. 2002, AJ, 123, 1002
Hinkle, K. H., Fekel, F. C., Joyce, R. R., & Wood, P. 2013, ApJ, 770, 28
Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27
Höfner, S., & Olofsson, H. 2018, A&ARv, 26, 1
Holl, B., Fabricius, C., Portell, J., et al. 2023, A&A, 674, A25
Huber, D., Bryson, S. T., Haas, M. R., et al. 2016, ApJS, 224, 2
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, AJ, 160, 120
Joy, A. H. 1926, ApJ, 63, 281
Joy, A. H. 1954, ApJS, 1, 39
Katz, D., Sartoretti, P., Guerrier, A., et al. 2023, A&A, 674, A5
Lasker, B. M., Lattanzi, M. G., McLean, B. J., et al. 2008, AJ, 136, 735
Lebzelter, T. 1999, A&A, 351, 644
Lebzelter, T., & Hinkle, K. H. 2002, ASP Conf. Ser., 259, 556
Lebzelter, T., Hinkle, K. H., & Hron, J. 1999, A&A, 341, 224
Lebzelter, T., Hinkle, K. H., Wood, P. R., Joyce, R. R., & Fekel, F. C. 2005a, A&A, 431, 623
Lebzelter, T., Wood, P. R., Hinkle, K. H., Joyce, R. R., & Fekel, F. C. 2005b, A&A, 432, 207
Lebzelter, T., Mowlavi, N., Marigo, P., et al. 2018, A&A, 616, L13
Lebzelter, T., Trabucchi, M., Mowlavi, N., et al. 2019, A&A, 631, A24
Lebzelter, T., Mowlavi, N., Lecoeur-Taibi, I., et al. 2023, A&A, 674, A15
Luo, A. L., Zhao, Y.-H., Zhao, G., et al. 2015, Res. Astron. Astrophys., 15, 1095
Magnier, E. A., Chambers, K. C., Flewelling, H. A., et al. 2020a, ApJS, 251, 3
Magnier, E. A., Schlafly, E. F., Finkbeiner, D. P., et al. 2020b, ApJS, 251, 6
Magnier, E. A., Sweeney, W. E., Chambers, K. C., et al. 2020c, ApJS, 251, 5
Mowlavi, N. 2014, A&A, 568, A78
Mowlavi, N., Lecoeur-Taïbi, I., Holl, B., et al. 2017, A&A, 606, A92
Mowlavi, N., Lecoeur-Taïbi, I., Lebzelter, T., et al. 2018, A&A, 618, A58
Mowlavi, N., Holl, B., Lecoeur-Taïbi, I., et al. 2023, A&A, 674, A16
Nicholls, C. P., Wood, P. R., & Cioni, M. R. L. 2010, MNRAS, 405, 1770
Nie, J. D., & Wood, P. R. 2014, AJ, 148, 118
Nie, J. D., Wood, P. R., & Nicholls, C. P. 2017, ApJ, 835, 209
Nowotny, W., Höfner, S., & Aringer, B. 2010, A&A, 514, A35

Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23
Onken, C. A., Wolf, C., Bessell, M. S., et al. 2019, PASA, 36, e033
Pawlak, M. 2021, A&A, 649, A110
Pérez, F., & Granger, B. E. 2007, Comput. Sci. Eng., 9, 21
Randich, S., Gilmore, G., Magrini, L., et al. 2022, A&A, in press
R Core Team 2013, R: A Language and Environment for Statistical Computing
    (Vienna, Austria: R Foundation for Statistical Computing)
Reid, M. J., & Dickinson, D. F. 1976, ApJ, 209, 505
Reid, M. J., & Goldston, J. E. 2002, ApJ, 568, 931
Ripepi, V., Clementini, G., Molinaro, R., et al. 2023, A&A, 674, A17
Roeser, S., Demleitner, M., & Schilbach, E. 2010, AJ, 139, 2440
Saio, H., Wood, P. R., Takayama, M., & Ita, Y. 2015, MNRAS, 452, 3863
Sartoretti, P., Blomme, R., David, M., & Seabroke, G. 2022, Gaia DR3 doc-
    umentation Chapter 6: Spectroscopy, *Gaia* DR3 documentation, European
    Space Agency; *Gaia* Data Processing and Analysis Consortium. Online at
    gea.esac.esa.int/archive/documentation/GDR3/index.html, id. 6
Sartoretti, P., Marchal, O., Babusiaux, C., et al. 2023, A&A, 674, A6
Seaquist, E. R., & Taylor, A. R. 1990, ApJ, 349, 313
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Steinmetz, M., Guiglion, G., McMillan, P. J., et al. 2020a, AJ, 160, 83
Steinmetz, M., Matijevič, G., Enke, H., et al. 2020b, AJ, 160, 82
Soszyński, I., & Udalski, A. 2014, ApJ, 788, 13
Soszyński, I., Udalski, A., Kubiak, M., et al. 2004, Acta Astron., 54, 347
Soszyński, I., Dziembowski, W. A., Udalski, A., et al. 2007, Acta Astron., 57,
    201
Soszyński, I., Udalski, A., Szymański, M. K., et al. 2009, Acta Astron., 59, 239
Soszyński, I., Olechowska, A., Ratajczak, M., et al. 2021, ApJ, 911, L22
Steinmetz, M., Matijevič, G., Enke, H., et al. 2020, AJ, 160, 82
Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series,
    347, Astronomical Data Analysis Software and Systems XIV, eds. P. Shopbell,
    M. Britton, & R. Ebert, 29
Taylor, M. B. 2006, in Astronomical Society of the Pacific Conference Series,
    351, Astronomical Data Analysis Software and Systems XV, eds. C. Gabriel,
    C. Arviset, D. Ponz, & S. Enrique, 666
Trabucchi, M., Wood, P. R., Montalbán, J., et al. 2017, ApJ, 847, 139
Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., & Mateo, M. 1992, Acta
    Astron., 42, 253
van Leeuwen, F. 2007, A&A, 474, 653
Wallerstein, G. 1985, PASP, 97, 994
Waters, C. Z., Magnier, E. A., Price, P. A., et al. 2020, ApJS, 251, 4
Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, A&AS, 143, 9
Wood, P. R. 2015, MNRAS, 448, 3829
Wood, P. R. 2000, PASA, 17, 18
Wood, P. R., & Sebo, K. M. 1996, MNRAS, 282, 958
Wray, J. J., Eyer, L., & Paczyński, B. 2004, MNRAS, 349, 1059
Yu, J., Bedding, T. R., Stello, D., et al. 2020, MNRAS, 493, 1388
Zacharias, N., Finch, C. T., Girard, T. M., et al. 2013, AJ, 145, 44
Zacharias, N., Finch, C., Subasavage, J., et al. 2015, AJ, 150, 101

[1] Department of Physics and Astronomy G. Galilei, University of
    Padova, Vicolo dell'Osservatorio 3, 35122 Padova, Italy
[2] Department of Astronomy, University of Geneva, Chemin Pegasi 51,
    1290 Versoix, Switzerland
[3] University of Vienna, Department of Astrophysics, Türkenschanz-
    straße 17, 1180 Vienna, Austria
[4] Department of Astronomy, University of Geneva, Chemin
    d'Ecogia 16, 1290 Versoix, Switzerland
[5] European Space Agency (ESA), European Space Astronomy Centre
    (ESAC), Camino bajo del Castillo s/n, Urbanización Villafranca del
    Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[6] RHEA for European Space Agency (ESA), Camino bajo del Castillo
    s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada,
    28692 Madrid, Spain
[7] Sednai Sàrl, 1204 Geneva, Switzerland
[8] GEPI, Observatoire de Paris, Université PSL, CNRS, 5 place Jules
    Janssen, 92190 Meudon, France
[9] Royal Observatory of Belgium, Ringlaan 3, 1180 Brussels, Belgium
[10] Université de Strasbourg, CNRS, Observatoire astronomique de
    Strasbourg, UMR 7550, 11 rue de l'Université, 67000 Strasbourg,
    France
[11] Institut d'Astrophysique et de Géophysique, Université de Liège, 19c,
    Allée du 6 Août, 4000 Liège, Belgium
[12] CRAAG – Centre de Recherche en Astronomie, Astrophysique et
    Géophysique, Route de l'Observatoire, BP 63, Bouzareah 16340,
    Algiers, Algeria
[13] Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA
    Leiden, The Netherlands
[14] CNES Centre Spatial de Toulouse, 18 avenue Édouard Belin, 31401
    Toulouse Cedex 9, France
[15] Mullard Space Science Laboratory, University College London,
    Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
[16] INAF – Osservatorio astronomico di Padova, Vicolo Osservatorio 5,
    35122 Padova, Italy
[17] European Space Agency (ESA), European Space Research and
    Technology Centre (ESTEC), Keplerlaan 1, 2201AZ, Noordwijk,
    The Netherlands
[18] Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France
[19] Astronomisches Rechen-Institut, Zentrum für Astronomie der
    Universität Heidelberg, Mönchhofstr. 12–14, 69120 Heidelberg,
    Germany
[20] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Lab-
    oratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice
    Cedex 4, France
[21] Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS,
    B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[22] Institute of Astronomy, University of Cambridge, Madingley Road,
    Cambridge CB3 0HA, UK
[23] Aurora Technology for European Space Agency (ESA), Camino bajo
    del Castillo s/n, Urbanización Villafranca del Castillo, Villanueva de
    la Cañada, 28692 Madrid, Spain
[24] Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona
    (UB), Martí i Franquès 1, 08028 Barcelona, Spain
[25] Departament de Física Quàntica i Astrofísica (FQA), Universi-
    tat de Barcelona (UB), c. Martí i Franquès 1, 08028 Barcelona,
    Spain
[26] Institut d'Estudis Espacials de Catalunya (IEEC), c. Gran Capità 2–
    4, 08034 Barcelona, Spain
[27] Lohrmann Observatory, Technische Universität Dresden, Momm-
    senstraße 13, 01062 Dresden, Germany
[28] Lund Observatory, Division of Astrophysics, Department of
    Physics, Lund University, Box 43, 22100 Lund, Sweden
[29] INAF – Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5,
    50125 Firenze, Italy
[30] Nicolaus Copernicus Astronomical Center, Polish Academy of
    Sciences, ul. Bartycka 18, 00-716 Warsaw, Poland
[31] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidel-
    berg, Germany
[32] INAF – Osservatorio Astrofisico di Torino, via Osservatorio 20,
    10025 Pino Torinese (TO), Italy
[33] University of Turin, Department of Physics, Via Pietro Giuria 1,
    10125 Torino, Italy
[34] DAPCOM Data Services, c. dels Vilabella 5–7, 80500 Vic,
    Barcelona, Spain
[35] ALTEC S.p.a, Corso Marche 79,10146 Torino, Italy
[36] Gaia DPAC Project Office, ESAC, Camino bajo del Castillo s/n,
    Urbanización Villafranca del Castillo, Villanueva de la Cañada,
    28692 Madrid, Spain
[37] Telespazio UK S.L. for European Space Agency (ESA), Camino
    bajo del Castillo s/n, Urbanización Villafranca del Castillo,
    Villanueva de la Cañada, 28692 Madrid, Spain
[38] SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne
    Université, LNE, 61 avenue de l'Observatoire, 75014 Paris, France
[39] IMCCE, Observatoire de Paris, Université PSL, CNRS, Sorbonne
    Université, Univ. Lille, 77 av. Denfert-Rochereau, 75014 Paris,
    France
[40] Serco Gestión de Negocios for European Space Agency (ESA),
    Camino bajo del Castillo s/n, Urbanización Villafranca del Castillo,
    Villanueva de la Cañada, 28692 Madrid, Spain
[41] INAF – Osservatorio di Astrofisica e Scienza dello Spazio di
    Bologna, via Piero Gobetti 93/3, 40129 Bologna, Italy

[42] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

[43] CIGUS CITIC – Department of Computer Science and Information Technologies, University of A Coruña, Campus de Elviña s/n, A Coruña, 15071, Spain

[44] ATG Europe for European Space Agency (ESA), Camino bajo del Castillo s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[45] Kavli Institute for Cosmology Cambridge, Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

[46] Department of Astrophysics, Astronomy and Mechanics, National and Kapodistrian University of Athens, Panepistimiopolis, Zografos, 15783 Athens, Greece

[47] Donald Bren School of Information and Computer Sciences, University of California, Irvine, CA 92697, USA

[48] CENTRA, Faculdade de Ciências, Universidade de Lisboa, Edif. C8, Campo Grande, 1749-016 Lisboa, Portugal

[49] INAF – Osservatorio Astrofisico di Catania, via S. Sofia 78, 95123 Catania, Italy

[50] Dipartimento di Fisica e Astronomia "Ettore Majorana", Università di Catania, Via S. Sofia 64, 95123 Catania, Italy

[51] INAF – Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monte Porzio Catone (Roma), Italy

[52] Space Science Data Center – ASI, Via del Politecnico SNC, 00133 Roma, Italy

[53] Department of Physics, University of Helsinki, PO Box 64, 00014 Helsinki, Finland

[54] Finnish Geospatial Research Institute FGI, Vuorimiehentie 5, 02150 Espoo, Finland

[55] Institut UTINAM CNRS UMR6213, Université de Franche-Comté, OSU THETA Franche-Comté Bourgogne, Observatoire de Besançon, BP1615, 25010 Besançon Cedex, France

[56] HE Space Operations BV for European Space Agency (ESA), Keplerlaan 1, 2201AZ, Noordwijk, The Netherlands

[57] Dpto. de Inteligencia Artificial, UNED, c/ Juan del Rosal 16, 28040 Madrid, Spain

[58] Institut d'Astronomie et d'Astrophysique, Université Libre de Bruxelles CP 226, Boulevard du Triomphe, 1050 Brussels, Belgium

[59] Leibniz Institute for Astrophysics Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany

[60] Konkoly Observatory, Research Centre for Astronomy and Earth Sciences, Eötvös Loránd Research Network (ELKH), MTA Centre of Excellence, Konkoly Thege Miklós út 15–17, 1121 Budapest, Hungary

[61] ELTE Eötvös Loránd University, Institute of Physics, 1117, Pázmány Péter sétány 1A, Budapest, Hungary

[62] Instituut voor Sterrenkunde, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium

[63] Department of Astrophysics/IMAPP, Radboud University, PO Box 9010, 6500 GL Nijmegen, The Netherlands

[64] Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland

[65] Quasar Science Resources for European Space Agency (ESA), Camino bajo del Castillo s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[66] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Edif. C6, Campo Grande, 1749-016 Lisboa, Portugal

[67] School of Physics and Astronomy , University of Leicester, University Road, Leicester LE1 7RH, UK

[68] School of Physics and Astronomy, Tel Aviv University, Tel Aviv 6997801, Israel

[69] Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE, UK

[70] Telespazio for CNES Centre Spatial de Toulouse, 18 avenue Édouard Belin, 31401 Toulouse Cedex 9, France

[71] National Observatory of Athens, I. Metaxa and Vas. Pavlou, Palaia Penteli, 15236 Athens, Greece

[72] Depto. Estadística e Investigación Operativa. Universidad de Cádiz, Avda. República Saharaui s/n, 11510 Puerto Real, Cádiz, Spain

[73] EURIX S.r.l., Corso Vittorio Emanuele II 61, 10128 Torino, Italy

[74] Porter School of the Environment and Earth Sciences, Tel Aviv University, Tel Aviv 6997801, Israel

[75] ATOS for CNES Centre Spatial de Toulouse, 18 avenue Édouard Belin, 31401 Toulouse Cedex 9, France

[76] HE Space Operations BV for European Space Agency (ESA), Camino bajo del Castillo s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[77] LFCA/DAS, Universidad de Chile, CNRS, Casilla 36-D, Santiago, Chile

[78] SISSA – Scuola Internazionale Superiore di Studi Avanzati, via Bonomea 265, 34136 Trieste, Italy

[79] University of Turin, Department of Computer Sciences, Corso Svizzera 185, 10149 Torino, Italy

[80] Thales Services for CNES Centre Spatial de Toulouse, 18 avenue Édouard Belin, 31401 Toulouse Cedex 9, France

[81] Dpto. de Matemática Aplicada y Ciencias de la Computación, Univ. de Cantabria, ETS Ingenieros de Caminos, Canales y Puertos, Avda. de los Castros s/n, 39005 Santander, Spain

[82] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain

[83] Port d'Informació Científica (PIC), Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain

[84] Instituto de Astrofísica, Universidad Andres Bello, Fernandez Concha 700, Las Condes, Santiago RM, Chile

[85] Centre for Astrophysics Research, University of Hertfordshire, College Lane, AL10 9AB, Hatfield, UK

[86] University of Turin, Mathematical Department "G. Peano", Via Carlo Alberto 10, 10123 Torino, Italy

[87] University of Antwerp, Onderzoeksgroep Toegepaste Wiskunde, Middelheimlaan 1, 2020 Antwerp, Belgium

[88] INAF – Osservatorio Astronomico d'Abruzzo, Via Mentore Maggini, 64100 Teramo, Italy

[89] Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Rua do Matão 1226, Cidade Universitaria, 05508-900 São Paulo, SP, Brazil

[90] Mésocentre de calcul de Franche-Comté, Université de Franche-Comté, 16 route de Gray, 25030 Besançon Cedex, France

[91] Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia

[92] Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, UK

[93] Data Science and Big Data Lab, Pablo de Olavide University, 41013 Seville, Spain

[94] Institute of Astrophysics, FORTH, Crete, Greece

[95] Barcelona Supercomputing Center (BSC), Plaça Eusebi Güell 1–3, 08034 Barcelona, Spain

[96] ETSE Telecomunicación, Universidade de Vigo, Campus Lagoas-Marcosende, 36310 Vigo, Galicia, Spain

[97] F.R.S.-FNRS, Rue d'Egmont 5, 1000 Brussels, Belgium

[98] Asteroid Engineering Laboratory, Luleå University of Technology, Box 848, 981 28 Kiruna, Sweden

[99] Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands

[100] IAC – Instituto de Astrofisica de Canarias, Via Láctea s/n, 38200 La Laguna S.C., Tenerife, Spain

[101] Department of Astrophysics, University of La Laguna, Via Láctea s/n, 38200 La Laguna S.C., Tenerife, Spain

[102] Astronomical Observatory, University of Warsaw, Al. Ujazdowskie 4, 00-478 Warszawa, Poland

[103] Research School of Astronomy & Astrophysics, Australian National University, Cotter Road, Weston, ACT 2611, Australia

[104] European Space Agency (ESA retired), European Space Research and Technology Centre (ESTEC), Keplerlaan 1, 2201AZ, Noordwijk, The Netherlands

[105] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92190 Meudon, France

[106] Université Rennes, CNRS, IPR (Institut de Physique de Rennes) – UMR 6251, 35000 Rennes, France

[107] INAF – Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131, Napoli, Italy

[108] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, PR China

[109] University of Chinese Academy of Sciences, No.19(A) Yuquan Road, Shijingshan District, Beijing 100049, PR China

[110] São Paulo State University, Grupo de Dinâmica Orbital e Planetologia, CEP 12516-410, Guaratinguetá, SP, Brazil

[111] Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark

[112] DXC Technology, Retortvej 8, 2500 Valby, Denmark

[113] Las Cumbres Observatory, 6740 Cortona Drive Suite 102, Goleta, CA 93117, USA

[114] CIGUS CITIC, Department of Nautical Sciences and Marine Engineering, University of A Coruña, Paseo de Ronda 51, 15071, A Coruña, Spain

[115] Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK

[116] IRAP, Université de Toulouse, CNRS, UPS, CNES, 9 Av. colonel Roche, BP 44346, 31028 Toulouse Cedex 4, France

[117] MTA CSFK Lendület Near-Field Cosmology Research Group, Konkoly Observatory, MTA Research Centre for Astronomy and Earth Sciences, Konkoly Thege Miklós út 15-17, 1121 Budapest, Hungary

[118] Pervasive Technologies s.l., c. Saragossa 118, 08006 Barcelona, Spain

[119] School of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK

[120] Villanova University, Department of Astrophysics and Planetary Science, 800 E Lancaster Avenue, Villanova, PA 19085, USA

[121] Departmento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid, 28040 Madrid, Spain

[122] INAF – Osservatorio Astronomico di Brera, via E. Bianchi, 46, 23807 Merate (LC), Italy

[123] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

[124] Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 7610001, Israel

[125] Centre de Données Astronomique de Strasbourg, Strasbourg, France

[126] University of Exeter, School of Physics and Astronomy, Stocker Road, Exeter EX2 7SJ, UK

[127] Departamento de Astrofísica, Centro de Astrobiología (CSIC-INTA), ESA-ESAC. Camino Bajo del Castillo s/n, 28692 Villanueva de la Cañada, Madrid, Spain

[128] naXys, Department of Mathematics, University of Namur, Rue de Bruxelles 61, 5000 Namur, Belgium

[129] INAF–Osservatorio Astronomico di Trieste, via G.B. Tiepolo 11, 34131 Trieste, Italy

[130] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., MS 15, Cambridge, MA 02138, USA

[131] H H Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK

[132] Escuela de Arquitectura y Politécnica – Universidad Europea de Valencia, Spain

[133] Escuela Superior de Ingeniería y Tecnología – Universidad Internacional de la Rioja, Spain

[134] Applied Physics Department, Universidade de Vigo, 36310 Vigo, Spain

[135] Instituto de Física e Ciencias Aeroespaciais (IFCAE), Universidade de Vigo, Á Campus de As Lagoas, 32004 Ourense, Spain

[136] Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, PR China

[137] Sorbonne Université, CNRS, UMR7095, Institut d'Astrophysique de Paris, 98bis bd. Arago, 75014 Paris, France

[138] Faculty of Mathematics and Physics, University of Ljubljana, Jadranska ulica 19, 1000 Ljubljana, Slovenia

[139] Institute of Mathematics, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Appendix A: Additional details on the classification of the catalog content

This FPR contains stellar sources displaying photometric and RV variability of different origins. In Sect. 3 we separated ellipsoidal variables from LPVs by comparing the amplitudes of their RV and *G*-band time series models. Then, among the LPV candidates, we found periods that result from pulsation as well as LSPs, and we tentatively distinguished the two types of variability in the *G*-band period-amplitude diagram. The separation of these two variability types in this diagram is not as clean as that between LPVs and ELLs in the *G* versus RV amplitude diagram. Therefore we examined in some more detail the two groups, in order to corroborate our classification.

We recall that, in absence of accurate and homogeneous knowledge about absolute brightness (i.e., distance), we have to rely on the variability parameters period and amplitude for this purpose. In addition to the variability parameters emerging from the analysis of the photometric time series, we can inspect the ones resulting from the RV curves. A combination of these four quantities is shown in Fig. A.1 (top panel), which displays the ratio between the semi-amplitude and the period of the *G*-band time series model against the same ratio for the RV time series model. For simplicity, we explicitly show the decimal logarithm of both quantities. There exists a clear correlation between the two ratios, indicating that the occurrence of a long period characterized by a small photometric amplitude is often associated with a relatively small RV amplitude as well.

The density contour lines in the top panel of Fig. A.1, as well as the histograms along the axes of the diagram, show that there are two regions of overdensity in the diagram. In order to better visualize them, in the bottom panel of Fig. A.1 we replaced the quantity on the vertical axis with its distance from the arbitrary reference line shown in the top panel. Fig. A.1 shows in two different colors the sources classified on the basis of Eq. (10). The latter is consistent with the distributions highlighted here with the inclusion of information related to RV variability, although it appears to lead to some contamination. It is possible that a classification that makes use of the periods and amplitudes resulting from both photometric and RV time series could help reduce such contamination. However, it would be more complex and less intuitive from a physical point of view, which supports the choice made in Sect. 3.

Appendix B: Median RV in unevenly sampled time series

In this section we present a few representative examples of sources showing relatively large differences between the median RV and the zero-point RV, so that the former is not necessarily an accurate indicator of the center-of-mass velocity. The RV curves displayed in Fig. B.1 belong to sources with relatively long periods, for which *Gaia* observations cover only a few cycles. They also have large RV amplitudes and highly sinusoidal variations, indicating that their variability is due to binarity, and leaving little doubt concerning the authenticity of its physical origin. In all the examples displayed, the distribution of RV epochs is asymmetric as the phase near minimum RV is oversampled compared with other phases of the variability cycle. As a result, the median RV is underestimated, and shows a large difference with respect to V_R^0.

Further examples are provided in Fig. B.2, illustrating how the median of the RV time series ceases to be representative of the systemic RV due to the occurrence of clusters of RV epochs
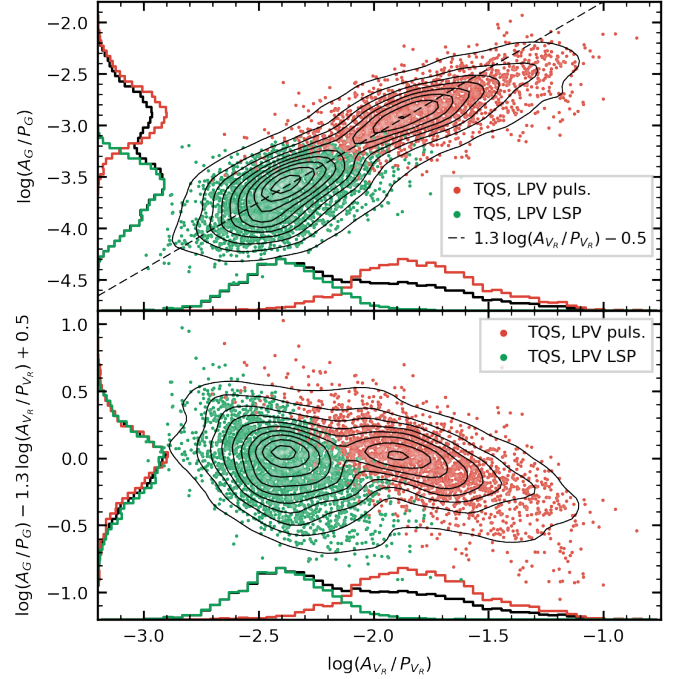
Fig. A.1. *G*-band versus RV period-amplitude ratios for the TQS sources classified as LPVs. Black contour lines represent a smoothed density map of the sample, and correspond to the black histograms on the sides of each panel. Each source is color-coded according to whether its period is identified as resulting from pulsation (red) or as an LSP (green). The same color-code is used for the histograms on the sides (normalized to their maximum). In the bottom panel the quantity $\log(A_G/P_G)$ is replaced with its distance from the dashed line displayed in the top panel, taken as an arbitrary reference.

spanning an interval of time much shorter than the variability period of a given source. The RV distributions of the example time series are dominated by the effect of these groups causing large discrepancies between $\langle V_R \rangle$ and V_R^0, except in the fortuitous circumstances in which the clusters end up near the mid-point of the RV curve or cancel out with each other, as illustrated in panels (d) and (e) of Fig. B.2, respectively.

Overall, there exist rather specific circumstances in which the median $\langle V_R \rangle$ cannot be considered a good indicator of the systemic RV. These arise primarily from the uneven and irregular time sampling of *Gaia* observations, as well as the relatively small number of epochs, and affect mainly the sources with large RV amplitudes and long periods. In particular, when only a few variability cycles are covered, even a relatively homogeneous phase coverage can be insufficient to obtain a well-centered median value.

Appendix C: Numerical differences with respect to *Gaia* DR3

The production of this FPR required running the LPV variability pipeline on a subset of the sample of LPV candidates published as part of *Gaia* DR3. Despite having run the same operations on the photometric time series, we detected numerical differences with respect to the results previously obtained and published in *Gaia* DR3. We traced the origin of these numerical differences to the execution of the Apache Math Commons `LevenbergMarquardtOptimizer`, which produces different results depending on the adopted Java Development Kit
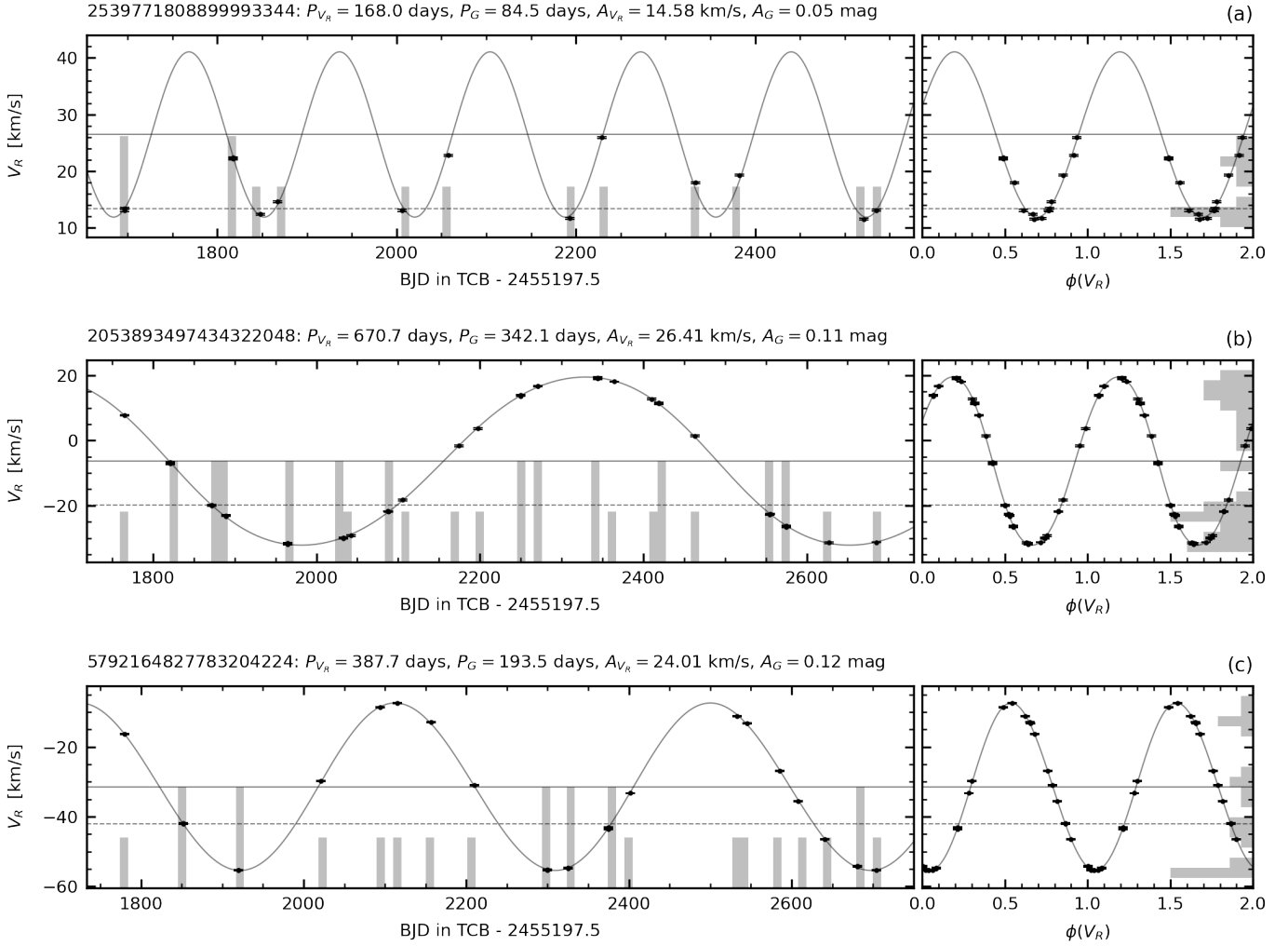
Fig. B.1. Examples of RV time series in which only a small number of cycles is covered, and the minimum phase is oversampled compared to other phases, leading the median RV to underestimate the systemic RV. In each row, the left panel shows the RV time series and best-fit model, whereas their folded counterparts are displayed in the right panel. Histograms in both panels aid to visualize the distribution of measurements both in time (to identify clustered measurements) and in RV. The solid and dashed lines mark the values of the zero-point and median RVs, respectively.

(JDK) version. The differences arised after updating from JDK-8 to JDK-17.

The differences are systematically reproducible[3], and a bug report was submitted to the Oracle bug submission system late in 2022, without receiving any feedback. The most plausible explanation we found for the numerical differences is that JDK-17 enforces the floating point standard IEEE 754, whereas in JDK-8 the runtime could decide to deviate from this standard in order to optimize the generated code[4].

In order to assess the impact of this bug we consider the periods of the G-band time series resulting from the variability pipeline before and after upgrading to JDK-17, and examine the differences between them. We use for this purpose the sample adopted for constructing the FPR catalog after pre-filtering that consists of 110 654 sources. From this set we exclude 1 413 sources for which either of the two periods is longer than the duration of the G-band time series. Within this reference set,

50% of the time series result in the exact same period regardless of the JDK version employed in the pipeline.

In Fig. C.1 we display the distribution of the relative difference between the JDK-8 and JDK-17 periods for the sources that have a non-null difference. It shows three well-distinct groups. About 60% of the time series display relative period differences that are at the level of machine precision, and thus entirely negligible ($< 10^{-13}$, and possibly zero). A second group contains sources that display relative differences between 10^{-13} and 10^{-5}, and consists of about 40% of the sample. In this case the absolute differences are on the order of a few minutes for periods of 1000 days, or a few seconds for periods of 10 days, hence they are negligible as well. Finally, there exists a third group containing fewer than 1% of the examined dataset, whose sources display relative period differences above 10^{-5}, but only 28 of them show a discrepancy above the 10% level (at most as large as 35%).

We examined the distribution of G-band time series statistics separately for these three groups, looking for systematic effects and features that might be triggering a strong effect of the bug. However, given the nature of this bug, one could expect it to impact different sources in a random fashion, and to be independent of the specific properties of the time series being processed.

---

[3] See the github repository https://github.com/gjevardat/ReproduceFloatingPointBug for an executable proof of the numerical differences.

[4] See https://openjdk.org/jeps/306 for more details.
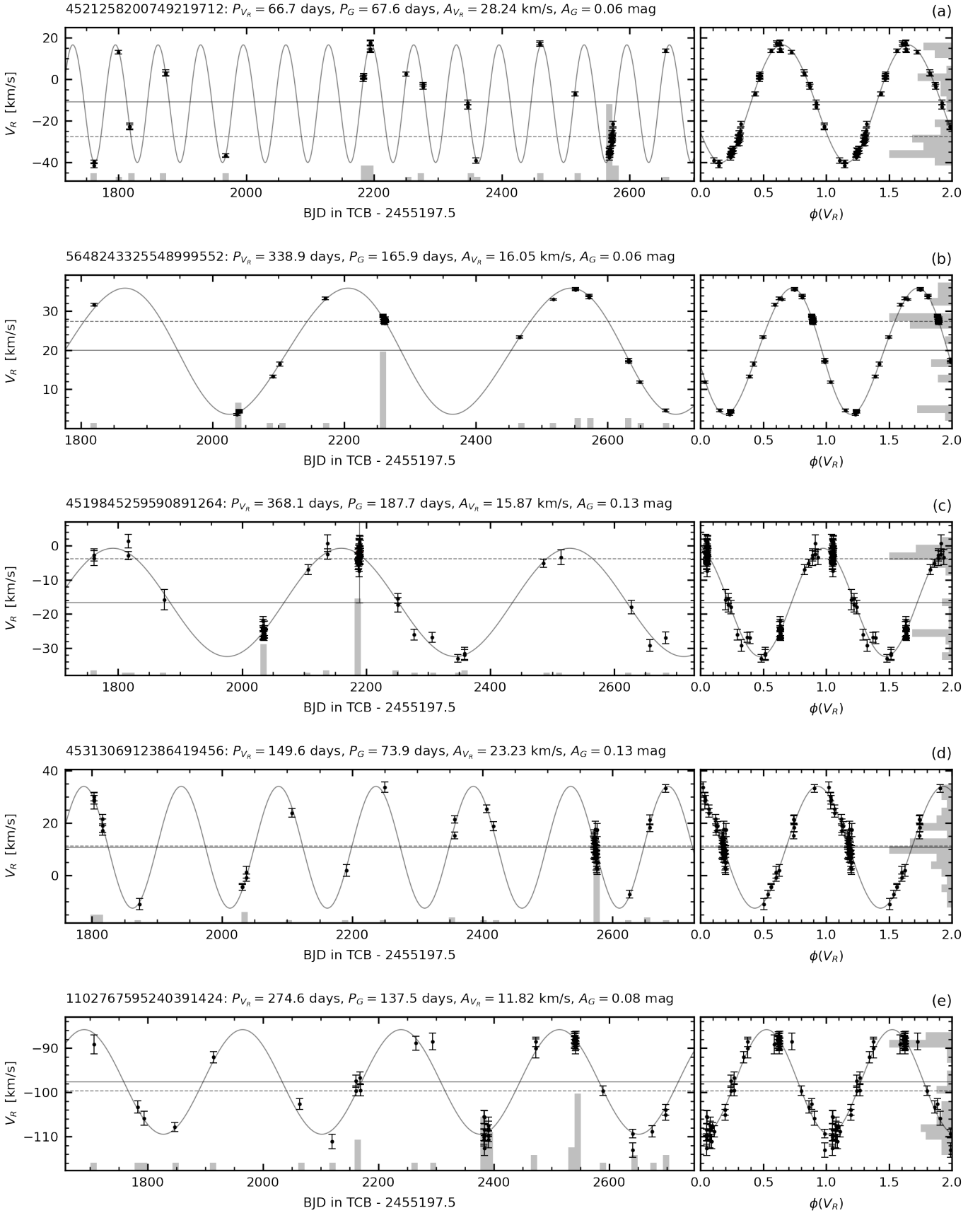
Fig. B.2. Similar to Fig. B.1, but showing examples of time series dominated by a large group of RV epochs spanning an interval of time much shorter than the typical period of the source, thereby distorting the time series statistics and causing a large difference between $\langle V_R \rangle$ and V_R^0. Cases (d) and (e) show examples where the clustered data points are, by chance, either located near the mid-point of the RV time series (case d), or cancel out with each other (case e).
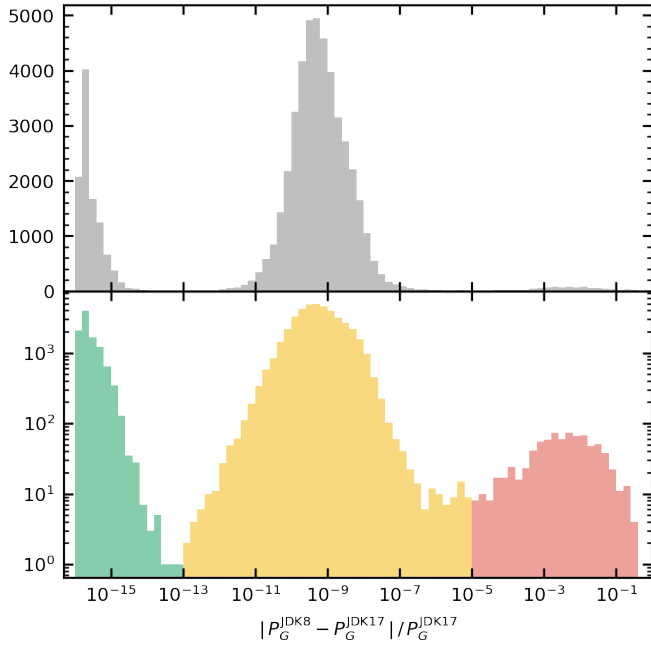
Fig. C.1. Distribution of relative differences between the period derived for *G*-band time series before and after the upgrade of the variability pipeline from JDK-8 to JDK-17. A linear scale is used in the top panel for the vertical axis, while a log-scale is used in the bottom panel. Three distinct groups are identified, colored in the bottom panel in green, orange and red, corresponding to various levels of difference. We note that only in the latter group the differences are non-negligible (being typically of order 1%). This group is so small that it is barely visible in the top panel.

Our analysis tends to confirm this expectation. Overall, all three groups of sources show similar distributions of number of observations, time series duration, and average brightness. Only the parameters related with the amplitude and the period itself display a connection with the strength of the effect of the bug. In particular, there is a slight tendency for the third group (shown in red in Figs. C.2 and C.3) to lack sources with relatively small amplitudes and short periods, whereas the opposite is true for the first (green) group. Moreover, there is a clear correlation between the relative difference and the period within each group. The most likely explanation is that these sources are more exposed to numerical instabilities in the period determination as their time series cover a small number of cycles, which makes the period difficult to constrain in any case.

Appendix D: Catalog retrieval

Here are two examples on how to retrieve LPV-related data from the *Gaia* archive. A query like the following allows for retrieving selected quantities from the source catalog and LPV FPR catalog for all LPV FPR sources.

```
SELECT gs.source_id,
    gs.ra, gs.dec,
    fprlpv.frequency_rv, fprlpv.frequency_rv_error,
    fprlpv.amplitude_rv, fprlpv.flag_rv,
    ...
FROM gaiadr3.gaia_source AS gs
LEFT JOIN gaiafpr.vari_long_period_variable AS fprlpv
  ON gs.source_id = fprlpv.source_id
WHERE fprlpv.source_id IS NOT NULL
```
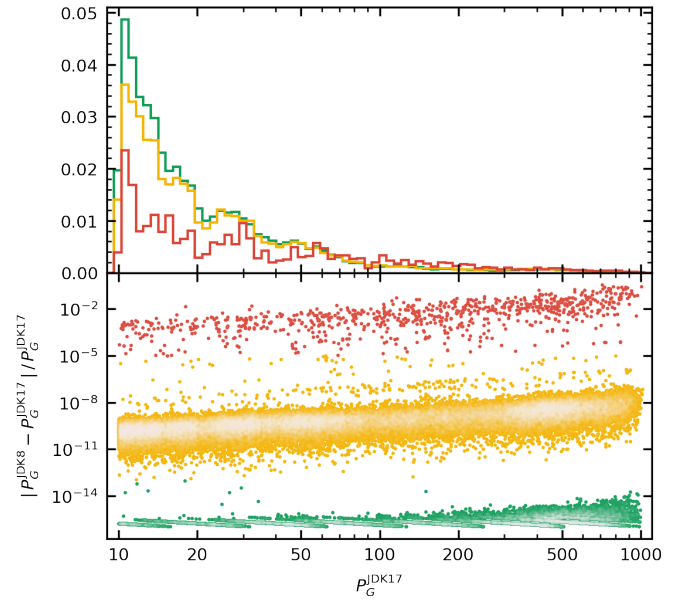

Fig. C.2. Period distribution for the three groups shown in Fig. C.1 (top panel), and relation of the period with the relative period difference. The period shown is the one obtained after the upgrade to JDK-17. The same color scheme as in Fig. C.1 is adopted. The histograms are normalized by area.
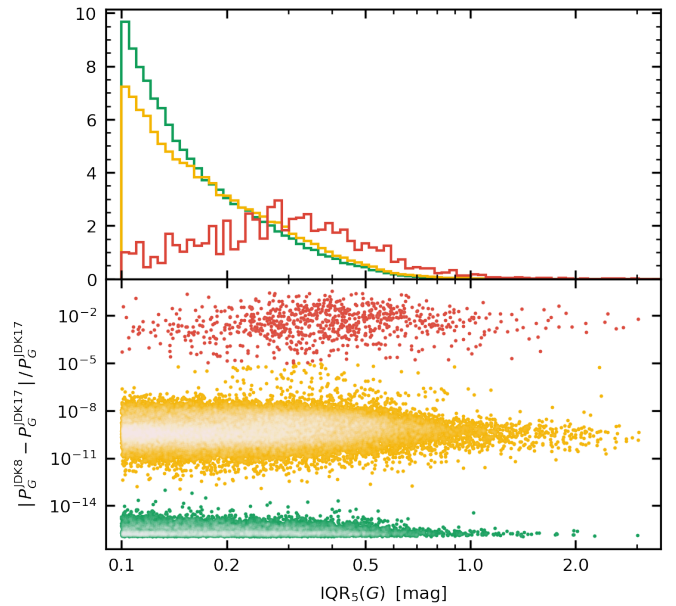
Fig. C.3. Similar to Fig. C.2, but for the 5–95% interquantile range of the *G*-band time series.

In order to retrieve all data from the source table, the LPV FPR table, the RV statistics table, and the tables of data on LPVs from DR3 for all sources published in the LPV FPR, a query like the following can be used.

```
SELECT src.*,
    fprlpv.*,
    rvstat.*,
    lpv.*,
    varistat.*
FROM gaiadr3.gaia_source AS src
LEFT JOIN gaiafpr.vari_long_period_variable AS fprlpv
```

```
    ON src.source_id = fprlpv.source_id
LEFT JOIN gaiafpr.vari_rad_vel_statistics AS rvstat
    ON src.source_id = rvstat.source_id
LEFT JOIN gaiadr3.vari_long_period_variable AS lpv
    ON src.source_id = lpv.source_id
LEFT JOIN gaiadr3.vari_summary AS varistat
    ON src.source_id = varistat.source_id
WHERE fprlpv.source_id IS NOT NULL
```

Finally, one can download all the RV time series published as part of this FPR using a source like the following.

```
SELECT fprlpv.source_id,
    varirv.transit_id,
    varirv.rv_obs_time,
    varirv.radial_velocity,
    varirv.radial_velocity_error
    varirv.rejected_by_variability
FROM gaiafpr.vari_long_period_variable AS fprlpv
LEFT JOIN gaiafpr.vari_epoch_radial_velocity AS varirv
    ON fprlpv.source_id = varirv.source_id
WHERE fprlpv.source_id IS NOT NULL
```

We also recall that it is possible to download ancillary data such as the RV and photometric time series using the *Gaia* datalink service[5].

Appendix E: Acknowledgements