Gaia Data Release 2

Gaia Data Release 2

Summary of the contents and survey properties

Gaia Collaboration, A. G. A. Brown[1,★], A. Vallenari[2], T. Prusti[3], J. H. J. de Bruijne[3], C. Babusiaux[4,5],
C. A. L. Bailer-Jones[6], M. Biermann[7], D. W. Evans[8], L. Eyer[9], F. Jansen[10], C. Jordi[11], S. A. Klioner[12],
U. Lammers[13], L. Lindegren[14], X. Luri[11], F. Mignard[15], C. Panem[16], D. Pourbaix[17,18], S. Randich[19], P. Sartoretti[4],
H. I. Siddiqui[20], C. Soubiran[21], F. van Leeuwen[8], N. A. Walton[8], F. Arenou[4], U. Bastian[7], M. Cropper[22],
R. Drimmel[23], D. Katz[4], M. G. Lattanzi[23], J. Bakker[13], C. Cacciari[24], J. Castañeda[11], L. Chaoul[16], N. Cheek[25],
F. De Angeli[8], C. Fabricius[11], R. Guerra[13], B. Holl[9], E. Masana[11], R. Messineo[26], N. Mowlavi[9],
K. Nienartowicz[27], P. Panuzzo[4], J. Portell[11], M. Riello[8], G. M. Seabroke[22], P. Tanga[15], F. Thévenin[15],
G. Gracia-Abril[28,7], G. Comoretto[20], M. Garcia-Reinaldos[13], D. Teyssier[20], M. Altmann[7,29], R. Andrae[6],
M. Audard[9], I. Bellas-Velidis[30], K. Benson[22], J. Berthier[31], R. Blomme[32], P. Burgess[8], G. Busso[8], B. Carry[15,31],
A. Cellino[23], G. Clementini[24], M. Clotet[11], O. Creevey[15,33], J. De Ridder[35], L. Delchambre[36],
A. Dell'Oro[19], C. Ducourant[21], J. Fernández-Hernández[37], M. Fouesneau[6], Y. Frémat[32], L. Galluccio[15],
M. García-Torres[38], J. González-Núñez[25,39], J. J. González-Vidal[11], E. Gosset[36,18], L. P. Guy[27,40],
J.-L. Halbwachs[41], N. C. Hambly[34], D. L. Harrison[8,42], J. Hernández[13], D. Hestroffer[31], S. T. Hodgkin[8],
A. Hutton[43], G. Jasniewicz[44], A. Jean-Antoine-Piccolo[16], S. Jordan[7], A. J. Korn[45], A. Krone-Martins[46],
A. C. Lanzafame[47,48], T. Lebzelter[49], W. Löffler[7], M. Manteiga[50,51], P. M. Marrese[52,53], J. M. Martín-Fleitas[43],
A. Moitinho[46], A. Mora[43], K. Muinonen[54,55], J. Osinde[56], E. Pancino[19,53], T. Pauwels[32], J.-M. Petit[57],
A. Recio-Blanco[15], P. J. Richards[58], L. Rimoldini[27], A. C. Robin[57], L. M. Sarro[59], C. Siopis[17], M. Smith[22],
A. Sozzetti[23], M. Süveges[6], J. Torra[11], W. van Reeven[43], U. Abbas[23], A. Abreu Aramburu[60], S. Accart[61],
C. Aerts[35,62], G. Altavilla[52,53,24], M. A. Álvarez[50], R. Alvarez[13], J. Alves[49], R. I. Anderson[63,9],
A. H. Andrei[64,65,29], E. Anglada Varela[37], E. Antiche[11], T. Antoja[3,11], B. Arcay[50], T. L. Astraatmadja[6,66],
N. Bach[43], S. G. Baker[22], L. Balaguer-Núñez[11], P. Balm[20], C. Barache[29], C. Barata[46], D. Barbato[67,23], F. Barblan[9],
P. S. Barklem[45], D. Barrado[68], M. Barros[46], M. A. Barstow[69], S. Bartholomé Muñoz[11], J.-L. Bassilana[61],
U. Becciani[48], M. Bellazzini[24], A. Berihuete[70], S. Bertone[23,29,71], L. Bianchi[72], O. Bienaymé[41],
S. Blanco-Cuaresma[9,21,73], T. Boch[41], C. Boeche[2], A. Bombrun[74], R. Borrachero[11], D. Bossini[2], S. Bouquillon[29],
G. Bourda[21], A. Bragaglia[24], L. Bramante[26], M. A. Breddels[75], A. Bressan[76], N. Brouillet[21], T. Brüsemeister[7],
E. Brugaletta[48], B. Bucciarelli[23], A. Burlacu[16], D. Busonero[23], A. G. Butkevich[12], R. Buzzi[23], E. Caffau[4],
R. Cancelliere[77], G. Cannizzaro[78,62], T. Cantat-Gaudin[2,11], R. Carballo[79], T. Carlucci[29], J. M. Carrasco[11],
L. Casamiquela[11], M. Castellani[52], A. Castro-Ginard[11], P. Charlot[21], L. Chemin[80], A. Chiavassa[15], G. Cocozza[24],
G. Costigan[1], S. Cowell[8], F. Crifo[4], M. Crosta[23], C. Crowley[74], J. Cuypers†[32], C. Dafonte[50], Y. Damerdji[36,81],
A. Dapergolas[30], P. David[31], M. David[82], P. de Laverny[15], F. De Luise[83], R. De March[26], D. de Martino[84],
R. de Souza[85], A. de Torres[74], J. Debosscher[35], E. del Pozo[43], M. Delbo[15], A. Delgado[8], H. E. Delgado[59],
P. Di Matteo[4], S. Diakite[57], C. Diener[8], E. Distefano[48], C. Dolding[22], P. Drazinos[86], J. Durán[56], B. Edvardsson[45],
H. Enke[87], K. Eriksson[45], P. Esquej[88], G. Eynard Bontemps[16], C. Fabre[89], M. Fabrizio[52,53], S. Faigler[90],
A. J. Falcão[91], M. Farràs Casas[11], L. Federici[24], G. Fedorets[54], P. Fernique[41], F. Figueras[11], F. Filippi[26],
K. Findeisen[4], A. Fonti[26], E. Fraile[88], M. Fraser[8,92], B. Frézouls[16], M. Gai[23], S. Galleti[24], D. Garabato[50],
F. García-Sedano[59], A. Garofalo[93,24], N. Garralda[11], A. Gavel[45], P. Gavras[4,30,86], J. Gerssen[87], R. Geyer[12],
P. Giacobbe[23], G. Gilmore[8], S. Girona[94], G. Giuffrida[53,52], F. Glass[9], M. Gomes[46], M. Granvik[54,95],
A. Gueguen[4,96], A. Guerrier[61], J. Guiraud[16], R. Gutiérrez-Sánchez[20], R. Haigron[4], D. Hatzidimitriou[86,30],
M. Hauser[7,6], M. Haywood[4], U. Heiter[45], A. Helmi[75], J. Heu[4], T. Hilger[12], D. Hobbs[14], W. Hofmann[7],
G. Holland[8], H. E. Huckle[22], A. Hypki[1,97], V. Icardi[26], K. Janßen[87], G. Jevardat de Fombelle[27], P. G. Jonker[78,62],
Á. L. Juhász[98,99], F. Julbe[11], A. Karampelas[86,100], A. Kewley[8], J. Klar[87], A. Kochoska[101,102]

★ Corresponding author: A. G. A. Brown, e-mail: brown@strw.leidenuniv.nl

R. Kohley[13], K. Kolenberg[103, 35, 73], M. Kontizas[86], E. Kontizas[30], S. E. Koposov[8, 104], G. Kordopatis[15], Z. Kostrzewa-Rutkowska[78, 62], P. Koubsky[105], S. Lambert[29], A. F. Lanza[48], Y. Lasne[61], J.-B. Lavigne[61], Y. Le Fustec[106], C. Le Poncin-Lafitte[29], Y. Lebreton[4, 107], S. Leccia[84], N. Leclerc[4], I. Lecoeur-Taibi[27], H. Lenhardt[7], F. Leroux[61], S. Liao[23, 108, 109], E. Licata[72], H. E. P. Lindstrøm[110, 111], T. A. Lister[112], E. Livanou[86], A. Lobel[32], M. López[68], S. Managau[61], R. G. Mann[34], G. Mantelet[7], O. Marchal[4], J. M. Marchant[113], M. Marconi[84], S. Marinoni[52, 53], G. Marschalkó[98, 114], D. J. Marshall[115], M. Martino[26], G. Marton[98], N. Mary[61], D. Massari[75], G. Matijevič[87], T. Mazeh[90], P. J. McMillan[14], S. Messina[48], D. Michalik[14], N. R. Millar[8], D. Molina[11], R. Molinaro[84], L. Molnár[98], P. Montegriffo[24], R. Mor[11], R. Morbidelli[23], T. Morel[36], D. Morris[34], A. F. Mulone[26], T. Muraveva[24], I. Musella[84], G. Nelemans[62, 35], L. Nicastro[24], L. Noval[6], W. O'Mullane[13, 40], C. Ordénovic[15], D. Ordóñez-Blanco[27], P. Osborne[8], C. Pagani[69], I. Pagano[48], F. Pailler[16], H. Palacin[61], L. Palaversa[8, 9], A. Panahi[90], M. Pawlak[116, 117], A. M. Piersimoni[83], F.-X. Pineau[41], E. Plachy[98], G. Plum[4], E. Poggio[67, 23], E. Poujoulet[118], A. Prša[102], L. Pulone[52], E. Racero[25], S. Ragaini[24], N. Rambaux[31], M. Ramos-Lerate[119], S. Regibo[35], C. Reylé[57], F. Riclet[16], V. Ripepi[84], A. Riva[23], A. Rivard[61], G. Rixon[8], T. Roegiers[120], M. Roelens[9], M. Romero-Gómez[11], N. Rowell[34], F. Royer[4], L. Ruiz-Dern[4], G. Sadowski[17], T. Sagristà Sellés[7], J. Sahlmann[13, 121], J. Salgado[122], E. Salguero[37], N. Sanna[19], T. Santana-Ros[97], M. Sarasso[23], H. Savietto[123], M. Schultheis[15], E. Sciacca[48], M. Segol[124], J. C. Segovia[25], D. Ségransan[9], I-C. Shih[4], L. Siltala[54, 125], A. F. Silva[46], R. L. Smart[23], K. W. Smith[6], E. Solano[68, 126], F. Solitro[26], R. Sordo[2], S. Soria Nieto[11], J. Souchay[29], A. Spagna[23], F. Spoto[15, 31], U. Stampa[7], I. A. Steele[113], H. Steidelmüller[12], C. A. Stephenson[20], H. Stoev[127], F. F. Suess[8], J. Surdej[36], L. Szabados[98], E. Szegedi-Elek[98], D. Tapiador[128, 129], F. Taris[29], G. Tauran[61], M. B. Taylor[130], R. Teixeira[85], D. Terrett[58], P. Teyssandier[29], W. Thuillot[31], A. Titarenko[15], F. Torra Clotet[131], C. Turon[4], A. Ulla[132], E. Utrilla[43], S. Uzzi[26], M. Vaillant[61], G. Valentini[83], V. Valette[16], A. van Elteren[1], E. Van Hemelryck[32], M. van Leeuwen[8], M. Vaschetto[26], A. Vecchiato[23], J. Veljanoski[75], Y. Viala[4], D. Vicente[94], S. Vogt[120], C. von Essen[133], H. Voss[11], V. Votruba[105], S. Voutsinas[34], G. Walmsley[16], M. Weiler[11], O. Wertz[134], T. Wevers[8, 62], Ł. Wyrzykowski[8, 116], A. Yoldas[8], M. Žerjal[101, 135], H. Ziaeepour[57], J. Zorec[136], S. Zschocke[12], S. Zucker[137], C. Zurbach[44], T. Zwitter[101]

(Affiliations can be found after the references)

ABSTRACT

Context. We present the second *Gaia* data release, *Gaia* DR2, consisting of astrometry, photometry, radial velocities, and information on astrophysical parameters and variability, for sources brighter than magnitude 21. In addition epoch astrometry and photometry are provided for a modest sample of minor planets in the solar system.

Aims. A summary of the contents of *Gaia* DR2 is presented, accompanied by a discussion on the differences with respect to *Gaia* DR1 and an overview of the main limitations which are still present in the survey. Recommendations are made on the responsible use of *Gaia* DR2 results.

Methods. The raw data collected with the *Gaia* instruments during the first 22 months of the mission have been processed by the *Gaia* Data Processing and Analysis Consortium (DPAC) and turned into this second data release, which represents a major advance with respect to *Gaia* DR1 in terms of completeness, performance, and richness of the data products.

Results. *Gaia* DR2 contains celestial positions and the apparent brightness in G for approximately 1.7 billion sources. For 1.3 billion of those sources, parallaxes and proper motions are in addition available. The sample of sources for which variability information is provided is expanded to 0.5 million stars. This data release contains four new elements: broad-band colour information in the form of the apparent brightness in the G_{BP} (330–680 nm) and G_{RP} (630–1050 nm) bands is available for 1.4 billion sources; median radial velocities for some 7 million sources are presented; for between 77 and 161 million sources estimates are provided of the stellar effective temperature, extinction, reddening, and radius and luminosity; and for a pre-selected list of 14 000 minor planets in the solar system epoch astrometry and photometry are presented. Finally, *Gaia* DR2 also represents a new materialisation of the celestial reference frame in the optical, the *Gaia*-CRF2, which is the first optical reference frame based solely on extragalactic sources. There are notable changes in the photometric system and the catalogue source list with respect to *Gaia* DR1, and we stress the need to consider the two data releases as independent.

Conclusions. *Gaia* DR2 represents a major achievement for the *Gaia* mission, delivering on the long standing promise to provide parallaxes and proper motions for over 1 billion stars, and representing a first step in the availability of complementary radial velocity and source astrophysical information for a sample of stars in the *Gaia* survey which covers a very substantial fraction of the volume of our galaxy.

Key words. catalogs – astrometry – techniques: radial velocities – stars: fundamental parameters – stars: variables: general – minor planets, asteroids: general

1. Introduction

We present the second intermediate *Gaia* data release (*Gaia* Data Release 2, *Gaia* DR2), which is based on the data collected during the first 22 months of the nominal mission lifetime (scientific data collection started in July 2014 and nominally lasts 60 months, see Gaia Collaboration 2016b). *Gaia* DR2 represents the planned major advance with respect to the first intermediate *Gaia* data release (*Gaia* DR1, Gaia Collaboration 2016a), making the leap to a high-precision parallax and proper motion catalogue for over 1 billion sources, supplemented by precise and homogeneous multi-band all-sky photometry and a large radial velocity survey at the bright ($G \lesssim 13$) end. The availability of precise fundamental astrophysical information required to map and understand the Milky Way is thus expanded to a very substantial fraction of the volume of our galaxy, well beyond the immediate solar neighbourhood. The data diversity of *Gaia* DR2 is also significantly enhanced with respect to *Gaia* DR1 through the availability of astrophysical parameters for a large sample of stars, the significant increase in the number and types of variable stars and their light curves, and the addition for the first time of solar system astrometry and photometry. This paper is structured as follows. In Sect. 2 we provide a short overview of the improvements and additions to the data processing that led to the production of *Gaia* DR2. We summarise the contents of the second data release in Sect. 3 and illustrate the quality of this release through all-sky maps of source counts and colours in Sect. 4. In Sect. 5 we discuss the major differences between *Gaia* DR2 and *Gaia* DR1, in particular pointing out the evolution of the source list and the need to always qualify *Gaia* source identifiers with the data release they refer to. The two releases should be treated as entirely independent catalogues. The known limitations of the second *Gaia* data release are presented in Sect. 6 and additional guidance on the use of the data is provided in Sect. 7. In Sect. 8 we provide updates to the *Gaia* data access facilities and documentation available to the astronomical community. We conclude with a look ahead at the next release in Sect. 9. Throughout the paper we make reference to other DPAC papers that provide more details on the data processing and validation for *Gaia* DR2. All these papers (together with the present article) can be found in the Astronomy & Astrophysics Special edition on *Gaia* DR2.

2. Data processing for *Gaia* DR2

To provide the context for the description of the data release contents in the next section, we provide here a summary of the input measurements used and the main additions and improvements implemented in the data processing for *Gaia* DR2. We recall that *Gaia* measurements are collected with three instruments. The astrometric instrument collects images in *Gaia*'s white-light *G*-band (330–1050 nm); the Blue (BP) and Red (RP) prism photometers collect low resolution spectrophotometric measurements of source spectral energy distributions over the wavelength ranges 330–680 nm and 630–1050 nm, respectively; and the radial velocity spectrometer (RVS) collects medium resolution ($R \sim 11\,700$) spectra over the wavelength range 845–872 nm centred on the Calcium triplet region. For more details on the *Gaia* instruments and measurements we refer to Gaia Collaboration (2016b). The RVS, from which results are presented in *Gaia* DR2 for the first time, is described in detail in Cropper et al. (2018). An important part of the pre-processing for all *Gaia* instruments is to remove the effect of non-uniformity of

the CCD bias levels, which is essential for achieving the ultimate image location and radial velocity determination performance. The details of this process are described in Hambly et al. (2018).

The timing of events on board *Gaia*, including the data collection, is given in terms of the on board mission timeline (OBMT) which is generated by the *Gaia* on board clock. By convention OBMT is expressed in units of 6 h (21 600 s) spacecraft revolutions (Gaia Collaboration 2016b). The approximate relation between OBMT (in revolutions) and the barycentric coordinate time (TCB, in Julian years) at *Gaia* is

$$\text{TCB} \simeq \text{J2015.0} + (\text{OBMT} - 1717.6256\ \text{rev})/(1461\ \text{rev yr}^{-1}). \tag{1}$$

The 22 month time interval covered by the observations used for *Gaia* DR2 starts at OBMT 1078.3795 rev = J2014.5624599 TCB (approximately 2014 July 25, 10:30:00 UTC), and ends at OBMT 3750.5602 rev = J2016.3914678 TCB (approximately 2016 May 23, 11:35:00 UTC). As discussed in Gaia Collaboration (2016a) this time interval contains gaps caused by both spacecraft events and by on-ground data processing problems. This leads to gaps in the data collection or stretches of time over which the input data cannot be used. Which data are considered unusable varies across the *Gaia* data processing systems (astrometry, photometry, etc) and as a consequence the effective amount of input data used differs from one system to the other. We refer to the specific data processing papers (listed below) for the details.

A broad overview of the data processing for *Gaia* is given in Gaia Collaboration (2016b) while the simplified processing for *Gaia* DR1 is summarised in Gaia Collaboration (2016a), in particular in their Fig. 10. With respect to *Gaia* DR1 the following major improvements were implemented in the astrometric processing (for details, see Lindegren et al. 2018):

– Creation of the source list: this process (also known as cross-matching; Fabricius et al. 2016) provides the link between the individual *Gaia* detections and the entries ("sources") in the *Gaia* working catalogue. For *Gaia* DR1 the detections were matched to the nearest source, using a match radius of 1.5 arcsec, and new sources were created when no match was found. Spurious detections and limitations of the initial source list resulted in many spurious sources but also the loss in *Gaia* DR1 of many real sources, including high proper motion stars. For *Gaia* DR2 the source list was created essentially from scratch, based directly on the detections and using a cluster analysis algorithm that takes into account a possible linear motion of the source. The source list for *Gaia* DR2 is therefore much cleaner and of higher angular resolution (Sect. 5.3), resulting in improved astrometry.

– Attitude modelling: in the astrometric solution, the pointing of the instrument is modelled as a function of time using splines. However, these cannot represent rapid variations caused by the active attitude control, micro-clanks (microscopic structural changes in the spacecraft), and micrometeoroid hits. In *Gaia* DR1 the accuracy of the attitude determination was limited by such effects. For *Gaia* DR2 the rapid variations are determined and subtracted by a dedicated process, using rate measurements from successive CCD observations of bright sources.

– Calibration modelling: optical aberrations in the telescopes and the wavelength-dependent diffraction create colour-dependent shifts of the stellar images (chromaticity). This will eventually be handled in the pre-processing of the raw data, by fitting colour-dependent PSFs or LSFs to the CCD samples. This procedure will only be in place for the next

release, and the effect was completely ignored for *Gaia* DR1. In the current astrometric solution chromaticity is handled by the introduction of colour-dependent terms in the geometric calibration model.

- Global modelling: the basic-angle variations are more accurately modelled thanks to an improved processing of the onboard measurements (using the Basic Angle Monitor) and the introduction of global corrections to these measurements as additional unknowns in the astrometric solution. This has been especially important for reducing large-scale systematics in the parallaxes.
- Celestial reference frame: establishing a link to the extragalactic reference frame was complicated and indirect in *Gaia* DR1, which relied on the Hipparcos and *Tycho*-2 catalogues for the determination of proper motions. By contrast, *Gaia* DR2 contains the positions and proper motions for about half a million identified quasars, which directly define a very accurate celestial reference frame (*Gaia*-CRF2), as described in Gaia Collaboration (2018e).

The various improvements in the astrometric models have reduced the RMS residual of typical observations of bright stars ($G \lesssim 13$) from about 0.67 mas in *Gaia* DR1 to 0.2–0.3 mas in *Gaia* DR2.

Additional improvements in the data processing for *Gaia* DR2 as well as the introduction of new elements facilitated the much expanded variety of data published in this second release. Although the photometric processing pipeline did treat the data from *Gaia*'s BP and RP photometers from the start of the mission operations, it was decided not to publish the results in *Gaia* DR1 (Evans et al. 2017) because of the still preliminary nature of the calibrations of these instruments. The processing for *Gaia* DR2 features enhancements in the photometric calibrations, including of the BP and RP prism spectra. The integrated light from these spectra is published in this release as the fluxes in the G_{BP} and G_{RP} passbands. In addition the photometric passbands for G, G_{BP}, and G_{RP} are published, both the versions used in the data processing and the revised versions (based on a deeper analysis involving the BP/RP spectra of standard stars). The photometric data processing and results validation for *Gaia* DR2 are described in Evans et al. (2018) and Riello et al. (2018).

The processing of RVS data was also in place from the start of mission operations but during the operations up to *Gaia* DR1 the adaptations necessary to the RVS pipeline to deal with the effects of the excess stray light on board *Gaia* prevented the publication of results. Hence *Gaia* DR2 features the first RVS results in the form of median radial velocities. The details of the RVS data processing and results validation are provided in Sartoretti et al. (2018), Katz et al. (2018), and Soubiran et al. (2018).

Epoch astrometry was determined for a list of 14 000 pre-selected small solar system bodies (henceforth referred to as Solar System Objects or SSOs). The data processing and validation for the *Gaia* DR2 SSO data are described in Gaia Collaboration (2018f).

Astrophysical parameters (T_{eff}, A_G, $E(G_{BP} - G_{RP})$, radius and luminosity) were determined for between 77 and 161 million stars from the *Gaia* broad-band photometry and parallaxes alone (no non-*Gaia* data was used). The details of the astrophysical parameter estimation and the validation of the results are described in Andrae et al. (2018).

Practically all sources present in *Gaia* DR2 were analysed for apparent brightness variations, resulting in a catalogue of about 0.5 million stars securely identified as variables and for which light curves and statistical information on the photometric time

Table 1. Number of sources of a given type or the number for which a given data product is available in *Gaia* DR2.

Data product or source type	Number of sources
Total	1 692 919 135
5-parameter astrometry	1 331 909 727
2-parameter astrometry	361 009 408
ICRF3 prototype sources	2820
Gaia-CRF2 sources	556 869
G-band	1 692 919 135
G_{BP}-band	1 381 964 755
G_{RP}-band	1 383 551 713
Radial velocity	7 224 631
Classified as variable	550 737
Variable type estimated	363 969
Detailed characterisation of light curve	390 529
Effective temperature T_{eff}	161 497 595
Extinction A_G	87 733 672
Colour excess $E(G_{BP} - G_{RP})$	87 733 672
Radius	76 956 778
Luminosity	76 956 778
SSO epoch astrometry and photometry	14 099

series are provided. The variability processing is described in Holl et al. (2018).

Finally, an overall validation of the *Gaia* DR2 catalogue is described in Arenou et al. (2018), which, as outlined in Gaia Collaboration (2016b), involves an extensive scientific validation of the combined data presented in this data release.
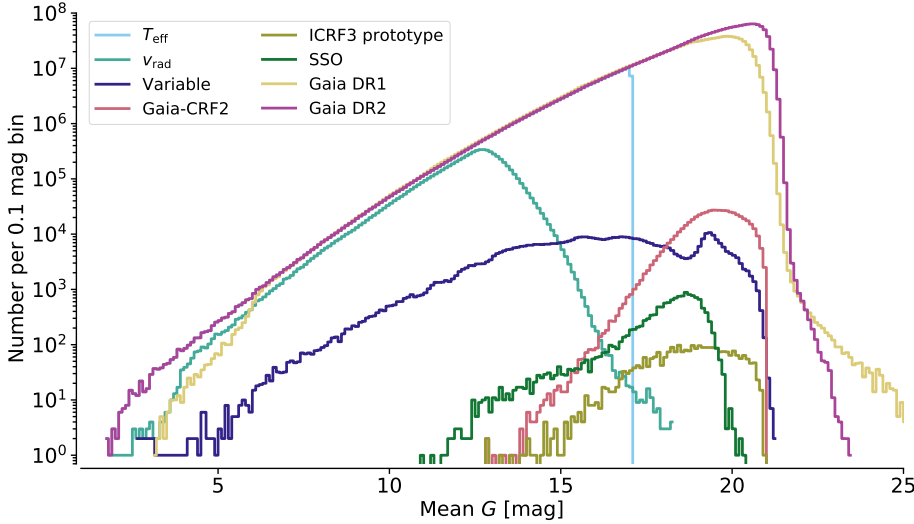
A number of important shortcomings remain in the data processing, leading to limitations in *Gaia* DR2 which require taking some care when using the data. In Sect. 6 we summarise the known limitations of the present *Gaia* data release and point out, where relevant, the causes. Section 7 provides additional guidance on the use of *Gaia* DR2 results. The reader is strongly encouraged to read the papers listed above and the online documentation[1] to understand the limitations in detail.

3. Overview of the contents of *Gaia* DR2

Gaia DR2 contains astrometry, broad-band photometry, radial velocities, variable star classifications as well as the characterisation of the corresponding light curves, and astrophysical parameter estimates for a total of 1 692 919 135 sources. In addition the epoch astrometry and photometry for 14 099 solar system objects are listed. Basic statistics on the source numbers and the overall distribution in G can be found in Table 1 and Table 2, where it should noted that 4 per cent of the sources are fainter than $G = 21$. The overall quality of *Gaia* DR2 results in terms of the typically achieved uncertainties is summarised in Table 3. The contents of the main components of the release, of which the magnitude distributions are shown in Figs. 1 and 2, are summarised in the following paragraphs. We defer the discussion on the known limitations of *Gaia* DR2 to Sect. 6.

---

[1] http://gea.esac.esa.int/archive/documentation/GDR2/index.html

Fig. 1. Distribution of the mean values of G for all *Gaia* DR2 sources shown as histograms with 0.1 mag wide bins. The distribution of the *Gaia* DR1 sources is included for comparison and illustrates the improved photometry at the faint end and the improved completeness at the bright end. The other histograms are for the main *Gaia* DR2 components as indicated in the legend. See text for further explanations on the characteristics of the histograms.

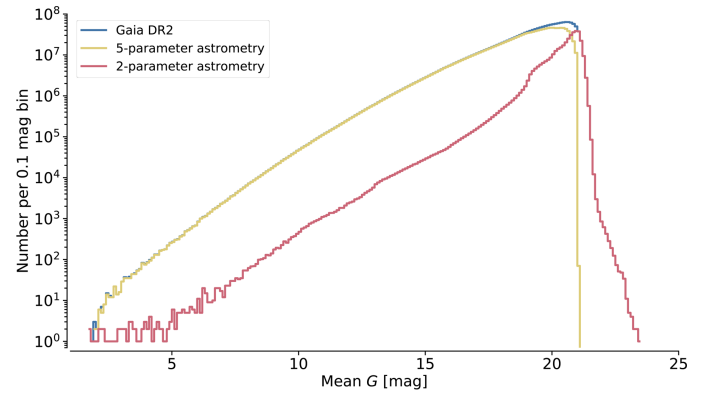**Table 2.** Distribution of the *Gaia* DR2 sources in G-band magnitude.

Magnitude distribution percentiles (G)			
Percentile	All	5-parameter	2-parameter
0.135%	11.6	11.4	15.3
2.275%	15.0	14.7	18.5
15.866%	17.8	17.4	19.8
50%	19.6	19.3	20.6
84.134%	20.6	20.3	21.0
97.725%	21.1	20.8	21.2
99.865%	21.3	20.9	21.4

Notes. The distribution percentiles are shown for all sources and for those with a 5-parameter and 2-parameter astrometric solution, respectively.

3.1. Astrometric data set

The astrometric data set consists of two subsets: for 1 331 909 727 sources the full five-parameter astrometric solution is provided ("5-parameter" in Table 1), hence including celestial position, parallax, and proper motion. For the remaining 361 009 408 sources ("2-parameter" in Table 1) only the celestial positions (α, δ) are reported. Figure 2 shows the distribution in G for the 5-parameter and 2-parameter sources compared to the overall magnitude distribution. The 2-parameter sources are typically faint (with about half those sources at $G > 20.6$, see Table 2), have very few observations, or very poorly fit the five-parameter astrometric model. All sources fainter than $G = 21$ have only positions in *Gaia* DR2. We refer to Lindegren et al. (2018) for the detailed criteria used during the data processing to decide which type of solution should be adopted.

For a 2-parameter source the position was computed using a special fall-back solution. Rather than ignoring the parallax and proper motion of the source (i.e. assuming that they are strictly zero), the fall-back solution estimates all five parameters but applies a prior that effectively constrains the parallax and proper motion to realistically small values, depending on the magnitude and Galactic coordinates of the source (Michalik et al. 2015b). The resulting position is usually more precise, and its uncertainty more realistic (larger), than if only the position had been solved for. The parallax and proper motion of the fall-back solution may however be strongly biased, which is why they are not published.

Fig. 2. Distribution of the mean values of G for the sources with a full astrometric solution in *Gaia* DR2 ("5-parameter") and for the sources for which only the celestial position is listed ("2-parameter") compared to the overall magnitude distribution for *Gaia* DR2.

The reference epoch for all (5- and 2-parameter) sources is J2015.5 (TCB). This epoch, close to the mid-time of the observations included in *Gaia* DR2, was chosen to minimise correlations between the position and proper motion parameters. This epoch is 0.5 yr later than the reference epoch for *Gaia* DR1, which must be taken into account when comparing the positions between the two releases.

As for *Gaia* DR1 all sources were treated as single stars when solving for the astrometric parameters. For a binary the parameters may thus refer to either component, or to the photocentre of the system, and the proper motion represents the mean motion of the component, or photocentre, over the 1.75 yr of data included in the solution. Depending on the orbital motion, this could be significantly different from the proper motion of the same object in *Gaia* DR1 (see Sect. 5).

The positions and proper motions are given in the second realisation of the *Gaia* celestial reference frame (*Gaia*-CRF2) which at the faint end ($G \sim 19$) is aligned with the International Celestial Reference Frame (ICRF) to about 0.02 mas RMS at epoch J2015.5 (TCB), and non-rotating with respect to the ICRF to within 0.02 mas yr^{-1} RMS. At the bright end ($G < 12$) the alignment can only be confirmed to be better than 0.3 mas while the bright reference frame is non-rotating to within 0.15 mas yr^{-1}. For details we refer to Lindegren et al. (2018). The *Gaia*-CRF2 is materialised by 556 869 QSOs and aligned

to the forthcoming version 3 of the ICRF through a subset of 2820 QSOs. It represents the first ever optical reference frame constructed on the basis of extragalactic sources only. The construction and properties of the *Gaia*-CRF2 as well as the comparison to the ICRF3 prototype are described in Gaia Collaboration (2018e).

3.2. Photometric data set

The photometric data set contains the broad band photometry in the G, G_{BP}, and G_{RP} bands, thus providing the major new element of colour information for *Gaia* DR2 sources. The mean value of the G-band fluxes is reported for all sources while for about 80 per cent of the sources the mean values of the G_{BP} and G_{RP} fluxes are provided (for a small fraction of these sources only the G_{RP} value is reported). The photometric data processing considered three types of sources, "Gold", "Silver", and "Bronze", which represent decreasing quality levels of the photometric calibration achieved, where in the case of the Bronze sources no colour information is available. The photometric nature of each source is indicated in the released catalogue by a numeric field (`phot_proc_mode`) assuming values 0, 1 and 2 for gold, silver, and bronze sources respectively. At the bright end the photometric uncertainties are dominated by calibration effects which are estimated to contribute 2, 5, and 3 mmag RMS per CCD observation, respectively for G, G_{BP}, and G_{RP} (Evans et al. 2018). For details on the photometric processing and the validation of the results we refer to Riello et al. (2018) and Evans et al. (2018).

The broad-band colour information suffers from strong systematic effects at the faint end of the survey ($G \gtrsim 19$), in crowded regions, and near bright stars. In these cases the photometric measurements from the blue and red photometers suffer from an insufficiently accurate background estimation and from the lack of specific treatment of the prism spectra in crowded regions, where the overlapping of images of nearby sources is not yet accounted for. This leads to measured fluxes that are inconsistent between the G and the G_{BP} and G_{RP} bands in the sense that the sum of the flux values in the latter two bands may be significantly larger than that in G (whereas it is expected that for normal spectral energy distributions the sum of fluxes in G_{BP} and G_{RP} should be comparable to that in G). A quantitative indication of this effect is included in *Gaia* DR2 in the form of the "flux excess factor" (the `phot_bp_rp_excess_factor` field in the data archive).

The distribution of the astrometric and photometric data sets in G is shown in purple in Fig. 1, where for comparison the distribution for *Gaia* DR1 is also shown in yellow. Note the improved completeness at the bright end of the survey and the improved photometry (less extremely faint sources) and completeness at the faint end. The distribution of the *Gaia*-CRF2 sources (pink-red line) shows a sharp drop at $G = 21$ which is because only QSOs at $G < 21$ were used for the construction of the reference frame.

3.3. Radial velocity data set

The radial velocity data set contains the median radial velocities, averaged over the 22 month time span of the observations, for 7 224 631 sources which are nominally brighter than 12th magnitude in the G_{RVS} photometric band. For the selection of sources to process, the provisional G_{RVS} magnitude as listed in the Initial Gaia Source List (Smart & Nicastro 2014) was used. The actual magnitudes in the G_{RVS} band differ from these

provisional values, meaning that the magnitude limit in G_{RVS} is not sharply defined. In practice the sources for which a median radial velocity is listed mostly have magnitudes brighter than 13 in G (see light green line in Fig. 1). The signal to noise ratio of the RVS spectra depends primarily on G_{RVS}, which is not listed in *Gaia* DR2. It was decided not to publish the G_{RVS} magnitude in *Gaia* DR2 because the processing of RVS data was focused on the production of the radial velocities, and the calibrations necessary for the estimation of the flux in the RVS passband (background light corrections and the knowledge of the PSF in the direction perpendicular to *Gaia*'s scanning direction) were only preliminary. As a result the G_{RVS} magnitudes were of insufficient quality for publication in *Gaia* DR2 (Sartoretti et al. 2018). The value of G_{RVS} as determined during the data processing was however used to filter out stars considered too faint ($G_{RVS} > 14$) for inclusion in the radial velocity data set. For convenience we provide here a relation which allows to predict the value of G_{RVS} from the $(G - G_{RP})$ colour.

$$G_{RVS} - G_{RP} = 0.042319 - 0.65124(G - G_{RP}) + 1.0215(G - G_{RP})^2$$
$$- 1.3947(G - G_{RP})^3 + 0.53768(G - G_{RP})^4$$

to within 0.086 mag RMS for $0.1 < (G - G_{RP}) < 1.4$, (2)

and

$$G_{RVS} - G_{RP} = 132.32 - 377.28(G - G_{RP}) + 402.32(G - G_{RP})^2$$
$$- 190.97(G - G_{RP})^3 + 34.026(G - G_{RP})^4$$

to within 0.088 mag RMS for $1.4 \leq (G - G_{RP}) < 1.7$. (3)

This relation was derived from a sample of stars for which the flux in the RVS band could be determined to a precision of 0.1 mag or better.

Radial velocities are only reported for stars with effective temperatures in the range 3550–6900 K (where these temperatures refer to the spectral template used in the processing, not to the T_{eff} values reported as part of the astrophysical parameter data set). The uncertainties of the radial velocities are summarised in Table 3. At the faint end the uncertainties show a dependency on stellar effective temperature, where the values are approximately 1.4 km s^{-1} and 3.6 km s^{-1} at $G_{RVS} = 11.75$ for stars with $T_{eff} \sim 5500$ K and $T_{eff} \sim 6500$ K, respectively. The distribution over G of the sources with radial velocities shown in Fig. 1 in light green reflects the fact that over the range $4 < G < 12$ the completeness of the radial velocity data set with respect to the *Gaia* DR2 data set varies from 60 to 80 per cent (Katz et al. 2018). At the faint end ($G > 13$) the shape of the distribution is determined by the selection of stars for which radial velocities were derived (using the provisional value of G_{RVS}) and the large differences between G and G_{RVS} that can occur depending on the effective temperature of the stars. For the details on the radial velocity data processing and the properties and validation of the resulting radial velocity catalogue we refer to Sartoretti et al. (2018) and Katz et al. (2018). The set of standard stars that was used to define the zeropoint of the RVS radial velocities is described in Soubiran et al. (2018).

3.4. Variability data set

The variability data set consists of 550 737 sources that are securely identified as variable (based on at least two transits of the sources across the fields of view of the two *Gaia* telescopes) and for which the photometric time series and corresponding statistics are provided. This number still represents

Table 3. Basic performance statistics for *Gaia* DR2.

Data product or source type	Typical uncertainty
Five-parameter astrometry (position & parallax)	0.02–0.04 mas at $G < 15$
	0.1 mas at $G = 17$
	0.7 mas at $G = 20$
	2 mas at $G = 21$
Five-parameter astrometry (proper motion)	0.07 mas yr^{-1} at $G < 15$
	0.2 mas yr^{-1} at $G = 17$
	1.2 mas yr^{-1} at $G = 20$
	3 mas yr^{-1} at $G = 21$
Two-parameter astrometry (position only)	1–4 mas
Systematic astrometric errors (averaged over the sky)	<0.1 mas
Gaia-CRF2 alignment with ICRF	0.02 mas at $G = 19$
Gaia-CRF2 rotation with respect to ICRF	<0.02 mas yr^{-1} at $G = 19$
Gaia-CRF2 alignment with ICRF	0.3 mas at $G < 12$
Gaia-CRF2 rotation with respect to ICRF	<0.15 mas yr^{-1} at $G < 12$
Mean *G*-band photometry	0.3 mmag at $G < 13$
	2 mmag at $G = 17$
	10 mmag at $G = 20$
Mean $G_{\rm BP}$- and $G_{\rm RP}$-band photometry	2 mmag at $G < 13$
	10 mmag at $G = 17$
	200 mmag at $G = 20$
Median radial velocity over 22 months	0.3 km s^{-1} at $G_{\rm RVS} < 8$
	0.6 km s^{-1} at $G_{\rm RVS} = 10$
	1.8 km s^{-1} at $G_{\rm RVS} = 11.75$
Systematic radial velocity errors	<0.1 km s^{-1} at $G_{\rm RVS} < 9$
	0.5 km s^{-1} at $G_{\rm RVS} = 11.75$
Effective temperature $T_{\rm eff}$	324 K
Extinction A_G	0.46 mag
Colour excess $E(G_{\rm BP} - G_{\rm RP})$	0.23 mag
Radius	10%
Luminosity	15%
Solar system object epoch astrometry	1 mas (in scan direction)

Notes. The astrometric uncertainties as well as the Gaia-CRF2 alignment and rotation limits refer to epoch J2015.5 TCB. The uncertainties on the photometry refer to the mean magnitudes listed in the main Gaia DR2 catalogue.
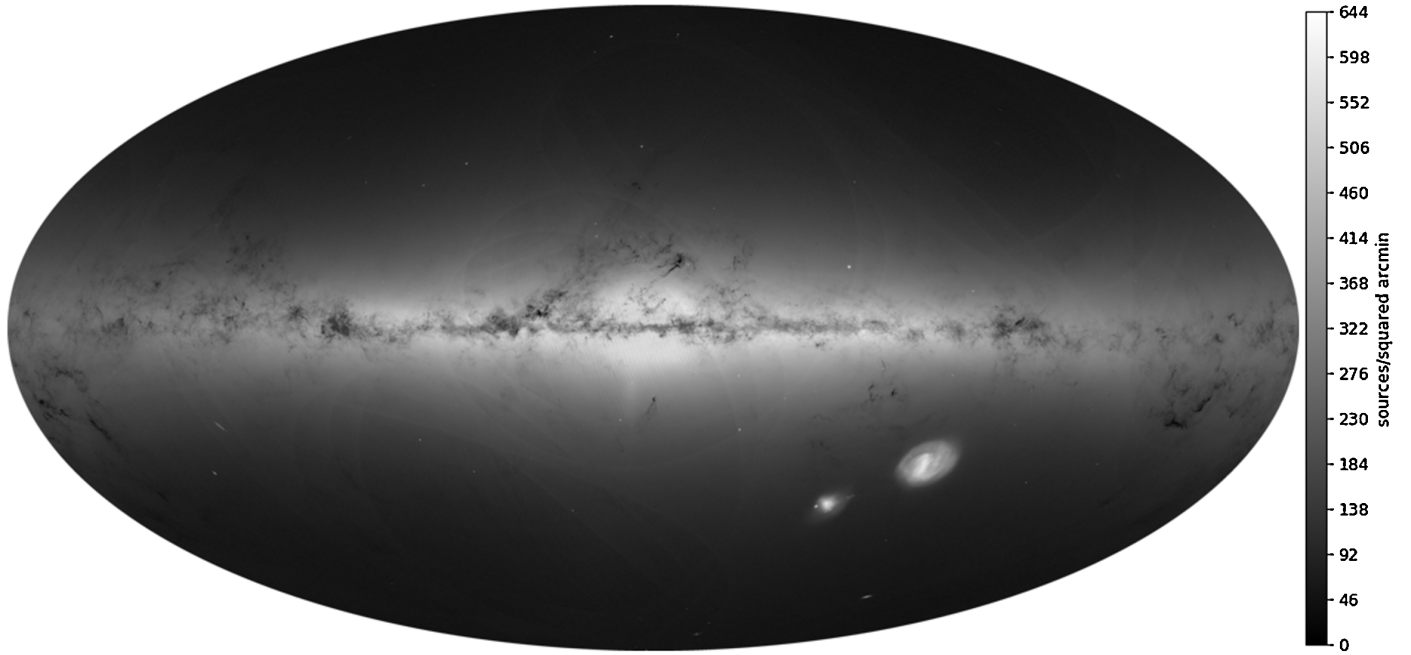
only a small subset of the total amount of variables expected in the *Gaia* survey and subsequent data releases will contain increasing numbers of variable sources. Of the sources identified as variable 363 969 were classified into one of nine variable types by a supervised light curve classifier. The types listed in the *Gaia* DR2 are: RR Lyrae (anomalous RRd, RRd, RRab, RRc); long period variables (Mira type and Semi-Regulars); Cepheids (anomalous Cepheids, classical Cepheids, type-II Cepheids); δ Scuti and SX Phoenicis stars. A second subset of 390 529 variable stars (largely overlapping with the variability type subset) was analysed in detail when at least 12 points were available for the light curve. These so-called "specific object studies" (SOS) were carried out for variables of the type Cepheid and RR Lyrae, long period variables, short time scale variables (with brightness variations on time scales of one day or less), and rotational modulation variables.

Figure 1 shows in dark blue the distribution over G of the sources identified as variable. The mean G value as determined in the photometric data processing (used in Fig. 1) may differ from the mean magnitude determined from the photometric time series where the variable nature of the source is properly

accounted for. Hence the distribution in Fig. 1 should be taken as illustrative only. For full details on the variable star processing and results validation we refer to Holl et al. (2018) and references therein.

3.5. Astrophysical parameter data set

The astrophysical parameter data set consists of estimated values of $T_{\rm eff}$, extinction A_G and reddening $E(G_{\rm BP} - G_{\rm RP})$ (both derived from the apparent dimming and reddening of a source), radius, and luminosity for stars brighter than $G = 17$. Table 1 contains the source counts for each of these astrophysical parameters. The magnitude distribution shown in Fig. 1 in cyan concerns all sources for which $T_{\rm eff}$ was estimated and indicates that this parameter is available for practically all sources at $G < 17$. Values of $T_{\rm eff}$ are only reported over the range 3000–10 000 K, which reflects the limits of the training data for the algorithm used to estimate $T_{\rm eff}$. Estimates of the other astrophysical parameters are published for about 50% of the sources for which $T_{\rm eff}$ is published. This is caused by the filtering of the pipeline results to remove parameter estimates for which the input data are too

Fig. 3. Sky distribution of all *Gaia* DR2 sources in Galactic coordinates. This image and the one in Fig. 4 are Hammer projections of the full sky. This projection was chosen in order to have the same area per pixel (not strictly true because of pixel discretisation). Each pixel is ~5.9 square arcmin. The colour scale is logarithmic and represents the number of sources per square arcmin.

poor or for which the assumptions made lead to invalid results. The details of the astrophysical parameter processing and the validation of the results are described in Andrae et al. (2018).

3.6. Solar system objects data set

The solar system objects data set features epoch astrometry and photometry for a pre-selected list of 14 099 known minor bodies in the solar system, primarily main belt asteroids. Epoch astrometry refers to the fact that the measured celestial position for a given SSO is listed for each instance in time when it passed across the field of view of one of *Gaia*'s telescopes. The celestial positions at each epoch are given as seen from *Gaia*. These measurements can be used to determine orbits for the SSOs and the results thereof are described in Gaia Collaboration (2018f). For details on the processing of SSOs we refer to the same paper.Over the apparent magnitude range $G \sim 12$–17 the typical focal plane transit level of uncertainty achieved for the instantaneous SSO celestial positions is 1 mas in the *Gaia* scanning direction. Figure 1 shows in dark green the magnitude distribution for the SSOs, where it should be noted that the magnitudes as can be measured by *Gaia* represent instantaneous measurements taken far from opposition. Hence the magnitude histogram is to be taken as illustrative only.

4. Scientific performance and potential of *Gaia* DR2

Gaia DR2 is accompanied by six papers that provide basic demonstrations of the scientific quality of the results included in this release. The topics treated by the papers are:

- the reference frame *Gaia*-CRF2 (Gaia Collaboration 2018e);
- orbital fitting of the epoch astrometry for solar system objects (Gaia Collaboration 2018f);
- variable stars as seen in the *Gaia* DR2 colour-magnitude diagram (Gaia Collaboration 2018b), where the motion of variables in colour-magnitude space is explored;
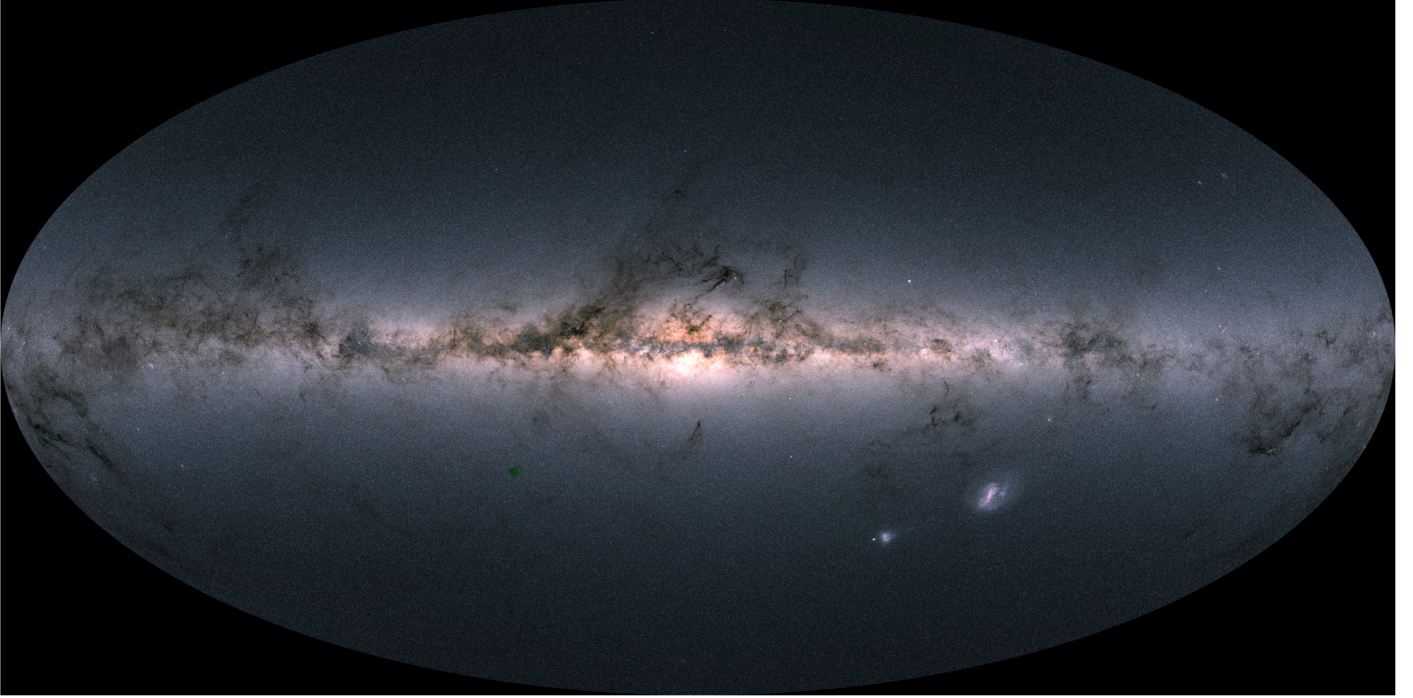
- the kinematics of the Milky Way disk (Gaia Collaboration 2018d), illustrating in particular the power of having radial velocities available in *Gaia* DR2;
- the kinematics of globular clusters, the LMC and SMC, and other dwarf galaxies around the Milky Way (Gaia Collaboration 2018c), showcasing the power of *Gaia* DR2 to study distant samples of stars;
- the observational Hertzsprung-Russell diagram is explored in Gaia Collaboration (2018a).

We strongly encourage the reader to consult these papers for a full impression of the enormous scientific potential of the second *Gaia* data release.

Here we restrict ourselves to illustrating both the improvement in the data quality and the expanded set of data products through the updated map of the *Gaia* sky. Figure 3 shows the sky distribution of all the sources present in *Gaia* DR2 in the form of source densities on a logarithmic scale. When comparing to the map produced from *Gaia* DR1 data (Gaia Collaboration 2016a) it is immediately apparent that there is a strong reduction in the artefacts caused by the combination of source filtering and the *Gaia* scanning law (see Gaia Collaboration 2016a, for a more detailed explanation of these artefacts), which is another illustration of the increased survey completeness of *Gaia* DR2. Nonetheless there are still source count variations visible, which clearly are imprints from the scanning law (as executed over the first 22 months of the mission). For example there are two arcs above and below the ρ Oph clouds that can be traced all the way down to and below the Galactic plane (these can best be seen in the electronic version of the figure). Such arcs occur all along the ecliptic plane and are regions on the sky that were scanned more frequently by *Gaia* and therefore contain relatively more sources that were observed often enough for inclusion in the published catalogue.

One newly visible (and real) feature in this map is the Sagittarius dwarf which can be noted as an excess in star counts in a strip below the bulge region, stretching to the R Corona Australis region.

Fig. 4. Map of the total flux measured in the G_{RP}, G, and G_{BP} bands, where the flux in these bands is encoded in the red, green, and blue channel, respectively. There is one easily visible artefact in this map, a "green" patch to the lower left of the bulge which is a region where G_{BP} and G_{RP} data are not available for a large number of sources, leading to the greenish colour which was used to encode the G-band fluxes (which are available for all sources). Such artefacts also occur (although not as visible) in the region to the upper left of the Small Magellanic Cloud and at high Galactic latitude to the right of the north Galactic pole region. The areas where green patches are likely to occur can be identified in Fig. 27 in Evans et al. (2018) which shows the celestial distribution of *Gaia* DR2 sources for which no BP/RP photometry is available.

Figure 4 shows a map that combines the integrated fluxes as observed in the G_{RP}, G, and G_{BP} bands, where the integrated flux map for each of the bands was used to colour code the image according to a red, green, and blue channel. The map illustrates the availability of homogeneous all-sky multi-band photometry in *Gaia* DR2 and offers a magnificent view of the Milky Way in colour. This flux map also reveals numerous open clusters which are not readily visible in the source count map (while on the other hand many faint source concentrations, such as distant dwarf galaxies are no longer visible). Complete details on the construction of the images in Figs. 3 and 4 are provided in Moitinho et al. (2018).

One aspect of the sky maps shown in Figs. 3 and 4 that is perhaps not as well appreciated is their effective angular resolution, which given the size of *Gaia*'s main telescope mirrors (1.45 m along the scanning direction, Gaia Collaboration 2016b) should be comparable to that of the *Hubble* Space Telescope. Gaia Collaboration (2016a) and Arenou et al. (2017) discuss how the effective angular resolution of *Gaia* DR1 is limited to about 2–4 arcsec owing to limitations in the data processing. This has much improved for *Gaia* DR2. The gain in angular resolution is illustrated in Fig. 5. The top panel shows the distribution of source pair distances in a small, dense field. For *Gaia* DR2 (upper, red curve) source pairs below 0.4–0.5 arcsec are rarely resolved, but the resolution improves rapidly and above 2.2 arcsec practically all pairs are resolved. For *Gaia* DR1 the fraction of resolved source pairs started to fail at separations of 3.5 arcsec, reaching very low values below 2.0 arcsec. The same, modest resolution is seen for *Gaia* DR2 if we only consider sources with G_{BP} and G_{RP} photometry. The reason is the angular extent of the prism spectra and the fact that *Gaia* DR1 only includes sources for which the integrated flux from the
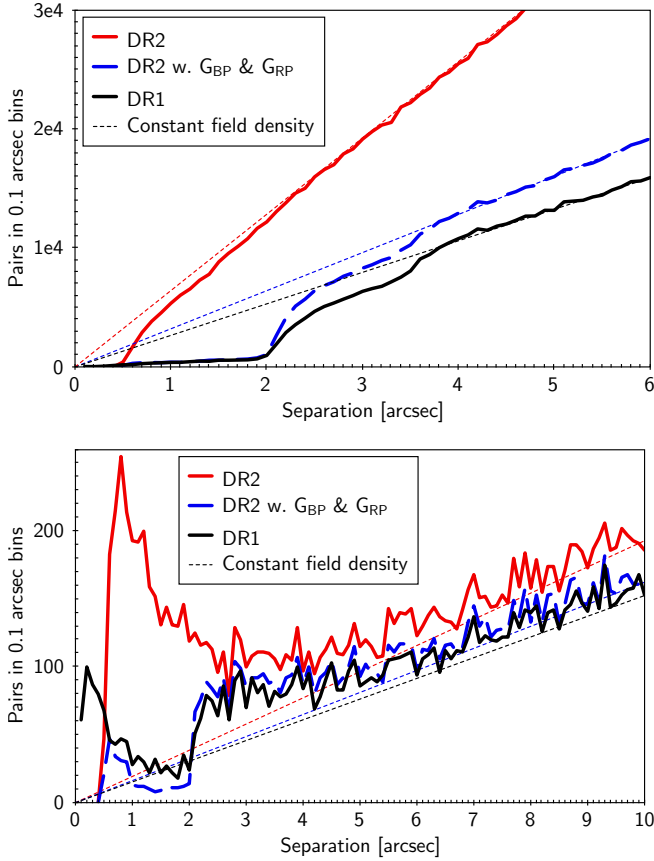
BP/RP spectra could be reliably determined. The lower panel shows in the same way the source pairs in the one hundred times larger, sparse field. The more remarkable feature here is the peak of resolved binaries at small separations, which was missed in *Gaia* DR1. A similar population must be present in the dense field, where it cannot be discerned because the field is dominated by distant sources. The figure also demonstrates that the gain in number of sources from *Gaia* DR1 to *Gaia* DR2 is mainly due to the close source pairs. Finally, Fig. 5 clearly demonstrates that the effective angular resolution of *Gaia* DR2 quite significantly exceeds that of all ground-based large-area optical sky surveys.

5. Treat *Gaia* DR2 as independent from *Gaia* DR1

Although *Gaia* DR1 and *Gaia* DR2 are based on observations from the same instruments, the discussion in the following subsections shows that the two releases should be treated as independent. In particular the tracing of sources from *Gaia* DR1 to *Gaia* DR2 (should this be needed for a particular application) must be done with care.

5.1. Gaia DR2 represents a stand-alone astrometric catalogue

Because the observational time baseline for *Gaia* DR2 is sufficiently long, parallax and proper motion can be derived from the *Gaia* observations alone. That is, the *Tycho-Gaia* Astrometric Solution (TGAS, Michalik et al. 2015a) as employed for the 2 million brightest stars in *Gaia* DR1 is no longer needed, and the astrometric results reported in *Gaia* DR2 are based solely on *Gaia* observations. For the TGAS subset from *Gaia* DR1 there is thus a large difference in the time baseline for the proper motions
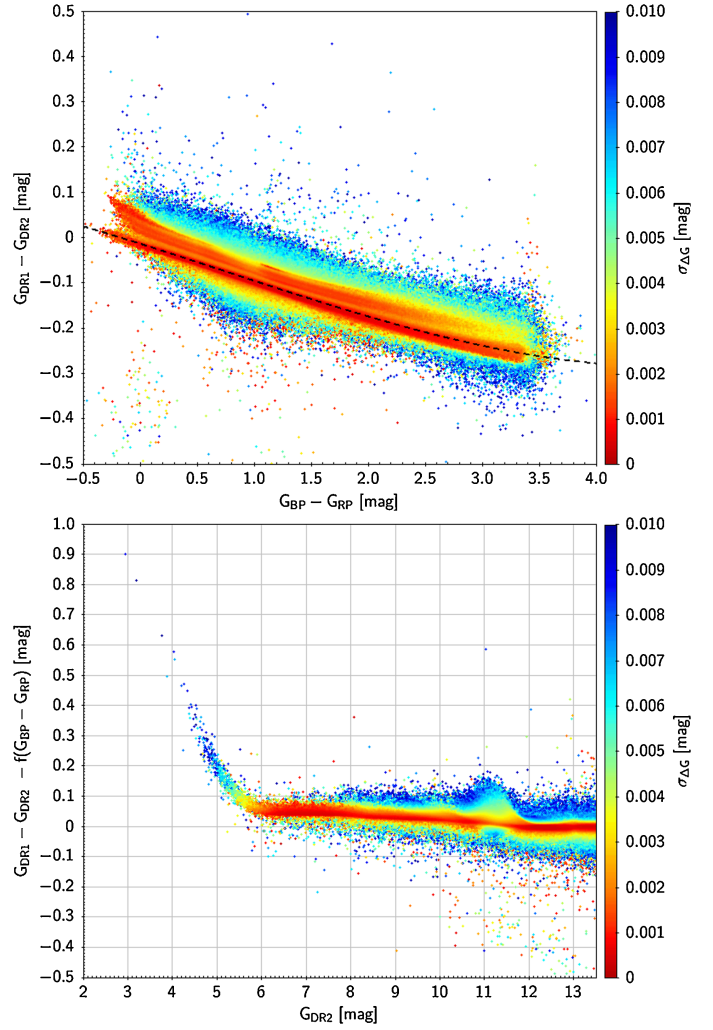
Fig. 5. Histograms from Arenou et al. (2018) of source pair separations in two circular test fields for *Gaia* DR2 sources (red lines); *Gaia* DR2 sources with G_{BP} and G_{RP} photometry (blue lines); and *Gaia* DR1 sources (black lines). *Top*: a dense field of radius $0.5°$ at $(\ell, b) = (-30°, -4°)$ with 456 142 sources, *bottom*: a sparse field of radius $5°$ at $(\ell, b) = (-100°, -60°)$ with 250 092 sources. The thin, dotted lines show the relations for a constant density across the field.

(~24 yr vs ~2 yr) which means there can be significant differences between TGAS and *Gaia* DR2 proper motions for binary stars with orbital periods comparable to 2 yr. The TGAS proper motions may be more reliable in such cases. However, discrepancies can also point to erroneous TGAS proper motions related to a mismatching between (components of) sources observed by *Gaia* and HIPPARCOS (see Makarov et al. 2017, for a discussion of this issue). In cases where proper motion discrepancies are of interest they should be carefully investigated before deciding which values to use or concluding that the discrepancy points to the source not being a single star.

5.2. Photometric system evolution

The photometric data processing for *Gaia* DR2 (Riello et al. 2018; Evans et al. 2018) features many improvements with respect to *Gaia* DR1 and represents a new photometric reduction. In particular more input data was used and the stretch of data selected for the initialisation of the photometric calibration was largely free of the effects of contamination by water ice (see Gaia Collaboration 2016b, for a summary of the contamination problem in the early phases of the *Gaia* mission). As a consequence the photometric system for *Gaia* DR2 is different from *Gaia* DR1. This is illustrated in Fig. 6 which shows the difference in *G*-band magnitude ($\Delta G = G_{DR1} - G_{DR2}$) for the same sources between the two data releases. The source pairs selected from

Fig. 6. *Top diagram*: difference in the value of *G* (with $\Delta G = G_{DR1} - G_{DR2}$) as listed for the same sources in *Gaia* DR1 and *Gaia* DR2 as a function of $(G_{BP} - G_{RP})$. The source pairs selected from the two releases match in celestial position to within 0.25 arcsec and the formal error on the magnitude differences is less than 0.01. All sources were selected to have a flux excess factor in *Gaia* DR2 of less than 1.6. The dashed line shows a polynomial relation between the difference in *G* and the colour. The colour scale indicates the estimated uncertainty on ΔG. *Bottom panel*: relation between ΔG and *G* after removing the colour dependency using the polynomial relation in the top panel.

the two releases match in celestial position to within 0.25 arcsec and the formal error on the magnitude differences is less than 0.01. All sources were selected to have a flux excess factor in *Gaia* DR2 of less than 1.6 (see Sect. 3.2 for a description of this quantity). The two panels in Fig. 6 show that there is a substantial difference in the *G* band values, with the mean of ΔG being about -0.1 mag, and a strong colour dependence which is indicated by the dashed line showing the polynomial relation

$$\begin{aligned} G_{DR1} - G_{DR2} = &-0.013612 - 0.079627(G_{BP} - G_{RP}) \\ &- 0.0040444(G_{BP} - G_{RP})^2 \\ &+ 0.0018602(G_{BP} - G_{RP})^3 . \end{aligned} \quad (4)$$

Removing the colour dependence and plotting ΔG vs. *G* (bottom panel of Fig. 6) reveals image saturation effects at the bright end which more strongly affect the *Gaia* DR1 magnitudes. Sources

with larger magnitude differences typically have large estimated uncertainties (blue points) whereas the majority of sources have smaller differences and small estimated errors (red points). The feature near $G \approx 11.5$ mag is due to the high and variable photometric uncertainties in *Gaia* DR1 for bright sources (see Fig. 9 in Evans et al. 2018).

This difference in photometric systems means that one should not apply photometric calibrations derived from *Gaia* DR1 (e.g., the calibration of the red-clump absolute G-band magnitude, Ruiz-Dern et al. 2018; Hawkins et al. 2017) to *Gaia* DR2 photometry. The G passband calibration also changes from *Gaia* DR1 to *Gaia* DR2. The passbands for G, G_{BP}, and G_{RP} are described in Evans et al. (2018). They are available in the version that was used for the *Gaia* DR2 data processing and in a revised version which was determined after the processing was finished (see Sect. 6.3.2). The revised passband should be used for precise photometric work based on the fluxes listed in *Gaia* DR2. The nominal (pre-launch) passband as provided on the *Gaia* science performance pages[2] and independent passband calibrations based on *Gaia* DR1 (Weiler et al. 2018; Maíz Apellániz 2017) should not be used. Likewise the nominal transformations between the *Gaia* broad-band photometry and other photometric systems listed in Jordi et al. (2010) should not be used. Refer to Evans et al. (2018) for the updated relations. To take full advantage of the high precision *Gaia* DR2 photometry, predictions of the *Gaia* broad-band magnitudes for stellar evolutionary tracks or isochrones in the colour-magnitude diagram (e.g. Choi et al. 2016; Marigo et al. 2017) should be updated.

5.3. Source list evolution

The processing for a given data release starts with a task that groups individual *Gaia* observations and links them to sources on the sky (see Lindegren et al. 2018; Fabricius et al. 2016, for a description of this process). The observations are linked to known sources, or sources are newly "created" from the clustering of the observations around a celestial position where previously no source was known to exist. This leads to a working catalogue of sources (hereafter called "the source list") and their corresponding observations, which forms the basis for the subsequent data processing. In this list the sources are assigned a *Gaia* source identifier which is intended to be stable for every source. The algorithm that carries out the grouping and linking was much improved at the beginning of the processing for *Gaia* DR2. The improved source list will lead to the following changes in linking the observations to the source identifiers for a substantial fraction of sources:

- The merging of groups of observations previously linked to more than one source will lead to a new source associated to the merged observations (with a new source identifier) and the disappearance of the original sources (along with their source identifiers).
- The splitting of groups of observations previously linked to one source will lead to new sources associated to the split groups of observations (with new source identifiers) and the disappearance of the original source (along with its source identifier).
- The list of observations linked to a source may change (and hence the source characteristics may change), while the source identifier remains the same.

In the processing for *Gaia* DR2 the number of changes of source identifiers (where the physical source remains the same) is large. At magnitudes brighter than $G \approx 16$ some 80–90 per cent of the sources changed source identifier. At $G \approx 18$ mag this reduces to some 20 per cent, going down to zero source identifier changes around $G = 20$ mag.

The consequence is that one should not blindly use the source identifier to look up sources from *Gaia* DR1 in *Gaia* DR2. Example applications we have in mind are the repeat of an analysis done with the first data release using the new data and the retrieval of a list *Gaia* DR1 sources, cross-matched against some other survey, from the *Gaia* DR2 tables. The recommendation is to treat the source lists from the two releases as completely independent. An additional field will be added to *Gaia* DR2 and subsequent releases which specifies the *Gaia source name* as "*Gaia* DR*n* source_id". The bare source identifier can be used for efficient queries of the large *Gaia* data base, while the source name should always be specified (i.e., including the data release number) when referring to the source in the literature. To facilitate the tracing of sources from *Gaia* DR1 to *Gaia* DR2 a table is provided which lists for each *Gaia* DR2 source the potential matching sources in *Gaia* DR1 (and vice versa). For the majority of sources (over 99 per cent) there is a one-to-one correspondence (although the source identifier can differ), but multiple matches may occur and then it is up to the user of the *Gaia* data to make a judgement as to which pair is the correct match (where the possible differences in the G-band magnitude should be kept in mind).

The source list is expected to stabilise in future *Gaia* data releases with much less change expected between *Gaia* DR2 and *Gaia* DR3. However some evolution of the source lists will take place up to the final data release and we stress that a change in source *character* can always occur as observations are added in future data releases (e.g., a stable source can turn into a variable from one data release to the next).

6. Using *Gaia* DR2 data: completeness and limitations

Gaia DR2 represents a major advance compared to *Gaia* DR1, featuring new data types and a much expanded and improved astrometric and photometric data set. Nevertheless this release is still intermediate, based on only a limited amount (~22 months) of input data, and still suffers from simplifications in the data processing that will introduce shortcomings in the calibrations which in turn can introduce systematic errors. We summarise here the main limitations of *Gaia* DR2 which the user of the data should be aware of.

6.1. Gaia DR2 validation and source filtering

The validation of the *Gaia* DR2 results followed the process described in Gaia Collaboration (2016a). We refer to the papers listed in Sects. 2 and 3 for full details on the validation of the data done at the level of the individual data processing systems. The overall validation, assessing the combined results is described in Arenou et al. (2018). As was the case for *Gaia* DR1 the results validation revealed no problems that prevented a timely release of *Gaia* DR2, but filtering of the available data processing outputs before their incorporation into *Gaia* DR2 was still necessary. The level of filtering is significantly reduced compared to that for *Gaia* DR1 as can be appreciated from the substantial increase in the number of sources for which astrometric and photometric data is published. We summarise the filtering that was

---

[2] https://www.cosmos.esa.int/web/gaia/science-performance

applied with the aim of providing a better understanding of some of the survey characteristics.

6.1.1. Astrometry

For the astrometric data set the results were filtered by requiring that a source was observed by *Gaia* at least five times (five focal plane transits), and that the astrometric excess noise and the semi-major axis of the position uncertainty ellipse are less than 20 and 100 mas, respectively. In addition within the astrometric solution pipeline the parallax and proper motions are determined only for sources satisfying the requirement that they are brighter than $G = 21$, that the number of "visibility periods" used is at least 6 (a visibility period represents a group of observations separated from other such groups by at least four days), and that the semi-major axis of the 5-dimensional uncertainty ellipse is below a magnitude dependent threshold. We refer to Lindegren et al. (2018) for the details. For sources that do not meet these requirements only the positions are reported in *Gaia* DR2.

6.1.2. Photometry

The photometric inputs were filtered as follows. Sources without a well-determined value for G do not appear in *Gaia* DR2. The photometry in the G, G_{BP}, or G_{RP} bands is only reported if the source was observed at least twice by *Gaia* in the respective bands. For the so-called "bronze" sources (see Sect. 3.2 and Riello et al. 2018) no colour information (i.e. no G_{BP} and G_{RP}) is reported. This also holds for sources fainter than $G = 21$ mag and sources for which the flux excess factor is above 5. Hence *Gaia* DR2 contains a substantial number of sources (~300 million) for which no colour information is available. Note however that the filtering on flux excess factor was not applied to the variable source time series tables, hence there may be sources that have no G_{BP} and/or G_{RP} value listed but for which a light curve in G_{BP} and/or G_{RP} is nevertheless reported.

6.1.3. Radial velocities

For sources satisfying the following conditions no radial velocity is reported in *Gaia* DR2. The source is fainter than $G_{RVS} = 14$ (the limit refers to the flux as actually measured in the RVS band, not the provisional G_{RVS} value mentioned in Sect. 3.3); the fraction of transits where the source was detected as having a double-lined spectrum was larger than 0.1 (this removes detected double-lined spectroscopic binaries); the uncertainty on the radial velocity is above 20 km s^{-1}; the effective temperature corresponding to the spectral template used to derive the radial velocity is outside the range 3550–6900 K. By construction the RVS data processing limited the range of possible radial velocities to $|v_{rad}| < 1000$ km s^{-1}. Special care was taken for the 613 sources that had measured radial velocities with absolute values above 500 km s^{-1}. Because this small subset can easily be contaminated by outliers caused by data processing limitations, their spectra were visually inspected. Of these 613 sources, 202 were included in *Gaia* DR2 as valid high velocity sources, while the remainder were removed from the published catalogue. For sources with radial velocities at absolute values below 500 km s^{-1} visual inspection was not possible due to the progressively (much) higher numbers. The users of *Gaia* DR2 should thus be aware of the specific selection applied to sources with $|v_{rad}| > 500$ km s^{-1}. We refer to Katz et al. (2018) for more details on this issue.

6.1.4. Variable stars

During the variability analysis a strict internal filtering was applied to the quality of the photometric time series (such as removing negative or unrealistically low flux values). This was followed by a filtering of the classification results to reduce the contamination due to data processing artefacts and confusion between variable types. The outputs from the specialised variable star characterisation pipelines were filtered to remove sources for which the results of the light curve analysis were not deemed reliable enough. This combination of filters reduced the number of sources flagged as variable to the numbers listed in Table 1. The reader interested in using the variable star data set is strongly advised to consult Holl et al. (2018) and references therein, as well as the online documentation.

6.1.5. Astrophysical parameters

The astrophysical parameter results are only presented for sources brighter than $G = 17$ (no fainter sources were processed) and only for sources for which G, G_{BP}, and G_{RP} are reported. Further filtering was applied based on the quality of the various inputs to the astrophysical parameter estimation, where particularly strict criteria were applied to the extinction and reddening estimations. The details of the filtering applied to the astrophysical parameters are best understood in conjunction with the description of how these parameters were estimated. Hence we refer to Andrae et al. (2018) for the details (see also Sect. 6.3.4).

6.1.6. Solar system objects

For the solar system data set the filtering on input data quality (internal to the processing pipeline) was followed only by the removal of some SSO observations for which the relative flux uncertainty in the G band was larger than 0.1. This mainly removes observations of the very "fast" SSOs for which the observation window may be badly placed (causing flux loss) toward the end of the focal plane transit. In addition a selection of the SSO observations was removed as well as some individual sources (see Gaia Collaboration 2018f, for details).

6.1.7. Duplicated sources

A global filter concerns the removal of duplicates of sources, which sometimes occur when the observation to source matching process creates two clusters of detections that later turn out to belong to the same source (see Gaia Collaboration 2016a; Fabricius et al. 2016). The 47 802 437 sources for which the duplicate was removed are indicated as such. The removal of duplicates is done after the completion of the data processing. Hence the observations corresponding to the removed component are effectively not used for, and do not appear in, the published catalogue. In future *Gaia* data releases the duplicates are expected to be merged into a single source.

6.2. Survey completeness

As can be appreciated from Fig. 1 the completeness of the *Gaia* survey has much improved for the second data release, being essentially complete between $G = 12$ and $G = 17$. The completeness at the bright end has improved, although a fraction of the bright stars at $G < 7$ is still missing with no stars brighter than $G = 1.7$ mag appearing in *Gaia* DR2. Gaia Collaboration (2016a) extensively explain how the combination of the *Gaia*

scan law coverage of the sky over the period covered by *Gaia* DR1 combined with the filtering applied to the astrometric and photometric results leads to strips or holes with a lack of sources (see Fig. 11 and 12 in that paper). Although much reduced (as seen in Fig. 3), these artefacts are still present in the *Gaia* DR2 source list and start appearing at $G > 17$.

We list here a number of more specific remarks on the completeness of *Gaia* DR2:

- The completeness for high proper motion stars has significantly improved with respect to *Gaia* DR1, but it is estimated that some 17 per cent of high proper motion stars (with $\mu > 0.6$ arcsec yr^{-1}) are still missing (for various reasons).
- In crowded regions the capability to observe all stars is reduced (Gaia Collaboration 2016b). In combination with the still limited data treatment in crowded areas (see section 6.2 in Gaia Collaboration 2016a) this means that the survey limit in regions with densities above a few hundred thousand stars per square degree can be as bright as $G = 18$.
- As described in Sect. 4 the effective angular resolution of the *Gaia* DR2 source list has improved to ~ 0.4 arcsec, with incompleteness in close pairs of stars starting below about 2 arcsec. Refer to Arenou et al. (2018) for details.
- We repeat that the radial velocity, astrophysical parameter and variable star data sets are far from complete with respect to the overall *Gaia* DR2 catalogue (see Sect. 3 above). In particular the radial velocities are only reported for a restricted range in effective temperatures (of the spectral templates, see Sect. 6.1.3) and the completeness of the radial velocity catalogue with respect to *Gaia* DR2 varies from 60 to 80 per cent (Katz et al. 2018) over the range $G = 4$ to $G = 12$.
- The solar system object sample processed for *Gaia* DR2 was pre-selected and is not a complete sample with respect to criteria like dynamics, type, category, etc. In addition bright SSOs ($G \lesssim 10$) were removed from the published results because the astrometry in that brightness range is limited in quality by calibration uncertainties and systematics related to the apparent source size and motion on the sky (leading to the use of inadequate PSF models for the image centroiding).

For more detailed information on the completeness of *Gaia* DR2 we refer to the individual data processing papers and the overall validation paper (Arenou et al. 2018). No attempt was made at deriving a detailed survey selection function.

6.3. Limitations

6.3.1. Astrometry

The astrometry in *Gaia* DR2 represents a major improvement over *Gaia* DR1 with an order of magnitude improvement in the uncertainties at the bright end and a vast expansion of available parallaxes and proper motions. In particular the individual uncertainties are much closer to having been drawn from Gaussian distributions and the systematics in the parallax uncertainties are now generally below the 0.1 mas level (as estimated from the analysis of QSO parallaxes, Lindegren et al. 2018). However, the users of the *Gaia* DR2 astrometry should be aware of the following. There is an overall parallax zeropoint of ~ -0.03 mas (as estimated from QSO parallaxes, in the sense of the *Gaia* DR2 parallaxes being too small) which the data have not been corrected for (see below), and the astrometry shows systematics correlated to celestial position, source colour, and source magnitude. Moreover the parallaxes and proper motions show significant spatial (i.e. source-to-source) correlations of up to 0.04 mas and 0.07 mas yr^{-1} over angular scales from <1 to

20 degrees (see Lindegren et al. 2018, for a more detailed characterisation of the spatial covariances). These regional systematics are visible in maps of average QSO parallaxes and in dense fields where the large amount of sources allows to average the astrometric parameters and visualise the systematic differences in, for example, the parallax zeropoint (Lindegren et al. 2018; Arenou et al. 2018).

One might expect that the published parallax values would have been adjusted according to the global zeropoint, however a deliberate choice was made not to apply any corrections to the *Gaia* DR2 astrometry. This is motivated by the fact that the value of the zeropoint depends on the sample used to estimate its value (Arenou et al. 2018). The differences are related to the dependence of the systematics in the astrometry on source position, colour, and magnitude, meaning that the zeropoint for QSOs (faint, blue) may not be representative of the zeropoint for a sample of bright red stars. In addition the correction of the global zeropoint would represent an arbitrary choice with respect to the regional systematics which would be left uncorrected.

The astrometric uncertainties listed in *Gaia* DR2 are derived from the formal uncertainties resulting from the astrometric data treatment, and unlike for *Gaia* DR1 these have not been externally calibrated (by comparison to the Hipparcos data, Lindegren et al. 2016). At a late stage during the preparation of *Gaia* DR2 a bug was discovered in the astrometric processing software. This did not significantly affect the astrometric parameters themselves but resulted in a serious underestimation of the uncertainties for the bright sources ($G \lesssim 13$). Rather than recomputing the full solution, with serious repercussions for the downstream processing and publication schedule, it was decided to apply an approximate ad hoc correction to the uncertainties. The details of this are described in Appendix A of Lindegren et al. (2018). While the corrected (published) uncertainties are thus approximately consistent with the residuals of the astrometric solution, comparisons with external data show that they are still underestimated (Arenou et al. 2018). The underestimation is moderate (~ 7–10%) for faint sources ($G > 16$) outside the Galactic plane, but may reach 30 to 50 per cent for sources of intermediate magnitude ($G \simeq 12$–15). At brighter magnitudes a comparison with Hipparcos data suggests that uncertainties are underestimated by no more than 25 per cent (Arenou et al. 2018). No additional correction was made in the published data based on these external comparisons, and users of the data may have to allow for it in their analyses.

The PSF model used in the pre-processing is essentially the same as that used for *Gaia* DR1, and the iterative loop between the astrometric and photometric data treatment and the pre-processing is not yet closed (see Sect. 6.1 and Fig. 10 in Gaia Collaboration 2016a). This implies that the PSF calibrations and the subsequent determination of the source flux and location have not benefited from better input astrometry and source colours. These inadequacies in the instrument calibration have a particularly large impact on the astrometry of bright stars ($G \lesssim 13$) which is visible in the uncertainties being larger than those for somewhat fainter stars. In addition there may be a systematic rotation of the proper motion system for the bright stars with respect to QSOs (see Table 3 and Lindegren et al. 2018), and the parallax zeropoint may be different.

6.3.2. Photometry

The strongly varying photometric uncertainty at the bright end in G and the bumps in the uncertainty around $G \sim 13$ and

$G \sim 16$ visible in the *Gaia* DR1 data (Gaia Collaboration 2016a; Evans et al. 2017) are still present although in much reduced form (Evans et al. 2018). The uncertainties on G_{BP} and G_{RP} as a function of magnitude are much smoother with the integrated prism photometry being much less sensitive to instrument configuration changes.

The flux excess factor can take extreme values and it was decided not to publish colour information for sources with a flux excess factor above 5 (this is a rather liberal filtering). We recommend that the value of the flux excess factor is used to clean samples of sources selected from *Gaia* DR2 from the most problematic cases, in particular if accurate colour information is important. The flux excess factor has a dependence on $(G_{BP} - G_{RP})$, which any filtering should take into account. We refer to Evans et al. (2018), Gaia Collaboration (2018a), and Lindegren et al. (2018) for more detailed recommendations on cleaning samples from the effects of the flux excess in the BP/RP bands.

Although not really a limitation in the photometric data, we nevertheless point out the following in relation to the photometric zeropoints and passbands. The photometric zeropoints used to convert the photometric fluxes into the magnitudes listed in *Gaia* DR2 are derived from the pass bands used internal to the processing for this release. The calibration of these passbands was done in a preliminary manner and they have been updated after the *Gaia* DR2 processing was completed through an analysis employing BP/RP spectra which were not available for the earlier calibrations.The magnitude zeropoints for the updated passbands differ by up to 3 mmag from those used to calculate the *Gaia* DR2 magnitudes (Evans et al. 2018). As remarked in Sect. 5, for precision photometric work the updated passbands should be used and then the difference in zeropoints should be accounted for (by recalculating the magnitudes from the fluxes listed in *Gaia* DR2).

6.3.3. Radial velocity data

When using the radial velocities from *Gaia* DR2 the following limitations should be taken into account. Single-lined spectroscopic binaries have been treated as single stars and only the median radial velocity, together with information on the scatter in the underlying (but unpublished) epoch radial velocities, is provided. Double lined spectroscopic binaries which were detected as such were not processed and are missing from the *Gaia* DR2 radial velocity data set. Double lined spectroscopic binaries with a weak secondary component are present in the catalogue and have also been treated as single stars. No radial velocities have been determined for stars with detected emission lines and there are no radial velocities for "cool" and "hot" stars (Sect. 6.1.3). Radial velocities with absolute values above 500 km s^{-1} should be treated with some care. Beyond this limit clearly dubious values were filtered out of the catalogue but it is not guaranteed that all remaining radial velocities above +500 km s^{-1} or below −500 km s^{-1} are reliable.

Through comparison with other radial velocity surveys it is concluded that the *Gaia* DR2 radial velocities are accurate to a few 100 m s^{-1}, where systematic differences can be due to both *Gaia* DR2 and the other surveys. Katz et al. (2018) show that while offsets are lower than 300 m s^{-1} for bright stars ($G_{RVS} < 10$), a trend with magnitude is seen in all the comparisons with other surveys, reaching ~500 m s^{-1} at the faint end.

Finally, we note that *Gaia* DR2 lists the atmospheric parameters (T_{eff}, $\log g$, [Fe/H]) of the spectral templates used in the derivation of the radial velocities through the cross-correlation technique. Their values should *not* be used as estimates of the actual atmospheric parameters of the stars, they are only provided as extra information to judge the quality of the radial velocities.

6.3.4. Astrophysical parameters

The values of T_{eff}, A_G, $E(G_{BP} - G_{RP})$, radius, and luminosity were determined only from the three broad-band photometric measurements and the parallax, on a star-by-star basis (where parallax was not used to estimate T_{eff}). The strong degeneracy between T_{eff} and extinction/reddening when using the broad band photometry necessitates rather extreme assumptions in order to estimate their values. This can lead to correspondingly strong systematics in the astrophysical parameters which are not accounted for in the uncertainties listed in *Gaia* DR2. We summarise here the most important caveats but refer to the online documentation and Andrae et al. (2018) for more extensive guidelines on the use of the astrophysical parameter estimates. The assessment of the quality of the astrophysical parameters from the perspective of the overall validation of *Gaia* DR2 can be found in Arenou et al. (2018).

The estimation of T_{eff}, A_G, and $E(G_{BP} - G_{RP})$ was done using a machine learning algorithm (specifically, the extremely randomised trees, or EXTRATREES algorithm Geurts et al. 2006). For the T_{eff} estimation the algorithm was trained on the photometry for *Gaia* sources for which T_{eff} estimates were available from existing independent surveys (see Andrae et al. 2018, Table 2). Only effective temperatures over the range 3000–10 000 K were considered and the training data shows strong peaks at specific T_{eff} values. The training set for the extinction and reddening estimation was based on synthetic photometry constructed using PARSEC 1.2S[3] stellar models which are accompanied by simulated photometry based on the Atlas 9 synthetic spectral library (Castelli & Kurucz 2003). No attempt was made at a realistic population of the synthetic colour magnitude diagrams in terms of the stellar initial mass function, the metallicity distribution, or the frequency of extinction values. All sources were treated as single stars and no attempt was made to filter out known galaxies, binaries, etc. Please refer to Andrae et al. (2018) for full details.

No T_{eff} values outside the range 3000–10 000 K are reported as these were not contained in the training data used for the estimation algorithm. This means that stars with effective temperatures outside the aforementioned range will have systematically too high or too low T_{eff} values listed in *Gaia* DR2. The distribution of T_{eff} values contains artefacts that reflect the distribution of the T_{eff} values in the training data. Effective temperature estimates in high extinction areas can be underestimated as the training data contained no examples of extincted stars.

The estimates of A_G and $E(G_{BP} - G_{RP})$ have such large uncertainties in general that their usefulness for individual stars is very limited. The extinction/reddening estimates should be used statistically only (for collections of stars) in which case the extinction maps shown in Andrae et al. (2018) demonstrate that on average the A_G estimates are reliable. The extinction estimates are strictly non-negative (with a model grid imposed maximum of $A_G = 4$) and have non-Gaussian posteriors, for which asymmetric uncertainties are listed in the catalogue. The non-negativity constraint can lead to apparent overestimation of the extinctions in regions, such as at high Galactic latitudes, where low extinction is expected on average. The effective temperature and extinction signals are degenerate in the broadband

---

[3] http://stev.oapd.inaf.it/cgi-bin/cmd

colours, which greatly limits the accuracy with which either can be estimated.

The radius and luminosity are estimated from the value of $T_{\rm eff}$ as determined from the *Gaia* photometry, including a bolometric correction obtained from synthetic spectra. The resulting estimates suffer from the naive use of $1/\varpi$ as a distance estimator and the assumption of zero extinction. Their uncertainties are probably underestimated.

6.3.5. Variability data

The variability data contained in *Gaia* DR2 is somewhat complex and consists of three data sets, as described in Sect. 3.4, that overlap to a large degree (for details refer to Holl et al. 2018). The mean G, $G_{\rm BP}$, and $G_{\rm RP}$ magnitudes and fluxes provided as part of the light curve statistics can differ from the values provided in the overall *Gaia* DR2 source table. In these cases the median or mean magnitudes and fluxes from the variability data set are to be preferred. There is a small number of stars with multiple entries in the SOS (Special Object Studies) tables and there are sources with a different type in the SOS and automated variability type estimation data sets. Classifications different from those of independent variable star surveys may occur (Holl et al. 2018; Arenou et al. 2018).

7. Using *Gaia* DR2 data: additional guidance

We briefly discuss a number of specific items that the users of *Gaia* DR2 should keep in mind. These concern issues inherent to the *Gaia* data (releases) and points to keep in mind when interpreting the results from analyses of *Gaia* DR2 data. More extensive examples of how to use the data responsibly are provided in the papers listed at the start of Sect. 4 and in Luri et al. (2018).

7.1. Time stamping in Gaia data releases

Gaia DR2 features photometric time series for sources varying in apparent magnitude and for solar system objects, as well as astrometric time series for the latter. Future releases will in addition contain time series for non-single star astrometry (such as binaries and stars with exoplanets), radial velocities, and the medium and low resolution spectra from the RVS and BP/RP instruments. As summarised in Lindegren et al. (2016) the primary coordinate system used for the *Gaia* (astrometric) data processing is the Barycentric Celestial Reference System (Soffel et al. 2003). The BCRS is a relativistic reference system that is physically adequate to describe both the motion of bodies in the solar system and the propagation of light from distant celestial sources. The time-like coordinate of the BCRS is the barycentric coordinate time (TCB). Consequently all the *Gaia* time series data are time-stamped using TCB. The numerical values in the *Gaia* DR2 tables are expressed JD−2455197.5(TCB) days, where by convention the origin for *Gaia* time-stamping is J2010.0(TCB) = JD 2455197.5(TCB).

7.2. Astrometric source model

All sources were treated as single stars in the astrometric solution for *Gaia* DR2 (Lindegren et al. 2018). This means that physical binaries and multiple systems as well as extended sources (galactic and extra-galactic, such as galaxies in the local universe), although present in *Gaia* DR2, received no special treatment. Moreover the sources that are not single stars are not marked as

such. For binaries with orbital periods of the order of 2 yr the proper motions or parallaxes listed in *Gaia* DR2 may be quite far from the true values for the system. The auxiliary information in *Gaia* DR2 can be used to isolate candidate non-single stars or galaxies but this should be done with care and the results validated against known samples.

7.3. Solar system object astrometry

The epoch astrometry for SSOs is provided with uncertainties (on α, δ) and correlations. These correlations are strong, reflecting the large difference in precision between the along-scan and across-scan astrometric uncertainties which project into the uncertainties in (α, δ) in a correlated manner. The correlations should be taken into account for any application in order to recover the full accuracy of the astrometry in the along-scan direction. A known limitation of asteroid astrometry in *Gaia* DR2 is that the relativistic light deflection is computed as for the stars (i.e., the source is considered to be at infinite distance). A correction corresponding to the difference with respect to the finite distance must be applied whenever mas or sub-mas precision is aimed at.

7.4. Interpretation of photometric colours

The problem of the excess flux in the BP/RP photometry manifests itself primarily at the faint ($G > 19$) end of the survey, in crowded regions and around bright stars. In all these cases when constructing colour magnitude diagrams one should be careful in interpreting them.

For example, open cluster sequences in non-crowded fields may manifest a turn towards the blue at the lower end of the main sequence, which is a consequence of a stronger flux excess in BP than in RP for faint sources. At the faint end one should be aware that the effects of zodiacal light are clearly visible in the distribution of the flux excess factor (Evans et al. 2018).

Care should be taken in the use of colour magnitude diagrams in crowded regions such as globular cluster cores or the Milky Way bulge. Examples of colour-magnitude diagrams affected by the flux excess problem are given in Arenou et al. (2018). Finally, around bright sources there may be a dependence in source colour on the distance from the bright source which will lead to spurious features in a colour magnitude diagram.

When faint red sources are being analysed it may be better to use the $(G - G_{\rm RP})$ colour instead of $(G_{\rm BP} - G_{\rm RP})$ as discussed in Gaia Collaboration (2018a) for the case of brown dwarfs.

7.5. Mean magnitudes of variable stars

If a source is flagged as variable the recommendation is to use the mean value for its photometry from the tables with variability information, as the varying brightness of the source can be more carefully accounted for in the variability analysis.

7.6. Use the astrophysical parameters with care

Andrae et al. (2018) provide extensive guidance on the use of the astrophysical parameter estimates, including how to select samples with the most reliable $T_{\rm eff}$, radius, and luminosity estimates, and examples of how to use the estimates of A_G responsibly. We strongly recommend that these guidelines are followed and encourage independent investigations into the quality and limitations of the astrophysical parameter estimates.

7.7. Filtering to create clean samples

Although the bulk of the data in *Gaia* DR2 is of excellent quality, specific analyses of the data may require further filtering on data quality. One can find examples of how to do such filtering, using the information contained in *Gaia* DR2, throughout the papers accompanying the release. However, in many cases some experimentation by the user of the data will be needed to establish the best ad-hoc filtering for a given application. Such filtering does come at the cost of introducing additional truncation of the data which will further complicate the survey/sample selection function and may in fact severely bias the interpretation of the results. For example, Gaia Collaboration (2018d) show how a seemingly innocuous selection on radial velocity error can lead to strong kinematic biases when studying the Milky Way disk. Further examples of biases induced by sample truncation are given in Luri et al. (2018). Finally, one should keep in mind that filtering on the observed values or uncertainties of source parameters can increase the imprint on the resulting sample of, for example, scanning law patterns.

7.8. Negative parallaxes

Gaia DR2 represents the largest parallax catalogue ever produced and contains parallaxes of faint objects observed relatively few times and of extragalactic objects. For many of such objects the value of the parallax listed in the catalogue may be negative. As explained in Luri et al. (2018) the presence of negative parallaxes is a natural consequence of the way the *Gaia* observations are described in terms of a linearised astrometric source model, with the parameters of the model solved for through a least-squares process. Perhaps this is most easily appreciated by considering the 0.5 million QSOs appearing in *Gaia* DR2 for which parallax solutions have been made. Given that the true parallax for these sources is close to zero it is to be expected that for half of them the observed parallax (as solved for from the observations) is negative (where in the case of *Gaia* DR2 the fraction of negative parallaxes for QSOs is higher because of the negative parallax zeropoint).

Hence negative parallaxes represent perfectly valid measurements and can be included in analyses of the *Gaia* DR2 data. Examples of how one can do this are given in Luri et al. (2018).

7.9. Known spurious results

There are a number of results listed in *Gaia* DR2 which are obviously wrong and which may surprise the user of the data. We point out two specific cases here.

For a small number of sources the parallaxes listed in *Gaia* DR2 have very large positive or negative values (with for example 59 sources having parallaxes larger than that of Proxima Centauri), where the negative values can be very far from zero when expressed in terms of the formal uncertainty on the parallax. These parallax values are spurious and caused by a close alignment (of order 0.2–0.3 arcsec) of sources, that are only occasionally resolved in the *Gaia* observations, depending on the scan direction. These cases show up typically in dense regions covered by only a few transits or an unfortunate distribution of scan directions and parallax factors. This is consistent with most of these sources being faint and concentrated in dense areas along the Galactic plane and toward the Galactic bulge (see Fig. C4 in Lindegren et al. 2018). Most likely the proper motions of these sources are also erroneous. This is consistent with the presence of a number of high-proper motion stars at $G > 19$ (104 243 at $\mu > 100$ mas yr^{-1}, 12 431 at $\mu > 200$ mas yr^{-1}, and

4459 at $\mu > 300$ mas yr^{-1}) which show a marked concentration toward the galactic bulge and galactic plane regions. These sources overlap to a large degree with the sources with spurious parallax values and their proper motions are thus likely to be unreliable. More details on this problem and guidance on how to clean samples from spurious parallaxes can be found in Lindegren et al. (2018, in particular their Appendix C).

Among the bright and well known (i.e. named) variable stars there are a number of cases where the mean *G*-band magnitude listed in *Gaia* DR2 is clearly wrong. One prominent case is the star RR Lyrae itself for which the mean magnitude is listed as $G = 17$. The wrong value is caused by the fact that the treatment of outliers, as implemented in the photometric processing for *Gaia* DR2, is not efficient in the case of variable sources that have an intrinsically large spread in the individual photometric observations. As a consequence of the wrong magnitude estimate, the parallax of RR Lyrae was determined to be -2.6 mas.

We stress here that the above problems concern only a very small number of cases which do not indicate overall problems with the quality of *Gaia* DR2.

7.10. Take into account uncertainties and correlations

The astrometric uncertainties are provided in the form of the full covariance matrix for the five astrometric parameters. The correlations between the uncertainties can be significant and they should always be accounted for to correctly calculate the standard uncertainties on linear combinations of (subsets of) the astrometric parameters and to correctly assess, for example, how far away a given set of astrometric parameters is from a model prediction. The mathematics involved in accounting for correlated uncertainties is summarised in Luri et al. (2018) and described more extensively in the *Gaia* DR2 online documentation.

In this context we point out that the longest principal axis of a scaled version of the covariance matrix is provided as the parameter `astrometric_sigma_5dmax` for both the 5-parameter and 2-parameter solutions. This parameter is equivalent to the semi-major axis of the position error ellipse and can be useful in filtering out sources for which one of the astrometric parameters, or a linear combination of several parameters, is particularly ill-determined. We refer to the online *Gaia* DR2 documentation for more details.

7.11. Dealing with underestimated uncertainties and/or systematic errors

As pointed out above the uncertainties quoted in *Gaia* DR2 on the various source parameters can be underestimated and there are also systematic errors with varying dependencies on source brightness, colour, and position on the sky, which moreover may be spatially correlated. We can provide no general recipe for taking these effects into account in scientific analyses of the *Gaia* DR2 data, but give a few recommendations here.

We strongly advise against attempts to "correct" the data themselves as a means to get rid of underestimated uncertainties or systematic errors. This would require a level of understanding and characterisation of these effects that would have allowed their removal during the data processing in the first place. We recommend (for studies where it matters) to include the presence of systematic effects in the uncertainties as part of the data analysis, for example in a forward modelling approach. The level of systematic errors (e.g. the size of the parallax zeropoint) or
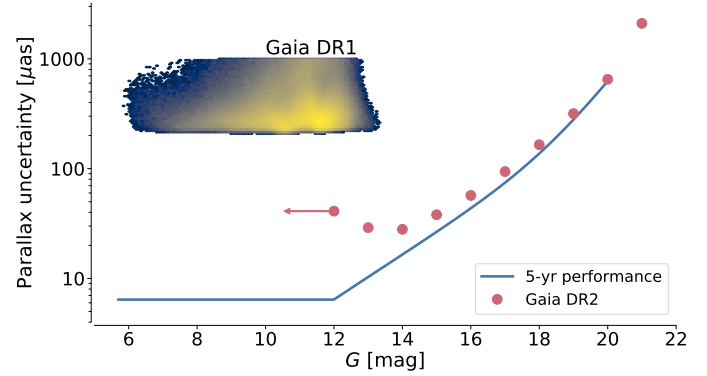
the factors by which uncertainties are under- or overestimated then become part of the model parameters to estimate. Examples of such analyses of *Gaia* DR1 parallax data can be found in Casertano et al. (2017) and Sesar et al. (2017), where the latter include both a parallax zeropoint and a scaling factor for the quoted uncertainties as part of their probabilistic model that fits a period luminosity relation to data for RR Lyrae stars. The spatial correlation parameters for the uncertainties and systematic errors can be included in a similar way as part of the modelling. Further guidance on the use of the astrometric data (in particular the parallaxes) from *Gaia* DR2 can be found in Luri et al. (2018).

8. *Gaia* DR2 access facilities

The main entry point to *Gaia* DR2 remains the ESA *Gaia* archive[4]. Access is also possible through a number of partner and affiliate data centres in Europe, the United States, Japan, Australia, and South Africa. These data centres provide their own access facilities, but do not necessarily host all data contained in the ESA *Gaia* archive. The services offered at the ESA *Gaia* archive remain as described in Gaia Collaboration (2016a) and we list here a few enhancements and changes:

- The access to the light curves for variable stars is now in the form of a URL that links from the main `gaia_source` table to the specific files that contain the light curves for the source in VOTable format[5].
- The astrometric and photometric time series for the SSOs are all collated into one large table containing multiple entries for each SSO. Note that the source identifiers for SSOs are negative numbers. To enable queries of SSOs based on orbital elements or absolute magnitude, an auxiliary table containing such data, plus ancillary quantities, is provided. In addition a table with the residuals of each *Gaia* observation with respect to an orbital fit is provided as a reference.
- The archive visualisation service (Moitinho et al. 2017) has been much expanded to allow for efficient preliminary exploration of the data in the entire *Gaia* DR2 catalogue. The service offers several pre-computed diagrams which can be explored through linked views and allows one to interactively define a query for a given data set. This serves in particular to narrow down queries for data to the exact samples one is interested in and thus save time and storage space for the actual query. Full details can be found in Moitinho et al. (2018).
- We provide pre-computed cross-matches between *Gaia* DR2 and a number of other large surveys. We recommend using these cross-matches as they have been carefully validated and their use facilitates reproducing analyses of *Gaia* DR2 data combined with other survey data. The details are provided in Marrese et al. (2018). The pre-computed cross-matches are provided for the following surveys: Hipparcos (new reduction, van Leeuwen 2007); *Tycho*-2 (Høg et al. 2000); 2MASS (Skrutskie et al. 2006); SDSS DR9 (Ahn et al. 2012); APASS DR9 (Henden et al. 2016, 2015); UCAC4 (Zacharias et al. 2013); Pan-STARRS1 (Chambers et al. 2016); AllWise (Wright et al. 2010); GSC2.3 (Lasker et al. 2008); PPMXL (Roeser et al. 2010); URAT1 (Zacharias et al. 2015); and RAVE DR5 (Kunder et al. 2017).

Fig. 7. Parallax uncertainties in *Gaia* DR2 (dots) as a function of *G* compared to the uncertainties quoted for *Gaia* DR1 (colour scale) and the expected end-of-mission parallax performance (solid line), as predicted after the commissioning of *Gaia*. Note how the performance for *Gaia* DR2 is still limited by calibration uncertainties for sources brighter than $G \sim 14$.

Finally we mention the creation of a *Gaia* Community forum[6] which is intended to facilitate discussion on the use of *Gaia* data. The principle is to let the users of the data discuss amongst themselves on this forum but the discussions will be monitored by members from the *Gaia* Data Processing and Analysis Consortium who may respond with comments and expert advice when necessary.

9. Conclusions

With the first *Gaia* data release in 2016 the astronomical community got an early taste of the potential of the *Gaia* mission results, in particular through the 2 million parallaxes and proper motions made available as part of the *Tycho-Gaia* Astrometric Solution. The science done with *Gaia* DR1 spans a wide range of topics and often features the powerful combination of *Gaia* and other surveys. *Gaia* DR1 was also quickly established as a reference for the astrometric and photometric calibration of other surveys, resulting among others in the rejuvenation of existing proper motion catalogues. For a brief review of the impact of *Gaia* DR1 we refer to Brown (2017).

With the release of *Gaia* DR2 the promise of the availability of fundamental astrophysical information for (over) a billion sources spread over a substantial fraction of the volume of the Milky Way starts to be fulfilled. The addition of the largest radial velocity survey to date, coupled with astrophysical information for 161 million sources and variability information for half a million sources will make *Gaia* DR2 a resource to be mined for stellar physics and galactic as well as extra-galactic astronomy for many years to come. Moreover, *Gaia* DR2 provides a first glimpse of the immense power of *Gaia* for solar system studies.

Nevertheless *Gaia* DR2 still represents an early data release based on only a limited amount (less than two years) of input data, partly inadequate calibrations, and an incomplete understanding of the behaviour of the spacecraft, payload, and instruments. These shortcomings manifest themselves as systematic errors which although much reduced in size from *Gaia* DR1 to *Gaia* DR2 will remain a limiting factor in scientific uses of the data, in particular at the bright end of the survey and, for example, for distant samples. This is illustrated in Fig. 7 which shows the parallax uncertainties as a function of *G* for *Gaia* DR2 (dots),

---

[4] It can be accessed at `http://archives.esac.esa.int/gaia`
[5] `http://www.ivoa.net/documents/latest/VOT.html`

[6] `https://www.cosmos.esa.int/web/gaia/forum`

Gaia DR1 (colour scale map), and the end of mission (solid line, as predicted after *Gaia* commissioning, Gaia Collaboration 2016b). The bright end ($G \lesssim 14$) performance for *Gaia* DR2 is still limited by calibration errors, while at the faint end the nominal end of mission performance is already being reached (this is probably due to a conservative assessment of the effect at the faint end of the excess stray light). The task for the *Gaia* data processing for the next data release will thus be to substantially reduce the systematics such that a real advantage can be gained, in particular at the bright end, from the increase in precision due to the longer time span of the input data. The main challenges will be the following. The PSF modelling used in the image location determination must be upgraded, such that for example astrometric colour terms are already accounted for at an early stage. This also requires the closing of the iterative loop shown in Fig. 10 in Gaia Collaboration (2016a). The modelling of the sky background (both astronomical and as caused by the excess stray light on board *Gaia*) has to be refined to further improve the image location process and to get rid of the flux excess in the BP/RP photometry. The latter will also benefit from an improvement in the treatment of crowded fields, specifically a better treatment of the effects of overlapping images in all of *Gaia*'s instruments and in particular for the BP/RP/RVS instruments where the measurement of spectra necessitates much larger images in the focal plane. Finally the origins of the systematic effects in the astrometry will be further investigated with much effort to be dedicated to the continued development of the possibility to calibrate the systematic effects from the observations.

The next *Gaia* data release will also feature new data products of which the BP/RP and RVS spectra and the non-single star astrometric and radial velocity solutions represent qualitative changes in the character with respect to *Gaia* DR2. Further enhancements include: epoch astrometry for non-single stars, an expanded radial velocity survey (to $G_{\rm RVS} \sim 14$) including the analysis of spectroscopic binaries, astrophysical parameter estimates based on BP/RP/RVS spectra, a further order of magnitude increase in the availability of variability information, the first results from eclipsing binary star processing, analyses of extended objects (galaxies, QSO hosts), and an expanded list of some hundred thousand solar system objects for which multi-colour photometry will also be provided. The latter opens up for investigation the powerful combination of precise orbits for SSOs combined with a homogeneous multi-colour photometric survey of these bodies.

Hence there is much more to come from *Gaia* but for now we invite the reader to start exploring the magnificent survey that is *Gaia* DR2.

References

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, ApJS, 203, 21
Altmann, M., Bouquillon, S., Taris, F., et al. 2014, in Observatory Operations: Strategies, Processes, and Systems V, Proc. SPIE, 9149, 91490P
Andrae, R., Fouesneau, M., Creevey, O., et al. 2018, A&A, 616, A8 (*Gaia* 2 SI)
Arenou, F., Luri, X., Babusiaux, C., et al. 2017, A&A, 599, A50
Arenou, F., Luri, X., Babusiaux, C., et al. 2018, A&A, 616, A17 (*Gaia* 2 SI)
Astropy Collaboration (Price-Whelan, et al.) 2018, ArXiv e-prints [arXiv:1801.02634]
Boch, T., & Fernique, P. 2014, in Astronomical Data Analysis Software and Systems XXIII, eds. N. Manset & P. Forshay, ASP Conf. Ser., 485, 277
Bonnarel, F., Fernique, P., Bienaymé, O., et al. 2000, A&AS, 143, 33
Brown, A. G. A. 2017, Proc. IAU Symp. 330, 13
Casertano, S., Riess, A. G., Bucciarelli, B., & Lattanzi, M. G. 2017, A&A, 599, A67
Castelli, F., & Kurucz, R. L. 2003, Proc. IAU Symp. 210, poster A20 [arXiv:astro-ph/0405087]
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints [arXiv:1612.05560]
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Cropper, M., Katz, D., Sartoretti, P., et al. 2018, A&A, 616, A5 (*Gaia* 2 SI)
Evans, D. W., Riello, M., De Angeli, F., et al. 2017, A&A, 600, A51
Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, A4 (*Gaia* 2 SI)
Fabricius, C., Bastian, U., Portell, J., et al. 2016, A&A, 595, A3
Gaia Collaboration (Brown, A. G. A., et al.) 2016a, A&A, 595, A2
Gaia Collaboration (Prusti, T., et al.) 2016b, A&A, 595, A1
Gaia Collaboration (Babusiaux, C., et al.) 2018a, A&A, 616, A10 (*Gaia* 2 SI)
Gaia Collaboration (Eyer, L., et al.) 2018b, A&A, submitted (*Gaia* 2 SI)
Gaia Collaboration (Helmi, A., et al.) 2018c, A&A, 616, A12 (*Gaia* 2 SI)
Gaia Collaboration (Katz, D., et al.) 2018d, A&A, 616, A11 (*Gaia* 2 SI)
Gaia Collaboration (Mignard, F., et al.) 2018e, A&A, 616, A14 (*Gaia* 2 SI)
Gaia Collaboration (Spoto, F., et al.) 2018f, A&A, 616, A13 (*Gaia* 2 SI)
Geurts, P., Ernst, D., & Wehenkel, L. 2006, Mach. Learn., 63, 3
Hambly, N., Cropper, M., Boudreault, S., et al. 2018, A&A, 616, A15 (*Gaia* 2 SI)
Hawkins, K., Leistedt, B., Bovy, J., & Hogg, D. W. 2017, MNRAS, 471, 722
Henden, A. A., Levine, S., Terrell, D., & Welch, D. L. 2015, AAS Meeting Abstracts, 225, 336.16
Henden, A. A., Templeton, M., Terrell, D., et al. 2016, VizieR Online Data Catalogue, II/336
Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27
Holl, B., Audard, M., Nienartowicz, K., et al. 2018, A&A, in press, DOI: 10.1051/0004-6361/201832892 (*Gaia* 2 SI)
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, A&A, 523, A48
Katz, D., Sartoretti, P., Cropper, M., et al. 2018, A&A, submitted (*Gaia* 2 SI)
Kunder, A., Kordopatis, G., Steinmetz, M., et al. 2017, AJ, 153, 75
Lasker, B. M., Lattanzi, M. G., McLean, B. J., et al. 2008, AJ, 136, 735
Lindegren, L., Lammers, U., Bastian, U., et al. 2016, A&A, 595, A4
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2 (*Gaia* 2 SI)
Luri, X., Brown, A., Sarro, L., et al. 2018, A&A, 616, A9
Maíz Apellániz, J. 2017, A&A, 608, L8
Makarov, V. V., Fabricius, C., & Frouard, J. 2017, ApJ, 840, L1
Marigo, P., Girardi, L., Bressan, A., et al. 2017, ApJ, 835, 77
Marrese, P., Marinoni, S., Fabrizio, M., & Altavilla, G. 2018, A&A, submitted (*Gaia* 2 SI)
Michalik, D., Lindegren, L., & Hobbs, D. 2015a, A&A, 574, A115
Michalik, D., Lindegren, L., Hobbs, D., & Butkevich, A. G. 2015b, A&A, 583, A68
Moitinho, A., Krone-Martins, A., Savietto, H., et al. 2017, A&A, 605, A52
Moitinho, A. et al. 2018, A&A, submitted (*Gaia* 2 SI)
Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23
Pérez, F., & Granger, B. E. 2007, Comput. Sci. Eng., 9, 21
Riello, M., De Angeli, F., Evans, D., et al. 2018, A&A, 616, A3 (*Gaia* 2 SI)
Roeser, S., Demleitner, M., & Schilbach, E. 2010, AJ, 139, 2440
Ruiz-Dern, L., Babusiaux, C., Arenou, F., Turon, C., & Lallement, R. 2018, A&A, 609, A116
Sartoretti, P., Katz, D., Cropper, M., et al. 2018, A&A, 616, A6 (*Gaia* 2 SI)
Sesar, B., Fouesneau, M., Price-Whelan, A. M., et al. 2017, ApJ, 838, 107
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Smart, R. L., & Nicastro, L. 2014, A&A, 570, A87
Soffel, M., Klioner, S. A., Petit, G., et al. 2003, AJ, 126, 2687
Soubiran, C., Jasniewicz, G., Chemin, L., et al. 2018, A&A, 616, A7 (*Gaia* 2 SI)
Taylor, M. B. 2005, in Astronomical Data Analysis Software and Systems XIV, eds. P. Shopbell, M. Britton, & R. Ebert, ASP Conf. Ser., 347, 29
van Leeuwen, F. 2007, Astrophys. Space Sci. Lib., 350
Weiler, M., Jordi, C., Fabricius, C., & Carrasco, J. M. 2018, A&A, 615, A24
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
Zacharias, N., Finch, C. T., Girard, T. M., et al. 2013, AJ, 145, 44
Zacharias, N., Finch, C., Subasavage, J., et al. 2015, AJ, 150, 101

[1] Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
[2] INAF - Osservatorio Astronomico di Padova, Vicolo Osservatorio 5, 35122 Padova, Italy
[3] Science Support Office, Directorate of Science, European Space Research and Technology Centre (ESA/ESTEC), Keplerlaan 1, 2201AZ, Noordwijk, The Netherlands
[4] GEPI, Observatoire de Paris, Université PSL, CNRS, 5 Place Jules Janssen, 92190 Meudon, France
[5] Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France
[6] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
[7] Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstr. 12-14, 69120 Heidelberg, Germany
[8] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[9] Department of Astronomy, University of Geneva, Chemin des Maillettes 51, 1290 Versoix, Switzerland
[10] Mission Operations Division, Operations Department, Directorate of Science, European Space Research and Technology Centre (ESA/ESTEC), Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[11] Institut de Ciències del Cosmos, Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
[12] Lohrmann Observatory, Technische Universität Dresden, Mommsenstraße 13, 01062 Dresden, Germany
[13] European Space Astronomy Centre (ESA/ESAC), Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
[14] Lund Observatory, Department of Astronomy and Theoretical Physics, Lund University, Box 43, 22100 Lund, Sweden
[15] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France
[16] CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France

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

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

[19] INAF - Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italy

[20] Telespazio Vega UK Ltd for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[21] Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France

[22] Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK

[23] INAF - Osservatorio Astrofisico di Torino, via Osservatorio 20, 10025 Pino Torinese, Italy

[24] INAF - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Piero Gobetti 93/3, 40129 Bologna, Italy

[25] Serco Gestión de Negocios for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[26] ALTEC, Corso Marche 79, 10146 Torino, Italy

[27] Department of Astronomy, University of Geneva, Chemin d'Ecogia 16, 1290 Versoix, Switzerland

[28] Gaia DPAC Project Office, ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[29] SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, LNE, 61 avenue de l'Observatoire, 75014 Paris, France

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

[31] IMCCE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Lille, 77 av. Denfert-Rochereau, 75014 Paris, France

[32] Royal Observatory of Belgium, Ringlaan 3, 1180 Brussels, Belgium

[33] Institut d'Astrophysique Spatiale, Université Paris XI, UMR 8617, CNRS, Bâtiment 121, 91405, Orsay Cedex, France

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

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

[36] Institut d'Astrophysique et de Géophysique, Université de Liège, 19c, Allée du 6 Août, 4000 Liège, Belgium

[37] ATG Europe for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[38] Área de Lenguajes y Sistemas Informáticos, Universidad Pablo de Olavide, Ctra. de Utrera, km. 1, 41013 Sevilla, Spain

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

[40] Large Synoptic Survey Telescope, 950 N. Cherry Avenue, Tucson AZ 85719, USA

[41] Observatoire Astronomique de Strasbourg, Université de Strasbourg, CNRS, UMR 7550, 11 rue de l'Université, 67000 Strasbourg, France

[42] Kavli Institute for Cosmology, University of Cambridge, Madingley Road, CB3 0HA, UK

[43] Aurora Technology for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[44] Laboratoire Univers et Particules de Montpellier, Université Montpellier, Place Eugène Bataillon, CC72, 34095 Montpellier Cedex 05, France

[45] Department of Physics and Astronomy, Division of Astronomy and Space Physics, Uppsala University, Box 516, 75120 Uppsala, Sweden

[46] CENTRA, Universidade de Lisboa, FCUL, Campo Grande, Edif. C8, 1749-016 Lisboa, Portugal

[47] Università di Catania, Dipartimento di Fisica e Astronomia, Sezione Astrofisica, Via S. Sofia 78, 95123 Catania, Italy

[48] INAF - Osservatorio Astrofisico di Catania, via S. Sofia 78, 95123 Catania, Italy

[49] University of Vienna, Department of Astrophysics, Türkenschanzstraße 17, 1180 Vienna, Austria

[50] CITIC - Department of Computer Science, University of A Coruña, Campus de Elviña s/n, 15071 A Coruña, Spain

[51] CITIC - Astronomy and Astrophysics, University of A Coruña, Campus de Elviña s/n, 15071 A Coruña, Spain

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

[53] Space Science Data Center - ASI, Via del Politecnico SNC, 00133 Roma, Italy

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

[55] Finnish Geospatial Research Institute FGI, Geodeetinrinne 2, 02430 Masala, Finland

[56] Isdefe for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[57] Institut UTINAM UMR6213, CNRS, OSU THETA Franche-Comté Bourgogne, Université Bourgogne Franche-Comté, 25000 Besançon, France

[58] STFC, Rutherford Appleton Laboratory, Harwell, Didcot OX11 0QX, UK

[59] Departamento de Inteligencia Artificial, UNED, c/ Juan del Rosal 16, 28040 Madrid, Spain

[60] Elecnor Deimos Space for ESA/ESAC, Camino bajo del Castillo s/n Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

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

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

[63] European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany

[64] ON/MCT, Rua Gal. José Cristino 77, Rio de Janeiro 20921-400, Brazil

[65] OV/UFRJ, Ladeira Pedro Antônio 43, Rio de Janeiro 20080-090, Brazil

[66] Department of Terrestrial Magnetism, Carnegie Institution for Science, 5241 Broad Branch Road, NW, Washington, DC 20015-1305, USA

[67] Università di Torino, Dipartimento di Fisica, via Pietro Giuria 1, 10125 Torino, Italy

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

[69] Leicester Institute of Space and Earth Observation and Department of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK

[70] Departamento de Estadística, Universidad de Cádiz, Calle República Árabe Saharawi s/n, 11510 Puerto Real, Cádiz, Spain

[71] Astronomical Institute Bern University, Sidlerstrasse 5, 3012 Bern, Switzerland (present address)

[72] EURIX Srl, Corso Vittorio Emanuele II 61, 10128 Torino, Italy

[73] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge MA 02138, USA

[74] HE Space Operations BV for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

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

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

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

[78] SRON, Netherlands Institute for Space Research, Sorbonnelaan 2, 3584CA Utrecht, The Netherlands

[79] Departamento 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] Unidad de Astronomía, Universidad de Antofagasta, Avenida Angamos 601, Antofagasta 1270300, Chile

[81] CRAAG - Centre de Recherche en Astronomie, Astrophysique et Géophysique, Route de l'Observatoire, Bp 63, Bouzareah 16340, Alger, Algeria

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

[83] INAF - Osservatorio Astronomico d'Abruzzo, Via Mentore Maggini, 64100 Teramo, Italy

[84] INAF - Osservatorio Astronomico di Capodimonte, Via Moiariello 16, 80131 Napoli, Italy

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

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

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

[88] RHEA for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

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

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

[91] UNINOVA - CTS, Campus FCT-UNL, Monte da Caparica, 2829-516 Caparica, Portugal

[92] School of Physics, O'Brien Centre for Science North, University College Dublin, Belfield, Dublin 4, Ireland

[93] Dipartimento di Fisica e Astronomia, Università di Bologna, Via Piero Gobetti 93/2, 40129 Bologna, Italy

[94] Barcelona Supercomputing Center - Centro Nacional de Supercomputación, c/ Jordi Girona 29, Ed. Nexus II, 08034 Barcelona, Spain

[95] Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Box 848, S-981 28 Kiruna, Sweden

[96] Max Planck Institute for Extraterrestrial Physics, High Energy Group, Gießenbachstraße, 85741 Garching, Germany

[97] Astronomical Observatory Institute, Faculty of Physics, Adam Mickiewicz University, Słoneczna 36, 60-286 Poznań, Poland

[98] Konkoly Observatory, Research Centre for Astronomy and Earth Sciences, Hungarian Academy of Sciences, Konkoly Thege Miklós út 15-17, 1121 Budapest, Hungary

[99] Eötvös Loránd University, Egyetem tér 1-3, 1053 Budapest, Hungary

[100] American Community Schools of Athens, 129 Aghias Paraskevis Ave. & Kazantzaki Street, Halandri, 15234 Athens, Greece

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

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

[103] Physics Department, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium

[104] McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

[105] Astronomical Institute, Academy of Sciences of the Czech Republic, Fričova 298, 25165 Ondřejov, Czech Republic

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

[107] Institut de Physique de Rennes, Université de Rennes 1, 35042 Rennes, France

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

[109] School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, PR China

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

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

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

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

[114] Baja Observatory of University of Szeged, Szegedi út III/70, 6500 Baja, Hungary

[115] Laboratoire AIM, IRFU/Service d'Astrophysique - CEA/DSM - CNRS - Université Paris Diderot, Bât 709, CEA-Saclay, 91191 Gif-sur-Yvette Cedex, France

[116] Warsaw University Observatory, Al. Ujazdowskie 4, 00-478 Warszawa, Poland

[117] Institute of Theoretical Physics, Faculty of Mathematics and Physics, Charles University in Prague, Czech Republic

[118] AKKA for CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France

[119] Vitrociset Belgium for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[120] HE Space Operations BV for ESA/ESTEC, Keplerlaan 1, 2201AZ Noordwijk, The Netherlands

[121] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

[122] QUASAR Science Resources for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[123] Fork Research, Rua do Cruzado Osberno, Lt. 1, 9 esq., Lisboa, Portugal

[124] APAVE SUDEUROPE SAS for CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France

[125] Nordic Optical Telescope, Rambla José Ana Fernández Pérez 7, 38711 Breña Baja, Spain

[126] Spanish Virtual Observatory, Spain

[127] Fundación Galileo Galilei - INAF, Rambla José Ana Fernández Pérez 7, 38712 Breña Baja, Santa Cruz de Tenerife, Spain

[128] INSA for ESA/ESAC, Camino bajo del Castillo s/n, Urbanizaciòn Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain

[129] Departamento de Arquitectura de Computadores y Automática, Facultad de Informática, Universidad Complutense de Madrid, C/ Prof. José García Santesmases s/n, 28040 Madrid, Spain

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

[131] Institut d'Estudis Espacials de Catalunya (IEEC), Gran Capitá 2-4, 08034 Barcelona, Spain

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

[133] Stellar Astrophysics Centre, Aarhus University, Department of Physics and Astronomy, 120 Ny Munkegade, Building 1520, 8000 Aarhus C, Denmark

[134] Argelander-Institut fü Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany

[135] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

[136] Sorbonne Universités, UPMC Univ. Paris 6 et CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd. Arago, 75014 Paris, France

[137] Department of Geosciences, Tel Aviv University, Tel Aviv 6997801, Israel

Appendix A: List of acronyms

Table A.1. List of acronyms used in this paper.

Acronym	Description
2MASS	Two-Micron All Sky Survey
AAVSO	American Association of Variable Star Observers
APASS	AAVSO Photometric All-Sky Survey
BCRS	Barycentric Celestial Reference System
BP	Blue Photometer
CCD	Charge-Coupled Device
DPAC	Data Processing and Analysis Consortium
ESA	European Space Agency
GBOT	Ground-Based Optical Tracking
GSC	Guide Star Catalog
ICRF	International Celestial Reference Frame
JD	Julian Date
LMC	Large Magellanic Cloud
OBMT	On-Board Mission Timeline
PSF	Point Spread Function
PPMXL	Position and Proper Motion Extended-L Catalog
QSO	Quasi-Stellar Object
RMS	Root-Mean-Square
RP	Red Photometer
RVS	Radial Velocity Spectrometer
SDSS	Sloan Digital Sky Survey
SMC	Small Magellanic Cloud (special, high-density area on the sky)
SOS	Specific Object Studies
SSO	Solar-System Object
TCB	Barycentric Coordinate Time
TGAS	Tycho-Gaia Astrometric Solution
UCAC	USNO CCD Astrograph Catalog
URAT	USNO Robotic Astrometric Telescope
URL	Uniform Resource Locator