Gaia Data Release 3: a golden sample of astrophysical parameters

Creevey, O.L.; Sarro, L.M.; Lobel, A.; Pancino, E.; Andrae, R.; Smart, R.L.; ... ; Liao, S. et al.

Astronomy & Astrophysics manuscript no. gold
June 14, 2022

Gaia Data Release 3: A Golden Sample of Astrophysical Parameters⋆

Gaia Collaboration, O.L. Creevey[1], L.M. Sarro[2], A. Lobel[3], E. Pancino[4,5], R. Andrae[6], R.L. Smart[7], G. Clementini[8], U. Heiter[9], A.J. Korn[9], M. Fouesneau[6], Y. Frémat[3], F. De Angeli[10], A. Vallenari[11], D.L. Harrison[10,12], F. Thévenin[1], C. Reylé[13], R. Sordo[11], A. Garofalo[8], A.G.A. Brown[14], L. Eyer[15], T. Prusti[16], J.H.J. de Bruijne[16], F. Arenou[17], C. Babusiaux[18,17], M. Biermann[19], C. Ducourant[20], D.W. Evans[10], R. Guerra[21], A. Hutton[22], C. Jordi[23], S.A. Klioner[24], U.L. Lammers[21], L. Lindegren[25], X. Luri[23], F. Mignard[1], C. Panem[26], D. Pourbaix†[27,28], S. Randich[4], P. Sartoretti[17], C. Soubiran[20], P. Tanga[1], N.A. Walton[10], C.A.L. Bailer-Jones[6], U. Bastian[19], R. Drimmel[7], F. Jansen[29], D. Katz[17], M.G. Lattanzi[7,30], F. van Leeuwen[10], J. Bakker[21], C. Cacciari[8], J. Castañeda[31], C. Fabricius[23], L. Galluccio[1], A. Guerrier[26], E. Masana[23], R. Messineo[32], N. Mowlavi[15], C. Nicolas[26], K. Nienartowicz[33,34], F. Pailler[26], P. Panuzzo[17], F. Riclet[26], W. Roux[26], G.M. Seabroke[35], G. Gracia-Abril[36,19], J. Portell[23], D. Teyssier[37], M. Altmann[19,38], M. Audard[15,34], I. Bellas-Velidis[39], K. Benson[35], J. Berthier[40], R. Blomme[3], P.W. Burgess[10], D. Busonero[7], G. Busso[10], H. Cánovas[37], B. Carry[1], A. Cellino[7], N. Cheek[41], Y. Damerdji[42,43], M. Davidson[44], P. de Teodoro[21], M. Nuñez Campos[22], L. Delchambre[42], A. Dell'Oro[4], P. Esquej[45], J. Fernández-Hernández[46], E. Fraile[45], D. Garabato[47], P. García-Lario[21], E. Gosset[42,28], R. Haigron[17], J.-L. Halbwachs[48], N.C. Hambly[44], J. Hernández[21], D. Hestroffer[40], S.T. Hodgkin[10], B. Holl[15,34], K. Janßen[49], G. Jevardat de Fombelle[15], S. Jordan[19], A. Krone-Martins[50,51], A.C. Lanzafame[52,53], W. Löffler[19], O. Marchal[48], P.M. Marrese[54,5], A. Moitinho[50], K. Muinonen[55,56], P. Osborne[10], T. Pauwels[3], A. Recio-Blanco[1], M. Riello[10], L. Rimoldini[34], T. Roegiers[57], J. Rybizki[6], C. Siopis[27], M. Smith[35], A. Sozzetti[7], E. Utrilla[22], M. van Leeuwen[10], U. Abbas[7], P. Ábrahám[58,59], A. Abreu Aramburu[46], C. Aerts[60,61,6], J.J. Aguado[2], M. Ajaj[17], F. Aldea-Montero[21], G. Altavilla[54,5], M.A. Álvarez[47], J. Alves[62], F. Anders[23], R.I. Anderson[63], E. Anglada Varela[46], T. Antoja[23], D. Baines[37], S.G. Baker[35], L. Balaguer-Núñez[23], E. Balbinot[64], Z. Balog[19,6], C. Barache[38], D. Barbato[15,7], M. Barros[50], M.A. Barstow[65], S. Bartolomé[23], J.-L. Bassilana[66], N. Bauchet[17], U. Becciani[52], M. Bellazzini[8], A. Berihuete[67], M. Bernet[23], S. Bertone[68,69,7], L. Bianchi[70], A. Binnenfeld[71], S. Blanco-Cuaresma[72], T. Boch[48], A. Bombrun[73], D. Bossini[74], S. Bouquillon[38,75], A. Bragaglia[8], L. Bramante[32], E. Breedt[10], A. Bressan[76], N. Brouillet[20], E. Brugaletta[52], B. Bucciarelli[7,30], A. Burlacu[77], A.G. Butkevich[7], R. Buzzi[7], E. Caffau[17], R. Cancelliere[78], T. Cantat-Gaudin[23,6], R. Carballo[79], T. Carlucci[38], M.I. Carnerero[7], J.M. Carrasco[23], L. Casamiquela[20,17], M. Castellani[54], A. Castro-Ginard[14], L. Chaoul[26], P. Charlot[20], L. Chemin[80], V. Chiaramida[32], A. Chiavassa[1], N. Chornay[10], G. Comoretto[37,81], G. Contursi[1], W.J. Cooper[82,7], T. Cornez[66], S. Cowell[10], F. Crifo[17], M. Cropper[35], M. Crosta[7,83], C. Crowley[73], C. Dafonte[47], A. Dapergolas[39], P. David[40], P. de Laverny[1], F. De Luise[84], R. De March[32], J. De Ridder[60], R. de Souza[85], A. de Torres[73], E.F. del Peloso[19], E. del Pozo[22], M. Delbo[1], A. Delgado[45], J.-B. Delisle[15], C. Demouchy[86], T.E. Dharmawardena[6], P. Di Matteo[17], S. Diakite[87], C. Diener[10], E. Distefano[52], C. Dolding[35], H. Enke[49], C. Fabre[88], M. Fabrizio[54,5], S. Faigler[89], G. Fedorets[55,90], P. Fernique[48,91], F. Figueras[23], Y. Fournier[49], C. Fouron[77], F. Fragkoudi[92,93,94], M. Gai[7], A. Garcia-Gutierrez[23], M. Garcia-Reinaldos[21], M. García-Torres[95], A. Gavel[9], P. Gavras[45], E. Gerlach[24], R. Geyer[24], P. Giacobbe[7], G. Gilmore[10], S. Girona[96], G. Giuffrida[54], R. Gomel[89], A. Gomez[47], J. González-Núñez[41,97], I. González-Santamaría[47], J.J. González-Vidal[23], M. Granvik[55,98], P. Guillout[48], J. Guiraud[26], R. Gutiérrez-Sánchez[37], L.P. Guy[34,99], D. Hatzidimitriou[100,39], M. Hauser[6,101], M. Haywood[17], A. Helmer[66], A. Helmi[64], T. Hilger[24], M.H. Sarmiento[22], S.L. Hidalgo[102,103], N. Hładczuk[21,104], D. Hobbs[25], G. Holland[10], H.E. Huckle[35], K. Jardine[105], G. Jasniewicz[106], A. Jean-Antoine Piccolo[26], Ó. Jiménez-Arranz[23], J. Juaristi Campillo[19], F. Julbe[23], L. Karbevska[34,107], P. Kervella[108], S. Khanna[64,7], G. Kordopatis[1], Á Kóspál[58,6,59], Z. Kostrzewa-Rutkowska[14,109], K. Kruszyńska[110], M. Kun[58], P. Laizeau[111], S. Lambert[38], A.F. Lanza[52], Y. Lasne[66], J.-F. Le Campion[20], Y. Lebreton[108,112], T. Lebzelter[62], S. Leccia[113], N. Leclerc[17], I. Lecoeur-Taibi[34], S. Liao[114,7,115], E.L. Licata[7], H.E.P. Lindstrøm[7,116,117], T.A. Lister[118], E. Livanou[100], A. Lorca[22], C. Loup[48], P. Madrero Pardo[23], A. Magdaleno Romeo[77], S. Managau[66], R.G. Mann[44], M. Manteiga[119], J.M. Marchant[120], M. Marconi[113], J. Marcos[37], M.M.S. Marcos Santos[41], D. Marín Pina[23], S. Marinoni[54,5], F. Marocco[121], D.J. Marshall[122], L. Martin Polo[41], J.M. Martín-Fleitas[22],

G. Marton[58], N. Mary[66], A. Masip[23], D. Massari[8], A. Mastrobuono-Battisti[17], T. Mazeh[89], P.J. McMillan[25], S. Messina[52], D. Michalik[16], N.R. Millar[10], A. Mints[49], D. Molina[23], R. Molinaro[113], L. Molnár[58,123,59], G. Monari[48], M. Monguió[23], P. Montegriffo[8], A. Montero[22], R. Mor[23], A. Mora[22], R. Morbidelli[7], T. Morel[42], D. Morris[44], T. Muraveva[8], C.P. Murphy[21], I. Musella[113], Z. Nagy[58], L. Noval[66], F. Ocaña[37,124], A. Ogden[10], C. Ordenovic[1], J.O. Osinde[45], C. Pagani[65], I. Pagano[52], L. Palaversa[125,10], P.A. Palicio[1], L. Pallas-Quintela[47], A. Panahi[89], S. Payne-Wardenaar[19], X. Peñalosa Esteller[23], A. Penttilä[55], B. Pichon[1], A.M. Piersimoni[84], F.-X. Pineau[48], E. Plachy[58,123,59], G. Plum[17], E. Poggio[1,7], A. Prša[126], L. Pulone[54], E. Racero[41,124], S. Ragaini[8], M. Rainer[4,127], C.M. Raiteri[7], P. Ramos[23,48], M. Ramos-Lerate[37], P. Re Fiorentin[7], S. Regibo[60], P.J. Richards[128], C. Rios Diaz[45], V. Ripepi[113], A. Riva[7], H.-W. Rix[6], G. Rixon[10], N. Robichon[17], A.C. Robin[13], C. Robin[66], M. Roelens[15], H.R.O. Rogues[86], L. Rohrbasser[34], M. Romero-Gómez[23], N. Rowell[44], F. Royer[17], D. Ruz Mieres[10], K.A. Rybicki[110], G. Sadowski[27], A. Sáez Núñez[23], A. Sagristà Sellés[19], J. Sahlmann[45], E. Salguero[46], N. Samaras[3,129], V. Sanchez Gimenez[23], N. Sanna[4], R. Santoveña[47], M. Sarasso[7], M. Schultheis[1], E. Sciacca[52], M. Segol[86], J.C. Segovia[41], D. Ségransan[15], D. Semeux[88], S. Shahaf[130], H.I. Siddiqui[131], A. Siebert[48,91], L. Siltala[55], A. Silvelo[47], E. Slezak[1], I. Slezak[1], O.N. Snaith[17], E. Solano[132], F. Solitro[32], D. Souami[108,133], J. Souchay[38], A. Spagna[7], L. Spina[11], F. Spoto[72], I.A. Steele[120], H. Steidelmüller[24], C.A. Stephenson[37,134], M. Süveges[135], J. Surdej[42,136], L. Szabados[58], E. Szegedi-Elek[58], F. Taris[38], M.B. Taylor[137], R. Teixeira[85], L. Tolomei[32], N. Tonello[96], F. Torra[31], J. Torra†[23], G. Torralba Elipe[47], M. Trabucchi[138,15], A.T. Tsounis[139], C. Turon[17], A. Ulla[140], N. Unger[15], M.V. Vaillant[66], E. van Dillen[86], W. van Reeven[141], O. Vanel[17], A. Vecchiato[7], Y. Viala[17], D. Vicente[96], S. Voutsinas[44], M. Weiler[23], T. Wevers[10,142], Ł. Wyrzykowski[110], A. Yoldas[10], P. Yvard[86], H. Zhao[1], J. Zorec[143], S. Zucker[71], and T. Zwitter[144]

(Affiliations can be found after the references)

Received ; accepted

ABSTRACT

Context. *Gaia* Data Release 3 (DR3) provides a wealth of new data products for the astronomical community to exploit, including astrophysical parameters for a half billion stars. In this work we demonstrate the high quality of these data products and illustrate their use in different astrophysical contexts.

Aims. We produce homogeneous samples of stars with high quality astrophysical parameters by exploiting *Gaia* DR3 while focusing on many regimes across the Hertzsprung-Russell (HR) diagram; spectral types OBA, FGKM, and ultra-cool dwarfs (UCDs). We also focus on specific sub-samples which are of particular interest to the community: solar analogues, carbon stars, and the Spectro Photometric Standard Stars (SPSS).

Methods. We query the astrophysical parameter tables along with other tables in *Gaia* DR3 to derive the samples of the stars of interest. We validate our results by using the *Gaia* catalogue itself and by comparison with external data.

Results. We have produced six homogeneous samples of stars with high quality astrophysical parameters across the HR diagram for the community to exploit. We first focus on three samples that span a large parameter space: young massive disk stars (OBA, ~3M), FGKM spectral type stars (~3M), and UCDs (~20 K). We provide these sources along with additional information (either a flag or complementary parameters) as tables that are made available in the *Gaia* archive. We furthermore identify 15 740 bone fide carbon stars, 5 863 solar-analogues, and provide the first homogeneous set of stellar parameters of the SPSS sample. We demonstrate some applications of these samples in different astrophysical contexts. We use a subset of the OBA sample to illustrate its usefulness to analyse the Milky Way rotation curve. We then use the properties of the FGKM stars to analyse known exoplanet systems. We also analyse the ages of some unseen UCD-companions to the FGKM stars. We additionally predict the colours of the Sun in various passbands (*Gaia*, 2MASS, WISE) using the solar-analogue sample.

Conclusions. *Gaia* DR3 contains a wealth of new high quality astrophysical parameters for the community to exploit.

Key words. catalogues; stars: fundamental parameters; stars: early-type; stars: low-mass; Galaxy: stellar content; Galaxy: kinematics and dynamics

1. Introduction

The knowledge of astrophysical parameters of stars (APs; effective temperatures, radii etc., see Sect. 2) is fundamental for understanding the structure, formation, and evolution of astrophysical systems. For example, exploring chemical distributions of populations of our Galaxy requires well-constrained stellar effective temperatures ($T_{\rm eff}$) and surface gravities ($\log g$) in order to derive precise and accurate abundances, see e.g. Nissen & Gustafsson (2018); Jofré et al. (2019) for reviews. If we want to place our solar system in the context of exoplanet system formation and evolution, we need to determine the radius, mass, and age of many exoplanets and their host stars, e.g. Kaltenegger & Selsis (2015); Rauer et al. (2014); Rando et al. (2020). *Gaia* DR3 contains a wealth of new data products. In particular, it provides us with stellar parameters derived from the analysis of the *Gaia* RVS spectra (Sartoretti et al. 2018), the low-resolution spectra produced by the Blue Photometer and the Red Photometer (BP and RP) (Carrasco et al. 2021; De Angeli et al. 2022), astrometry (Lindegren et al. 2021c), and integrated photometry (Riello et al. 2021a) for up to 470 million stars (Andrae et al. 2022; Creevey et al. 2022; Fouesneau et al. 2022; Lanzafame et al.

⋆ This paper contains tables that are made available only through the *Gaia* archive under *Performance Verification*. Table 8 is only available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/

2022; Recio-Blanco & et al. 2022). The accuracy and precision of these parameters vary, as expected, with brightness, distance, stellar types and the number of observations. These parameters can be exploited in many ways from detailed studies of individual stars, to statistical studies of large samples of stars or populations. This catalogue, based uniquely on *Gaia* data, also has a long-term legacy value as a rich database for target selection for future follow-up studies and missions.

In this work we demonstrate the potential of the new data products in *Gaia* DR3 by producing very high quality samples of astrophysical parameters of stars all across the HR diagram. We aim to make clean samples of stars based on severe quality cuts. We consider these sources to have the most accurate and precise stellar properties in this catalogue that can be used on a star-by-star basis. These quality cuts, however, have an important impact on the selection function and completeness. Our selection criteria will not be optimal for many user's specific scientific cases, and we fully encourage the exploration of the full catalogue of APs in *Gaia* DR3.

This paper is laid out as follows. In Sect. 2 we describe the data products that are used in this work. Then, in the first part of this analysis, we focus on three main stellar regimes and we produce large high quality samples of stars covering the hot O-, B-, and A-type stars (OBA, Sect. 3), the cooler F-, G-, K, and M-type stars (FGKM, Sect. 4), and the sub-stellar ultra-cool dwarfs (UCDs, Sect. 5). We then focus on specific objects of interest: carbon stars (Sect. 6), solar analogues (Sect. 7), and finally the *Gaia* Spectro-Photometric Standards (SPSS, Sect. 8, Pancino et al. 2021a). Sects. 3 – 8 are entirely independent sections and a reader can choose to focus only on their section of choice without missing important information for the rest of the paper. In Sect. 9 we describe the six tables from this work that are made available in *Gaia* DR3, and then in Sect. 10 we illustrate some applications of the various samples in different astrophysical domains.

2. Data description

To define our samples of stars, we use primarily the astrophysical parameters (APs) from the *Gaia* DR3 catalogue. These data provide us with a uniformly-derived, all-sky catalogue of APs: atmospheric properties ($T_{\rm eff}$, log g, [M/H], [α/Fe], activity index, emission lines, rotation), abundance estimates for 13 chemical species, evolution characteristics (radius, age, mass, bolometric luminosity), distance, and dust extinction. The APs are found in two tables of the archive: `astrophysical_parameters` and `astrophysical_parameters_supp`, and a subset of these are also copied to the `gaia_source` table for convenience to the user.

These data were produced by the *Gaia* Data Processing and Analysis Consortium (DPAC) – Coordination Unit 8 (CU8) using the Astrophysical Parameters Inference Software (Apsis, Bailer-Jones, C. A. L. et al. 2013; Creevey et al. 2022) and a series of three papers describe the methodologies and content of the APs in *Gaia* DR3. Creevey et al. (2022) presents an overview of the processing, the architecture and the modules of Apsis, along with a summary of the data products. Fouesneau et al. (2022) focuses on the stellar content, its description, and quality assessments, and Delchambre et al. (2022) details the non-stellar content, in particular object classification, extinction and extra-galactic objects.

The DPAC data processing chain also uses these APs, for example, to identify the best template spectrum for the extraction of the radial velocities from the RVS spectra, the identification of quasars used to fix the astrometric reference frame, and the optimization of the BP and RP calibration.

In this work we focus on the data products produced by six modules of the Apsis chain; the General Stellar Parametrizer from Photometry, GSP-Phot, the General Stellar Parametrizer from Spectroscopy, GSP-Spec, Extended Stellar Parametrizer for Emission-Line Stars, ESP-ELS, Extended Stellar Parametrizer for Hot Stars, ESP-HS, Extended Stellar Parametrizer for Ultra-Cool Dwarfs, ESP-UCD, and the Final Luminosity Age Mass Estimator, FLAME. These are described in detail in Creevey et al. (2022) and in the online documentation. Further details on GSP-Phot and GSP-Spec are also found in the dedicated module papers (Andrae et al. 2022; Recio-Blanco & et al. 2022).

Briefly, GSP-Phot processes all sources with mean BP and RP spectra (De Angeli et al. 2022; Montegriffo et al. 2022) to produce spectroscopic parameters and extinction estimates. It also uses parallaxes and photometry[1]. It processes the sources considering four stellar libraries and the individual results for each of these libraries are found in the `astrophysical_parameters_supp` table. The results from the library responsible for the highest log posterior for that source (see `libname_gspphot`) are those that appear in the main `astrophysical_parameters` table. GSP-Spec processes sources with mean RVS spectra (Seabroke & et al. 2022) and produces not only atmospheric parameters but also chemical abundances and the diffuse interstellar band characterisation. These latter products are not the focus of this work, we instead refer readers to Gaia Collaboration et al. (2022d) and Gaia Collaboration et al. (2022e) respectively. The results from GSP-Spec used in this work are found in the `astrophysical_parameters` table. ESP-HS processes both the BP and RP and the RVS spectra when available and by default just the BP and RP spectra. It produces stellar parameters for stars hotter than 7 500 K along with a spectral type for all stars. The ESP-ELS module analyses emission-line stars and provides class probabilities and labels, along with a measurement of the H-α equivalent width. ESP-UCD is a module dedicated to the analysis of UCDs and it produces a $T_{\rm eff}$. All of these results are found in the `astrophysical_parameters` table. Finally, FLAME processes the output spectroscopic parameters from GSP-Phot and GSP-Spec along with astrometry and photometry to derive evolutionary parameters (R, L, M, age). The FLAME results based on the GSP-Phot input are found in the `astrophysical_parameters` table, while those based on the GSP-Spec input are found in the `astrophysical_parameters_supp` table. These six modules (GSP-Phot, GSP-Spec, ESP-ELS, ESP-HS, ESP-UCD, and FLAME) produce the data that are the focus of this paper. For further details on the methods we refer readers to the above references.

This work also exploits other data products from *Gaia* DR3; the astrometry (parallaxes errors and proper motions) and properties of the photometry and spectroscopy are found in the main `gaia_source` table and these were also available in *Gaia* EDR3, see also Damerdji et al. 2022; Lindegren et al. 2021b,a; Riello et al. 2021b; Seabroke & et al. 2022. We additionally exploit the variability analysis performed by the Coordination Unit 7 (Eyer et al. 2022; Clementini et al. 2022; Mowlavi et al. 2022) and the analysis of binary and multiple systems by the Coordination

[1] Within the Apsis software, the parallaxes are corrected for the known zero-point biases as a function of latitude, magnitude and colour, see Lindegren et al. (2021a).

Unit 4 (Gaia Collaboration et al. 2022a; Halbwachs et al. 2022; Holl et al. 2022; Siopis 2022) to further define our samples.

3. OBA stars

3.1. Scientific motivation

O- and B- and A-type (OBA) stars are intermediate to large mass stars that evolve rapidly and usually do not migrate very far away from their birth association or cluster. For this reason, they are the best targets to study the structure and dynamics of star forming regions, as well as of the Galactic spiral arms (e.g. Gaia Collaboration et al. 2022b). Hot stars also play an important role in the evolution of the Galaxy: they are the main contributors to its enrichment in elements heavier than carbon and their strong ultra-violet radiation is the main source of ionisation of the interstellar medium. The most massive OB stars have strong stellar winds and explode as supernovae at the end of their lives, and therefore are an important contributor to the chemical composition of the interstellar medium. Here we focus on the construction of a sample of OBA-type stars which are part of the Milky Way young disk population. Hence older stellar populations covering the same effective temperature range (e.g., white dwarfs and blue horizontal branch stars) are excluded from this sample as much as possible. Because young OBA stars are significantly less numerous than cooler stars, their identification can be considered as a key issue in large surveys. Their spectral energy distribution is less sensitive to effective temperature in comparison to later type stars which hampers the accurate and non-biased determination of $T_{\rm eff}$. We fixed the lower $T_{\rm eff}$ threshold at 7500 K. Moreover, although we have taken care to eliminate the most significant contaminants (e.g. white dwarfs, RR Lyrae, Sect. 3.2), our sample still includes stars that are not young OBA stars, such as blue horizontal branch (HB) stars (Sect. 3.3). In the rest of this section we will continue to refer to the 'OBA star sample' as a shorthand for young OBA stars in the disk of the Milky Way.

3.2. Sample selection

GSP-Phot and ESP-HS are the two main Apsis modules that derive the APs of OBA stars. While GSP-Phot processed all targets with $G \leq 19$, ESP-HS only processed OBA stars brighter than $G = 17.65$[2], and additionally it only processed those stars that received a `spectraltype_esphs` tag of $\in$ ['A','B','O']. This tag is derived from a random forest classification of the BP and RP spectra, see Sect. 11.3.7 of the online documentation for details. In terms of effective temperature this is equivalent to selecting targets hotter than 7500 K. The same lower $T_{\rm eff}$ limit is applied to the GSP-Phot stellar sample. Because GSP-Phot processes targets down to $G = 19$, the corresponding sample initially contains more (11 156 449 stars) candidate-OBA targets than ESP-HS (2 344 484). The GSP-Phot parametrisation partly relies on the use of parallax, and more outliers (e.g. misclassified cool objects, white dwarfs, ...) are included when the astrometry is less reliable (Fouesneau et al. 2022). To exclude a significant fraction of these we removed all targets based on the ratio of the parallax ϖ to its uncertainty σ_{ϖ} `parallax_over_error` ≤ 15, as illustrated in Fig. 1 (panels a and c). ESP-HS does not use information that allows, for example, to remove white dwarfs. Therefore, we applied a lower luminosity threshold to both samples and removed all sub-luminous objects. The limit was fixed by computing the dispersion around the running median of $M_{\rm G}$ as a function of $T_{\rm eff}$ and by using the AP determinations obtained by ESP-HS in its BP/RP+RVS processing mode. The gray shading in Fig. 1 (panels b and c) shows the area of the HR diagram from which targets were excluded. Ideally the observed de-reddened $(G_{\rm BP} - G_{\rm RP})$ color vs. $T_{\rm eff}$ follows the same relation as the one found from synthetic spectra (Fig. 1, panels d and e, blue curve) used to derive the APs. All outliers at more than 6 standard deviations from the theoretical relation were discarded from the sample, as shown by the gray shading in Fig. 1 (panels d and e). The Kiel diagram of each sample is shown in the bottom row of the same figure, with the corresponding number of remaining stars. We noticed that the modules were misclassifiying some RR-Lyrae stars as OBA stars, therefore the list was cross-matched with the RR-Lyrae table `vari_rrlyrae` in *Gaia* DR3 (Clementini et al. 2022). After filtering, 3 023 388 unique sources remained in the list of candidate-OBA stars. Among these, 1 661 459 and 843 324 have ESP-HS or GSP-Phot APs, respectively, while 518 605 have both. Among those targets with GSP-Phot parameters, all but 889 received a `spectraltype_esphs` tag.

The corresponding `gold_sample_oba_stars` table (this will appear as `gaiadr3.gold_sample_oba_stars`) has two columns: one lists the `source_id` and the other a `flag` that provides information on the kinematics of the targets (Sect. 3.3).

We tested the completeness of the GDR3 OBA sample by cross-matching it with the Galactic open cluster members identified by Cantat-Gaudin et al. (2020). The selection of the *expected* OBA stars in each cluster is based on the $(G_{\rm BP} - G_{\rm RP})_0$ color at $T_{\rm eff} = 7500$ K, estimated by taking into account the published cluster extinction A_0. Their number, $N_{\rm OBA}^{\rm expected}$, was used to estimate the completeness fraction as follows

$$\text{fraction} = \frac{N_{\rm OBA}^{\rm GDR3}}{N_{\rm OBA}^{\rm expected}} \quad (1)$$

where $N_{\rm OBA}^{\rm GDR3}$ is the number of the OBA open cluster targets found in our sample. We expect the fraction to vary with magnitude and, due to the extinction/temperature degeneracy, with interstellar extinction. We show in Fig. 2 how the completeness varies with A_0. The fraction of targets we have in common with the LAMOST OBA (Xiang et al. 2021) and GOSC (Maíz Apellániz et al. 2013) catalogues are 0.55 and 0.41, respectively. The $T_{\rm eff}$ distributions provided by both modules confirm that above 10 000 K, the ESP-HS APs should be preferred over the GSP-Phot estimates in the `astrophysical_parameters` table, whose temperature scale tends to be underestimated in this regime. This is especially true at higher interstellar extinction.

A Simbad query of the proposed OBA sample provides 34 055 targets with a confirmed main object type not equal to "*". Among these, 27% have types not compatible with what would be expected for hot young stars, and of which 79% are known HB stars. This high density of hot HB stars can be seen, for example, in the bottom panels of Fig. 1 where their presence produces a significant overdensity of stars with $T_{\rm eff}$ ranging from 8,000 to 10 000 K and log g lower than 3.5. As explained in the following section, a number of these lower mass evolved targets can be flagged by studying their kinematics. Furthermore, 134 498 targets in our list have a spectral type recorded in Simbad, which in 96% of the cases starts with the letter "O", "B" or "A".

[2] This limitation was imposed during operations in order to remain within the processing schedule, see Sect. 11.1.4 of the online documentation for details

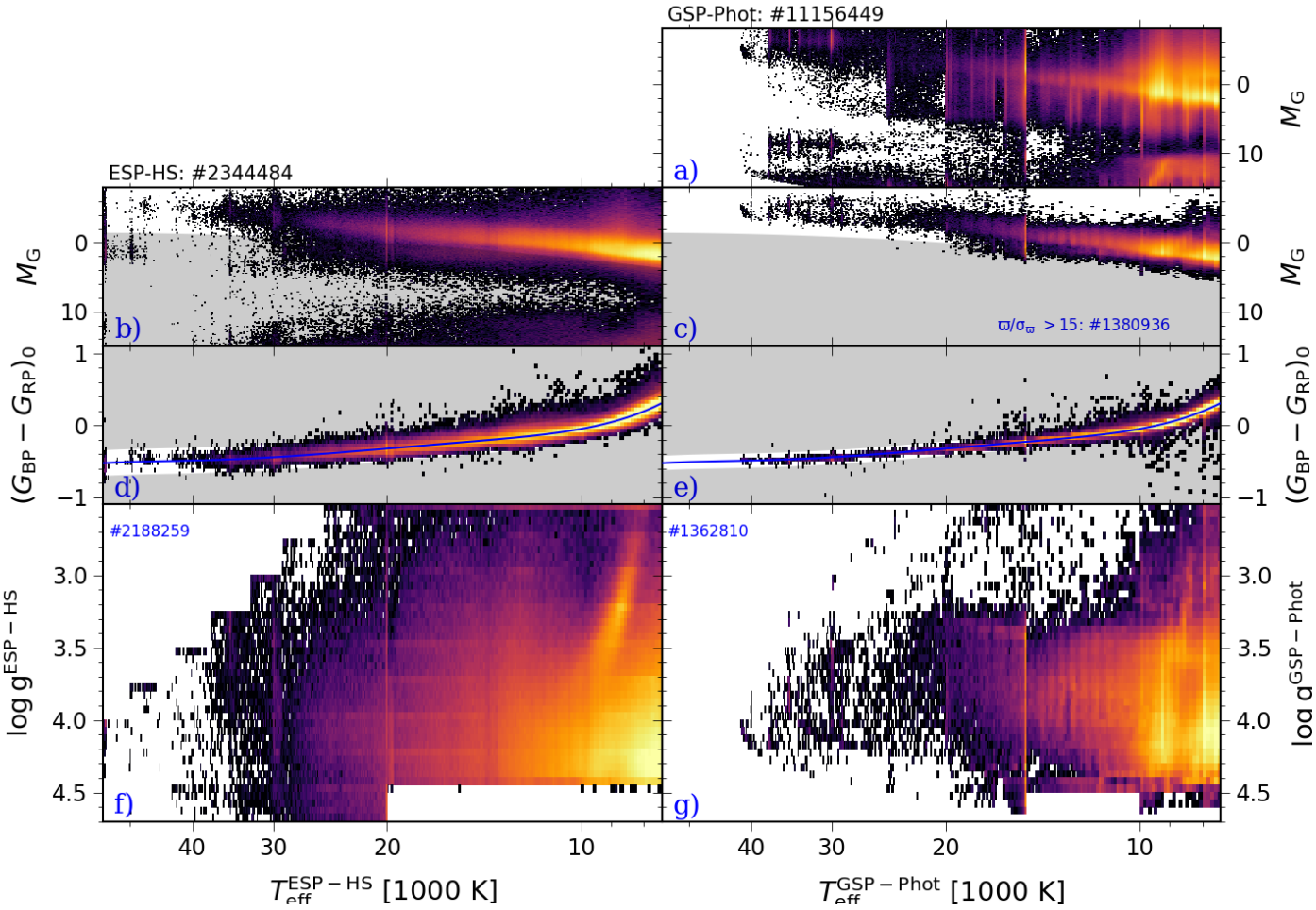

Fig. 1. Selection of the OBA sample. Left and right panels show the OBA samples from ESP-HS and GSP-Phot, respectively. A first filter on the parallax SNR is applied to the GSP-Phot targets (from panel a to c). For both samples sub-luminous targets are removed (gray shading in panels b and c), and then the outliers at 6 standard deviations from the expected colour vs. $T_{\rm eff}$ relation (blue line) are filtered out (gray shading in panels d and e). The absolute magnitude $M_{\rm G}$ is computed using the measured parallax and the estimated interstellar extinction $A_{\rm G}$ provided by both modules. The de-reddened colour, $(G_{\rm BP} - G_{\rm RP})_0$, is derived using the value of $E(G_{\rm BP} - G_{\rm RP})$. The resulting Kiel diagrams are shown in the bottom rows (panels f and g). The over-densities seen at $T_{\rm eff}$ = 15 000 K, 20 000 K, and 30 000 K are linked to the temperature limits of the adopted synthetic spectra libraries.

3.3. Using kinematics to remove halo contaminants

To further clean the sample of (young) OBA stars from contaminating populations we propose a simple kinematic filter which removes what are presumably blue horizontal branch stars from the halo, which occupy the same colour-brightness space in the colour-magnitude diagram as the OBA stars, as well as the same $T_{\rm eff}$-log g space in the Kiel diagram. We filter on the tangential velocity $v_{\rm tan} = A_v(\mu_{\alpha*}^2 + \mu_\delta^2)^{1/2}/\varpi$, where $\mu_{\alpha*}$ and μ_δ are the proper motions in the right ascension and declination, and A_v = 4.74074... km yr s^{-1}, using similar limits for the thin disk, thick disk, and halo as in Gaia Collaboration et al. (2018). Thin disk stars are defined as having $v_{\rm tan} < 40$ km s^{-1}, thick disk stars as having $40 \leq v_{\rm tan} \leq 180$ km s^{-1}, and halo stars have $v_{\rm tan} > 180$ km s^{-1}. We next illustrate the effects of this kinematic selection and thereby focus on stars for which $\varpi/\sigma_\varpi > 10$. This parallax quality cut ensures a reliable calculation of the tangential velocities.

Figure 3 shows the distribution of tangential velocities for the OBA star sub-sample for which $\varpi/\sigma_\varpi > 10$. The vertical dashed lines indicate the above limits on $v_{\rm tan}$ and these correspond well to the inflections in the histograms for the full, B, and A star samples. The O star sample contains almost no sources with $v_{\rm tan} > 180$ km s^{-1}. To further explore the tangential velocity selection we show in Fig. 4 the Toomre diagram, which shows $\sqrt{V_R^2 + V_z^2}$ along the vertical axis and V_ϕ along the horizontal axis, where (V_R, V_ϕ, V_z) are the velocity components of the stars in the Galactocentric cylindrical coordinate system, with R pointing from the Galactic centre to the Sun, z along the axis perpendicular to the Galactic plane, and ϕ along the azimuthal direction in the Milky Way disk plane (where a left handed coordinate system is used such that the value of V_ϕ is positive for prograde stars in the disk). The values of (V_R, V_ϕ, V_z) are calculated assuming the local circular velocity from the `MWPotential2014` Milky Way model (Bovy 2015), which is 219 km s^{-1} at the distance of the Sun from the Galactic center (8277 pc, Gravity Collaboration et al. 2022). The height of the Sun above the disk plane is assumed to be 20.8 pc (Bennett & Bovy 2019) and the peculiar motion of the Sun is assumed to be $(U, V, W) = (11.1, 12.24, 7.25)$ km s^{-1} (Schönrich et al. 2010). Figure 4 only contains stars for which the radial velocity is available in *Gaia* DR3 and the colour coding indicates the value of $v_{\rm tan}$. The two half-circles indicate the limits on to-

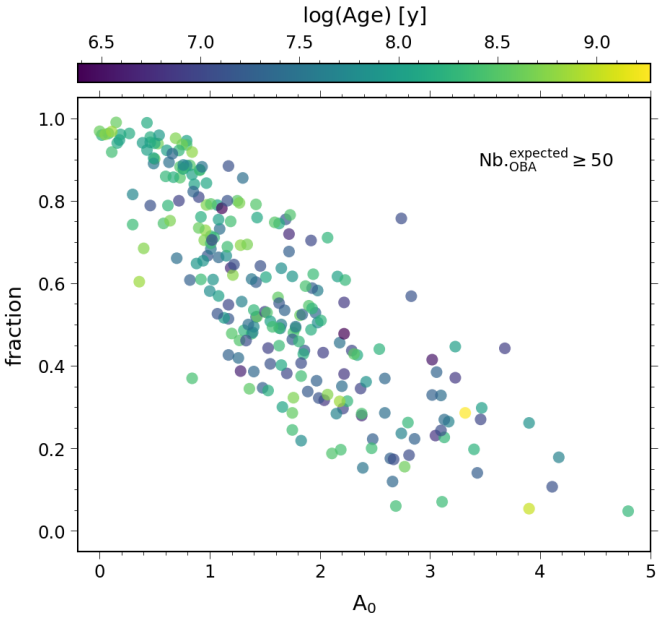

Fig. 2. Completeness of the OBA list in various open clusters (Cantat-Gaudin et al. 2020) as a function of the interstellar extinction. The fraction corresponds to the ratio between the number of cluster members present in our list, and the number of expected OBA stars. The color code follows the cluster age provided by Cantat-Gaudin et al. (2020).

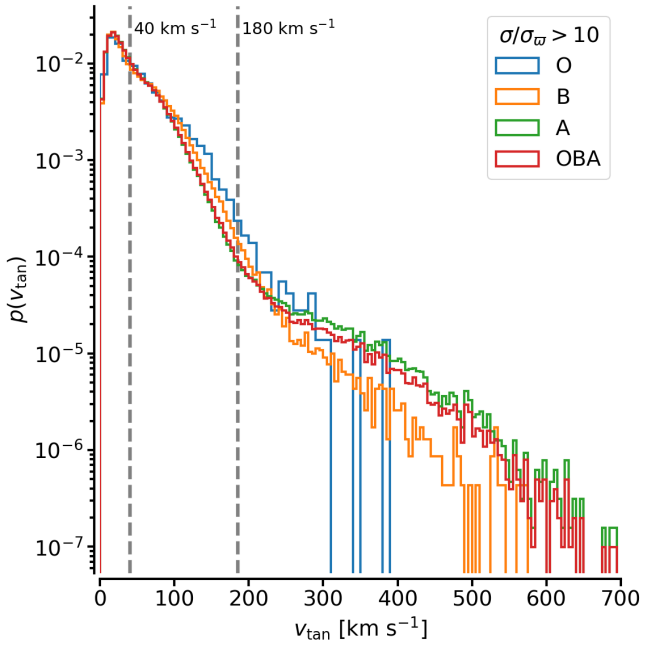

Fig. 3. Histogram of tangential velocities of the stars in the OBA sample with $\varpi/\sigma_\varpi > 10$. The combined OBA star sample is shown as well as the individual O, B, and A star samples (based on the classifications from the ESP-HS module). The limits in tangential velocity separating the thin disk, thick disk, and halo populations are shown as vertical dashed lines.

tal velocity $v_{\rm tot} = \sqrt{V_R^2 + V_\phi^2 + V_z^2}$ of 50 and 180 km s^{-1} which separate thin disk, thick disk, and halo populations (Gaia Collaboration et al. 2018). In this figure a population of stars can be seen at total velocities of more than 180 km s^{-1} from the local circular velocity and these are most probably halo stars, in

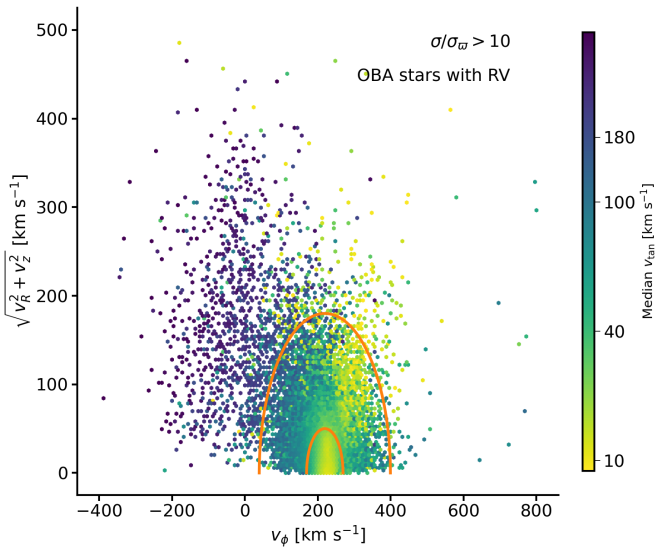

Fig. 4. Toomre diagram for the OBA stars for which a radial velocity is available in *Gaia* DR3. See text for explanations on the diagram. The colour coding indicates the median value of $v_{\tan}$ at a given location on this diagram. The half-circles indicate limits on the total velocity with respect to the local circular velocity of 50 and 180 km s^{-1}.

particular the population at negative V_ϕ which is associated with merger debris in the halo (e.g., Helmi et al. 2018).

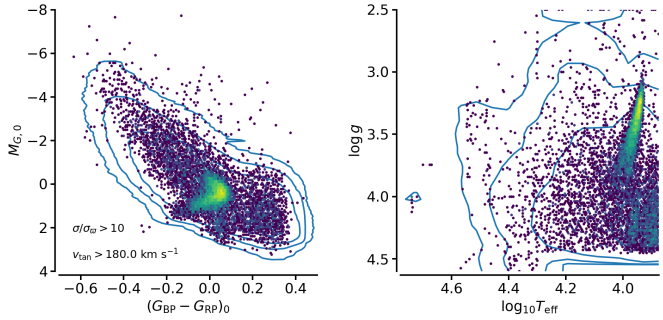

Fig. 5. Left: observational Hertzsprung-Russell diagram for the stars in the OBA sample with $\varpi/\sigma_\varpi > 10$. Right: Kiel diagram for the same sample of stars. The contours indicate the distribution of the full sample. The colour coded density images show the stars for which $v_{\tan} > 180$ km s^{-1}.

The colour coding in Fig. 4 suggests that the halo contaminants in the OBA sample can be filtered out by demanding $v_{\tan} < 180$ km s^{-1}, although clearly there will be stars left at low tangential velocities which have a total velocity which puts them in the halo. Figure 5 shows the observational Hertzsprung-Russell and the Kiel diagrams for the sample of OBA stars with $\varpi/\sigma_\varpi > 10$. Extinction corrections using A_G and $E(G_{\rm BP} - G_{\rm RP})$ from the ESP-HS module were applied. The contours show the distribution of the full sample, while the colour coded density images show the distribution of stars selected according to $v_{\tan} > 180$ km s^{-1}. The velocity-filtered sample mostly occupies the colour magnitude space where blue horizontal branch stars are expected, around $(G_{\rm BP} - G_{\rm RP})_0 \sim 0.05$ and $M_{G,0} \sim 0.5$ (compare to the rightmost panel of figure 21 in Gaia Collaboration et al. 2018). The Kiel diagram shows a prominent feature at $\log_{10} T_{\rm eff} \sim 4$, from $\log g \sim 4$ to $\log g \sim 2$, corresponding to the known location of the horizontal branch stars in this diagram. These same stars are also primarily located at high galactic lat-

itude as expected for a halo population. A search in SIMBAD (Wenger et al. 2000) results in 8124 matches for which there is information on the stellar type, of which 5770 are incompatible with stellar types corresponding to hot young stars, including 5499 sources classified as horizontal branch star. This further supports using $v_{\rm tan} > 180$ km s^{-1} as a filter to clean the OBA sample from halo star contamination.

One might consider further filtering on $v_{\rm tan}$, however we caution that because of the large reach of the OBA sample this can lead to significant spatial selection effects. This is illustrated in Fig. 6. The figure shows the OBA stars with $\varpi/\sigma_\varpi > 10$ projected on the Galactic plane. The full sample is shown in the leftmost panel and the other panels show the effects of filtering on $v_{\rm tan}$. The star positions in Galactocentric coordinates were calculated using the same Milky Way parameters as listed above. The red contours show the limits of 40 km s^{-1} and 180 km s^{-1} on the observed tangential velocities, predicted from a simplistic model of the Milky Way disk kinematics. In this model it is assumed that all stars are located in the disk and follow perfectly circular orbits according to the rotation curve from the `MWPotential2014` Milky Way model (Bovy 2015). The expected values of $v_{\rm tan}$ are then calculated over a grid of (X, Y) positions, using the method outlined in Brunetti & Pfenniger (2010). The contours indicate the boundaries between smaller $v_{\rm tan}$ values to the left and larger ones to the right. The contours show that due to the large reach of *Gaia* even for stars moving at zero velocity dispersion on circular orbits in the disk one can still expect to observe tangential velocities out to values normally associated with the thick disk and halo. The rightmost panel in Fig. 6 again confirms that $v_{\rm tan} > 180$ km s^{-1} can be used to clean the OBA sample from halo stars, as the stars are all located to the left of the 180 km s^{-1} contour, where these would be expected on the right (in the simple model used) if they were disk stars. The middle panels illustrate the spatial selection bias introduced when further restricting the tangential velocity. The second panel from the right shows that limiting the sample to $v_{\rm tan} < 40$ km s^{-1} leads to the exclusion of a significant fraction of young OBA stars which occupy regions of the Galactic disk where the values of $v_{\rm tan}$ are expected to be larger than 40 km s^{-1}. In addition there is a lack of stars in the sample along the $X = 0$ line which roughly follows the shape of the 40 km s^{-1} contour. The simple disk model predicts zero stars there, thus the shape of the gap shows that the model is useful in assessing the spatial selection biases induced by the kinematic selection. In the third panel from the right the shape of the sample distribution also roughly follows the 40 km s^{-1} contour.

In conclusion, we provide a table of 3 023 388 young OBA disk stars, cleaned as much as possible to remove older stellar populations, for exploitation by the community. We recommend to further clean the OBA star sample by applying the kinematic filter $v_{\rm tan} \leq 180$ km s^{-1}. Sources with $v_{\rm tan} > 180$ km s^{-1} have the flag in the table `gold_sample_oba_stars` set to 1, all other sources have the flag set to 0. We have only used the simple Galactic disk kinematic model to make the point that one should be careful not to introduce spatial biases when selecting on kinematics. Zari et al. (2021) describe a more sophisticated way of employing a simple disk kinematic model to select a clean sample of OBA stars. By assuming the stars follow disk kinematics they can use the observed proper motions to infer distances. Stars with kinematic distances inconsistent with distances based on the parallaxes and photometric information can then be analyzed further to see if they should be removed from the OBA star sample. Further filtering can of course be done on the various data quality indicators available in *Gaia* DR3 (see the following section for examples), and one can also use the astrometric fidelity indicator from Rybizki et al. (2022).

4. FGKM stars

4.1. Scientific motivation

F, G, K, and M stars form the majority of the stars of our Milky Way. These stars inform us of how our Galaxy was formed and how it has evolved and are thus the targets of many Milky Way surveys. These stars are also the targets of the future ESA PLATO mission (Rauer et al. 2014) which promises to help answer our questions about the formation and evolution of our own Solar System by studying other exoplanet systems. In this section we focus on F, G, K and M star types (FGKM) to provide a clean sample of stars with the following astrophysical parameters : $T_{\rm eff}$, log g, [M/H], R, M, age, evolutionary stage, and spectral type. Our final sample contains 3 273 041 stars after vigorous quality cuts based on astrometric, photometric, and astrophysical parameters, along with other *Gaia*-based criteria.

Our sample selection is described in Sects. 4.2 and 4.3 where we analyse the GSP-Phot-based and GSP-Spec-based atmospheric parameters individually. For both samples we also report on evolutionary parameters from FLAME and the spectral type from ESP-HS. We then perform some additional filtering by removing variables and binaries. We also further filter on individual parameters from FLAME and ESP-HS for some sources. We validate the target list using open clusters and comparisons with external survey catalogues. In Sect. 10 we illustrate two applications of this sample by analysing known transiting exoplanets and studying unseen UCD-companions in the *Gaia* data.

4.2. GSP-Phot sample selection

GSP-Phot provides stellar and extinction parameters, along with distances, radii, and an absolute magnitude for 470 million stars with $G \leq 19$. We performed our initial query on the full *Gaia* archive by selecting sources with a parallax signal-to-noise ratio (SNR) better than 10, along with a number of other initial quality cuts based on astrometric and photometric parameters. These criteria were based on an analysis of a random set of 2 million sources. This resulted in a total of 70.4 million stars which we refer to as sample fgkm_1, and which is described by the following Astronomical Data Query Language (ADQL) query:

```
parallax_over_error > 10
ipd_frac_multi_peak < 6
phot_bp_n_blended_transits < 10
teff_gspphot > 2500
```

in addition to a quality cut on `bp_rp_error < 0.06` ($\sigma_{\rm (BP-RP)}$). This latter quantity is calculated from a standard propagation of errors using the parameters `phot_bp_mean_flux_over_error` $\left(\frac{f_{\rm BP}}{\sigma f_{\rm BP}}\right)$ and `phot_rp_mean_flux_over_error` $\left(\frac{f_{\rm RP}}{\sigma f_{\rm RP}}\right)$ from the archive[3].

We then refined this selection by considering the number of photometric transits, colour-colour and colour-$T_{\rm eff}$ correlations, ensuring that the source is classified as a star by DSC, along with further constraints based on the GSP-Phot parameters themselves. These are described in the following paragraphs.

[3] $\sigma^2_{\rm (BP-RP)} = \sqrt{\sigma^2_{\rm BP} + \sigma^2_{\rm RP}}$, where $\sigma^2_{\rm BP} = \sqrt{\left(\frac{-2.5}{\ln(10)}\frac{\sigma f_{\rm BP}}{f_{\rm BP}}\right)^2 + \sigma^2_{\rm BP,0}}$ and $\sigma^2_{\rm RP} = \sqrt{\left(\frac{-2.5}{\ln(10)}\frac{\sigma f_{\rm RP}}{f_{\rm RP}}\right)^2 + \sigma^2_{\rm RP,0}}$, and the *Gaia* EDR3 passband zeropoint errors are $\sigma_{\rm BP,0} = 0.00279$ and $\sigma_{\rm RP,0} = 0.00231$.

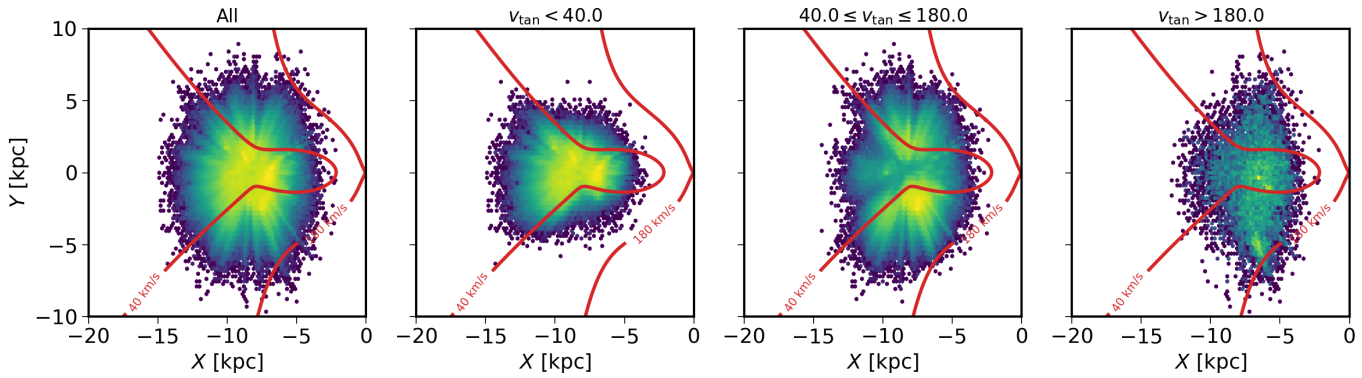

Fig. 6. Distribution of stars in the OBA sample projected on the Galactic plane. The galactic centre is to the right at $(X, Y) = (0, 0)$ and the Sun is at $(-8, 0)$. The panels show from left to right the full sample (with $\varpi/\sigma_\varpi > 10$) and the samples selected according to the v_{tan} ranges indicated. The red contours indicate lines of constant v_{tan}, calculated with the simple kinematic disk model as explained in the text.

Table 1. The coefficients of the polynomials used to fit the T_{eff} versus $(G_{BP} - G)_0$ and $(G - G_{RP})_0$ versus $(G_{BP} - G)_0$ relations in order to remove outliers from the fgkm_1 sample.

y	a_0	a_1	a_2	a_3	a_4	a_5	fitted range	Δy
T_{eff}	9255.55	-17911.0	27241.4	-23103.4	9659.18	-1480.37	3 500-7 500	203
$(G - G_{RP})_0$	0.000373747	2.19368	-2.95122	3.21155	-1.87172	0.418337		0.0267

Notes. The independent variable x is $(G_{BP} - G)_0$, y is the fitted parameter, and a_i are the coefficients of the fit: $y = \sum_{i=0}^{5} a_i x^i$. Δy is the difference between the fit and the values within which we kept the source.

We retained sources whose parameters are within the FGKM regime: T_{eff} < 7500 K, M_G< 12, R < 100 $R_\odot$, and had a log posterior > −4000 (goodness-of-fit indicator). We also retained sources with [M/H] > –0.8 which excludes low-metallicity sources with unreliable metallicities (Andrae et al. 2022; Creevey et al. 2022; Fouesneau et al. 2022).

GSP-Phot provides four results for each source based on different stellar libraries: MARCS, PHOENIX, A and OB. Only MARCS and PHOENIX are applicable to the stellar regime considered here. The results for all libraries are found in the `astrophysical_parameters` table, and we used the difference between `teff_gspphot_marcs` and `teff_gspphot_phoenix` (below called ΔT_{eff} or `dteff`) as a criterion to further refine the sample. There is a small bias of up to 100 K between results from these two libraries, due in part to the different spectral energy distributions (SEDs) of the different models and this bias varies with stellar parameters. We therefore selected those sources where the two values modulo the peak offset were in agreement, i.e. $|\Delta T_{eff} + 65| < 150$ K. Additionally, we only retained sources when their uncertainties (upper − lower) are < 150 K, and the sources for which the "best" model is the MARCS one (75% of sample), i.e. `libname_gspphot` = "MARCS" in the `astrophysical_parameters` table. These strict criteria based on T_{eff} removed about 70% of the sources. We also imposed that `distance_gspphot` was less than the distance corresponding to the `parallax` decreased by 4 times `parallax_error` (and vice versa). Fig. 7 top panel illustrates the impact of the cut based on distance. The sources in fgkm_1 are shown in the background, and those with the distance criteria applied (40%) are shown in the foreground. We also show the one-to-one line to guide the eye.

We corrected the G_{BP}, G_{RP}, and G observed colours for the interstellar extinction provided by GSP-Phot: $G_{BP0} = G_{BP} - A_{BP}$, $G_{RP0} = G_{RP} - A_{RP}$, and $G_0 = G - A_G$. Then we fitted polynomials to the T_{eff} versus $(G_{BP} - G)_0$ (difference between fit and values denoted as `dtb`) and $(G - G_{RP})_0$ versus $(G_{BP} - G)_0$ (difference denoted as `dgb`), and used these polynomial fits to remove sources further than 3σ (~7% of fgkm_1). The coefficients of the polynomials are given in Table 1. The bottom panel of Fig. 7 illustrates the $(G-G_{RP})_0-(G_{BP}-G_{RP})_0$ relation for sample fgkm_1 in the background and the sample with the 3σ constraints on the colour-colour and the colour-T_{eff} relations in the foreground.

All of the above criteria along with further constraints on DSC class probabilities and number of transits (`n_obs` below) were used to define the sample fgkm_2 which resulted in a total of 6.3M sources i.e. 12.5% of the fgkm_1 sample. A projection of the retained sources on the Galactic plane is shown in Fig. 8. We note that the criteria on the number of transits were adjusted in order to ensure a full-sky coverage. The full list of constraints for sample fgkm_2 is summarised as follows:

```
|dgb| < 203,
|dtb| < 0.0267
|dteff + 65| < 150
libname_gspphot = "MARCS"
teff_gspphot_upper-teff_gspphot_lower < 150
teff_gspphot < 7500
mh_gspphot > -0.8
distance_gspphot < 1e3/(parallax-4*parallax_error)
distance_gspphot > 1e3/(parallax+4*parallax_error)
radius_gspphot < 100
mg_gspphot < 12
logposterior_gspphot > -4000
classprob_dsc_combmod_star > 0.9
phot_bp_n_obs > 19
phot_rp_n_obs > 19
phot_g_n_obs > 150
```

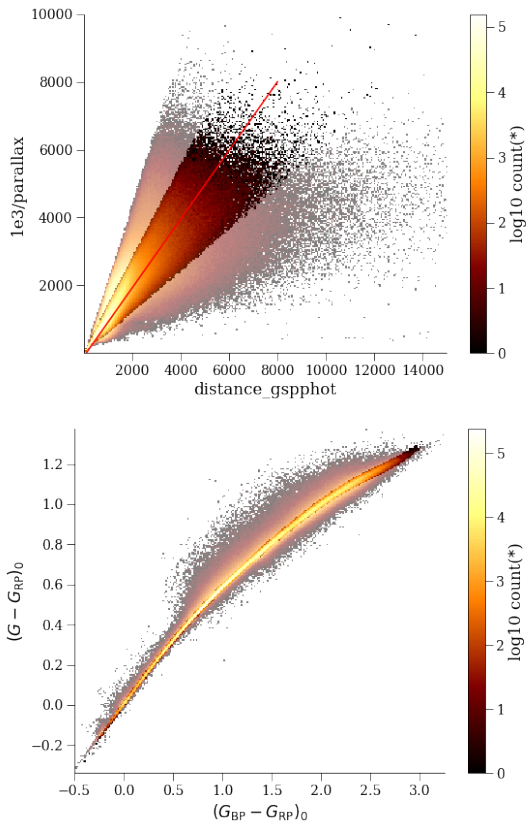

Fig. 7. Comparisons between sample fgkm_1 and intermediate samples based on some of the criteria used to define sample fgkm_2. The top panel illustrates the distance-parallax-error constraint and the lower panel shows the $(G - G_{RP})_0 - (G_{BP} - G_{RP})_0$ relation after imposing the colour-T_{eff} and colour-colour cuts described in Sect. 4.2. In both panels the sources in fgkm_1 are shown in the background, while those satisfying the criteria are illustrated in the foreground, colour-coded by logarithmic count.

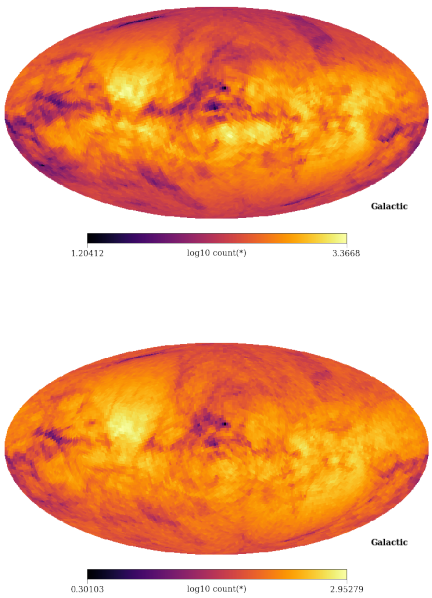

Fig. 8. Galactic plane projections illustrating the density of sources of the samples fgkm_2 (top) and fgkm_3 (bottom).

The final selection, fgkm_3, was done by applying different quality cuts based on the position of the star in the HR diagram. Giants were defined as $\log g < 3.6$ and $T_{eff} < 5900$ K, and outliers were removed by retaining sources with $\log g < 0.34M_G + 2.45$. Subgiants were defined as $3.6 \leq \log g \leq 4.0$ and $T_{eff} < 5900$ K, and outliers were removed by retaining sources with $\log g < 0.75M_G + 1.13$. Main sequence stars were defined as $\log g > 4.0$ and $T_{eff} < 7450$ K, and we imposed a further constraint of `parallax_over_error` > 33.34 in order to have sources with relative errors on R and L with contributions of parallax errors at 3% or less. To further refine the main sequence sample we applied different criteria in three different colour regimes. For $x < 0.98$ where $x = (G_{BP} - G_{RP})_0$ no further selection was done. For $0.98 \leq x \leq 1.8$ we removed the sequence of young pre-main sequence stars and binaries by retaining sources that satisfied $\log L < 2.32 - 3.20x + 0.78x^2$ where L is `lum_flame`. For $x > 1.8$ we retained sources that satisfied $\log g < 8.525 - 6.950x + 3.680x^2 - 0.584x^3$. This final refinement resulted in a total sample size of 3 530 174 sources.

We illustrate the different selection criteria in the HR diagram in Fig. 9. The top left panel shows the HR diagram using a random sample of data from the *Gaia* archive and imposing only that T_{eff} and L exist. The top right panel shows the selection of sources after applying the ADQL search criteria (fgkm_1) which is dominated by the criterion on parallax SNR. One can see that already many outliers and artefacts have been removed with this cut. The bottom left panel shows the sample fgkm_2 where constraints were based on the GSP-Phot parameters, along with further constraints on the photometry and DSC class probabilities. The HR diagram has not changed drastically, but the quality of the data in fgkm_2 is much higher than in fgkm_1. Finally, the bottom right panel illustrates sample fgkm_3 where the sample was separated into five parts (giants, subgiants, upper, middle, and lower main sequence as described above) and different polynomial cuts were applied based on $\log g$, M_G, L, and $(G_{BP} - G_{RP})_0$.

The galactic projection of density of sources is illustrated in the bottom panel of Fig. 8. We also illustrate the distribution of the observable parameters, G, parallax, and $(G_{BP} - G_{RP})_0$ in Fig. 10. The main sequence stars occupy the dense triangular region and extend to approximately 1900 parsec for the hottest stars.

4.3. GSP-Spec sample selection

The selection described in the previous section relies entirely on the BP and RP spectra and their parametrisation, apart from a few criteria on astrometric and photometric parameters. BP and RP spectra have important degeneracies between T_{eff} and A_G, and by imposing our strict selection criteria, we not only inevitably remove sources with excellent parameters derived from the RVS spectra by GSP-Spec, but we can not guarantee either that they fulfill the 'gold' criteria. We therefore made an independent selection by first querying the archive for sources with `flags_gspspec LIKE '0000000000000%'`, i.e., sources for which the first 13 characters of the 41-character long quality flag provided by GSP-Spec are equal to '0', see Recio-Blanco & et al. (2022). These flag settings indicate low potential biases on T_{eff}, $\log g$, [M/H], and to some extent $[\alpha/Fe]$ due to rotational velocity, macroturbulence, uncertainties in the radial velocity shift correction and in the RVS flux, and extrapolation, absence of undefined or negative flux values or emission lines, non-zero uncertainties in the parameters, as well as high quality parameters for KM-type giants (see online documentation).

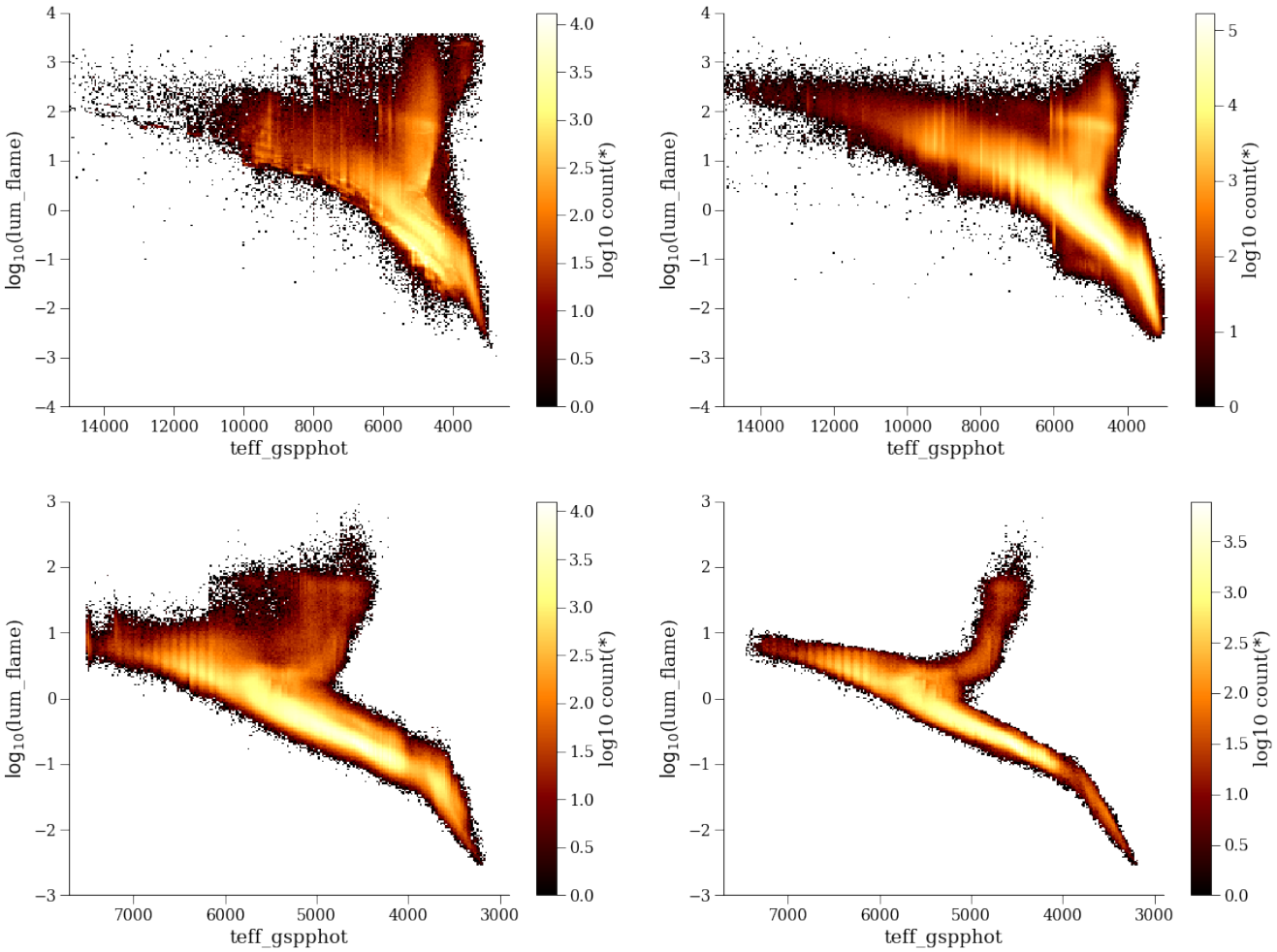

Fig. 9. HR diagram based on GSP-Phot and FLAME for the definition of the FGKM sample. The top left panel illustrates the HR diagram before any selection is made using a random sample of 2M stars. The rest of the panels show the various quality cuts. Top right is fgkm_1, bottom left is fgkm_2, and bottom right is fgkm_3 before cleaning for variables and binaries.

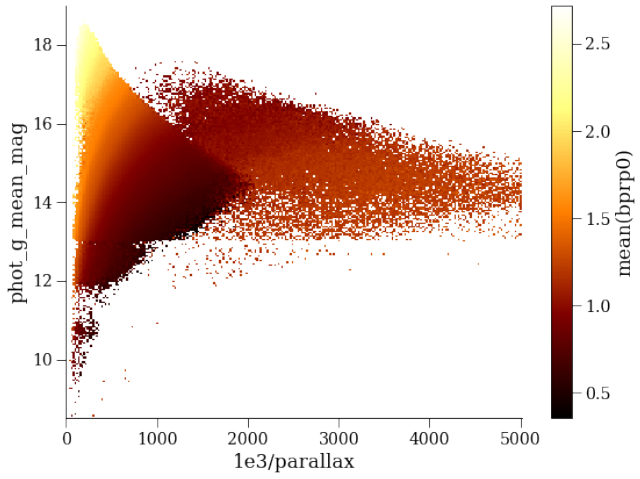

Fig. 10. Distribution of the final sample fgkm_3 of the observed parameters G and parallax, colour-coded by $(G_{\rm BP} - G_{\rm RP})_0$.

The remaining flag characters are related to element abundances and CN and Diffuse Interstellar Band (DIB) features and were not taken into account for the current selection. This resulted in about 1.9 million sources.

For the further selection we considered the quality parameters `parallax_over_error` and `rvs_spec_sig_to_noise`. The latter contains the signal-to-noise ratio in the mean RVS spectrum and is provided only for stars for which the mean RVS spectrum is published in *Gaia* DR3. We produced HR diagrams (`lum_flame_spec` versus `teff_gspspec`) and Kiel diagrams (`logg_gspspec` versus `teff_gspspec`) by imposing different lower limits on `rvs_spec_sig_to_noise`[4]. Visual inspection of the HR diagrams showed a group of sources at $T_{\rm eff}$ ~4000 K clustered around unrealistically high luminosities. Applying the criterion `rvs_spec_sig_to_noise` $\geq$ 150 removed 99% of these sources. We combined this with the criterion `parallax_over_error` > 33.34, similarly to what was applied to main sequence stars in the GSP-Phot based sample, resulting in 22 143 sources (~1% of the flag-selected sources), hereafter "fgkm_spec".

The HR and Kiel diagrams for this selection are shown in Fig. 11. The HR diagram displays a distinct giant branch and red clump as well as a region with turn-off stars and a clear main-sequence. However, as can be seen in the Kiel diagram, the log g

[4] 0, 50, 100, 150, 200, 250, 400, 500

values for main sequence stars show a large spread. This is addressed by further filtering described in the next section.

We also compared the distributions of uncertainties in $T_{\rm eff}$, log g, [M/H], and [α/Fe] from GSP-Spec for the flag-selected sample and the fgkm_spec sample, where the uncertainty was defined as half of the difference between the upper and lower confidence levels. We found that the distributions for the latter sample have a smaller width by a factor 3 to 9 and peak at about half the uncertainty compared with the former sample.

4.4. Final sample and table description

We merged the two samples described in Sects. 4.2 and 4.3 with the objective to provide one unique FGKM gold sample. As both sample definitions contain criteria that are not applicable to the other sample, we publish an independent table in the *Gaia* archive, `gold_sample_fgkm_stars`, that also accounts for additional filtering on specific parameters. The description of the published table is given in Table 2. Further filtering is also done based on other archive products. This is described in this section, and it results in a total of 3 273 041 sources.

4.4.1. Filtering of FLAME, GSP-Spec and ESP-HS parameters in samples fgkm_3 and fgkm_spec

The fgkm_3 GSP-Phot sample includes 3 529 613 sources with FLAME parameters, and 3 313 190 with at least one model-dependent parameter (mass, age, evolutionary stage). Figure 12 shows an HR diagram using $T_{\rm eff}$ and L, colour-coded by evolutionary stage (ϵ). There is a region on the giant branch that has low evolutionary stages compared to the bulk of the giant branch. These could be red clump stars that have been incorrectly assigned, because the models that were used to produce these parameters only span from the Zero-Age-Main-Sequence (ZAMS) to the tip of the giant branch. These targets also have masses larger than 2 $M_{\odot}$. Validation of FLAME parameters has shown that the model values are inaccurate when $M > 2\ M_{\odot}$ for giants (Babusiaux et al. 2022; Creevey et al. 2022). We therefore retained `mass_flame`, `age_flame`, and `evolstage_flame` for giants, only if the following conditions were met: log $g < 3.5$ and $M < 2.0\ M_{\odot}$ and age > 1.0 Gyr. For log $g > 3.5$ no filtering was done. This same criteria was applied to the FLAME parameters in the fgkm_spec sample.

The fgkm_spec sample shown in Fig. 11 shows some problems with log g, below a certain threshold. Validation of these values have indicated a systematic offset on the order of 0.3 with respect to external catalogues for main sequence stars, see e.g. Creevey et al. (2022); Fouesneau et al. (2022); Recio-Blanco & et al. (2022). We therefore removed log g when log $g > 4.0$ in order to retain a 'gold' status, and kept all of the other parameters. As explained in the above references, a calibration of this parameter has been provided and a user can safely use the archive values with or without the calibration, depending on their use case.

We retained the `spectraltype` tag from ESP-HS in our table only if it had a quality flag of rank 1 or 2 (out of 5). This is given in the `flags_esphs` field in the `astrophysical_parameters` table as the second character in that string field.

4.4.2. Further filtering of the merged sample

To ensure that our samples are as clean as possible, we further exploited other *Gaia* DR3 products. We removed all sources that were considered variable or non-single stars, by cross-matching our final source list with the source lists given in the `vari_summary` table, which removed 249 020 sources, 4 873 of which are eclipsing binaries. We also removed the sources appearing in any of the non-single star tables `nss_two_body_orbit`, `nss_acceleration_astro`, `nss_non_linear_spectro` or `nss_vim_fl` which removed a further 28 896 sources. We then used the DPAC-Source Environment Analysis Pipeline (SEAPipe) to further check for any new binary contaminants, and this removed a further 16 sources[5].

4.5. Validation of the sample

4.5.1. Validation with clusters

We take advantage of the properties of open clusters to assess the global quality of the FGKM sample. From the FGKM sample, we selected those stars classified as cluster members in the Cantat-Gaudin et al. (2020) catalogue as refined by Gaia Collaboration et al. (2022b). The cross-match between those stars and our sample corresponds to 4 132 stars and contains only cross-matches with the GSP-Phot sample. Using the full set of cluster members, we approximated each cluster with an isochrone and derived reference values of $T_{\rm eff}$ and log g. Using this $T_{\rm eff}$ we derived A_G adopting the literature value of A_V as a proxy of A_0. We made use of the PARSEC isochrone data set (Bressan et al. 2012). Differential extinction was assumed to be negligible inside the clusters for this validation work. This is justified by the fact that our sample excluded clusters younger than 100 Myr.

We compared the $T_{\rm eff}$, log g, M, A_G, and [M/H] reference values with those from GSP-Phot and FLAME in our sample. We adopted the average values of the member's [M/H] as the cluster [M/H]. Table 3 and Figs. 13 and 14 present the results, which show good agreement with the reference values. Stars cooler than $T_{\rm eff} \sim 4500$ K have GSP-Phot parameters that show the largest differences with reference values. This overestimation of $T_{\rm eff}$ at low temperatures often has higher increased extinction in this regime.

4.5.2. Validation with other galactic surveys

We compared the FGKM sample parameters with the ones of the major spectroscopic surveys, using a cross-match computation specifically performed, using the *Gaia* DR2 cross-match software (Marrese et al. 2017, 2019), for the Survey of Surveys project (SoS Tsantaki et al. 2022). The used surveys are APOGEE (DR16, Ahumada et al. 2020), GALAH (DR2, Buder et al. 2018), Gaia-ESO (DR3, hereafter GES, Gilmore et al. 2012), RAVE (DR6, Steinmetz et al. 2020), and LAMOST

[5] The aim of SEAPipe is to combine the transit data for each source and to identify any additional sources in the local vicinity. Its first operation is image reconstruction, where a two-dimensional image is formed from the mostly one-dimensional transit data ($G > 13$ mag), see Harrison (2011). These images are then analysed and classified, based on whether the source is extended, whether additional sources are present or whether the source is an isolated point source within the reconstructed image area (radius of $\sim 2''$). It is this classification which is used to reject sources not found to be isolated point sources, from our sample. The full SEAPipe analysis will be described in Harrison et. al. (in preparation).

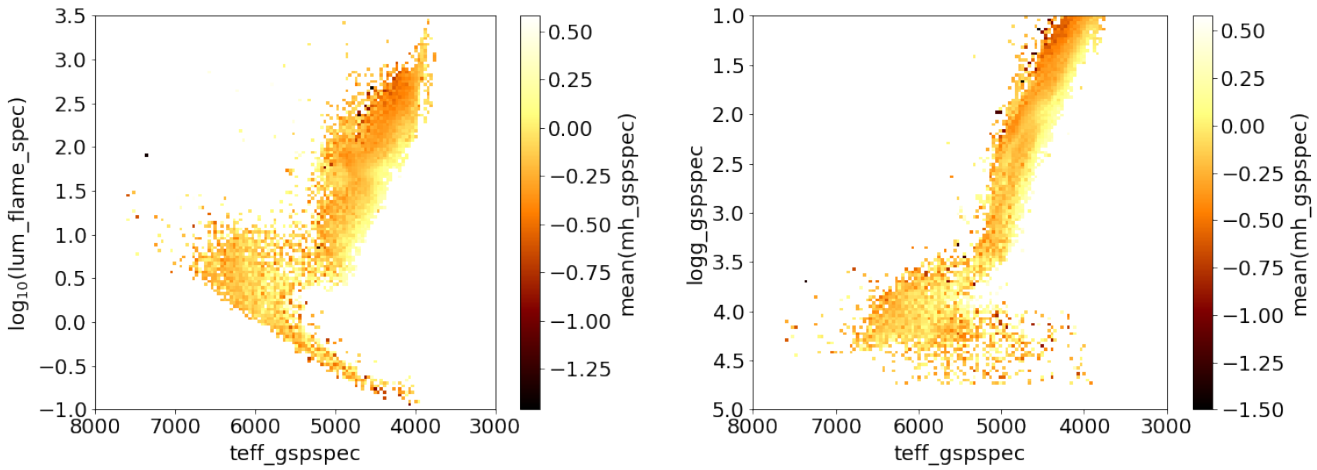

Fig. 11. HR and Kiel diagrams using GSP-Spec-based parameters for the fgkm_spec sample described in Sect. 4.3 colour-coded by the metallicity from GSP-Spec, with `parallax_over_error` ≥ 33.34 and `rvs_spec_sig_to_noise` ≥ 150.

Table 2. Content of the table `gold_sample_fgkm_stars` in the *Gaia* archive. The final column lists the section where the sample is defined in this work. Further filtering on all of these samples is described in Sects. 4.4.1 and 4.4.2.

table	CU8 module	fields	sample definition
ap	GSP-Phot	`teff, logg, mh, ag, ebpminrp, mg`	Sect. 4.2
ap	GSP-Spec	`teff, logg, mh, alphafe`	Sect. 4.3
ap	FLAME	`radius, lum, mass, age, evolstage`	Sect. 4.2
aps	FLAME	`mass, age, evolstage` (`_spec`)	Sect. 4.3
ap	ESP-HS	`spectraltype, flags_esphs`	Sects. 4.2 – 4.3

Notes. The first column indicates the archive table from which the parameters were taken, where ap = `astrophysical_parameters` and aps = `astrophysical_parameters_supp`. The second column indicates the CU8 module responsible for producing the data. The third column indicates the parameter name that is copied from the ap and aps tables to the `gold_sample_fgkm_stars` table. The fourth column indicates the section where the sample definition is described.

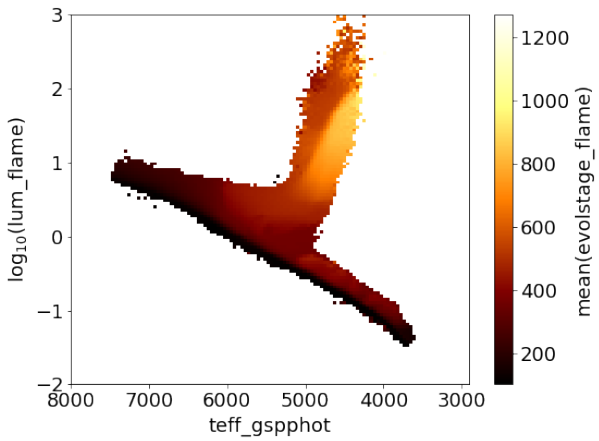

Fig. 12. HR diagram using sample fgkm_3 colour-coded by `evolstage_flame`. The low values of evolution stage on the giant branch correspond to the FLAME parameters that were removed from the table, see Sect. 4.4.1 for details.

Table 3. Differences in GSP-Phot and FLAME parameters from isochrone-fitted values for stars of the FGKM sample in clusters. ΔP is given in the sense of *Gaia* DR3 value minus the cluster value. MD and MAD indicate the median and median absolute deviation of the differences, respectively.

P	ΔP.MD$^{(1)}$	ΔP.MAD$^{(2)}$	Units
$T_{\rm eff}$	-94	136	K
log g	-0.09	0.04	dex
[M/H]	-0.20	0.07	dex
A_G	0.05	0.09	mag
M	-0.04	0.05	$M_\odot$

(DR5, Deng et al. 2012). For each survey, we applied the quality selection criteria suggested in the relevant survey papers and summarized by Tsantaki et al. (2022)[6]. We further removed all the confirmed and candidate spectroscopic binaries identified in the surveys (Merle et al. 2017; Birko et al. 2019; Qian et al. 2019;

[6] The SoS is based on *Gaia* DR2, thus we used the cross-match between DR2 and EDR3 to find the updated source IDs. We removed all sources with a DR2-DR3 magnitude difference higher than 0.5 mag, angular difference higher than 0.5", and all sources with more than one neighbour or mate.

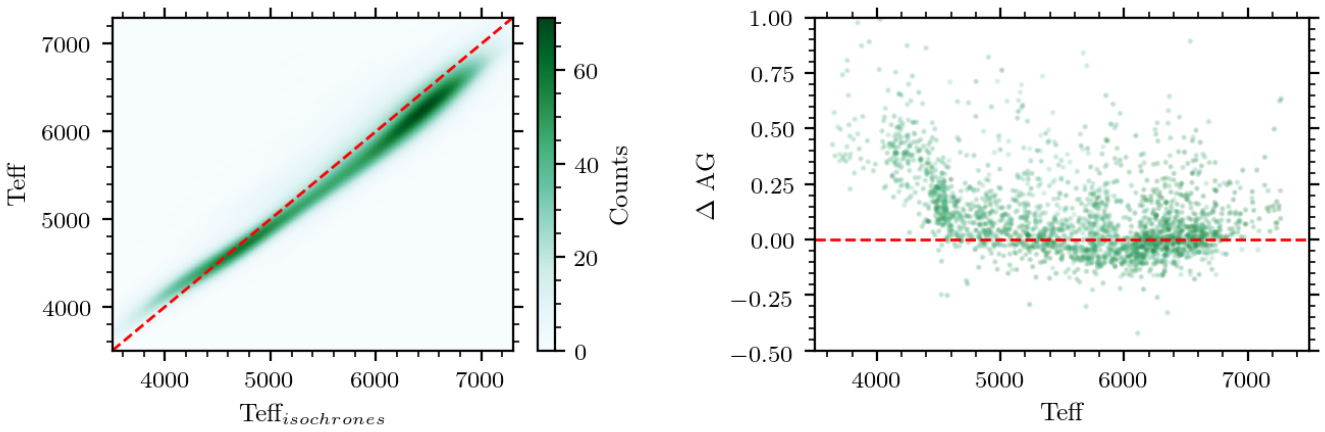

Fig. 13. Comparison of $T_{\rm eff}$ and A_G from GSP-Phot compared to the reference values from isochrones for stars of the FGKM sample in clusters. *Left:* Comparison of $T_{\rm eff}$, with colour indicating the density of sources. The red line indicates the one-to-one values. *Right:* $\Delta A_G = A_{G,\rm GSP-Phot} - A_{G,\rm isochrones}$ versus $T_{\rm eff\,GSP-Phot}$.

Table 4. Comparison of the GSP-Phot and GSP-Spec parameters with the ones from the five main spectroscopic surveys, for the FGKM sample.

Survey	Set	$\Delta T_{\rm eff}$ (K)	# $T_{\rm eff}$	$\Delta\log g$ (dex)	# $\log g$	Δ[Fe/H] (dex)	# [Fe/H]
APOGEE	GSP-Spec	–45 ± 79	2942	–0.38 ± 0.13	2484	–0.08 ± 0.10	2942
	GSP-Phot	–33 ± 100	22160	–0.03 ± 0.09	22155	0.04 ± 0.14	22149
GALAH	GSP-Spec	–105 ± 73	1572	–0.57 ± 0.31	1498	–0.18 ± 0.10	1571
	GSP-Phot	–64 ± 108	28545	0.05 ± 0.13	28545	0.00 ± 0.16	28545
Gaia-ESO	GSP-Spec	55 ± 0	1	–0.03 ± 0.00	1	0.00 ± 0.00	1
	GSP-Phot	–115 ± 92	745	–0.01 ± 0.12	650	0.01 ± 0.15	680
RAVE	GSP-Spec	–55 ± 77	15108	–0.38 ± 0.13	13955	–0.12 ± 0.09	15108
	GSP-Phot	–81 ± 117	21374	–0.01 ± 0.04	21374	–0.07 ± 0.14	21374
LAMOST	GSP-Spec	–29 ± 71	1260	–0.50 ± 0.25	1072	–0.09 ± 0.10	1260
	GSP-Phot	–83 ± 105	299148	–0.02 ± 0.11	299148	0.02 ± 0.16	299148

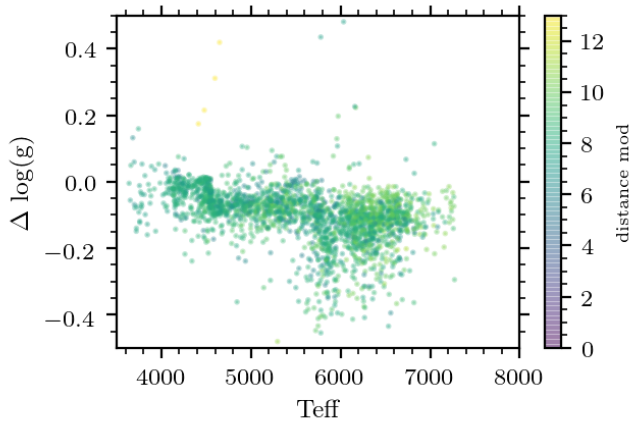

Fig. 14. $\Delta\log g = \log g_{\rm GSP-Phot} - \log g_{\rm isochrones}$ versus $T_{\rm eff\,GSP-Phot}$ for stars of the FGKM sample in clusters. The color indicates the distance modulus $(m - M)$ as derived from the GSP-Phot distance.

Price-Whelan et al. 2020; Tian et al. 2020; Traven et al. 2020; Kounkel et al. 2021). The summary of the number of FGKM stars from the golden sample found in each survey is given in Table 4, where the median differences of the main parameters, computed in the sense *Gaia* minus the surveys, are reported together with their MAD (median absolute deviation). A graphical comparison for the main parameters can be found in Figure 15.

The $T_{\rm eff}$ comparison shows agreement with all surveys, both in GSP-Phot and GSP-Spec, within uncertainties. The median offsets for GSP-Spec are generally negative and of the order of –50—100 K, and the same is true for the GSP-Phot offsets. The spreads range from roughly ±70 to ±120 K, in line with expectations. We note that the surveys agree with each other within a few tens of K, at least in the central portion of the $T_{\rm eff}$ range. Figure 15 shows some systematic substructures in the comparisons. For GSP-Spec, we find good agreement in $T_{\rm eff}$. At the extremes of the $T_{\rm eff}$ range, some discrepancies occur between GSP-Phot and the comparison with LAMOST, which has the lowest resolution among the surveys.

The GSP-Spec $\log g$ comparison shows an offset of about –0.3 dex, which is a known feature, as reported in Section 9.1.1

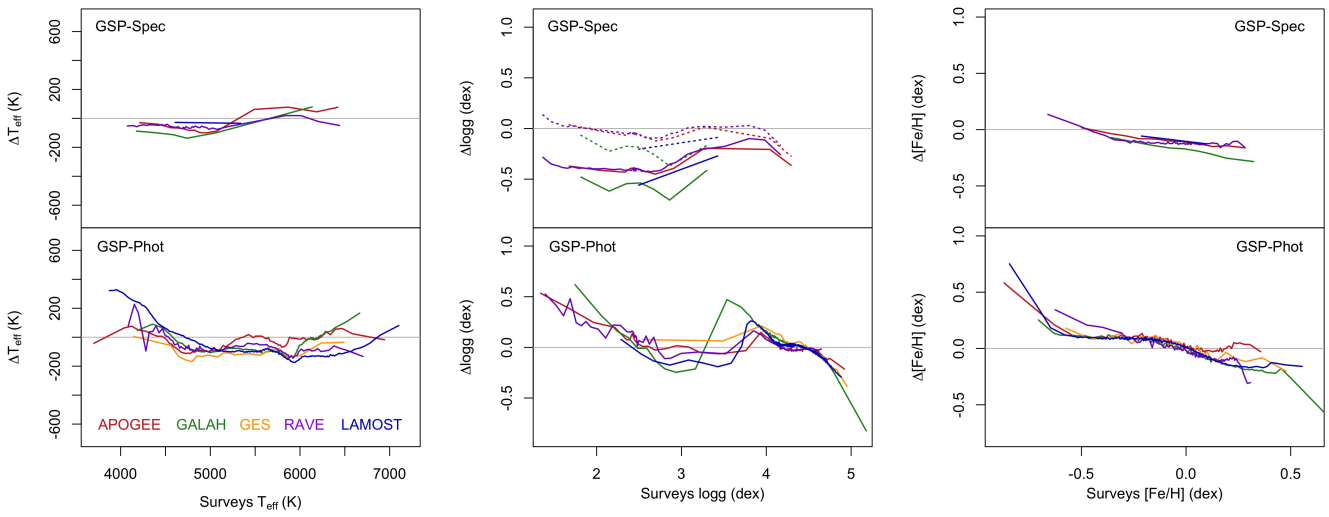

Fig. 15. Comparison of atmospheric parameters with the spectroscopic surveys for the FGKM sample. The top panels show the comparison of GSP-Spec parameters and the bottom panels GSP-Phot. The left panels show the case of $T_{\rm eff}$, the middle ones of log g and the right ones of [Fe/H]. The differences on the y-axes are in the sense of *Gaia* minus the surveys. Each survey is represented as the medians of equally populated bins (solid lines, colored according to the legend in the bottom-left panel). The dotted lines for the GSP-Spec log g are obtained after the corrections recommended by Recio-Blanco & et al. (2022).

by Recio-Blanco & et al. (2022, see their equations 1 and 2), while the GSP-Phot comparison shows excellent agreement with the surveys. When applying the recommended correction to the GSP-Spec log g (dotted lines in Figure 15), the offsets and main trends are highly mitigated. The spreads in the comparisons are roughly around ±0.1 dex in GSP-Phot and up to ±0.2–0.3 dex in GSP-Spec (before correction). The GSP-Phot estimates show good agreement for the subgiants, and most of the dwarfs, and disagreements at the level of up to 0.3 dex is found for the very high (> 4.5) and low (< 2) log g stars. Again, we note that the surveys agree with each other to 0.1–0.2 dex, approximately over most of the log g range.

For metallicity, we use [Fe/H] as an indicator, to be more in line with what is commonly measured by the surveys, which we computed from [M/H] and [α/Fe] using the formula by Salaris et al. (1993). Again, we note a better agreement of the GSP-Phot parameters with the surveys than for the GSP-Spec ones in terms of median offset, which is about zero dex for GSP-Phot and 0.1 dex for GSP-Spec. This was also reported by Recio-Blanco & et al. (2022, see their equations 3 and 4). The spreads are in both cases of about 0.10–0.15 dex, which is more than reasonable. We note that the surveys themselves tend to agree with each other to 0.1 dex or better. There is a tendency of both the GSP-Spec and GSP-Phot parameters to overestimate the [Fe/H] of metal-poor stars and to underestimate it for metal-rich ones. This effect has been commonly observed in several other projects where the parameters were derived from low- or medium-resolution spectroscopy or photometry.

In conclusion, the overall agreement with the main spectroscopic surveys is good, but there are substructures in the comparisons that need to be kept in mind. Additionally, depending on the type of stars, we note that the GSP-Spec parameters do not necessarily produce a better agreement with the survey results compared with the GSP-Phot ones, and the use of the GSP-Spec log g and [M/H] corrections (Recio-Blanco & et al. 2022) is recommended. This is in part due to the fact that the RVS spectral range extent and resolution is limited, but also the fact that we are dealing with a high S/N regime, free from major systematic problems, where both the GSP-Spec and GSP-Phot perform close to optimal.

4.5.3. Validation with the PLATO input catalogue

We cross-matched our source list with the PLATO input catalogue (PIC) version 1.1 (Montalto et al. 2021) and obtained 10 828 common sources. In Fig. 16 we compare $T_{\rm eff}$, R, and M (in the sense *Gaia* – PIC) normalised by the combined uncertainties added in quadrature. We also show the $\pm 3\sigma$ lines, which shows good agreement between the catalogues, but some insignificant artefacts for the comparison of masses. While we show the differences in terms of σ, we report the median and the median absolute deviation of their differences in absolute values on each panel. The agreement with $T_{\rm eff}$ is similar to that reported in the previous sections, where the GSP-Phot one is on average 50 K smaller. There are no matches with the GSP-Spec sources. Radius and mass differences are on the order of 1% and 6%, respectively.

In conclusion, we have made a clean sample of 3 273 041 FGKM stars, comprising main sequence, subgiant, and giant stars. This sample was selected using many *Gaia* based indicators along with GSP-Phot- and GSP-Spec-based astrophysical parameters. The APs of interest are $T_{\rm eff}$, log g, [M/H], A_G, $E(G_{\rm BP} - G_{\rm RP})$, L, R, M, age and spectral type, and we provide a separate table of these parameters in the *Gaia* archive. We have not applied any calibration or correction to the values in *Gaia* DR3, but we have filtered some parameters for some sources. We validated our selection by comparing with parameters from clusters and other surveys which show typical offsets of < 100 K in $T_{\rm eff}$ with other surveys. In Sect. 10.2 we exploit this sample's $T_{\rm eff}$, radius, and mass to analyse known exoplanet systems, and in Sect. 10.4 we analyse the ages of 11 unseen UCD-companions. A user could further filter by selecting in a specific $T_{\rm eff}$ range, or by excluding distances further than a certain distance, or by providing an upper limit to the amount to extinction between the observer and the star.

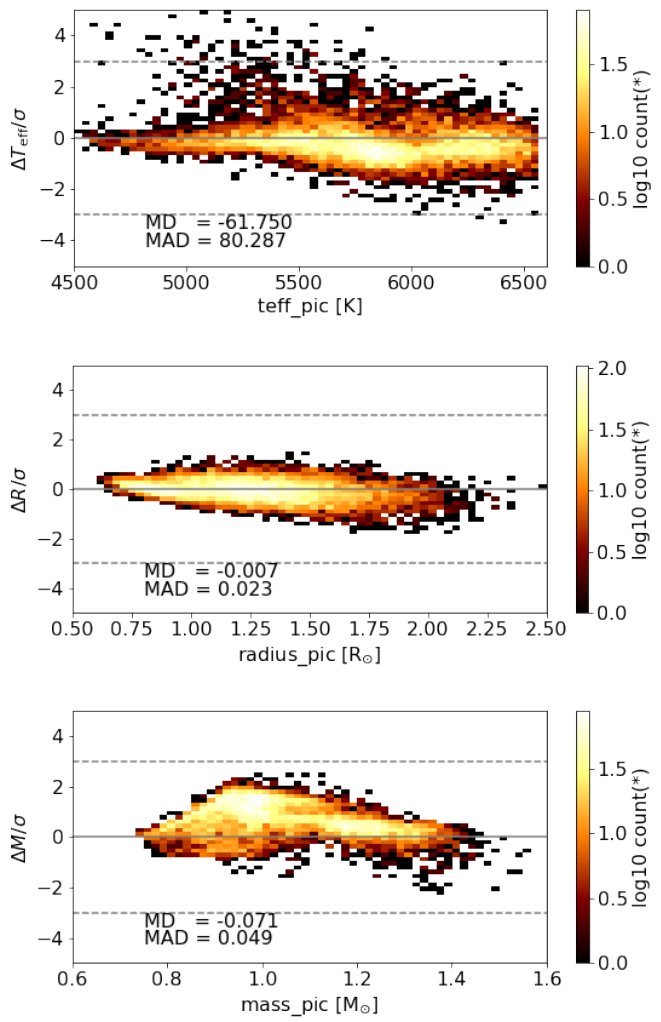

Fig. 16. Difference between $T_{\rm eff}$ (top), R (middle), and M (bottom) from *Gaia* for the FGKM sample and the PICv1.1 catalogue values normalised by their combined uncertainties for stars in common. We overlay the $\pm 3\sigma$ lines. On each panel we also give the median difference ('MD') and the median absolute deviation ('MAD') in K, $R_\odot$ and $M_\odot$, respectively.

5. UltraCool dwarfs

5.1. Scientific motivation

Ultracool dwarfs (UCDs) are objects at the faint end of the main sequence. They were defined in Kirkpatrick et al. (1997) as sources with spectral types M7 or later. This definition includes the coolest hydrogen burning stars and brown dwarfs. Even though brown dwarfs can sustain lithium or deuterium fusion at their cores for a short period of time in the early phases of their evolution (Burrows et al. 2001), the nuclear reactions stop by the time they reach the main sequence and they keep cooling and fading thereafter. Despite the fundamental differences in the internal structure across the stellar/sub-stellar regime, the atmospheric properties overlap at this boundary and it becomes very difficult to distinguish between the two regimes based on photometric or spectroscopic properties. In this section we define a high quality sample of UCDs which we propose as excellent candidates to advance our knowledge of these low mass objects. To complement the $T_{\rm eff}$ of UCDs in *Gaia* DR3 we provide a catalogue of radii and luminosities by complementing the *Gaia* data with infra-red photometry and we explore the existence of a minimum in the mass-radius relation slope (e.g. Dieterich et al. 2014; Smart et al. 2018; Cifuentes et al. 2020).

5.2. Sample selection

Our initial sample of UCD candidates is from the `astrophysical_parameters` table where a total of 94 158 sources have been processed as UCD candidates and estimate `teff_espucd`[7]. We also imposed that the first digit of `flags_espucd` = 0 or 1 (the most reliable categories), which gives a total of 67 428 candidates. We then require that the *Gaia* astrometric flags fulfil the following conditions: `ruwe` < 1.4, `ipd_frac_multi_peak` = 0 and `ipd_gof_harmonic_amplitude` < 0.1 to reduce contamination by unresolved binaries. We then select sources with a cross-match (as provided in the *Gaia* archive) in the 2MASS (Skrutskie et al. 2006) and AllWISE (Wright et al. 2010; Mainzer et al. 2011) catalogues, with available measurements in the J, H and K_s 2MASS bands, and the $W1$ and $W2$ AllWISE bands, all with quality A flags. The $W3$ band was not included as a requirement because the lack of measurement uncertainties reduces drastically the number of sources. Finally, we remove sources above the $G + 5\log_{10}(\varpi) + 5 = 3 + 2.5\,(G - J)$ line to avoid including suspected low gravity UCDs that have not yet contracted and reached equilibrium. This gives a total of 31 822 candidates for this study.

We used the virtual observatory VOSA (Bayo et al. 2008) to calculate the minimum reduced χ^2 fit to the spectral energy distributions constructed using the *Gaia* G and $G_{\rm RP}$ bands and the near- and mid-infrared photometry listed above to CIFIST 2011_2015 BT Settl models (Allard et al. 2012). We retain the sources whose reduced $\chi^2_r < 100$. We allow for rather large values of χ^2_r in order to account for the known discrepancies between the models and observations, and the discrete nature of the model library. The distribution of $\log_{10}(\chi^2_r)$ is approximately normal and 96.5% of all the values are below the imposed threshold which therefore only removes obvious pathological fits. The final sample has a total of 21 068 sources.

5.3. Combining Gaia with external data to derive R and L

R and L are parameters that are also calculated by the FLAME module and available in the `astrophysical_parameters` table, however these are only available for sources with $T_{\rm eff} > 2\,500$ K. A comparison of the values for the sources in common is discussed in the next section. We computed bolometric fluxes using *Gaia* and IR photometry. To account for the unobserved flux outside the observed wavelength bands we needed to calculate bolometric corrections. We used again the CIFIST 2011_2015 BT Settl models in order to calculate the ratio of observed to total flux for the aforementioned set of photometric bands[8]. This produces a theoretical flux correction factor f_{obs}/f_{total} for the $T_{\rm eff}$ range between 1200 and 2700 K in steps of 100 K. For each of the UCD candidates with full photometry we obtain the correction factor by interpolating the $T_{\rm eff}$ value derived by the ESP-

[7] We remind readers that the main difference with respect to existing compilations of UCD candidates (for example Reylé 2018) is the use of the *Gaia* DR3 RP spectra to produce $T_{\rm eff}$ and to help define the selection criteria as described in the *Gaia* DR3 online documentation.

[8] See the help and documentation of VOSA for an updated description of how the band wavelength overlaps are handled in computing the observed flux

UCD module in this grid. The resulting corrections are in the range between 0.48 and 0.54 mag with a median value of 0.53 mag. We use this ratio to infer the total flux that would be observed at the Earth and derive the bolometric luminosity using the *Gaia* parallax measurement. Finally, using the ESP-UCD $T_{\rm eff}$ estimate and the bolometric luminosity, we inferred radii for the UCD candidates using the Stefan-Boltzmann law. Figures 17 and 18 show the scatter plot of the inferred radii and luminosities as a function of the ESP-UCD $T_{\rm eff}$. The uncertainties (represented as error bars only for sources cooler than 1900 K to aid readability) were calculated using a simple Taylor expansion and neglecting correlations amongst the intervening variables.

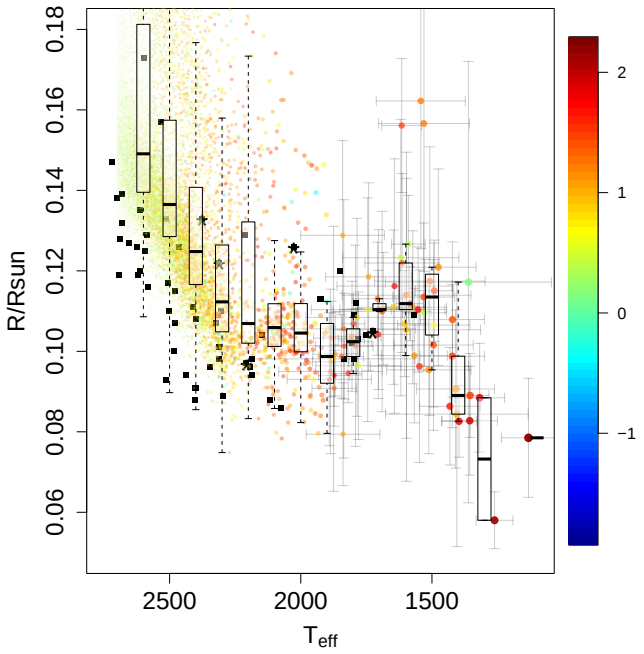

Fig. 17. Radii of candidate UCDs in the *Gaia* golden sample. The colour code indicates the logarithm of the VOSA fits χ^2 values, squares represent the data points in Table 1 of Dieterich et al. (2014) and black asterisks denote unresolved binaries therein. The boxplots are calculated within bins of 100 K.

To fully exploit this UCD golden sample, we provide an accompanying table in the *Gaia* archive `gold_sample_ucd` which lists `source_id`, the correction factor to calculate the bolometric flux, radius, luminosity, and uncertainties, along with the χ^2_r value. This table can be used with the `teff_espucd` provided in the `astrophysical_parameters` table.

5.4. Validation

In Figure 19 we compare the radii values of sources with estimates from the FLAME and ESP-UCD modules. It shows a remarkable agreement for the lowest temperature regime ($T_{\rm eff}$ < 2600 K) but it also shows evidence for a systematic difference in the sense of larger FLAME radii above. This is due in part to a difference of approximately 85 K in the temperatures used for the derivation of radii (in the sense of the $T_{\rm eff}$ used by FLAME– from GSP-Phot– are hotter than the ones produced by the ESP-UCD module).

Figure 17 shows the expected decrease in radius as the temperature decreases down to temperatures of the order of ≈ 2200 – 2000 K. Then, the radii increases for even cooler temperatures until $T_{\rm eff}$ ≈ 1400 K where the trend reverses and the slope becomes positive again.

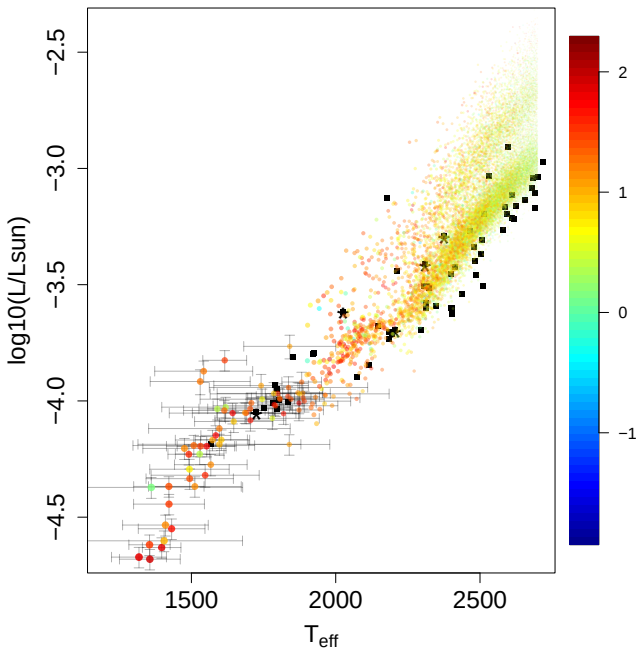

Fig. 18. Bolometric luminosities of candidate UCDs in the *Gaia* golden sample. The colour code indicates the logarithm of the VOSA fits χ^2 values, squares represent the data points in Table 1 of Dieterich et al. (2014) and black asterisks denote unresolved binaries therein.

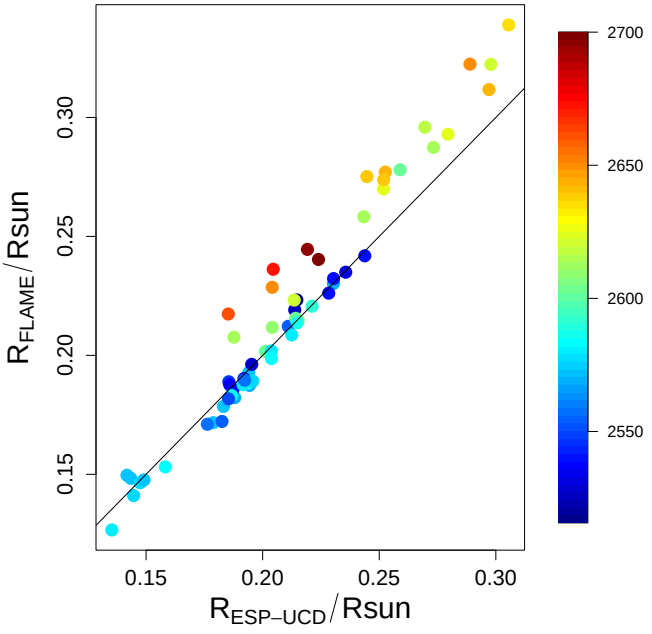

Fig. 19. Comparison of the radii estimated for the UCD sample by the ESP-UCD (x axis) and FLAME (y axis) modules for the sources in common. The colour code reflects the effective temperature used by the FLAME module to estimate the radii.

In Fig. 18 we can see a systematic difference between the luminosities estimated by Dieterich et al. (2014) (represented by the black squares) and the ones from this work, in the range of temperatures $T_{\rm eff}$ > 2000 K. This difference translates into an

offset in radii in Figure 17. The offset in luminosity can be due either to (1) a difference in the $T_{\rm eff}$ estimates if our temperatures were systematically cooler than those of Dieterich et al. (2014) in that regime and/or (2) a difference in the calculation of the bolometric correction (BC) if BCs derived by Dieterich et al. (2014) produce bolometric luminosities systematically fainter than the ones derived here. We examine the two alternatives more closely in the following paragraphs.

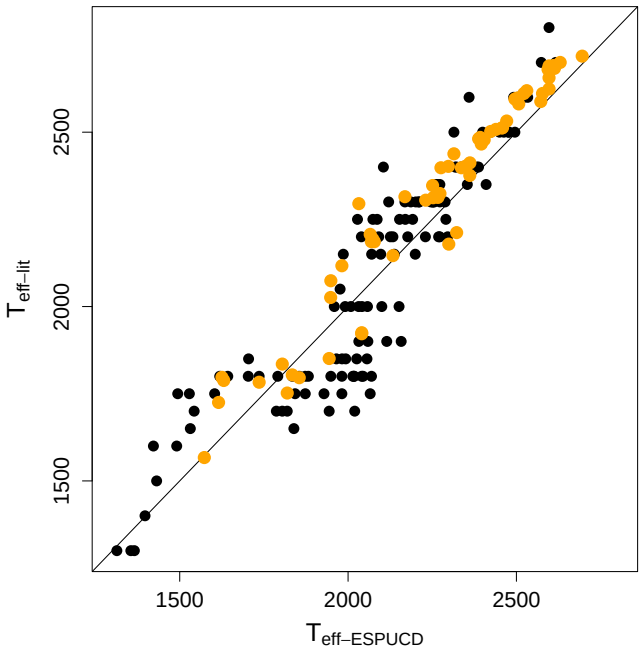

Fig. 20. Comparison of the effective temperatures used to derive radii in this work (x axis) and those used in the literature (y axis) for the UCD sample. Black filled circles denote sources from Cifuentes et al. (2020) and orange filled circles, those from Dieterich et al. (2014).

Figure 20 shows a comparison of the temperatures used by Dieterich et al. (2014) and Cifuentes et al. (2020) to infer radii with those estimated by the ESP-UCD module. It shows hints of a systematic difference of approximately 65 K above $T_{\rm eff} \approx$ 2200 K. This is different from, but consistent, with the difference encountered in the comparison with the FLAME outputs.

The ESP-UCD $T_{\rm eff}$ are based on an empirical training set built from the the *Gaia* UltraCool Dwarf Sample (GUCDS; Smart et al. 2017, 2019) and the spectral type-$T_{\rm eff}$ relation by Stephens et al. (2009). The values derived by the ESP-UCD regression module were calibrated as described in the *Gaia* DR3 online documentation to account for a discrepancy that was found with respect to the regression module trained on BT-Settl models. The RP spectra, simulated from the BT-Settl library of synthetic spectra, were found to reproduce well the observed RP spectra in this $T_{\rm eff}$ regime. Also, the calibrated temperatures were found to produce relatively good agreement with the SIMBAD spectral types where available (again, using the relations by Stephens et al. 2009) as illustrated in the validation of the ESP-UCD module in Fouesneau et al. (2022). However, in view of the comparisons described above, it is not implausible that the correction applied in the calibration of the results from the empirical training set was overestimated by an amount of the order of 65 K. In any case, the systematic difference in effective temperatures explains part but not all of the discrepancy in the luminosities/radii. Hence, we suspect that this discrepancy may also be caused by differences in the corrections applied to the observed fluxes to derive bolometric luminosities. Our procedure to estimate the bolometric luminosity is different from that used by Dieterich et al. (2014) and this can be the source of the systematic difference in the luminosities above 2000 K apparent in Figure 18. While we interpolate directly the fraction of the total flux emitted in the photometric bands on a grid of BT Settl models, Dieterich et al. (2014) apply a wavelength dependent correction to the BT Settl models such that they agree with the observed photometric magnitudes before that fraction is estimated. Since a direct comparison is not possible due to the unavailability of their correction factors, we cannot discard this different procedure as a potential explanation of the difference.

The overall trends of decreasing radii down to $\sim$ 2000 K and slowly increasing radii for even cooler temperatures are confirmed with the *Gaia* data although the associated uncertainties are large. The final positive slope in the regime $T_{\rm eff} < 1400$ K is also compatible with that shown in Cifuentes et al. (2020) but not predicted (to the best of our knowledge) by theoretical studies. The sample of UCDs used here can be expected to be a combination of different ages, masses and metallicities (all of them with an impact on the effective temperatures and radii) and hence, no direct conclusion about these fundamental parameters can be easily drawn from Figure 17 .

In summary, we provide a catalogue of 21 068 UCDs that we consider to be of very high quality from the available sources in the `astrophysical_parameters`. We derive their luminosities and radii by calculating bolometric corrections and make these new parameters available in the accompanying `gold_sample_ucd` table.

6. Carbon stars

6.1. Scientific motivation

A high number of Asymptotic Giant Branch (AGB) stars have carbon enriched atmospheres and show C_2 and CN molecular bands stronger than usual in stars cooler than 3 800 K (i.e. $G_{\rm BP} - G_{\rm RP} \geq 2$). The origin of the enrichment can be due to mass transfer in binary systems or due to the pollution by nuclear He fusion products from the inner to the outer layers. Because they belong to a late stage of stellar evolution where mass loss occurs and which precedes the formation of the planetary nebula, carbon stars are important contributors to the interstellar medium and provide good reference cases to study the physical processes affecting the end of the life of low mass stars. During the *Gaia* DR3 development and processing, no synthetic spectra showing such high carbon abundances were included in the simulations that are used in the Apsis software to produce APs (Creevey et al. 2022). Hence, the spectral libraries used as templates to derive the astrophysical parameters from BP and RP, as well as those adopted to measure the radial velocities, are not fully adapted to analyze the data of carbon stars. Therefore, an attempt was made by the ESP-ELS module to flag suspected carbon stars.

6.2. Sample selection

The identification of candidate carbon stars by ESP-ELS is based on a random forest classifier trained on the synthetic BP and RP spectra as well as on the observed *Gaia* data obtained for a sample of galactic carbon stars (Abia et al. 2020). This identification is saved in the `spectraltype_esphs` field of the `astrophysical_parameters` table. In total, 386 936 targets re-

Table 5. Molecular band head strength used to identify the most probable carbon stars

Strength (molecule)	λ_1 [nm]	λ_2 [nm]	λ_3 [nm]
$R_{482.3}$ (C_2)	462.2345	482.3455	505.3195
$R_{527.1}$ (C_2)	505.3195	527.1080	546.5995
$R_{773.3}$ (CN)	716.5865	773.2905	810.7805
$R_{895.0}$ (CN)	806.8910	894.9855	936.6820

ceived the "CSTAR" tag. While most of these stars are M stars, only a smaller fraction of the sample exhibit significant C_2 and CN molecular bands. To identify these cases, we measured the band head strength as follows:

$$R_{\lambda_2} = \frac{f(\lambda_2)}{g_{\lambda_1,\lambda_3}(\lambda_2)} \quad (2)$$

where $f(\lambda_2)$ is the flux measured at the top of the band head of the molecular band, and g_{λ_1,λ_3} the value linearly interpolated between wavelengths λ_1 and λ_3. The four band heads we considered are described in Table 5. These were computed for a random sample of 27 528 stars having $G_{BP} - G_{RP}$ (not dereddened) colours uniformly distributed between 1 and 5, in order to locate the range of $R_{773.3}$ and $R_{895.0}$ values occupied by non-carbon stars. The upper limit of the interquantile dispersion (2.7 % and 97.3 %) is the threshold below which the targets providing the weakest values are excluded (i.e. it provides one lower threshold on $R_{773.3}$, and one on $R_{895.0}$).

In Figs. 21 and 22, the results obtained for known carbon stars, and for the candidate carbon stars flagged by the ESP-ELS module are reported, respectively. Most of the 386 936 candidate carbon stars (upper panels of Fig. 22) flagged by the algorithm have $G_{BP} - G_{RP} > 2$ mag, and have colors consistent with M stars. However, the known carbon stars, especially in the Magellanic clouds have colours down to ∼1 mag. A significant fraction of these have therefore not been detected and are not part of the golden sample. Our proposed sample of carbon stars is obtained after applying the lower thresholds on both $R_{773.3}$ and $R_{895.0}$ ratios.

6.3. Validation of sample

The sample we propose includes 15 740 stars exhibiting the strongest CN molecular bands. Their spatial distribution is shown in Fig. 24. As previously noted, most of the remaining carbon stars have $G_{BP} - G_{RP} > 2$ mag, which is consistent with what is expected from M-type stars (Fig. 23). From a cross-match with the 3 main catalogues of carbon stars, about two thirds are known cases. The magnitude and color distributions of the targets found in the literature and in common with the proposed sample are shown in Fig. 23. Most of the carbon stars that have not been identified correspond to targets bluer than $G_{BP} - G_{RP} = 2$ or/and fainter than $G = 17.65$ mag. Taking magnitude and color/T_{eff} constraints into account, the fractions of detected known carbon stars are shown in Table 6.

Carbon stars are located at the very cool edge of GSP-Phot's T_{eff} domain ($T_{eff} > 2500$ K). In addition, no synthetic spectra adapted for the accurate AP determination of carbon stars were available, and only a fraction of the carbon stars have their astrophysical parameters published in GDR3. Hence, it is not surprising that the T_{eff} that is obtained tends to be overestimated (by

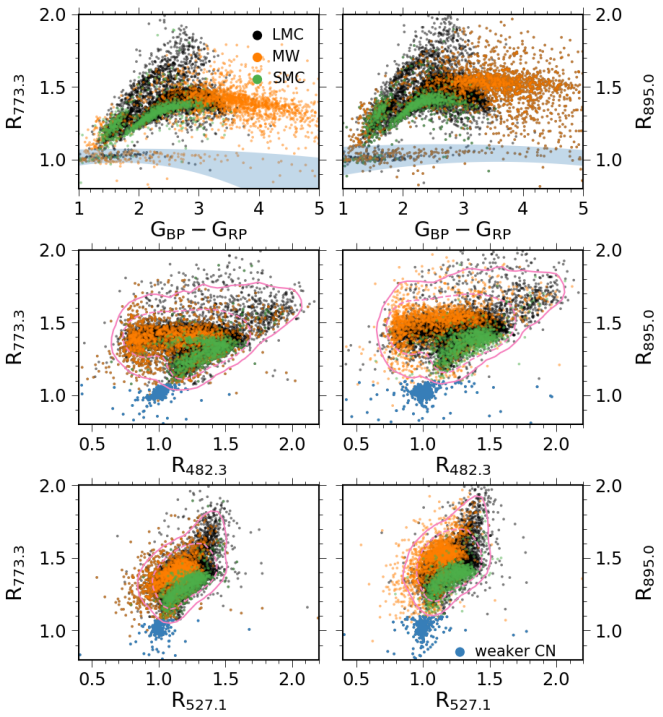

Fig. 21. Band head strengths (see Eq. 2 and Table 5) measured in the BP and RP spectra of known Galactic (MW, orange points, Alksnis et al. 2001), Large Magellenic Cloud (LMC, black points, Kontizas et al. 2001), and Small Magellenic Cloud (SMC, green points, Morgan & Hatzidimitriou 1995) carbon stars. Only targets within 1 arcsecond of a *Gaia* DR3 source_id are taken into account. Upper panels: the locus occupied by non-carbon stars is represented by the blue shaded area. Middle and lower panels: targets with weaker or non-existing CN features are shown with blue points (i.e. they fall in the shaded areas of the upper panels). The pink broken and full lines delimit the domain occupied by 87 % and 98 % of the carbon stars with strong CN features.

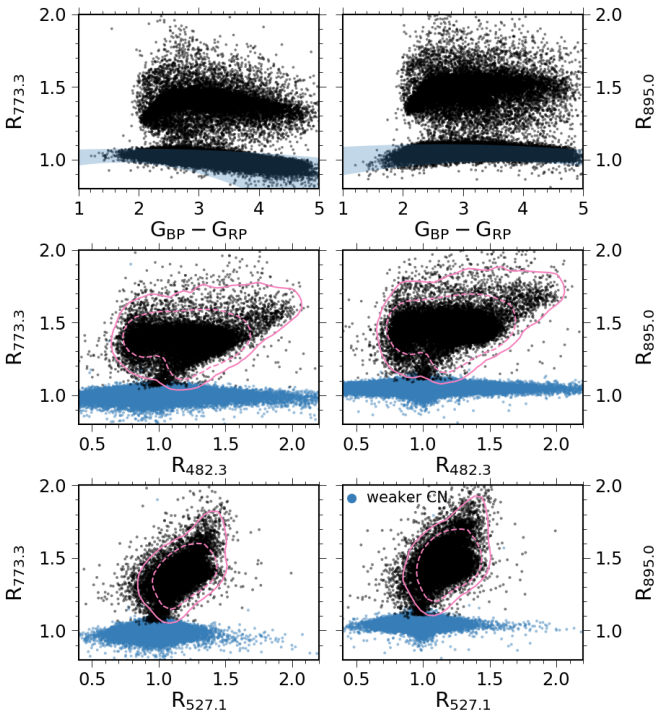

Fig. 22. Same as Fig. 21 but for the 386 936 candidate carbon stars flagged by ESP-ELS. Pink curves represent the domain occupied by the carbon stars found in the literature (Fig. 21, and Sect. 6.2).

Table 6. Fractions of detected known carbon stars.

galaxy	$G \leq 17.65$ & $G_{BP} - G_{RP} \geq 2$	$G \leq 17.65$
MW	0.82	0.70
LMC	0.61	0.54
SMC	0.41	0.27

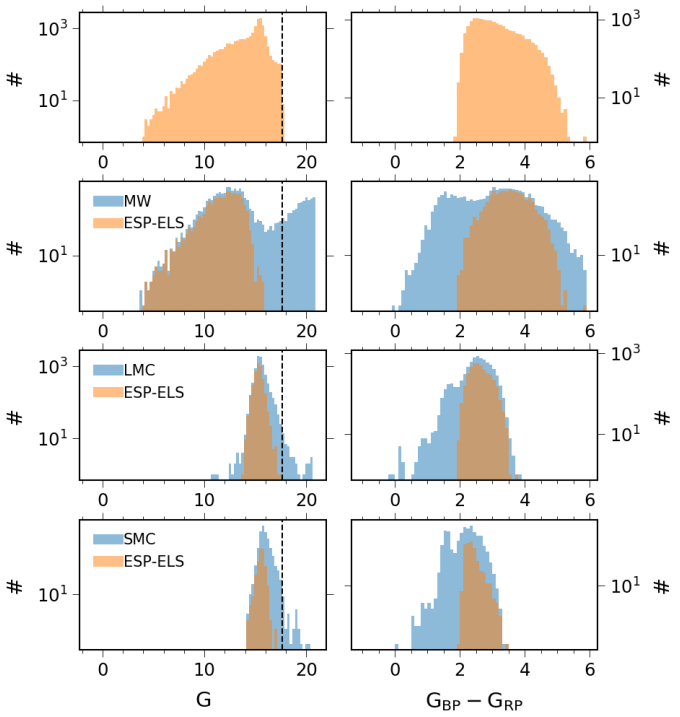

Fig. 23. Magnitude and color distribution of carbon stars. Left panels: The vertical black dashed line shows the upper magnitude limit of the data processed by ESP-ELS. Upper panels: All the targets belonging to the golden sample of carbon stars are taken into account. Other panels: Distributions obtained for the known MW (Alksnis et al. 2001), LMC (Kontizas et al. 2001), and SMC (Morgan & Hatzidimitriou 1995) carbon stars are shown in blue. In orange, we show the distribution of the targets in common with the sample we propose in this work.

500 to 1500 K) and should be considered with caution. However, the Kiel diagrams obtained for the known carbon stars (Fig. 25, left panel) and those from our list (same Figure, right panel) are consistent with each other. Notwithstanding the estimated $T_{\rm eff}$ and their location in the diagram is also consistent with AGB stars. Note that a few targets (254) have $T_{\rm eff}$ hotter than 6 000 K, while the corresponding SEDs are typical of AGB carbon stars (showing typical CN bands in the RP) as shown in Fig. 26.

To exploit this sample, the list of `source_id` are made available as a separate table in the archive `gold_sample_carbon_stars` for the 15 740 bone fide carbon stars, which were also flagged in the main `astrophysical_parameters` table (see `flags_esphs` for details). The initial set of 386 936 carbon-candidate stars can still be found in the same table, as these remain tagged "CSTAR" in the `spectraltype_esphs` field.

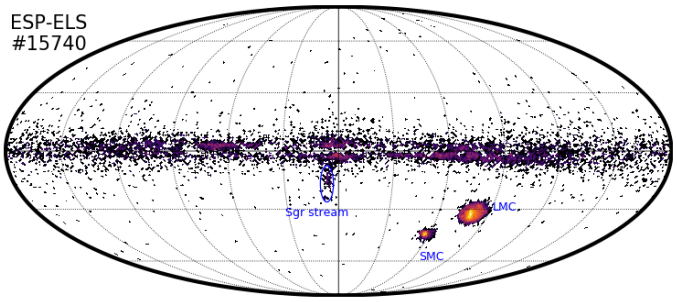

Fig. 24. Mollweide view in Galactic coordinates of the carbon stars sample described in this work. The locations of the Magellanic Clouds, and of the Sagittarius stream are shown in blue.

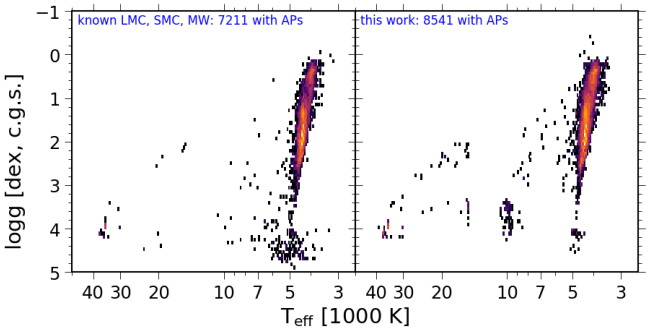

Fig. 25. Kiel diagram of carbon stars with published GSP-Phot parameters. Left panel: density plot for known MW, SMC, and LMC C stars. Right panel: density plot for the carbon stars in our sample.

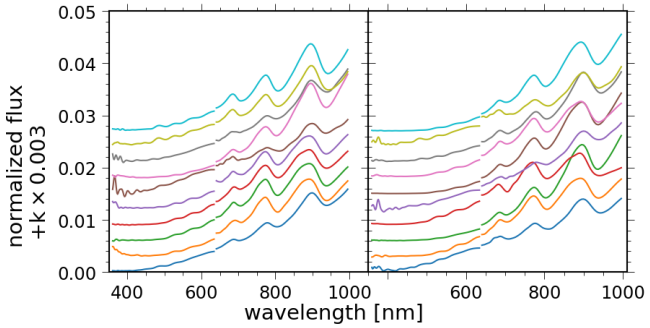

Fig. 26. BP and RP spectra of the 20 randomly chosen (amongst 254) carbon stars (this work) with $T_{\rm eff\ GSP-Phot} > 6000$ K. The ordinate axis provide the flux normalized to the total flux, and shifted by $k \times 0.003$ (where k is an integer that varies from 0 to 9 from the bottom to the top spectrum).

7. Solar analogues

7.1. Scientific motivation

The Sun is the reference point in much of stellar astronomy and astrophysics. Solar analogues are stars which in a restricted set of parameters resemble the Sun. In contrast to the Sun, they can be observed in the night sky and with the very same instruments used to study stars in the Milky Way. There is no strict definition of what constitutes a solar analogue, both the set of parameters and allowed parameter ranges vary in the literature. For astrophysical purposes one often aims to constrain the photometrically/spectroscopically accessible parameters $T_{\rm eff}$, $\log g$ and the overall metallicity [M/H] to within typical measurement uncertainties. Depending on data quality and analysis technique ap-

plied, uncertainties as small as 10 K in $T_{\rm eff}$, 0.03 in $\log g$ and 0.01 in [M/H] are achievable[9] (Yana Galarza et al. 2021), but 50 K, 0.15 and 0.05 are more typical values. These small errors are the result of line-by-line differential analyses relative to the Sun, a technique which cancels many of the systematic sources of errors that stellar analyses otherwise often suffer from.

The most accurate analyses have revealed systematic differences of the chemical composition of the Sun relative to solar analogues in the solar neighbourhood: When selected to be good matches in [Fe/H] (iron abundance), the Sun is among the 10-15% of stars rich in volatile elements (Meléndez et al. 2009). A tight (broken) trend of abundance with condensation temperature of the various elements is found with an amplitude of 0.08 dex (20% in linear abundance). The reason for this effect is still unknown, but is speculated to be related to selective accretion of gas over dust due to the presence of planets. This finding potentially opens up new avenues for systematic evolutionary studies of solar-type stars and their planets.

Solar analogues have also been used to identify the abundance ratios which depend most sensitively on stellar age and can thus serve as precise spectroscopic clocks. One such study identifies the [Y/Mg] abundance ratio as particularly age-sensitive (Nissen 2015). Working with ages rather than metallicity as a proxy for age puts chemical-evolution studies on a much firmer footing. Loosening constraints on the stellar parameters, one can also study "the Sun as a star" and its evolution.

Finally, solar analogues also serve a purpose in the study of minor bodies of the solar system. In this context, they are used to subtract the solar spectrum (and earth-atmospheric contributions in the case of ground-based observations) from reflectance spectroscopy of e.g. asteroids with the aim of a more uniform classification, see for example the `sso_reflectance_spectrum` table in *Gaia* DR3 and Gaia Collaboration et al. (2022c). Note that this type of science case asks for stars whose spectral energy distributions resemble that of the Sun as closely as possible. This requirement does not necessarily ask for a perfect match in stellar parameters, especially if one considers fainter G dwarfs which may suffer from extinction and associated reddening.

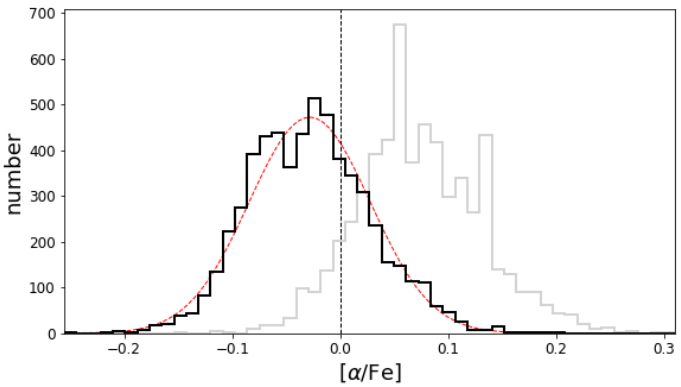

Fig. 27. Distribution of [α/Fe] abundances from GSP-Spec for solar-analogue candidates. Grey shows the raw `alphafe_gspspec` values and black shows the calibrated values (Recio-Blanco & et al. 2022). The dashed red line shows a Gaussian distribution with mean of –0.028 and standard deviation of 0.056.

A scientific application of solar analogues is presented in Section 10.

[9] Stars with parameters indistinguishable from the Sun are sometimes also referred to as solar twins.

7.2. Candidate selection

In order to identify candidates for solar analogues from the full *Gaia* sample, we need to define selection criteria. We apply two general criteria:

1. Apparent magnitude brighter than $G < 16$ since fainter sources would be difficult to follow up efficiently with ground-based spectroscopy.
2. Excellent parallax quality, $\varpi/\sigma_\varpi > 20$, in order to reliably place sources in the HR diagram.

From these basic criteria, we continue to select candidates from GSP-Spec results. On a known sample of solar analogues and twins (Porto de Mello et al. 2014; Ramírez et al. 2014; Nissen 2015; Mahdi et al. 2016; Tucci Maia et al. 2016; Lorenzo-Oliveira et al. 2018; Giribaldi et al. 2019; Casali et al. 2020; Yana Galarza et al. 2021), GSP-Spec on average deviates from solar values by 14.4K in $T_{\rm eff}$, by −0.071 in $\log g$ and by −0.05 in [M/H] (Fouesneau et al. 2022, Sect. 4.1 therein). Taking those average differences into account, we require that GSP-Spec results agree with 5772 K to within 100 K, to $\log g = 4.44$ to within 0.25, and to [M/H] = 0 to within 0.1. Furthermore, we require good GSP-Spec flags[10]. Finally, we combine GSP-Spec results with FLAME estimates to further weed out possible contamination: First, we require that `mass_flame_spec` is between 0.95 $M_\odot$ and 1.05 $M_\odot$. Second, we require that `radius_flame_spec` is between 0.8 $R_\odot$ and 1.2 $R_\odot$. This results in a total of 5863 GSP-Spec candidates for solar analogues, of which 916 have RVS spectra published in *Gaia* DR3[11]. The list of *Gaia* DR3 source IDs for the 5863 solar-analogue candidates from GSP-Spec is provided in the *Gaia* archive as a separate table `gold_sample_solar_analogues`.

Due to the selection on very high parallax quality ($\varpi/\sigma_\varpi > 20$), the candidates tend to be nearby and thus scatter more or less uniformly over the whole sky. Yet, the sky distribution shows the imprint of the *Gaia* scanning law, because high parallax quality is easiest to achieve for sources with many transits.

In Fig. 27, we check and verify that the solar-analogues candidates have [α/Fe] abundances that are statistically consistent with the solar value of zero. The standard deviation of [α/Fe] for this particular subset of solar-like stars is 0.056, which is lower than the global uncertainty reported for all stellar types in Recio-Blanco & et al. (2022).

7.3. RVS spectra of candidates

As a visual confirmation of the candidate selection, we inspect the published RVS spectra of the candidates. For comparison, we also take the RVS spectra of 13 known solar analogues which have RVS spectra published in *Gaia* DR3. Figure 28a shows

[10] `gspspec_flags` equal to 0 in characters 1 to 13, except for 8, and equal to 0 or 1 in character 8. These flag characters are related to the fundamental spectral parameters, see Sect. 4.3 for details. All other flag characters relate to specific elemental abundances and we ignore them in this context.

[11] GSP-Phot results can also be used to select candidates. Here, we only consider PHOENIX and MARCS in the context of solar analogues. We require that GSP-Phot results *for both libraries* agree with 5772 K within 100 K, to $\log g = 4.44$ within 0.25, and to [M/H] = 0 within 0.1, where we correct results from each library for its mean differences to known solar analogues (Fouesneau et al. 2022, Sect. 4.1 therein). This results in a total of 234 779 GSP-Phot candidates for solar analogues, 7884 of which have RVS spectra. However, we do not publish this candidate list. Interested readers may contact the authors.

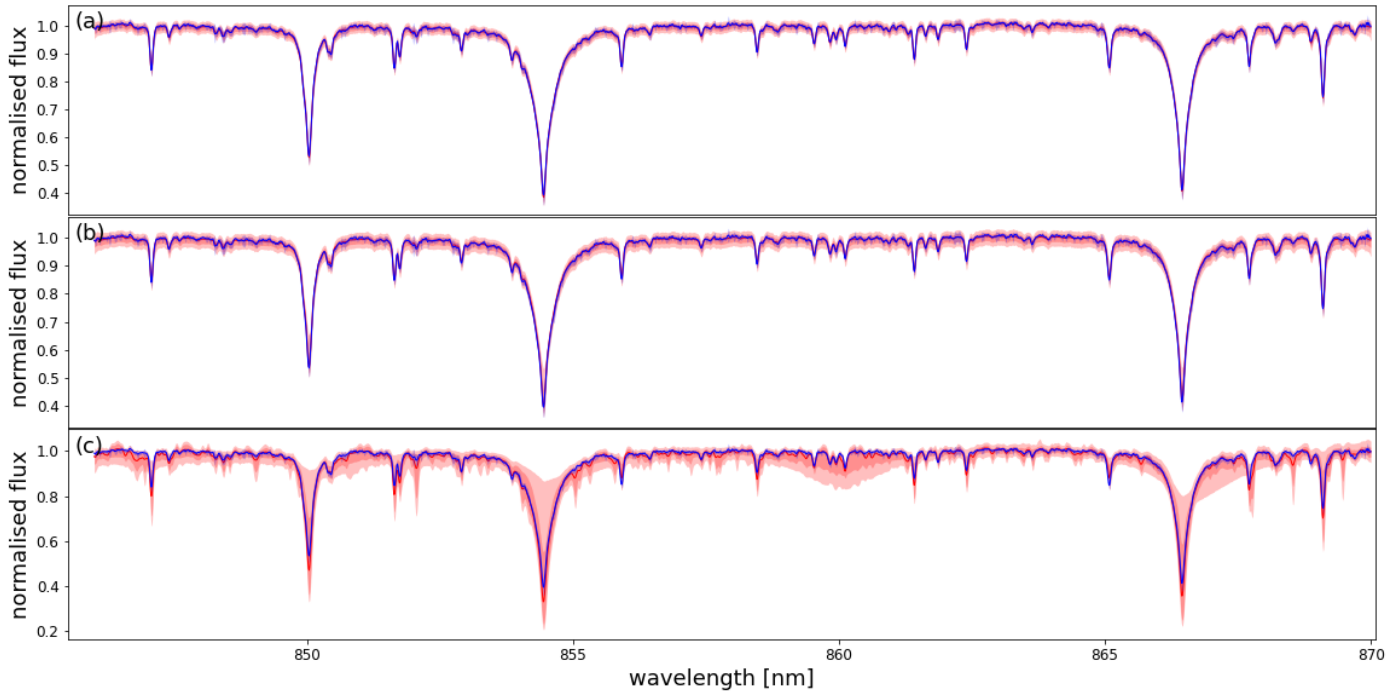

Fig. 28. RVS spectra of 916 solar-analogue candidates from GSP-Spec (panel a) where 95% of GSP-Spec candidates satisfy $G < 11.7$. We also show the solar-analogue candidates obtained from a possible selection from GSP-Phot results in panels b, but we only show 1985 GSP-Phot candidates with $G < 11.7$. For comparison, panel c shows RVS spectra of 7589 randomly selected stars (i.e. no solar-analogue candidates) also with $G < 11.7$. In all panels, the red line shows the median in each pixel and the shaded red contours show the pixel-wise central 68% and 90% intervals. The solid blue line is identical in all three panels and shows the mean RVS spectrum of 13 solar analogues known from the literature.

that the 916 GSP-Spec candidates with RVS spectra are in excellent agreement with the mean RVS spectrum of known solar analogues. Most of these 916 GSP-Spec candidates are brighter than $G < 11.7$. The 1985 candidates with RVS spectra one would obtain from GSP-Phot and that are similarly bright ($G < 11.7$) are shown in Fig. 28b. They show equally good agreement with the known solar analogues as Fig. 28a. This demonstrates that GSP-Phot results are also very reliable under these selection criteria. For orientation, Fig. 28c shows RVS spectra of 7589 random stars with $G < 11.7$ and here we see clear differences, e.g. the Ca lines vary in depth, where for hot stars in particular the Ca lines are usually weak and instead Paschen lines start to appear. In Fig. 28c, we can also see the DIB around 860nm (Gaia Collaboration et al. 2022e).

7.4. Candidates with extinction

Solar analogues with notable extinction would be of particular scientific interest, e.g. for inferring the extinction law. In Fig. 29, we show colours of GSP-Spec candidates including photometry from *Gaia* and AllWISE (Cutri et al. 2021) as a function of GSP-Phot's A_0 estimate. The $G-W_1$ colour clearly reddens with increasing A_0 in Fig. 29a whereas the W_1-W_2 colour remains virtually constant in Fig. 29b.[12] In Fig. 29c, we further investigate the reddening of the $G_{BP}-W_2$ colour, which has the largest wavelength coverage from the near ultra-violet (320-670 nm for G_{BP}) to the near infrared (4.6 μm for W_2). In particular, G_{BP} will be much more affected by extinction than W_2, in fact $A_{BP} \gg A_{W2}$, such that we can take GSP-Phot's A_{BP} estimate as an approximation for the reddening of the $G_{BP} - W_2$ colour. Indeed, Fig. 29c shows a linear relation with a low RMS deviation of 0.087 mag across an A_{BP} range of 1.75 mag. This attests to the quality of the A_{BP} estimate from GSP-Phot (at least for bright sources with high-quality parallax measurements).

Having established in Fig. 29 that the GSP-Phot extinction estimates agree with the reddening of colours of the candidates, we inspect how the low-resolution BP and RP spectra themselves vary with GSP-Phot extinction. For the 5863 GSP-Spec candidates, Fig. 30 shows that the BP and RP spectra clearly redden and dim as the GSP-Phot estimate of A_0 increases. While BP and RP spectra at low extinction show much more flux in BP than in RP, BP and RP spectra at $A_0 \sim 1.5$ mag already show equally high peak fluxes in both BP and RP while their overall flux is reduced by a factor of ~5 in BP and ~3 in RP with respect to a zero-extinction solar-like BP/RP spectrum.

8. SPSS

The *Gaia* Spectro-Photometric Standard Stars[13] (SPSS, Pancino et al. 2012, 2021b) are a grid of flux tables specifically designed to calibrate *Gaia* photometry and BP and RP spectra. They are the result of a dedicated set of ground-based observing campaigns to collect spectrophotometry (Altavilla et al. 2015), light curves for constancy monitoring (Marinoni et al. 2016), and absolute photometry for validation (Altavilla et al. 2021), over more than ten years. The latest version of the grid, SPSS V2, was used to calibrate the *Gaia* photometry in EDR3 and the BP and RP spectra in DR3. It contains 111 stars[14], based

[12] We also inspected the variation of these colours with GSP-Spec's DIB measurements (Gaia Collaboration et al. 2022e) and find qualitatively similar results. Unfortunately, only very few DIB measurements are available for our candidates.

[13] http://gaiaextra.ssdc.asi.it:8900/

[14] Pancino et al. (2021b) list 112 stars, but one (SPSS 192, see their Figure 11) was found to have a close companion at about 0.25" with SEApipe (Harrison 2011).

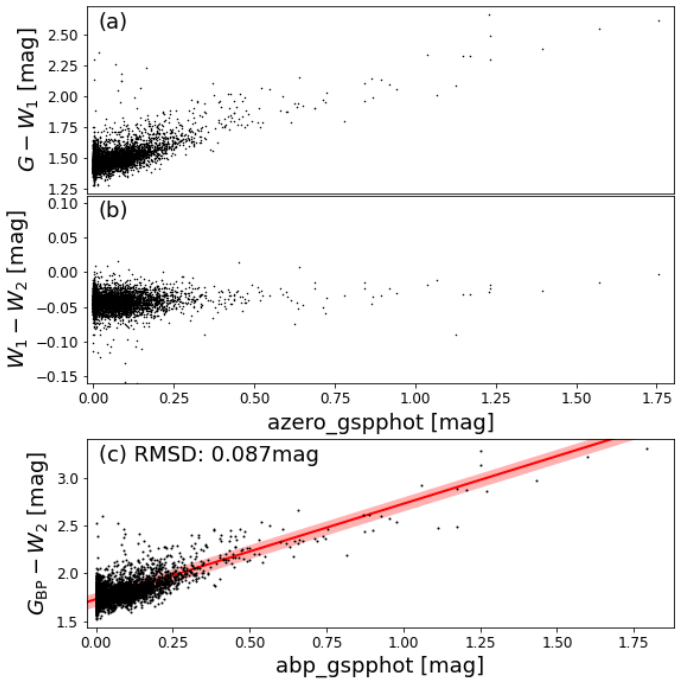

Fig. 29. Colours of GSP-Spec solar-analogue candidates as a function of GSP-Phot extinction estimates. W_1 and W_2 denote AllWISE photometry (Cutri et al. 2021). We restrict to candidates with $W_1 > 8$ mag, since AllWISE photometry suffers from saturation for brighter sources. In panel c, the red line is a linear increase with `abp_gspphot` offset by the mean $G_{BP} - W_2$ colour of 589 stars where $A_0 < 0.001$ mag according to GSP-Phot. The red interval marks the uncertainty from the standard deviation of the mean. The quoted root-mean-square (RMS) difference is between the $G_{BP} - W_2$ colour and `abp_gspphot` plus the mean.

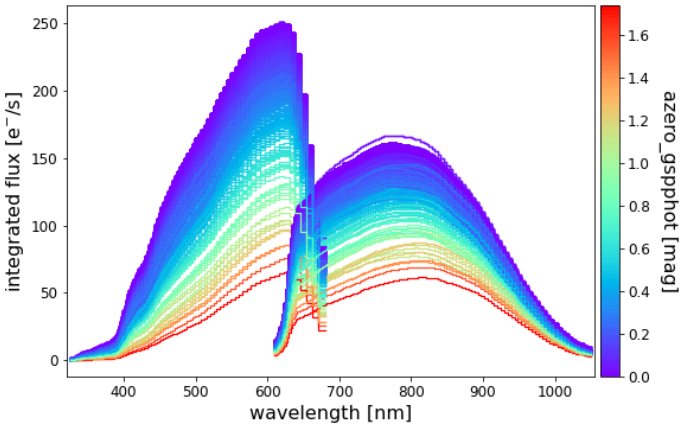

Fig. 30. Variation of low-resolution BP and RP spectra of GSP-Spec solar-analogue candidates with GSP-Phot's A_0 estimate. In order to make the BP and RP spectra comparable, they have been rescaled to an apparent magnitude of $G = 15 + A_G$ with A_G taken from GSP-Phot.

on ≃1 500 spectra, and it is calibrated on the 2013 version of the CALSPEC[15] grid (Bohlin et al. 1995; Bohlin 2014; Bohlin et al. 2019) with a zero-point accuracy of better than 1%. The SPSS grid is designed to cover those areas of the stellar parameter space that are not well sampled by CALSPEC, in particular the FGKM star types, and to cover the entire *Gaia* wavelength range (330–1050 nm). The final release, SPSS V3, will be used to calibrate *Gaia* DR4, it will contain about 200 stars and will make full use of all the ≃6 500 spectra collected in the observing campaigns. It will be calibrated on the latest version of the CALSPEC grid (Bohlin 2014; Bohlin et al. 2019), which differs by about <0.5% from the 2013 one in the grid zero-point. The S/N ratio of the ground-based SPSS spectra is generally well above 100, with the exception of the blue and red extremes of the *Gaia* wavelength range. The SPSS flux tables were thus extended with theoretical spectra, adjusted to match the central, high-S/N ratio region of the observed spectra (see Pancino et al. 2021b, for details). It is therefore of the utmost importance, for the next SPSS release, to have a robust estimate of the spectral type, atmospheric parameters ($T_{\rm eff}$, log$g$, [Fe/H], and [α/Fe]), and of the interstellar absorption for as many of the SPSS as possible.

To this aim, we explored and selected relevant information from the `astrophysical_parameters` table of *Gaia* DR3 for the SPSS V2 stars. In particular, whenever available, we selected the GSP-Spec parameters over the GSP-Phot ones for the $T_{\rm eff}$, log$g$, [Fe/H][16]. Similarly, for the choice of the FLAME parameters, i.e., mass, age, luminosity, and radius, we always selected the corresponding FLAME-spec determinations when available (in the `astrophysical_parameters_supp` table). Parameters were available for all the SPSS in the sample, except for the 56 white dwarfs. For hot stars, a handful of parameters from ESP-HS were available that were not parametrized by GSP-Phot and GSP-Spec. The two binarity estimators available (`specmod` and `combmod`) agreed in indicating SPSS 028 (SA105-663) as a binary, while 15 different SPSS were indicated as photometrically variable (`phot_variable_flag`) and will be carefully re-evaluated in the preparation of the SPSS V3 release.

To explore the quality of the results, we compared them with the two sets of parameters presented by Pancino et al. (2021b): (1) a collection of literature estimates and (2) the best-fit parameters obtained by extending the SPSS V2 flux tables with model libraries. First, we compared the spectral type determinations and found that only 16 SPSS out of 111 had discrepant spectral types, and in all cases the discrepancies never spanned more than one spectral class (e.g., an F star classified as G). For one star, SPSS 313 (M5–S1490), discordant previous literature spectral type determination (from A to F) was available, and we found it to be a K giant, about 500 K cooler than the coolest literature determination. We then compared the three main atmospheric parameters (Figure 31) with both reference sets. As can be seen, apart from very few outliers, the agreement with the two sets of reference parameters appears good, especially when considering the heterogeneity of the literature estimates. There is an indication that a few stars with $A_0 \gtrsim 1$ mag have problems in some of the parameters. However, for the majority of stars, the agreement for $T_{\rm eff}$ and log g is excellent, with median differences of $\Delta T_{\rm eff} = -4 \pm 322$ K and $\Delta \log g = -0.04 \pm 0.59$ dex. The comparison of [Fe/H] is still good if one includes metal-poor stars, with Δ[Fe/H] = 0.15 ± 0.61 dex. When excluding stars below [Fe/H] ≃ –2 dex, which appear to have an overestimated iron metallicity, the comparison improves, with Δ[Fe/H] = –0.09 ± 0.44 dex. We note that an overestimate for metal-poor stars is a common problem when metallicities or iron

[15] https://www.stsci.edu/hst/instrumentation/reference-data-for-calibration-and-tools/astronomical-catalogs/calspec

[16] To obtain [Fe/H] from the GSP-Phot [M/H] estimates, we used the formula by Salaris et al. (1993) and assumed an α-enhancement of +0.35 for metal-poor stars, +0.15 for intermediate metallicities, and zero for solar or higher metallicity. Note that we did not apply any recalibration to the log g and [Fe/H] GSP-Phot and GSP-Spec values.

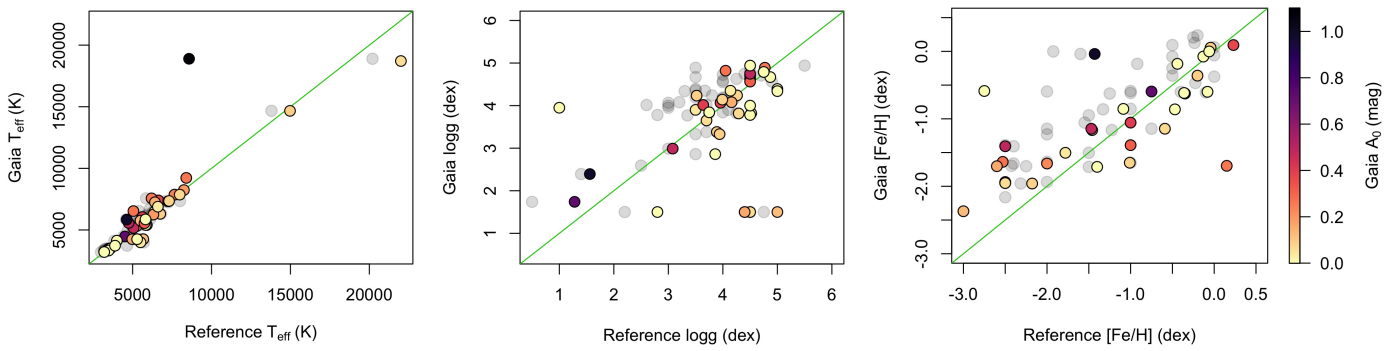

Fig. 31. Comparison of the SPSS sample main parameters derived here with the two reference sets by Pancino et al. (2021b): the best-fit parameters to the SPSS flux tables are shown in grey, while a collection of literature spectroscopic estimates is colored according to the interstellar absorption A_0 obtained here. The left panel shows the case of $T_{\rm eff}$, the middle one of log$g$, the right one of [Fe/H]. The 1:1 line is shown in green in all panels.

abundances are derived from photometric data or low resolution spectra (see also Miller 2015; Anders et al. 2022; Xu et al. 2022).

In *Gaia* DR3 we present a table `gold_sample_spss` which contains the 111 SPSS stars, and for each source, their *Gaia* DR3 source_id, name, spectral type, binary and variability flags, along with the stellar parameters, extinction, distance, radial velocity, and $v \sin i$ (where available) for the 52 non-subdwarf/white dwarf stars of the sample.

9. Summary of golden samples

In Sect. 3 to Sect. 8 we defined several samples of stars, carefully selected to be homogeneous and with the highest quality, that can be used in many different astrophysical contexts. Complementary data tables have been made available in the *Gaia* DR3 archive to help exploiting these samples, see here in the online documentation. In Table 7 we summarise the names, sizes, and contents of these tables, and here we provide an overview.

The six tables are all entitled `gold_sample_`*name* where *name* is specific to the sample, i.e. `oba_stars`, `fgkm_stars`, `carbon_stars`, `solar_analogues`, `spss`, and `ucd`. These can be called in an ADQL query in *Gaia* DR3 as `gaiadr3.gold_sample_`*name*.

The tables for the solar analogues and the carbon stars contain the `source_id` only. The OBA table also includes a flag that allows one to apply a kinematic filter. The table for the UCDs contains, along with `source_id`, the newly derived radii and luminosities from the analysis of the *Gaia* and infrared data, and the bolometric flux correction. The SPSS sample table contains all 111 SPSS sources along with information such as binary and variability flags, radial velocity, and $v \sin i$. The stellar parameters and extinction are given for the non-white dwarf stars, some of which are based on GSP-Spec parameters and others on GSP-Phot or on ESP-HS, this is indicated by the `notes` in that table. Finally, for the FGKM sample, a table with `source_id`, the atmospheric parameters, the evolutionary parameters and the spectral type is provided, where specific parameters for some sources have been removed (compared to the `astrophysical_parameters` table).

10. Exploitation of the golden samples

In this section we demonstrate four applications of the golden samples presented in this paper. For the first application we exploit the OBA sample to derive the parameters of the Milky Way rotation curve and the peculiar motion of the Sun. We then use the FGKM sample to characterise known transiting exoplanets. This is followed by an exploitation of the solar analogue sample to derive the colours of the Sun, and finally we use the stellar companions of unseen UCDs to explore the ages of these substellar systems.

10.1. Milky Way rotation curve

A classical application for the OBA star sample is to infer the parameters of the Milky Way rotation curve near the Sun. Young disk stars have often been used for this purpose because of the low dispersion of their velocities around the overall differential rotation of stars in the thin disk (for a recent example based on *Gaia* EDR3 data see Bobylev & Bajkova 2022). We illustrate this application with a very simple modelling of the proper motions in terms of the Milky Way disk rotation curve. The rotation curve is described with the circular velocity and the slope of the circular velocity as a function of Galactocentric cylindrical distance R, both evaluated at the position of the Sun (or equivalently, the Oort constants A and B for an axisymmetric Milky Way, see e.g. Olling & Dehnen 2003). We use a sub-sample of the OBA stars, namely those with `spectraltype_esphs` equal to 'B', with $\varpi/\sigma_\varpi > 10$ and $v_{\rm tan} < 180$ km s^{-1}. The sample is further restricted to $(1000/\varpi) \times \sin b < 250$ pc and $6.5 < R < 15$ kpc. Over this range in R the above approximation to the rotation curve is reasonable (see e.g. Eilers et al. 2019, their Fig. 3). Figure 32 shows the proper motions in ℓ and b as a function of Galactic longitude for the 385 423 B-stars in this sample. The figure beautifully reveals the variation of $\mu_{\ell *}$ with cos(2ℓ), a consequence of Galactic differential rotation, and shows the slight offset of the proper motions in latitude from zero, reflecting the Sun's motion perpendicular to the Galactic plane. The width of the proper motion distributions mainly reflects the range of distances to the stars in the sample.

To derive the rotation curve parameters we use a Bayesian model for the proper motions in Fig. 32. The model has parameters similar to the simple kinematic model described in Sect. 3: the circular velocity at the position of the Sun $V_{\rm circ,\odot}$, the slope of the circular velocity curve $dV_{\rm circ}/dR$, the peculiar motion vector of the Sun $\mathbf{v}_{\odot,\rm pec} = (U_\odot, V_\odot, W_\odot)$, and the velocity dispersions in the plane and perpendicular to the plane, σ_{xy} and σ_z. The position of the Sun is fixed at a Galactocentric distance of 8277 pc (Gravity Collaboration et al. 2022), while the height above the Galactic plane is taken as the me-

Table 7. Summary of the tables in the *Gaia* DR3 archive to help in the exploitation of the samples presented in this work.

star type	table name	N	field contents of tables	section	notes
OBA	gold_sample_oba_stars	3 023 388	source_id, flag	3	
FGKM	gold_sample_fgkm_stars	3 273 041	source_id, all APs	4	Table 2
UCD	gold_sample_ucd	21 068	source_id, R, L	5	
Carbon	gold_sample_carbon_stars	386 936	source_id	6	
Solar analogues	gold_sample_solar_analogues	5863	source_id	7	
SPSS	gold_sample_spss	111	source_id and all APs	8	

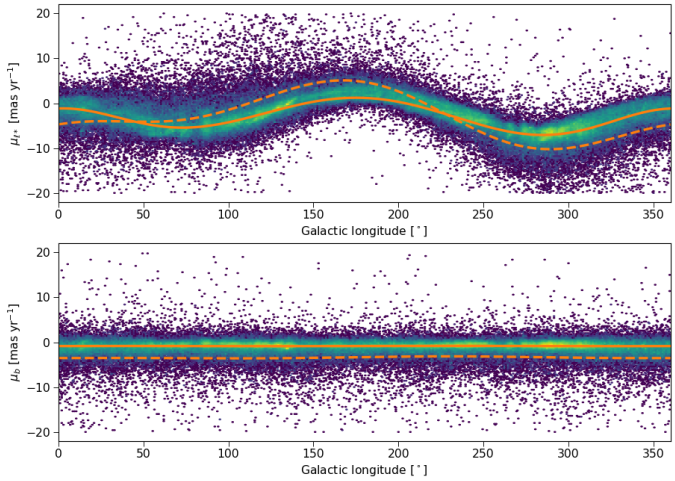

Fig. 32. Proper motions in Galactic longitude (top) and latitude (bottom) as a function of Galactic longitude for the sample of 385 423 B-stars described in the text. The lines show the proper motions predicted from the rotation curve model parameters, resulting from the fit to the data, for stars at 500 pc from Sun (dashed line) and at 2000 pc (solid line, close to the median distance of stars in the sample).

dian of $-(1000/\varpi) \times \sin b$ for the B-star sample, which is 16 pc. The model velocities $\mathbf{v}$ of the stars are calculated from the azimuthal component of the velocity in Galactocentric cylindrical coordinates, $V_\phi = -(V_{\mathrm{circ},\odot} + dV_{\mathrm{circ}}/dR \times (R - R_\odot))$, as $\mathbf{v} = (-V_\phi \sin\phi, V_\phi \cos\phi, 0)$[17]. The model proper motions $\boldsymbol{\mu}_{\mathrm{pred}}$ are then calculated from $\mathbf{v} - \mathbf{v}_\odot$, the parallaxes, and celestial positions of the stars (with $\mathbf{v}_\odot = (U_\odot, V_\odot + V_{\mathrm{circ},\odot}, W_\odot)$). The parallaxes were used as error-free observables. The velocity of the local standard of rest with respect to the rotational standard of rest is assumed to be 0 km s^{-1} (Bland-Hawthorn & Gerhard 2016) and not included in the model.

The 7 model parameters are optimized through a Markov-Chain Monte Carlo sampling of the posterior. The likelihood for the observed proper motions is a normal distribution centered on $\boldsymbol{\mu}_{\mathrm{pred}}$ with a covariance matrix that accounts for the covariance matrix of the observed proper motions and the velocity dispersions, using the appropriate form of equation (16) in Lindegren et al. (2000). The priors on the model parameters are broad normal distributions centred on 220, 11, 12, 7 km s^{-1} for $V_{\mathrm{circ},\odot}$, $U_\odot$, $V_\odot$, $W_\odot$, respectively. The prior on dV_{circ}/dR is a normal distribution centred on 0 km s^{-1} kpc^{-1}. The priors on the velocity dispersions are Gamma distributions with parameters $\alpha = 2$ and $\beta = 0.1$. The model was implemented in Stan (Stan Development Team 2022), using the CmdStanPy interface. The posterior was sampled with 4 Markov Chain Monte Carlo (MCMC) chains for 1500 steps each, and the first 500 steps were discarded as 'burn-in'. To keep the required computational resources within bounds, the Stan model was run using a random subset of 20 000 stars chosen from the B-star sample above.

The resulting model parameters are: $V_{\mathrm{circ},\odot} = 234 \pm 0.5$ km s^{-1}, $dV_{\mathrm{circ}}/dR = -3.6 \pm 0.1$ km s^{-1} kpc^{-1}, $U_\odot = 8.1 \pm 0.1$ km s^{-1}, $V_\odot = 11.2 \pm 0.2$ km s^{-1}, $W_\odot = 8.1 \pm 0.1$ km s^{-1}, $\sigma_{xy} = 14.2 \pm 0.1$ km s^{-1}, and $\sigma_z = 7.3 \pm 0.1$ km s^{-1}. These numbers are consistent with results from the literature (e.g., as compiled by Bland-Hawthorn & Gerhard 2016). The corresponding Oort parameters are $A = 16$ km s^{-1} kpc^{-1}, $B = -12$ km s^{-1} kpc^{-1}, and $A - B = 28$ km s^{-1} kpc^{-1}. The total velocity of the Sun translates to an apparent proper motion at the position of Sgr A* of -6.25 mas yr^{-1} along the plane and -0.21 mas yr^{-1} perpendicular to the plane. This is consistent with the most recent evaluation of the proper motion of Sgr A* by Reid & Brunthaler (2020).

The uncertainties quoted for the above results should be interpreted as the precision achieved in the context of the model and the sub-sample used. The uncertainties are underestimated. They do not account for the variance due to the choice of the specific random sub-sample of 20 000 B stars. More importantly, the obvious model deficiencies are not accounted for, such as ignoring the effects of the Milky Way disk warp and the motions induced by spiral arms (Olling & Dehnen 2003), as well as deviations of the true rotation curve from the simple model. The 'mode-mixing' effect discussed in Olling & Dehnen (2003) is not an issue here because of the precise knowledge of the parallaxes of the stars in the sample. The model deficiencies are apparent in Fig. 33 which shows a comparison between the observed and model proper motion distributions. As noted above the modelling here is a mere illustration of the possibilities offered in analyzing the proper motions for a sample of young disk stars covering a large range in R. For a much more in-depth look at a sample of young disk stars, selected slightly differently from what we presented in Sect. 3, we refer to the *Gaia* DR3 paper on mapping the asymmetric disk of the Milky Way (Gaia Collaboration et al. 2022b). The paper presents maps showing rich structure in the velocity field of OB stars which can be traced to the spiral arms, something which the above model obviously does not capture. On the other hand the average V_ϕ curve shown in that paper for the OB stars (calculated from the proper motions, parallaxes, and radial velocities) shows that the description of the rotation curve used above is accurate in an average sense.

[17] Here we use a right handed coordinate system, so V_ϕ is negative for the stars in the disk of the Milky Way

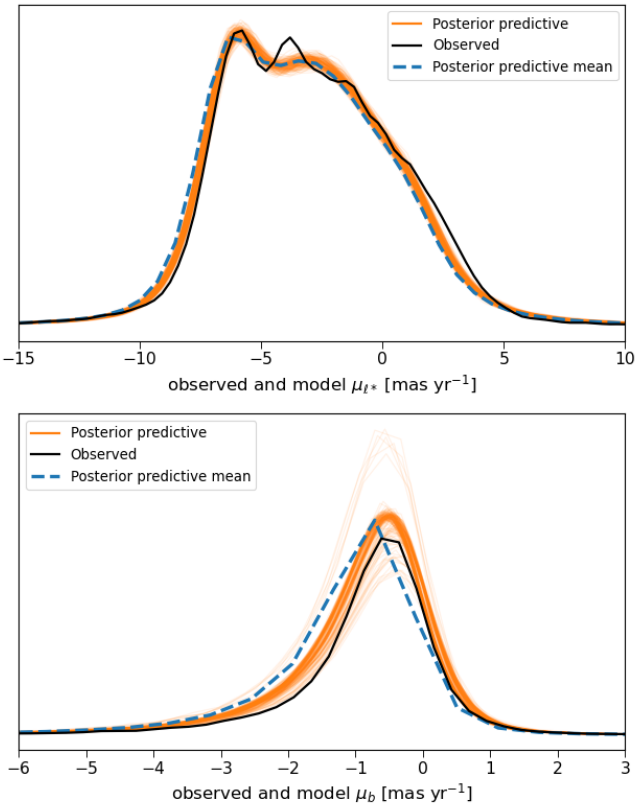

Fig. 33. Distributions of the observed and model proper motions for the sample of 385 423 B-stars described in the text, with proper motions in Galactic longitude and latitude in the upper and lower panels, respectively. The black lines show the observed proper motion distributions. The thin orange lines are the predicted proper motion distributions for 200 randomly sampled MCMC model parameters. The mean of all such sample distributions is indicated by the thick dashed line.

10.2. Exoplanet characterisation

The search for and characterisation of exoplanet systems is at the forefront of scientific research, with many current and future ground- and space-based projects dedicated to this quest, e.g. Gardner et al. (2006); Borucki et al. (2008); Rauer et al. (2014); Ricker (2014); Tinetti et al. (2018). The characterisation of the planet itself relies on the knowledge of the planet host. In particular, the planet's radius and mass depends directly on the stars's radius and mass through the following equations

$$M_p \sin(i) = \frac{M_\star^{2/3} P^{1/3} K (1-e^2)^{0.5}}{(2\pi G)^{1/3}} \quad (3)$$

and

$$d_{\rm tr} = \left(\frac{R_p}{R_\star}\right)^2 \quad (4)$$

where M_p, $M_\star$ are the mass of the planet and star, respectively, i and e are the inclination and eccentricity of the orbital system, P the orbital period, K is the semi-amplitude of the radial velocity curve, and $d_{\rm tr}$ is the transit depth due to the planet with radius R_p passing in front of the star with radius $R_\star$ and blocking a part of its light. In reality the relationship between the transit depth and relative radii is a little more complicated than Eq. 4, see e.g Heller (2019), but for this illustration purpose we keep things simple. Additionally, we consider only transiting systems, so the inclination of the system is very close to 90° and $\sin(i) \sim 1$.

We obtained a list of the known transiting planets and light curve parameters from `exoplanets.org`[18]. This catalogue contains (as of March 2022) 2651 confirmed transiting exoplanets. We cross-matched these sources with the FGKM sample and obtained 593 planet matches. Of these, 354 contain transiting parameters to estimate the planetary radius while 108 entries contain parameters to estimate both the planet mass and radius, but only 94 have a valid stellar mass in our sample.

We calculated the radii of the exoplanets using `radius_flame`, along with the available transit depth parameter. To evaluate the uncertainties we performed a bootstrap method where we perturbed each of the input observations 1000 times and used the resulting standard distribution of the evaluated parameters to estimate the uncertainties. We show the distribution of the planetary radii as a function of orbital separation of the planet-star system in Fig. 34. We colour-coded the planet symbols according to `teff_gspphot` and the symbol size indicates the orbital period of the system, which ranges from 0.57 days to just under 365 days. We also show the position of the Earth and Jupiter as grey squares, which highlights how different other planetary systems are to our own. In particular, many of these planets are well inside the inner limit of the habitable zone and the Jupiter-sized planets are equally close to their host star.

We furthermore calculated the mass of the planets for the 94 sources with radial velocity parameters and stellar masses. Of these, four did not have a reported eccentricity, and of the other 90, only 24 have non-zero values. For the planets with no reported eccentricity we assumed circular orbits. 10 of the planets also did not have a reported inclination and so we assumed $i = 90°$ ($\sin i = 1$) which is reasonable for a transiting system. The median value of the inclinations of the other 84 planetary systems is 87.2° ($\sin i = 0.9988 \sim 1$).

We show our results in the planet radius – planet mass diagram in Fig. 35. We also show some models corresponding to model mass–radius relationships for different Earth-like planet compositions from Zeng et al. (2016) and Jupiter-like planet compositions from Guillot & Gautier (2015). The black lines represent Earth-like planet mass–radius relations assuming an ice-like (dashed), rocky Earth-like (dashed-dotted) and iron (dashed) composition. The coloured lines represent models of an isolated planet of solar composition at 5 Gyr (like for Jupiter, blue), a heavily irradiated planet with an equilibrium temperature of 1000 K with no core (red) and one with a 100 $M_{\rm Earth}$ central core (green).

The precision on our results (the error bars are shown although they are not always visible) does allow one to distinguish between different bulk compositions of these planets provided we have full control of the potential systematic errors. We provide the mass, radius and age properties of the planet and its host in Table 8.

As these figures highlight, there is a dearth of knowledge of Earth-size exoplanets in the habitable zone, along with their accurate characterisation. The upcoming ESA PLATO mission promises to populate the habitable zone by observing (at least) one large field over a two-to-three year period, which will al-

[18] For four of the planets we adopted the values from the reference paper due to errors or inconsistencies: XO-6b from Crouzet et al. (2017), KELT-8b from Fulton et al. (2015), *Kepler*-407b from Marcy et al. (2014), and *Kepler*-68b Gilliland et al. (2013).

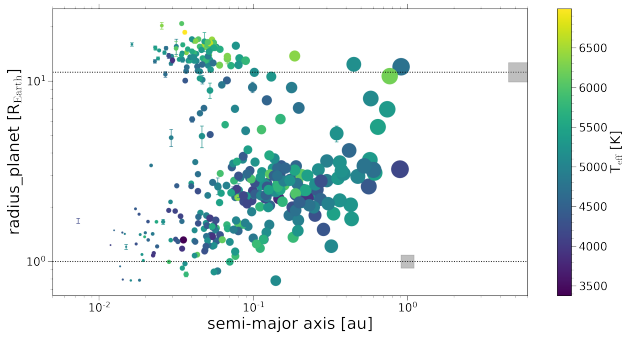

Fig. 34. Distribution of planetary radii compared to the separation from their host star (orbital semi-major axis a) for planetary systems in the FGKM sample. Colour-code indicates the $T_{\rm eff}$ of the host star, while the symbol size indicates the orbital period in $\log_{10}$ scale (range = 0.57 - 364.8 days). The dotted lines indicate 1 $R_{\rm Earth}$ and 1 $R_{\rm Jup}$ and the Earth and Jupiter are denoted by the square symbols.

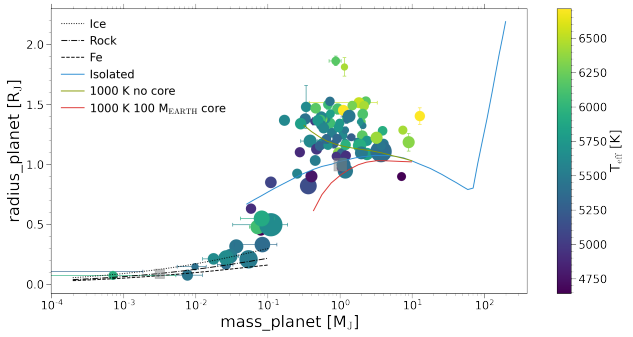

Fig. 35. Planet mass and radius (in Jupiter units) of 94 planets with radial velocity and stellar parameters in the FGKM sample. Colour-coding is as in Fig. 34. The symbol size corresponds to the semi-major axis. Some radius–mass models of Earth-like (Zeng et al. 2016) and Jupiter-like (Guillot & Gautier 2015) planets are also shown, see text for details.

low us to detect and confirm Earth-like orbits around Earth-like planet hosts.

10.3. The colours of the Sun

The colours of the Sun are not as well known as one would think and desire, neither observationally nor from modelling. Solar analogues offer the possibility to validate, and if necessary calibrate, our understanding of the solar flux as a function of wavelength (Holmberg et al. 2006; Casagrande & VandenBerg 2018). They have also been used to estimate the solar bolometric correction in *Gaia*'s photometric system. Below we make a new attempt to determine precise and accurate solar colours.

We use the sample of solar-analogue candidates selected from GSP-Spec from Sect. 7 in order to estimate the colours of the Sun. As we demonstrated in Sect. 7.4, these stars have reliable extinction estimates from GSP-Phot. Consequently, the BP/RP spectra with very low extinction (according to GSP-Phot) can be used to indirectly estimate the intrinsic continuum shape of the Sun. Among GSP-Spec's solar-analogue candidates, there are 682 with GSP-Phot's $A_0 < 0.001$ mag. Given these, we obtain an absolute magnitude of

$$M_{G,\odot} = (4.614 \pm 0.179)\,\mathrm{mag} \quad (5)$$

For this, we adopt the inverse parallax as a distance estimator because our candidate selection requires very high parallax quality ($\frac{\varpi}{\sigma_\varpi} > 20$). For comparison, a value of $M_{G,\odot} = 4.66$ is adopted for FLAME (Creevey et al. 2022, Sect. 4.3 therein). Given the 682 candidates with $A_0 < 0.001$ mag, we also obtain mean colours and standard deviations of

$$(G_{\rm BP} - G_{\rm RP})_\odot = (0.818 \pm 0.029)\,\mathrm{mag} \quad (6)$$
$$(G_{\rm BP} - G)_\odot = (0.324 \pm 0.016)\,\mathrm{mag} \quad (7)$$
$$(G - G_{\rm RP})_\odot = (0.494 \pm 0.020)\,\mathrm{mag} \quad (8)$$
$$(G - J)_\odot = (0.969 \pm 0.578)\,\mathrm{mag} \quad (9)$$
$$(G - H)_\odot = (1.292 \pm 0.401)\,\mathrm{mag} \quad (10)$$
$$(G - K_s)_\odot = (1.371 \pm 0.351)\,\mathrm{mag} \quad (11)$$
$$(G - W_1)_\odot = (1.449 \pm 0.066)\,\mathrm{mag} \quad (12)$$
$$(G - W_2)_\odot = (1.405 \pm 0.065)\,\mathrm{mag} \quad (13)$$

where we restrict the AllWISE comparison to cases with $W_1 > 8$ mag in order to avoid saturation. These colours are in excellent agreement with the values $(G_{\rm BP} - G_{\rm RP})_\odot = 0.82$ mag, $(G_{\rm BP} - G)_\odot = 0.33$ mag and $(G - G_{\rm RP})_\odot = 0.49$ mag obtained in Casagrande & VandenBerg (2018) from *Gaia* DR2 passbands and synthetic as well as observed spectra for the Sun. Their absolute magnitude of $M_{G,\odot} = 4.67$ mag is also consistent with our estimate. In order to do this comparison with *Gaia* DR3 passbands and also include near-infrared photometry, we take the Kurucz model `sun_mod_001.fits` from the CALSPEC library[19] (Bohlin et al. 1995, 2014, 2020) and simulate its photometry using the `pyphot` package[20]. We obtain synthetic colours of $(G_{\rm BP} - G_{\rm RP})_\odot = 0.813$ mag, $(G_{\rm BP} - G)_\odot = 0.324$ mag and $(G - G_{\rm RP})_\odot = 0.490$ mag, which are again in excellent agreement with our estimated colours. For color combinations with 2MASS, we obtain $(G-J)_\odot = 0.992$ mag, $(G-H)_\odot = 1.320$ mag, and $(G - Ks)_\odot = 1.360$ mag, which are again in excellent agreement with our candidates. Concerning AllWISE (Cutri et al. 2021), we obtain $(G - W_1)_\odot = 1.380$ mag and $(G - W_2)_\odot = 1.301$ mag. These values are slightly bluer than the values we estimate from the GSP-Spec candidates, but are still within 1σ and 1.6σ, respectively.

10.4. Ages of UCDs not seen by Gaia

Another application of the *Gaia* astrophysical parameters is to constrain the characteristics of faint UCDs that are beyond the mission magnitude limit but in binary systems with brighter objects that are observed by *Gaia*. Once we have identified a multiple system we assume that the UCD has the same chemical composition, age, distance and, after allowing for orbital motion, proper motions. In addition, if the movement due to the orbital motion is detected by *Gaia* this will provide a constraint on the mass of the various components. Brown dwarfs evolve and cool over time and their observational properties are degenerate with age, mass and metallicity; binary systems are therefore benchmarks for understanding these processes. *Gaia* will provide a large homogeneous multi-parametric sample with intersecting constraints that will tie down the UCD regime. For this illustrative discussion we concentrate on the age parameter[21]. We note

[19] https://www.stsci.edu/hst/instrumentation/reference-data-for-calibration-and-tools/astronomical-catalogs/calspec

[20] http://mfouesneau.github.io/docs/pyphot

[21] The BaSTI models (Hidalgo et al. 2018) were used to derive the age and they span from the ZAMS to the tip of the red giant branch for stellar masses between 0.5 and 10 $M_\odot$.

Table 8. Mass, radius, and age of known exoplanets and their host stars in the FGKM sample. The full table is made available online as an electronic table.

	planet properties			host properties		
Gaia DR3 source_id	planet	radius	mass	radius	mass	age
		[$R_\oplus$]	[$M_\oplus$]	[$R_\odot$]	[$M_\odot$]	[Ga]
1424011082893734272	WASP-92 b	1.423 ± 0.034	0.731 ± 0.069	1.306 ± 0.027	1.042 ± 0.045	7.74 ± 1.13
2101243789577188736	Kepler-103 c	0.498 ± 0.017	0.110 ± 0.083	1.562 ± 0.032	1.037 ± 0.040	9.30 ± 0.90
4285511294172309504	CoRoT-11 b	1.527 ± 0.037	2.274 ± 0.336	1.466 ± 0.033	1.225 ± 0.041	3.49 ± 0.55

Table 9. Identification and parameters of UCDs without *Gaia* solutions in binary systems with full-solution companions. The mass and ages indicate the median, lower and upper confidence intervals, while the parallax shows the value and uncertainty.

UCD	Parallax	SpT	Mass	Companion	SpT	Age	Lit. age
Code, Name	[mas]		[$M_\oplus$]	Source_ID		[Ga]	[Ga]
A, 2MASSI J0025036+475919	22.8 ± 0.9[2]	L4:+L4:[2]	$84.4^{87.4}_{83.7}$	392562179817077120	F8	$5.1^{5.9}_{4.2}$	$4.4^{5.3}_{1.6}$
B, 2MASS J02233667+5240066	12.2 ± 1.1[1]	L1.5[3]	$80.6^{80.7}_{80.3}$	452046549154458880	F5	$2.9^{3.4}_{2.5}$	$2.0^{2.8}_{1.7}$
C, 2MASS J06462756+7935045	53.6 ± 2.2[1]	L9[4]	$59.4^{64.4}_{51.9}$	1141280704422528128	F7V	$2.2^{2.9}_{1.6}$	$4.3^{6.0}_{2.5}$
D, HD 49197B	29.9 ± 2.1[1]	L4[5]	$79.3^{79.6}_{78.9}$	952346742338146176	F5	$4.0^{4.9}_{3.1}$	$2.8^{4.7}_{0.3}$
E, 2MASS J12173646+1427119	16.1 ± 0.7[1]	L1[3]	$82.3^{84.5}_{82.3}$	3921176983720146560	F8	$4.9^{5.7}_{4.2}$†	$1.8^{3.8}_{0.5}$
F, HD 118865B	21.7 ± 0.9[1]	T5[6]	$67.5^{69.1}_{65.2}$	3663438298389132416	F7V	$4.6^{5.4}_{3.9}$†	$3.2^{3.5}_{2.5}$
G, 2MASS J14165987+5006258	22.1 ± 0.8[7]	L5.5[8]	$75.6^{75.8}_{75.3}$	1508557582834745088	G5	$6.3^{7.3}_{5.3}$	$5.8^{9.0}_{4.2}$
H, ULAS J142320.79+011638.2	29.4 ± 1.0[1]	T8p[9]	$34.9^{38.4}_{30.3}$	3654496279558010624	G1.5V	$5.6^{7.2}_{4.0}$	$5.9^{11.2}_{4.6}$
I, Gl 564 C	74.7 ± 21.3[1]	L4:[10]	$77.4^{77.6}_{77.1}$	1265976524286377856	F9IV-V	$5.3^{6.2}_{4.4}$†	$0.9^{4.7}_{0.6}$
J, Gl779B 244691	47.1 ± 6.1[1]	L4.5[11]	$71.0^{73.3}_{39.9}$	1821708351374312064	G0V	$2.8^{3.8}_{0.5}$	$3.2^{5.2}_{1.7}$
K, eps Indi C	275.3 ± 3.0[12]	T6[13]	$36.5^{50.2}_{11.0}$	6412595290592307840	K5V	$2.0^{4.3}_{0.2}$	$4.6^{5.7}_{0.9}$

Notes: 1: Spectrophotometric distance, 2: Cruz et al. (2007), 3: Deacon et al. (2014), 4: Loutrel et al. (2011), 5: Metchev & Hillenbrand (2004), 6: Burningham et al. (2013), 7: Faherty et al. (2012), 8: Chiu et al. (2006), 9: Pinfield et al. (2012), 10: Goto et al. (2002), 11: Liu et al. (2002), 12: Weinberger et al. (2016), 13: Burgasser et al. (2006). The † indicates FLAME ages derived using the GSP-Phot-$T_{\rm eff}$, while other ages are derived using the GSP-Spec-$T_{\rm eff}$.

however, that more precise ages can be obtained by combining *Gaia* with other observational data such as asteroseismology.

To identify a potential list of objects with a high probability to be in a binary system we used the positional and kinematical criteria given by Eq. 2 in Smart et al. (2019) and the list of known UCDs from that study. When the faint UCD did not have a measured parallax we used its spectro-photometric distance. We found 8 UCDs without *Gaia* DR3 5 parameter solutions that are in binary systems in the FGKM sample, while also in the regime of reliable ages (see Fouesneau et al. 2022). We added a further three interesting targets with reliable ages here because they were rejected from the FGKM sample for failing on only one of the criteria: A `ipd_frac_multi_peak` = 22; and B and F `classprob_dsc_combmod_binary` > 0.99.

These 11 UCDs are listed in Table 9 with name, adopted parallax, spectral type and mass along with the companion *Gaia* `source_id`, age and the median published ages with 16% and 84% percentiles.

The number of literature age estimates vary from 6 to 46 for each target and are from varied sources: model comparisons (Holmberg et al. 2009; Casagrande et al. 2011), chromospheric activity (Pace 2013; Metchev & Hillenbrand 2004), or Galactic kinematics (Gontcharov 2012). The published age percentiles often indicate uncertainties of a factor of 2 or a large portion of the age of the Galaxy indicating the current difficulty in determining ages for stars. In Figure 36 we show the *Gaia* vs the median published values from Table 9. When available, we used the values based on the GSP-Spec $T_{\rm eff}$: `age_flame_spec` these are denoted by the filled circles. The open circles are `age_flame` which are based on GSP-Phot $T_{\rm eff}$. For most of the stars we find general agreement with the literature, with the worst agreements for systems E, I and K. For E, the $T_{\rm eff}$ from both GSP-Phot and GSP-Spec agree to within 25 K and we would therefore trust its age if the star is within the regime of models that were used. For I and K we find significant disagreements between the GSP-Phot and GSP-Spec $T_{\rm eff}$, and this could indicate a possible issue with `age_flame`. We discuss each of the systems individually in the next section.

The interpolated masses of the UCDs are estimated from a comparison to the illustrated tracks in Figure 37 taken from Baraffe et al. (2015) for stars and Phillips et al. (2020) for brown dwarfs assuming the age of the companion star from this work, and these are reported in Table 9.

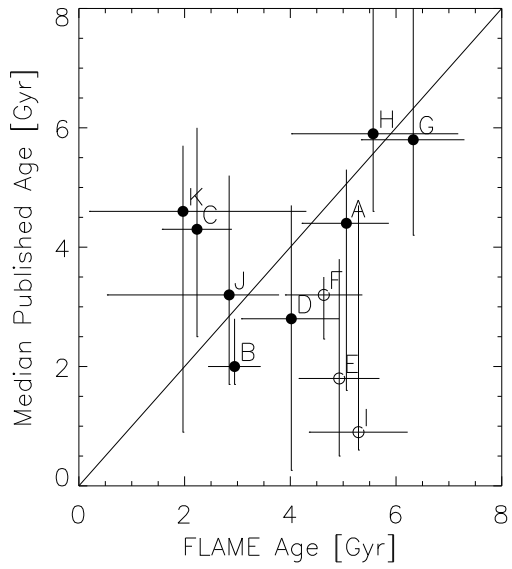

Fig. 36. Literature ages versus FLAME ages for companions of UCDs not seen by *Gaia*. Filled circles correspond to `age_flame_spec` (i.e. using the $T_{\rm eff}$ from GSP-Spec) and open circles are `age_flame`. The error bars represent the 16% and 84% percentiles.

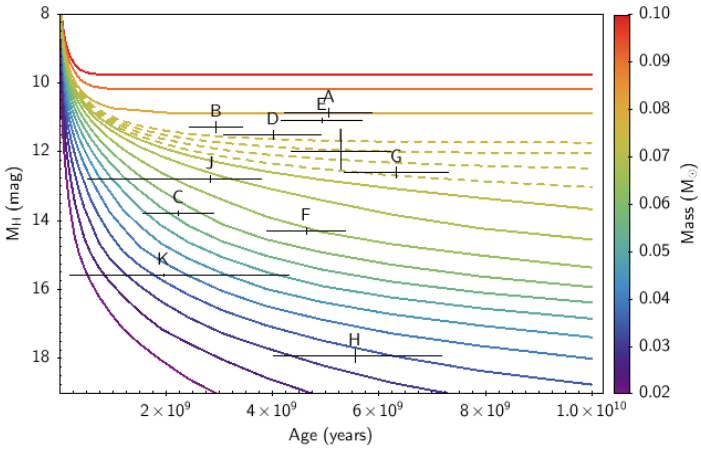

Fig. 37. Evolutionary tracks and UCD locations in the H-band absolute magnitude versus age diagram, adopting the companion age. The tracks are colour-coded by mass. The dashed lines indicate the stellar to substellar transition zone (from 0.072 to 0.075 $M_\odot$).

10.4.1. Notes on individual systems

2MASSI J0025036+475919 (A) is an L4+L4 binary in a multiple system with the spectroscopic binary HD 2057 (Reid et al. 2006, and references therein) and another component 11″ from the primary (*Gaia* EDR3 3925621798172975 36). Lithium absorption has been detected in the combined spectra of the secondary indicating it has an age less than 1.0 Ga (Cruz et al. 2007; Faherty et al. 2010; Filippazzo et al. 2015) which is much lower than the primary age indicated here. This is one of the widest binary systems (~10 000 AU) with an ultracool component but in the range of other systems of similar total mass. The difference in age estimates of the primary and secondary is not easily reconcilable. One possible solution is that it is not a binary system; the agreement of high proper motions is a strong constraint, however, the spectroscopic distance is very uncertain as the binary nature of the secondary requires an assumption of the component flux contributions. Another possibility is that the primary age estimate is high due to its binary nature.

2MASS J02233667+5240066 (B) was first noted to be in a common proper motion system with HIP 11161 in Deacon et al. (2014). The primary has been shown to have acceleration terms (Kervella et al. 2022; Brandt 2021) but the separation with the UCD is large (41″) and the primary has now been resolved by *Gaia* into two components and it is listed as a spectroscopic binary in the non-single stars orbital solution results. It also has a very high `classprob_dsc_combmod_binary` (>0.99). The observed acceleration is therefore due to the primary binarity and not the UCD. Using `age_flame_spec` we find a mass of 80 $M_{\rm Jup}$ which defines the end of the stellar main sequence.

2MASS J06462756+7935045 (C) was indicated as being in a binary system with HD 46588 based on a high common proper motion (Loutrel et al. 2011). It is an L9 brown dwarf, one of the few known at the L/T transition in wide binary systems. These allow constraints on their astrophysical properties. The `age_flame_spec` is lower by 1σ than the primary literature age. Since this is one of the few L9s where an independent age is known it is important to clarify this discrepancy. Assuming the literature age and distance from the primary, Loutrel et al. find a $T_{\rm eff} = 1360^{1410}_{1280}\,K$, which is an important constraint for the temperature at the the L/T boundary. If we assume the lower `age_flame_spec` this will increase the temperature estimate at this boundary.

HD 49197 B (D) has been studied extensively since its first discovery by Metchev & Hillenbrand (2004) using high resolution observing techniques. It is at a separation of 0.95″ from the primary. There are ongoing adaptive optics projects to try to determine a binary solution (Bowler et al. 2020; Tokovinin 2014). With a magnitude difference of greater than 10 *Gaia* will not be able to resolve the system. If we adopt the low end of the literature age range, HD 49197 B is a brown dwarf, if we adopt the high end - for example that indicated by `age_flame_spec` - the object becomes a star. Since there is also a possibility of finding the mass of this companion either through high resolution imaging or the detection of acceleration terms in the *Gaia* primary solution (proper motion anomalies between the Hipparcos and *Gaia* results have already been detected in Kervella et al. 2022), knowing its age will be crucial for constraining the stellar-substellar boundary.

2MASS J12173646+1427119 (E) was first discovered in the Pan-STARRS survey as a companion to HIP 59933 at 40″. The secondary is detected by *Gaia* (EDR3 3921177219942653696) but with only a two-parameter solution. The primary, EDR3 3921176983720146560, has a non-single star solution which indicates a companion of 0.09$M_\odot$ with a period of 1 yr and a corresponding separation of 1 AU; given the small separation there must be a third component in the system. Any age above 0.5 Ga would indicate that this is a stellar object but very close to the stellar-substellar boundary as indicated by our 82 $M_{\rm Jup}$.

HD 118865B (F) is a T5 in a system with an F4 spectral type first noted in Burningham et al. (2013) where they find an age range of 1.5–4.9 Ga and mass of 45–60 $M_{\rm Jup}$. We find a primary age that is at the top end of their range and hence a slightly larger mass. If confirmed it will provide a high mass for this T5 compared to other similar type brown dwarfs.

2MASS J14165987+5006258 (G) is noted as a binary system in Faherty et al. (2010) with a very large separation of ~26 000 AU. The primary `age_flame_spec` estimate is consistent with published values and also with blue near-IR colors observed in Faherty et al. (2010) where they also re-evaluate its

spectral type from L5.5 to L4. The estimated mass indicates that this object is of stellar and not sub-stellar type and further characterisation will contribute to our understanding of very old borderline stellar objects.

ULAS J142320.79+011638.2 (H) is the coolest object in this sample in a system with an early-G dwarf, HIP 70319. There are a significant number of age estimates from very young to very old and the `age_flame_spec` is in agreement with the median. This age is consistent with a low metallicity primary and also with a broader Y-band peak and more depressed K-band peak than other T8s (Kirkpatrick et al. 2021). This is an important benchmark for metal poor T dwarfs.

Gl 564 B/C (I) is an L4+L4 binary in a triple system with Gl 564, a G2 V star. The majority of the published age ranges are very young because Gl 564 is chromospherically active with a high lithium abundance and fast rotation (Potter et al. 2002). The space motion also puts the object in the Ursa Major moving group from the Banyan Σ tool (Gagné et al. 2018), which has an age of around 500 Myr (King et al. 2003). This is in contrast to the high `age_flame` which is difficult to reconcile given the high lithium abundance and space motion. A possible explanation for this discrepancy is in the limitations of the models that were used, for example, they do not include rotation. If the system is in the first 0.5 Ga they will be contracting brown dwarfs. The orbital period of the UCD binary system is around 10 yr (Potter et al. 2002) and we will therefore soon have dynamical masses with *Gaia*. These objects will provide a well-constrained calibration point for the theoretical models describing low-mass, ultracool objects.

Gl779B (J) is an L4.5 UCD at 0.7″ from GJ 779, a G0 star. High levels of chromospheric activity suggest a young age, lithium abundance indicates a slightly older than the Hyades age but kinematics indicate an old disk star. The `age_flame_spec` is consistent with the published estimates. The orbit is such that it should be visible in the future *Gaia* observations which will lead to a dynamical mass estimate (Crepp et al. 2014). A comparison of the accelerations found from comparisons of Hipparcos and *Gaia* DR2 results indicate a mass of around $0.07 M_{\odot}$ (Brandt et al. 2019) and this is therefore on the stellar-substellar boundary currently defining the end of the main sequence, and in agreement with our estimated mass.

Eps Ind C (K) is the second closest brown dwarf binary T1+T6 system in a triple system with the K5V star eps Ind. One of the brown dwarfs has a *Gaia* solution (*Gaia* EDR3 6412596012146801152) which we assume is the T1. Later releases should provide a dynamical solution for the component masses. There is a significant history of publications for both the primary and the secondary system. With a period of around 11 yr and an observed separation that varies from 0.6 to 2″, it is a defining system for parameters of early T dwarfs. The `age_flame_spec` is at the low end of the published age range for the primary, and the masses of the secondaries from Dieterich et al. (2018) also imply an inconsistency with such a young age. A dynamical mass determination from the *Gaia* observations should resolve this inconsistency.

We have seen that the results of *Gaia* can be brought to bear on our understanding of objects fainter than its magnitude limit. Indeed there will probably be less than 1000 brown dwarfs brighter than the *Gaia* magnitude limit (Smart et al. 2019) while we expect there to be tens of thousands in binary systems or detected from astrometric and radial velocity perturbations. Therefore, the contribution of *Gaia* to brown dwarf studies will be predominantly due to indirectly detected objects rather than direct detections.

11. Conclusion

In this work we defined homogeneous samples of high quality astrophysical parameters by exploiting many *Gaia* data products that appear in *Gaia* DR3, while focusing on the sources and data products in the `astrophysical_parameters` and the `astrophysical_parameters_supp` tables which were produced by the Apsis software (Creevey et al. 2022; Delchambre et al. 2022; Fouesneau et al. 2022). We considered different regimes of stars all across the HR diagram. In the first part of this work we considered large samples of young massive disk OBA stars (Sect. 3), FGKM spectral type stars (Sect. 4), and faint ultra-cool dwarfs (UCDs, Sect. 5). Then we focussed on smaller samples of specific object types; carbon stars (Sect. 6), solar analogues (Sect. 7), and the Spectro Photometric Standard stars (SPSS, Pancino et al. 2021a, Sect. 8). Concerning the latter, this paper provides the first homogeneous determination of the SPSS dataset to date. We validated each of the samples using the *Gaia* data itself and external catalogues, and our results are published in six tables that will appear alongside *Gaia* DR3, see Sect. 9 and Table 7.

In Sect. 10, we demonstrated some use cases of these samples of stars. We used a subset of the OBA sample to illustrate its usefulness to analyse the Milky Way rotation curve (Sect. 10.1). We then used the properties of the FGKM stars to analyse known exoplanet systems including the determination of planet radii and masses (Sect. 10.2). We then predicted the colours of the Sun in various passbands using the solar analogue sample (Sect. 10.3). Finally, we analysed the ages of some unseen UCD-companions to the FGKM stars (Sect. 10.4).

The aim of this work was to highlight the science that can be done with *Gaia* DR3. We focused on specific types of stars using strict quality criteria on many of the data products, which sometimes included some ad hoc filtering criteria tuned with a particular science case in mind. We emphasize that our strict personal selections may not be applicable to a user's specific science case, and the user should acknowledge this before exploiting these samples. We fully encourage all users to exploit all of the astrophysical parameters in *Gaia* DR3 independent of our specific selection criteria highlighted in this work. Indeed there are up to 470 million stars with stellar parameters derived using the mean BP and RP spectra, up to 6 million stellar parameters and abundances derived from the mean RVS spectra, along with up to 130 million masses and ages, and many other new stellar products that were not the focus of this work, such as DIB estimates, activity index of active stars, and Hα emission. As illustrated in this work, many science cases can be explored with these data.

Acknowledgements. We thank the referee for their constructive comments on the manuscript. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`), processed by the Gaia Data Processing and Analysis Consortium (DPAC, `https://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, in particular, the institutions participating in the Gaia Multilateral Agreement. The full list of funding agencies and grants is listed after the references.
This research has used NASA's Astrophysics Data System, and the VizieR catalogue access tool (CDS, Strasbourg, France).
The data processing and analysis made use of matplotlib (Hunter 2007), NumPy (Harris et al. 2020), the IPython package (Pérez & Granger 2007), Vaex (Breddels & Veljanoski 2018), TOPCAT (Taylor 2005), pyphot (`http://github.com/mfouesneau/pyphot`), R (R Core Team 2013), Astropy (Astropy Collaboration et al. 2013, 2018), CmdStanPy (`https://github.com/stan-dev/cmdstanpy`), and ArviZ (Kumar et al. 2019).
In case of errors or omissions, please contact the *Gaia* Helpdesk. The full list of acknowledgements can also be found in the official online documentation for *Gaia* DR3.

References

Abia, C., de Laverny, P., Cristallo, S., Kordopatis, G., & Straniero, O. 2020, A&A, 633, A135
Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3
Alksnis, A., Balklavs, A., Dzervitis, U., et al. 2001, Baltic Astronomy, 10, 1
Allard, F., Homeier, D., & Freytag, B. 2012, Philosophical Transactions of the Royal Society of London Series A, 370, 2765
Altavilla, G., Marinoni, S., Pancino, E., et al. 2021, MNRAS, 501, 2848
Altavilla, G., Marinoni, S., Pancino, E., et al. 2015, Astronomische Nachrichten, 336, 515
Anders, F., Khalatyan, A., Queiroz, A. B. A., et al. 2022, A&A, 658, A91
Andrae, R., Fouesneau, M., Sordo, R., Bailer-Jones, C., & et al. 2022, A&A, accepted
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Babusiaux, C., Fabricius, C., & et al. 2022, A&A, accepted
Bailer-Jones, C. A. L., Andrae, R., Arcay, B., et al. 2013, A&A, 559, A74
Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, A&A, 577, A42
Bayo, A., Rodrigo, C., Barrado Y Navascués, D., et al. 2008, A&A, 492, 277
Bennett, M. & Bovy, J. 2019, MNRAS, 482, 1417
Birko, D., Zwitter, T., Grebel, E. K., et al. 2019, AJ, 158, 155
Bland-Hawthorn, J. & Gerhard, O. 2016, ARA&A, 54, 529
Bobylev, V. V. & Bajkova, A. T. 2022, Astronomy Reports, 66, 269
Bohlin, R. C. 2014, AJ, 147, 127
Bohlin, R. C., Colina, L., & Finley, D. S. 1995, AJ, 110, 1316
Bohlin, R. C., Deustua, S. E., & de Rosa, G. 2019, AJ, 158, 211
Bohlin, R. C., Gordon, K. D., & Tremblay, P. E. 2014, PASP, 126, 711
Bohlin, R. C., Hubeny, I., & Rauch, T. 2020, AJ, 160, 21
Borucki, W. J., Koch, D., Basri, G., et al. 2008, in A Decade of Extrasolar Planets around Normal Stars Proceedings of the Space Telescope Science Institute Symposium, 36–49
Bovy, J. 2015, ApJS, 216, 29
Bowler, B. P., Blunt, S. C., & Nielsen, E. L. 2020, AJ, 159, 63
Brandt, T. D. 2021, ApJS, 254, 42
Brandt, T. D., Dupuy, T. J., & Bowler, B. P. 2019, AJ, 158, 140
Breddels, M. A. & Veljanoski, J. 2018, A&A, 618, A13
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Brunetti, M. & Pfenniger, D. 2010, A&A, 510, A34
Buder, S., Asplund, M., Duong, L., et al. 2018, MNRAS, 478, 4513
Burgasser, A. J., Burrows, A., & Kirkpatrick, J. D. 2006, ApJ, 639, 1095
Burningham, B., Cardoso, C. V., Smith, L., et al. 2013, MNRAS, 433, 457
Burrows, A., Hubbard, W. B., Lunine, J. I., & Liebert, J. 2001, Reviews of Modern Physics, 73, 719
Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, 640, A1
Carrasco, J. M., Weiler, M., Jordi, C., et al. 2021, A&A, 652, A86
Casagrande, L., Schönrich, R., Asplund, M., et al. 2011, A&A, 530, A138
Casagrande, L. & VandenBerg, D. A. 2018, MNRAS, 479, L102
Casali, G., Spina, L., Magrini, L., et al. 2020, A&A, 639, A127
Chiu, K., Fan, X., Leggett, S. K., et al. 2006, AJ, 131, 2722
Cifuentes, C., Caballero, J. A., Cortés-Contreras, M., et al. 2020, A&A, 642, A115
Clementini, G., Ripepi, V., Garofalo, A., et al. 2022, A&A, submitted
Creevey, O., Sordo, R., Pailler, F., et al. 2022, A&A, accepted
Crepp, J. R., Johnson, J. A., Howard, A. W., et al. 2014, ApJ, 781, 29
Crouzet, N., McCullough, P. R., Long, D., et al. 2017, AJ, 153, 94
Cruz, K. L., Reid, I. N., Kirkpatrick, J. D., et al. 2007, AJ, 133, 439
Cutri, R. M., Wright, E. L., Conrow, T., et al. 2021, VizieR Online Data Catalog, II/328
Damerdji, Y. Gosset, E., Morel, T., Blomme, R., & et al. 2022, A&A, in prep.
De Angeli, F., Weiler, M., Montegriffo, P., & et al. 2022, A&A, submitted
Deacon, N. R., Liu, M. C., Magnier, E. A., et al. 2014, ApJ, 792, 119
Delchambre, L., Bailer-Jones, C., Bellas-Velidis, I., Drimmel, R., & et al. 2022, A&A, accepted
Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, Research in Astronomy and Astrophysics, 12, 735
Dieterich, S. B., Henry, T. J., Jao, W.-C., et al. 2014, AJ, 147, 94
Dieterich, S. B., Weinberger, A. J., Boss, A. P., et al. 2018, ApJ, 865, 28
Eilers, A.-C., Hogg, D. W., Rix, H.-W., & Ness, M. K. 2019, ApJ, 871, 120
Eyer, L., Audard, M., Holl, B., Rimoldini, L., & et al. 2022, A&A, in prep.
Faherty, J. K., Burgasser, A. J., Walter, F. M., et al. 2012, ApJ, 752, 56
Faherty, J. K., Burgasser, A. J., West, A. A., et al. 2010, AJ, 139, 176
Filippazzo, J. C., Rice, E. L., Faherty, J., et al. 2015, ApJ, 810, 158
Fouesneau, M., Frémat, Y., Andrae, R., Korn, A., & et al. 2022, A&A, in prep.
Fulton, B. J., Collins, K. A., Gaudi, B. S., et al. 2015, ApJ, 810, 30
Gagné, J., Mamajek, E. E., Malo, L., et al. 2018, ApJ, 856, 23
Gaia Collaboration, Arenou, F., Babusiaux, C., et al. 2022a, A&A, accepted
Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018, A&A, 616, A10
Gaia Collaboration, Drimmel, R., Romero-Gomez, M., et al. 2022b, A&A, in prep.
Gaia Collaboration, Galluccio, L., Delbo, M., et al. 2022c, A&A, accepted
Gaia Collaboration, Recio-Blanco, A., Kordopatis, G., et al. 2022d, A&A, accepted
Gaia Collaboration, Schultheis, M., Zhao, H., et al. 2022e, A&A, accepted
Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, Space Sci. Rev., 123, 485
Gilliland, R. L., Marcy, G. W., Rowe, J. F., et al. 2013, ApJ, 766, 40
Gilmore, G., Randich, S., Asplund, M., et al. 2012, The Messenger, 147, 25
Giribaldi, R. E., Porto de Mello, G. F., Lorenzo-Oliveira, D., Amôres, E. B., & Ubaldo-Melo, M. L. 2019, A&A, 629, A33
Gontcharov, G. A. 2012, Astronomy Letters, 38, 771
Goto, M., Kobayashi, N., Terada, H., et al. 2002, ApJ, 567, L59
Gravity Collaboration, Abuter, R., Aimar, N., et al. 2022, A&A, 657, L12
Guillot, T. & Gautier, D. 2015, in Treatise on Geophysics, ed. G. Schubert, 529–557
Halbwachs, J.-L., Pourbaix†, D., Frédéric Arenou, F., Galluccio, L., & et al. 2022, A&A, in prep.
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Harrison, D. L. 2011, Experimental Astronomy, 31, 157
Heller, R. 2019, A&A, 623, A137
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Nature, 563, 85
Hidalgo, S. L., Pietrinferni, A., Cassisi, S., et al. 2018, ApJ, 856, 125
Holl et al., B. 2022, A&A in prep.
Holmberg, J., Flynn, C., & Portinari, L. 2006, MNRAS, 367, 449
Holmberg, J., Nordström, B., & Andersen, J. 2009, A&A, 501, 941
Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90
Jofré, P., Heiter, U., & Soubiran, C. 2019, ARA&A, 57, 571
Kaltenegger, L. & Selsis, F. 2015, in Encyclopedia of Astrobiology, ed. M. Gargaud, W. M. Irvine, R. Amils, I. Cleaves, Henderson James (Jim), D. L. Pinti, J. C. Quintanilla, D. Rouan, T. Spohn, S. Tirard, & M. Viso, 1039–1048
Kervella, P., Arenou, F., & Thévenin, F. 2022, A&A, 657, A7
King, J. R., Villarreal, A. R., Soderblom, D. R., Gulliver, A. F., & Adelman, S. J. 2003, AJ, 125, 1980
Kirkpatrick, J. D., Gelino, C. R., Faherty, J. K., et al. 2021, ApJS, 253, 7
Kirkpatrick, J. D., Henry, T. J., & Irwin, M. J. 1997, AJ, 113, 1421
Kontizas, E., Dapergolas, A., Morgan, D. H., & Kontizas, M. 2001, A&A, 369, 932
Kounkel, M., Covey, K. R., Stassun, K. G., et al. 2021, AJ, 162, 184
Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. 2019, Journal of Open Source Software, 4, 1143
Lanzafame, A., Brugaletta, E., Frémat, Y., Sordo, R., & et al. 2022, A&A, in prep.
Lindegren, L., Bastian, U., Biermann, M., et al. 2021a, A&A, 649, A4
Lindegren, L., Klioner, S., Hernández, J., et al. 2021b, A&A, 649, A2
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021c, A&A, 649, A2
Lindegren, L., Madsen, S., & Dravins, D. 2000, A&A, 356, 1119
Liu, M. C., Wainscoat, R., Martín, E. L., Barris, B., & Tonry, J. 2002, ApJ, 568, L107
Lorenzo-Oliveira, D., Freitas, F. C., Meléndez, J., et al. 2018, A&A, 619, A73
Loutrel, N. P., Luhman, K. L., Lowrance, P. J., & Bochanski, J. J. 2011, ApJ, 739, 81
Mahdi, D., Soubiran, C., Blanco-Cuaresma, S., & Chemin, L. 2016, A&A, 587, A131
Mainzer, A., Bauer, J., Grav, T., et al. 2011, ApJ, 731, 53
Maíz Apellániz, J., Sota, A., Morrell, N. I., et al. 2013, in Massive Stars: From alpha to Omega, 198
Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, VizieR Online Data Catalog, J/ApJS/210/20
Marinoni, S., Pancino, E., Altavilla, G., et al. 2016, MNRAS, 462, 3616
Marrese, P. M., Marinoni, S., Fabrizio, M., & Altavilla, G. 2019, A&A, 621, A144
Marrese, P. M., Marinoni, S., Fabrizio, M., & Giuffrida, G. 2017, A&A, 607, A105
Meléndez, J., Asplund, M., Gustafsson, B., & Yong, D. 2009, ApJ, 704, L66
Merle, T., Van Eck, S., Jorissen, A., et al. 2017, A&A, 608, A95
Metchev, S. A. & Hillenbrand, L. A. 2004, ApJ, 617, 1330
Miller, A. A. 2015, ApJ, 811, 30
Montalto, M., Piotto, G., Marrese, P. M., et al. 2021, VizieR Online Data Catalog, J/A+A/653/A98
Montegriffo, P., De Angeli, F., Andrae, R., Riello, M., & et al. 2022, A&A, accepted
Morgan, D. H. & Hatzidimitriou, D. 1995, A&AS, 113, 539
Mowlavi, N., Holl, B., Lecoeur-Tabi, I., Barbian, F., & et al. 2022, A&A, in prep.
Nissen, P. E. 2015, A&A, 579, A52
Nissen, P. E. & Gustafsson, B. 2018, A&A Rev., 26, 6
Olling, R. P. & Dehnen, W. 2003, ApJ, 599, 275
Pace, G. 2013, A&A, 551, L8
Pancino, E., Altavilla, G., Marinoni, S., et al. 2012, MNRAS, 426, 1767
Pancino, E., Sanna, N., Altavilla, G., et al. 2021a, MNRAS, 503, 3660

Pancino, E., Sanna, N., Altavilla, G., et al. 2021b, MNRAS, 503, 3660
Pérez, F. & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21
Phillips, M. W., Tremblin, P., Baraffe, I., et al. 2020, A&A, 637, A38
Pinfield, D. J., Burningham, B., Lodieu, N., et al. 2012, MNRAS, 422, 1922
Porto de Mello, G. F., da Silva, R., da Silva, L., & de Nader, R. V. 2014, A&A, 563, A52
Potter, D., Martín, E. L., Cushing, M. C., et al. 2002, ApJ, 567, L133
Price-Whelan, A. M., Hogg, D. W., Rix, H.-W., et al. 2020, ApJ, 895, 2
Qian, S.-B., Shi, X.-D., Zhu, L.-Y., et al. 2019, Research in Astronomy and Astrophysics, 19, 064
R Core Team. 2013, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria
Ramírez, I., Meléndez, J., Bean, J., et al. 2014, A&A, 572, A48
Rando, N., Asquier, J., Corral Van Damme, C., et al. 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11443, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1144314
Rauer, H., Catala, C., Aerts, C., et al. 2014, Experimental Astronomy, 38, 249
Recio-Blanco, A. & et al. 2022, A&A, accepted
Reid, I. N., Lewitus, E., Allen, P. R., Cruz, K. L., & Burgasser, A. J. 2006, AJ, 132, 891
Reid, M. J. & Brunthaler, A. 2020, ApJ, 892, 39
Reylé, C. 2018, A&A, 619, L8
Ricker, G. R. 2014, *JAAVSO*, 42, 234
Riello, M., De Angeli, F., Evans, D. W., et al. 2021a, A&A, 649, A3
Riello, M., De Angeli, F., Evans, D. W., et al. 2021b, A&A, 649, A3
Rybizki, J., Green, G. M., Rix, H.-W., et al. 2022, MNRAS, 510, 2597
Salaris, M., Chieffi, A., & Straniero, O. 1993, ApJ, 414, 580
Sartoretti, P., Katz, D., Cropper, M., et al. 2018, A&A, 616, A6
Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829
Seabroke, G. & et al. 2022, A&A, in prep.
Siopis, C. 2022, A&A in prep.
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Smart, R. L., Bucciarelli, B., Jones, H. R. A., et al. 2018, Monthly Notices of the Royal Astronomical Society, 481, 3548
Smart, R. L., Marocco, F., Caballero, J. A., et al. 2017, MNRAS, 469, 401
Smart, R. L., Marocco, F., Sarro, L. M., et al. 2019, MNRAS, 485, 4423
Stan Development Team. 2022, Stan Modeling Language Users Guide and Reference Manual, V2.29
Steinmetz, M., Guiglion, G., McMillan, P. J., et al. 2020, AJ, 160, 83
Stephens, D. C., Leggett, S. K., Cushing, M. C., et al. 2009, ApJ, 702, 154
Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29
Tian, H.-J., El-Badry, K., Rix, H.-W., & Gould, A. 2020, ApJS, 246, 4
Tinetti, G., Drossart, P., Eccleston, P., et al. 2018, Experimental Astronomy, 46, 135
Tokovinin, A. 2014, AJ, 147, 86
Traven, G., Feltzing, S., Merle, T., et al. 2020, A&A, 638, A145
Tsantaki, M., Pancino, E., Marrese, P., et al. 2022, A&A, 659, A95
Tucci Maia, M., Ramírez, I., Meléndez, J., et al. 2016, A&A, 590, A32
Weinberger, A. J., Boss, A. P., Keiser, S. A., et al. 2016, AJ, 152, 24
Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, A&AS, 143, 9
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Xiang, M., Rix, H.-W., Ting, Y.-S., et al. 2021, arXiv e-prints, arXiv:2108.02878
Xu, S., Yuan, H., Niu, Z., et al. 2022, ApJS, 258, 44
Yana Galarza, J., López-Valdivia, R., Lorenzo-Oliveira, D., et al. 2021, MNRAS, 504, 1873
Zari, E., Rix, H. W., Frankel, N., et al. 2021, A&A, 650, A112
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127

[1] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France
[2] Dpto. de Inteligencia Artificial, UNED, c/ Juan del Rosal 16, 28040 Madrid, Spain
[3] Royal Observatory of Belgium, Ringlaan 3, 1180 Brussels, Belgium
[4] INAF - Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italy
[5] Space Science Data Center - ASI, Via del Politecnico SNC, 00133 Roma, Italy
[6] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
[7] INAF - Osservatorio Astrofisico di Torino, via Osservatorio 20, 10025 Pino Torinese (TO), Italy
[8] INAF - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Piero Gobetti 93/3, 40129 Bologna, Italy
[9] Observational Astrophysics, Division of Astronomy and Space Physics, Department of Physics and Astronomy, Uppsala University, Box 516, 751 20 Uppsala, Sweden
[10] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, United Kingdom
[11] INAF - Osservatorio astronomico di Padova, Vicolo Osservatorio 5, 35122 Padova, Italy
[12] Kavli Institute for Cosmology Cambridge, Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA
[13] Institut UTINAM CNRS UMR6213, Université Bourgogne Franche-Comté, OSU THETA Franche-Comté Bourgogne, Observatoire de Besançon, BP1615, 25010 Besançon Cedex, France
[14] Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
[15] Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland
[16] European Space Agency (ESA), European Space Research and Technology Centre (ESTEC), Keplerlaan 1, 2201AZ, Noordwijk, The Netherlands
[17] GEPI, Observatoire de Paris, Université PSL, CNRS, 5 Place Jules Janssen, 92190 Meudon, France
[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] Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[21] European Space Agency (ESA), European Space Astronomy Centre (ESAC), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[22] Aurora Technology for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[23] Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
[24] Lohrmann Observatory, Technische Universität Dresden, Mommsenstraße 13, 01062 Dresden, Germany
[25] Lund Observatory, Department of Astronomy and Theoretical Physics, Lund University, Box 43, 22100 Lund, Sweden
[26] CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[27] Institut d'Astronomie et d'Astrophysique, Université Libre de Bruxelles CP 226, Boulevard du Triomphe, 1050 Brussels, Belgium
[28] F.R.S.-FNRS, Rue d'Egmont 5, 1000 Brussels, Belgium
[29] European Space Agency (ESA, retired)
[30] University of Turin, Department of Physics, Via Pietro Giuria 1, 10125 Torino, Italy
[31] DAPCOM for Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
[32] ALTEC S.p.a, Corso Marche, 79,10146 Torino, Italy
[33] Sàrl, Geneva, Switzerland
[34] Department of Astronomy, University of Geneva, Chemin d'Ecogia 16, 1290 Versoix, Switzerland
[35] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, United Kingdom
[36] Gaia DPAC Project Office, ESAC, Camino bajo del Castillo, s/n, Urbanizacion 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, Urbanizacion 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] National Observatory of Athens, I. Metaxa and Vas. Pavlou, Palaia Penteli, 15236 Athens, Greece

[40] IMCCE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Lille, 77 av. Denfert-Rochereau, 75014 Paris, France
[41] Serco Gestión de Negocios for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[42] Institut d'Astrophysique et de Géophysique, Université de Liège, 19c, Allée du 6 Août, B-4000 Liège, Belgium
[43] CRAAG - Centre de Recherche en Astronomie, Astrophysique et Géophysique, Route de l'Observatoire Bp 63 Bouzareah 16340 Algiers, Algeria
[44] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, United Kingdom
[45] RHEA for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[46] ATG Europe for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[47] CIGUS CITIC - Department of Computer Science and Information Technologies, University of A Coruña, Campus de Elviña s/n, A Coruña, 15071, Spain
[48] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 11 rue de l'Université, 67000 Strasbourg, France
[49] Leibniz Institute for Astrophysics Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
[50] CENTRA, Faculdade de Ciências, Universidade de Lisboa, Edif. C8, Campo Grande, 1749-016 Lisboa, Portugal
[51] Department of Informatics, Donald Bren School of Information and Computer Sciences, University of California, Irvine, 5226 Donald Bren Hall, 92697-3440 CA Irvine, United States
[52] INAF - Osservatorio Astrofisico di Catania, via S. Sofia 78, 95123 Catania, Italy
[53] Dipartimento di Fisica e Astronomia ""Ettore Majorana"", Università di Catania, Via S. Sofia 64, 95123 Catania, Italy
[54] INAF - Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monte Porzio Catone (Roma), Italy
[55] Department of Physics, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland
[56] Finnish Geospatial Research Institute FGI, Geodeetinrinne 2, 02430 Masala, Finland
[57] HE Space Operations BV for European Space Agency (ESA), Keplerlaan 1, 2201AZ, Noordwijk, The Netherlands
[58] 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
[59] ELTE Eötvös Loránd University, Institute of Physics, 1117, Pázmány Péter sétány 1A, Budapest, Hungary
[60] Instituut voor Sterrenkunde, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium
[61] Department of Astrophysics/IMAPP, Radboud University, P.O.Box 9010, 6500 GL Nijmegen, The Netherlands
[62] University of Vienna, Department of Astrophysics, Türkenschanzstraße 17, A1180 Vienna, Austria
[63] Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
[64] Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands
[65] School of Physics and Astronomy / Space Park Leicester, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom
[66] Thales Services for CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[67] Depto. Estadística e Investigación Operativa. Universidad de Cádiz, Avda. República Saharaui s/n, 11510 Puerto Real, Cádiz, Spain
[68] Center for Research and Exploration in Space Science and Technology, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore MD, USA
[69] GSFC - Goddard Space Flight Center, Code 698, 8800 Greenbelt Rd, 20771 MD Greenbelt, United States
[70] EURIX S.r.l., Corso Vittorio Emanuele II 61, 10128, Torino, Italy
[71] Porter School of the Environment and Earth Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
[72] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., MS 15, Cambridge, MA 02138, USA
[73] HE Space Operations BV for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanizacion Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[74] Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal
[75] LFCA/DAS,Universidad de Chile,CNRS,Casilla 36-D, Santiago, Chile
[76] SISSA - Scuola Internazionale Superiore di Studi Avanzati, via Bonomea 265, 34136 Trieste, Italy
[77] Telespazio for CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[78] University of Turin, Department of Computer Sciences, Corso Svizzera 185, 10149 Torino, Italy
[79] 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
[80] Centro de Astronomía - CITEVA, Universidad de Antofagasta, Avenida Angamos 601, Antofagasta 1270300, Chile
[81] DLR Gesellschaft für Raumfahrtanwendungen (GfR) mbH Münchener Straße 20 , 82234 Weßling
[82] Centre for Astrophysics Research, University of Hertfordshire, College Lane, AL10 9AB, Hatfield, United Kingdom
[83] University of Turin, Mathematical Department ""G.Peano"", Via Carlo Alberto 10, 10123 Torino, Italy
[84] INAF - Osservatorio Astronomico d'Abruzzo, Via Mentore Maggini, 64100 Teramo, Italy
[85] Instituto de Astronomia, Geofisica e Ciências Atmosféricas, Universidade de São Paulo, Rua do Matão, 1226, Cidade Universitaria, 05508-900 São Paulo, SP, Brazil
[86] APAVE SUDEUROPE SAS for CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[87] Mésocentre de calcul de Franche-Comté, Université de Franche-Comté, 16 route de Gray, 25030 Besançon Cedex, France
[88] ATOS for CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[89] School of Physics and Astronomy, Tel Aviv University, Tel Aviv 6997801, Israel
[90] Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, UK
[91] Centre de Données Astronomique de Strasbourg, Strasbourg, France
[92] Institute for Computational Cosmology, Department of Physics, Durham University, Durham DH1 3LE, UK
[93] European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany
[94] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85748 Garching, Germany
[95] Data Science and Big Data Lab, Pablo de Olavide University, 41013, Seville, Spain
[96] Barcelona Supercomputing Center (BSC), Plaça Eusebi Güell 1-3, 08034-Barcelona, Spain
[97] ETSE Telecomunicación, Universidade de Vigo, Campus Lagoas-Marcosende, 36310 Vigo, Galicia, Spain
[98] Asteroid Engineering Laboratory, Space Systems, Luleå University of Technology, Box 848, S-981 28 Kiruna, Sweden
[99] Vera C Rubin Observatory, 950 N. Cherry Avenue, Tucson, AZ 85719, USA
[100] Department of Astrophysics, Astronomy and Mechanics, National and Kapodistrian University of Athens, Panepistimiopolis, Zografos, 15783 Athens, Greece

[101] TRUMPF Photonic Components GmbH, Lise-Meitner-Straße 13, 89081 Ulm, Germany
[102] IAC - Instituto de Astrofisica de Canarias, Via Láctea s/n, 38200 La Laguna S.C., Tenerife, Spain
[103] Department of Astrophysics, University of La Laguna, Via Láctea s/n, 38200 La Laguna S.C., Tenerife, Spain
[104] Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands
[105] Radagast Solutions
[106] Laboratoire Univers et Particules de Montpellier, CNRS Université Montpellier, Place Eugène Bataillon, CC72, 34095 Montpellier Cedex 05, France
[107] Université de Caen Normandie, Côte de Nacre Boulevard Maréchal Juin, 14032 Caen, France
[108] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 Place Jules Janssen, 92190 Meudon, France
[109] SRON Netherlands Institute for Space Research, Niels Bohrweg 4, 2333 CA Leiden, The Netherlands
[110] Astronomical Observatory, University of Warsaw, Al. Ujazdowskie 4, 00-478 Warszawa, Poland
[111] Scalian for CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
[112] Université Rennes, CNRS, IPR (Institut de Physique de Rennes) - UMR 6251, 35000 Rennes, France
[113] INAF - Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131, Napoli, Italy
[114] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, People's Republic of China
[115] University of Chinese Academy of Sciences, No.19(A) Yuquan Road, Shijingshan District, Beijing 100049, People's Republic of China
[116] Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark
[117] DXC Technology, Retortvej 8, 2500 Valby, Denmark
[118] Las Cumbres Observatory, 6740 Cortona Drive Suite 102, Goleta, CA 93117, USA
[119] CIGUS CITIC, Department of Nautical Sciences and Marine Engineering, University of A Coruña, Paseo de Ronda 51, 15071, A Coruña, Spain
[120] Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, United Kingdom
[121] IPAC, Mail Code 100-22, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
[122] IRAP, Université de Toulouse, CNRS, UPS, CNES, 9 Av. colonel Roche, BP 44346, 31028 Toulouse Cedex 4, France
[123] 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
[124] Departmento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid, 28040 Madrid, Spain
[125] Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia
[126] Villanova University, Department of Astrophysics and Planetary Science, 800 E Lancaster Avenue, Villanova PA 19085, USA
[127] INAF - Osservatorio Astronomico di Brera, via E. Bianchi, 46, 23807 Merate (LC), Italy
[128] STFC, Rutherford Appleton Laboratory, Harwell, Didcot, OX11 0QX, United Kingdom
[129] Charles University, Faculty of Mathematics and Physics, Astronomical Institute of Charles University, V Holesovickach 2, 18000 Prague, Czech Republic
[130] Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 7610001, Israel
[131] Department of Astrophysical Sciences, 4 Ivy Lane, Princeton University, Princeton NJ 08544, USA
[132] Departamento de Astrofísica, Centro de Astrobiología (CSIC-INTA), ESA-ESAC. Camino Bajo del Castillo s/n. 28692 Villanueva de la Cañada, Madrid, Spain
[133] naXys, University of Namur, Rempart de la Vierge, 5000 Namur, Belgium
[134] CGI Deutschland B.V. & Co. KG, Mornewegstr. 30, 64293 Darmstadt, Germany
[135] Institute of Global Health, University of Geneva
[136] Astronomical Observatory Institute, Faculty of Physics, Adam Mickiewicz University, Poznań, Poland
[137] H H Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom
[138] Department of Physics and Astronomy G. Galilei, University of Padova, Vicolo dell'Osservatorio 3, 35122, Padova, Italy
[139] CERN, Geneva, Switzerland
[140] Applied Physics Department, Universidade de Vigo, 36310 Vigo, Spain
[141] Association of Universities for Research in Astronomy, 1331 Pennsylvania Ave. NW, Washington, DC 20004, USA
[142] European Southern Observatory, Alonso de Córdova 3107, Casilla 19, Santiago, Chile
[143] Sorbonne Université, CNRS, UMR7095, Institut d'Astrophysique de Paris, 98bis bd. Arago, 75014 Paris, France
[144] Faculty of Mathematics and Physics, University of Ljubljana, Jadranska ulica 19, 1000 Ljubljana, Slovenia

Gaia funding institutions

This work presents results from the European Space Agency (ESA) space mission *Gaia*. *Gaia* data are processed by the *Gaia* Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the *Gaia* MultiLateral Agreement (MLA). The *Gaia* mission website is https://www.cosmos.esa.int/gaia. The *Gaia* archive website is https://archives.esac.esa.int/gaia.

The *Gaia* mission and data processing have financially been supported by, in alphabetical order by country:
– the Algerian Centre de Recherche en Astronomie, Astrophysique et Géophysique of Bouzareah Observatory;
– the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (FWF) Hertha Firnberg Programme through grants T359, P20046, and P23737;
– the BELgian federal Science Policy Office (BELSPO) through various PROgramme de Développement d'Expériences scientifiques (PRODEX) grants, the Research Foundation Flanders (Fonds Wetenschappelijk Onderzoek) through grant VS.091.16N, the Fonds de la Recherche Scientifique (FNRS), and the Research Council of Katholieke Universiteit (KU) Leuven through grant C16/18/005 (Pushing AsteRoseismology to the next level with TESS, GaiA, and the Sloan DIgital Sky SurvEy – PARADISE);
– the Brazil-France exchange programmes Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Coordenação de Aperfeicoamento de Pessoal de Nível Superior (CAPES) - Comité Français d'Evaluation de la Coopération Universitaire et Scientifique avec le Brésil (COFECUB);
– the Chilean Agencia Nacional de Investigación y Desarrollo (ANID) through Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) Regular Project 1210992 (L. Chemin);
– the National Natural Science Foundation of China (NSFC) through grants 11573054, 11703065, and 12173069, the China Scholarship Council through grant 201806040200, and the Natural Science Foundation of Shanghai through grant 21ZR1474100;
– the Tenure Track Pilot Programme of the Croatian Science Foundation and the École Polytechnique Fédérale de Lausanne and the project TTP-2018-07-1171 'Mining the Variable Sky', with the funds of the Croatian-Swiss Research Programme;
– the Czech-Republic Ministry of Education, Youth, and Sports through grant LG 15010 and INTER-EXCELLENCE grant LTAUSA18093, and the Czech Space Office through ESA PECS contract 98058;
– the Danish Ministry of Science;
– the Estonian Ministry of Education and Research through grant IUT40-1;
– the European Commission's Sixth Framework Programme through the European Leadership in Space Astrometry (ELSA) Marie Curie Research Training Network (MRTN-CT-2006-033481), through Marie Curie project PIOF-GA-2009-255267 (Space AsteroSeismology & RR Lyrae stars, SAS-RRL), and through a Marie Curie Transfer-of-Knowledge (ToK) fellowship (MTKD-CT-2004-014188); the European Commission's Seventh Framework Programme through grant FP7-606740 (FP7-SPACE-2013-1) for the *Gaia* European Network for Improved data User Services (GENIUS) and through grant 264895 for the *Gaia* Research for European Astronomy Training (GREAT-ITN) network;
– the European Cooperation in Science and Technology (COST) through COST Action CA18104 'Revealing the Milky Way with *Gaia* (MW-Gaia)';
– the European Research Council (ERC) through grants 320360, 647208, and 834148 and through the European Union's Horizon 2020 research and innovation and excellent science programmes through Marie Skłodowska-Curie grant 745617 (Our Galaxy at full HD – Gal-HD) and 895174 (The build-up and fate of self-gravitating systems in the Universe) as well as grants 687378 (Small Bodies: Near and Far), 682115 (Using the Magellanic Clouds to Understand the Interaction of Galaxies), 695099 (A sub-percent distance scale from binaries and Cepheids – CepBin), 716155 (Structured ACCREtion Disks – SACCRED), 951549 (Sub-percent calibration of the extragalactic distance scale in the era of big surveys – UniverScale), and 101004214 (Innovative Scientific Data Exploration and Exploitation Applications for Space Sciences – EXPLORE);
– the European Science Foundation (ESF), in the framework of the *Gaia* Research for European Astronomy Training Research Network Programme (GREAT-ESF);
– the European Space Agency (ESA) in the framework of the *Gaia* project, through the Plan for European Cooperating States (PECS) programme through contracts C98090 and 4000106398/12/NL/KML for Hungary, through contract 4000115263/15/NL/IB for Germany, and through PROgramme de Développement d'Expériences scientifiques (PRODEX) grant 4000127986 for Slovenia;
– the Academy of Finland through grants 299543, 307157, 325805, 328654, 336546, and 345115 and the Magnus Ehrnrooth Foundation;
– the French Centre National d'Études Spatiales (CNES), the Agence Nationale de la Recherche (ANR) through grant ANR-10-IDEX-0001-02 for the 'Investissements d'avenir' programme, through grant ANR-15-CE31-0007 for project 'Modelling the Milky Way in the *Gaia* era' (MOD4Gaia), through grant ANR-14-CE33-0014-01 for project 'The Milky Way disc formation in the *Gaia* era' (ARCHEOGAL), through grant ANR-15-CE31-0012-01 for project 'Unlocking the potential of Cepheids as primary distance calibrators' (UnlockCepheids), through grant ANR-19-CE31-0017 for project 'Secular evolution of galxies' (SEGAL), and through grant ANR-18-CE31-0006 for project 'Galactic Dark Matter' (GaDaMa), the Centre National de la Recherche Scientifique (CNRS) and its SNO *Gaia* of the Institut des Sciences de l'Univers (INSU), its Programmes Nationaux: Cosmologie et Galaxies (PNCG), Gravitation Références Astronomie Métrologie (PNGRAM), Planétologie (PNP), Physique et Chimie du Milieu Interstellaire (PCMI), and Physique Stellaire (PNPS), the 'Action Fédératrice *Gaia*' of the Observatoire de Paris, the Région de Franche-Comté, the Institut National Polytechnique (INP) and the Institut National de Physique nucléaire et de Physique des Particules (IN2P3) co-funded by CNES;
– the German Aerospace Agency (Deutsches Zentrum für Luft- und Raumfahrt e.V., DLR) through grants 50QG0501, 50QG0601, 50QG0602, 50QG0701, 50QG0901, 50QG1001, 50QG1101, 50QG1401, 50QG1402, 50QG1403, 50QG1404, 50QG1904, 50QG2101, 50QG2102, and 50QG2202, and the Centre for Information Services and High Performance Computing (ZIH) at the Technische Universität Dresden for generous allocations of computer time;
– the Hungarian Academy of Sciences through the Lendület Programme grants LP2014-17 and LP2018-7 and the Hungarian National Research, Development, and Innovation Office (NKFIH) through grant KKP-137523 ('SeismoLab');
– the Science Foundation Ireland (SFI) through a Royal Society - SFI University Research Fellowship (M. Fraser);

– the Israel Ministry of Science and Technology through grant 3-18143 and the Tel Aviv University Center for Artificial Intelligence and Data Science (TAD) through a grant;
– the Agenzia Spaziale Italiana (ASI) through contracts I/037/08/0, I/058/10/0, 2014-025-R.0, 2014-025-R.1.2015, and 2018-24-HH.0 to the Italian Istituto Nazionale di Astrofisica (INAF), contract 2014-049-R.0/1/2 to INAF for the Space Science Data Centre (SSDC, formerly known as the ASI Science Data Center, ASDC), contracts I/008/10/0, 2013/030/I.0, 2013-030-I.0.1-2015, and 2016-17-I.0 to the Aerospace Logistics Technology Engineering Company (ALTEC S.p.A.), INAF, and the Italian Ministry of Education, University, and Research (Ministero dell'Istruzione, dell'Università e della Ricerca) through the Premiale project 'MIning The Cosmos Big Data and Innovative Italian Technology for Frontier Astrophysics and Cosmology' (MITiC);
– the Netherlands Organisation for Scientific Research (NWO) through grant NWO-M-614.061.414, through a VICI grant (A. Helmi), and through a Spinoza prize (A. Helmi), and the Netherlands Research School for Astronomy (NOVA);
– the Polish National Science Centre through HARMONIA grant 2018/30/M/ST9/00311 and DAINA grant 2017/27/L/ST9/03221 and the Ministry of Science and Higher Education (MNiSW) through grant DIR/WK/2018/12;
– the Portuguese Fundação para a Ciência e a Tecnologia (FCT) through national funds, grants SFRH/BD/128840/2017 and PTDC/FIS-AST/30389/2017, and work contract DL 57/2016/CP1364/CT0006, the Fundo Europeu de Desenvolvimento Regional (FEDER) through grant POCI-01-0145-FEDER-030389 and its Programa Operacional Competitividade e Internacionalização (COMPETE2020) through grants UIDB/04434/2020 and UIDP/04434/2020, and the Strategic Programme UIDB/00099/2020 for the Centro de Astrofísica e Gravitação (CENTRA);
– the Slovenian Research Agency through grant P1-0188;
– the Spanish Ministry of Economy (MINECO/FEDER, UE), the Spanish Ministry of Science and Innovation (MICIN), the Spanish Ministry of Education, Culture, and Sports, and the Spanish Government through grants BES-2016-078499, BES-2017-083126, BES-C-2017-0085, ESP2016-80079-C2-1-R, ESP2016-80079-C2-2-R, FPU16/03827, PDC2021-121059-C22, RTI2018-095076-B-C22, and TIN2015-65316-P ('Computación de Altas Prestaciones VII'), the Juan de la Cierva Incorporación Programme (FJCI-2015-2671 and IJC2019-04862-I for F. Anders), the Severo Ochoa Centre of Excellence Programme (SEV2015-0493), and MICIN/AEI/10.13039/501100011033 (and the European Union through European Regional Development Fund 'A way of making Europe') through grant RTI2018-095076-B-C21, the Institute of Cosmos Sciences University of Barcelona (ICCUB, Unidad de Excelencia 'María de Maeztu') through grant CEX2019-000918-M, the University of Barcelona's official doctoral programme for the development of an R+D+i project through an Ajuts de Personal Investigador en Formació (APIF) grant, the Spanish Virtual Observatory through project AyA2017-84089, the Galician Regional Government, Xunta de Galicia, through grants ED431B-2021/36, ED481A-2019/155, and ED481A-2021/296, the Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC), funded by the Xunta de Galicia and the European Union (European Regional Development Fund – Galicia 2014-2020 Programme), through grant ED431G-2019/01, the Red Española de Supercomputación (RES) computer resources at MareNostrum, the Barcelona Supercomputing Centre - Centro Nacional de Supercomputación (BSC-CNS) through activities AECT-2017-2-0002, AECT-2017-3-0006, AECT-2018-1-0017, AECT-2018-2-0013, AECT-2018-3-0011, AECT-2019-1-0010, AECT-2019-2-0014, AECT-2019-3-0003, AECT-2020-1-0004, and DATA-2020-1-0010, the Departament d'Innovació, Universitats i Empresa de la Generalitat de Catalunya through grant 2014-SGR-1051 for project 'Models de Programació i Entorns d'Execució Parallels' (MPEXPAR), and Ramon y Cajal Fellowship RYC2018-025968-I funded by MICIN/AEI/10.13039/501100011033 and the European Science Foundation ('Investing in your future');
– the Swedish National Space Agency (SNSA/Rymdstyrelsen);
– the Swiss State Secretariat for Education, Research, and Innovation through the Swiss Activités Nationales Complémentaires and the Swiss National Science Foundation through an Eccellenza Professorial Fellowship (award PCEFP2_194638 for R. Anderson);
– the United Kingdom Particle Physics and Astronomy Research Council (PPARC), the United Kingdom Science and Technology Facilities Council (STFC), and the United Kingdom Space Agency (UKSA) through the following grants to the University of Bristol, the University of Cambridge, the University of Edinburgh, the University of Leicester, the Mullard Space Sciences Laboratory of University College London, and the United Kingdom Rutherford Appleton Laboratory (RAL): PP/D006511/1, PP/D006546/1, PP/D006570/1, ST/I000852/1, ST/J005045/1, ST/K00056X/1, ST/K000209/1, ST/K000756/1, ST/L006561/1, ST/N000595/1, ST/N000641/1, ST/N000978/1, ST/N001117/1, ST/S000089/1, ST/S000976/1, ST/S000984/1, ST/S001123/1, ST/S001948/1, ST/S001980/1, ST/S002103/1, ST/V000969/1, ST/W002469/1, ST/W002493/1, ST/W002671/1, ST/W002809/1, and EP/V520342/1.

The GBOT programme uses observations collected at (i) the European Organisation for Astronomical Research in the Southern Hemisphere (ESO) with the VLT Survey Telescope (VST), under ESO programmes 092.B-0165, 093.B-0236, 094.B-0181, 095.B-0046, 096.B-0162, 097.B-0304, 098.B-0030, 099.B-0034, 0100.B-0131, 0101.B-0156, 0102.B-0174, and 0103.B-0165; and (ii) the Liverpool Telescope, which is operated on the island of La Palma by Liverpool John Moores University in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias with financial support from the United Kingdom Science and Technology Facilities Council, and (iii) telescopes of the Las Cumbres Observatory Global Telescope Network.