

Gaia Data Release 2

Observations of solar system objects

Gaia Collaboration, F. Spoto^{1,2,*}, P. Tanga¹, F. Mignard¹, J. Berthier², B. Carry^{1,2}, A. Cellino³, A. Dell’Oro⁴, D. Hestroffer², K. Muinonen^{5,6}, T. Pauwels⁷, J.-M. Petit⁸, P. David², F. De Angeli⁹, M. Delbo¹, B. Frézouls¹⁰, L. Galluccio¹, M. Granvik^{5,11}, J. Guiraud¹⁰, J. Hernández¹², C. Ordénovic¹, J. Portell¹³, E. Poujoulet¹⁴, W. Thuillot², G. Walmsley¹⁰, A. G. A. Brown¹⁵, A. Vallenari¹⁶, T. Prusti¹⁷, J. H. J. de Bruijne¹⁷, C. Babusiaux^{18,19}, C. A. L. Bailer-Jones²⁰, M. Biermann²¹, D. W. Evans⁹, L. Eyer²², F. Jansen²³, C. Jordi¹³, S. A. Klioner²⁴, U. Lammers¹², L. Lindegren²⁵, X. Luri¹³, C. Panem¹⁰, D. Pourbaix^{26,27}, S. Randich⁴, P. Sartoretti¹⁸, H. I. Siddiqui²⁸, C. Soubiran²⁹, F. van Leeuwen⁹, N. A. Walton⁹, F. Arenou¹⁸, U. Bastian²¹, M. Cropper³⁰, R. Drimmel³, D. Katz¹⁸, M. G. Lattanzi³, J. Bakker¹², C. Cacciari³¹, J. Castañeda¹³, L. Chaoul¹⁰, N. Cheek³², C. Fabricius¹³, R. Guerra¹², B. Holl²², E. Masana¹³, R. Messineo³³, N. Mowlavi²², K. Nienartowicz³⁴, P. Panuzzo¹⁸, M. Riello⁹, G. M. Seabroke³⁰, F. Thévenin¹, G. Gracia-Abril^{35,21}, G. Comoretto²⁸, M. Garcia-Reinaldos¹², D. Teyssier²⁸, M. Altmann^{21,36}, R. Andrae²⁰, M. Audard²², I. Bellas-Velidis³⁷, K. Benson³⁰, R. Blomme⁷, P. Burgess⁹, G. Busso⁹, G. Clementini³¹, M. Clotet¹³, O. Creevey¹, M. Davidson³⁸, J. De Ridder³⁹, L. Delchambre⁴⁰, C. Ducourant²⁹, J. Fernández-Hernández⁴¹, M. Fouesneau²⁰, Y. Frémat⁷, M. García-Torres⁴², J. González-Núñez^{32,43}, J. J. González-Vidal¹³, E. Gosset^{40,27}, L. P. Guy^{34,44}, J.-L. Halbwachs⁴⁵, N. C. Hambly³⁸, D. L. Harrison^{9,46}, S. T. Hodgkin⁹, A. Hutton⁴⁷, G. Jasiewicz⁴⁸, A. Jean-Antoine-Piccolo¹⁰, S. Jordan²¹, A. J. Korn⁴⁹, A. Krone-Martins⁵⁰, A. C. Lanzafame^{51,52}, T. Lebzelter⁵³, W. Löffler²¹, M. Manteiga^{54,55}, P. M. Marrese^{56,57}, J. M. Martín-Fleitas⁴⁷, A. Moitinho⁵⁰, A. Mora⁴⁷, J. Osinde⁵⁸, E. Pancino^{4,57}, A. Recio-Blanco¹, P. J. Richards⁵⁹, L. Rimoldini³⁴, A. C. Robin⁸, L. M. Sarro⁶⁰, C. Siopis²⁶, M. Smith³⁰, A. Sozzetti³, M. Süveges²⁰, J. Torra¹³, W. van Reeven⁴⁷, U. Abbas³, A. Abreu Aramburu⁶¹, S. Accart⁶², C. Aerts^{39,63}, G. Altavilla^{56,57,31}, M. A. Álvarez⁵⁴, R. Alvarez¹², J. Alves⁵³, R. I. Anderson^{64,22}, A. H. Andrei^{65,66,36}, E. Anglada Varela⁴¹, E. Antiche¹³, T. Antoja^{17,13}, B. Arcay⁵⁴, T. L. Astraatmadja^{20,67}, N. Bach⁴⁷, S. G. Baker³⁰, L. Balaguer-Núñez¹³, P. Balm²⁸, C. Barache³⁶, C. Barata⁵⁰, D. Barbato^{68,3}, F. Barblan²², P. S. Barklem⁴⁹, D. Barrado⁶⁹, M. Barros⁵⁰, M. A. Barstow⁷⁰, S. Bartholomé Muñoz¹³, J.-L. Bassilana⁶², U. Becciani⁵², M. Bellazzini³¹, A. Berihuete⁷¹, S. Bertone^{3,36,72}, L. Bianchi⁷³, O. Bienaymé⁴⁵, S. Blanco-Cuaresma^{22,29,74}, T. Boch⁴⁵, C. Boeche¹⁶, A. Bombrun⁷⁵, R. Borrachero¹³, D. Bossini¹⁶, S. Bouquillon³⁶, G. Bourda²⁹, A. Bragaglia³¹, L. Bramante³³, M. A. Breddels⁷⁶, A. Bressan⁷⁷, N. Brouillet²⁹, T. Brüsemeister²¹, E. Brugaletta⁵², B. Bucciarelli³, A. Burlacu¹⁰, D. Busonero³, A. G. Butkevich²⁴, R. Buzzi³, E. Caffau¹⁸, R. Cancellerie⁷⁸, G. Cannizzaro^{79,63}, T. Cantat-Gaudin^{16,13}, R. Carballo⁸⁰, T. Carlucci³⁶, J. M. Carrasco¹³, L. Casamiquela¹³, M. Castellani⁵⁶, A. Castro-Ginard¹³, P. Charlot²⁹, L. Chemin⁸¹, A. Chiavassa¹, G. Cocozza³¹, G. Costigan¹⁵, S. Cowell⁹, F. Crifo¹⁸, M. Crosta³, C. Crowley⁷⁵, J. Cuypers¹⁷, C. Dafonte⁵⁴, Y. Damerjji^{40,82}, A. Dapergolas³⁷, M. David⁸³, P. de Laverny¹, F. De Luise⁸⁴, R. De March³³, R. de Souza⁸⁵, A. de Torres⁷⁵, J. Debosscher³⁹, E. del Pozo⁴⁷, A. Delgado⁹, H. E. Delgado⁶⁰, S. Diakite⁸, C. Diener⁹, E. Distefano⁵², C. Dolding³⁰, P. Drazinos⁸⁶, J. Durán⁵⁸, B. Edvardsson⁴⁹, H. Enke⁸⁷, K. Eriksson⁴⁹, P. Esquej⁸⁸, G. Eynard Bontemps¹⁰, C. Fabre⁸⁹, M. Fabrizio^{56,57}, S. Faigler⁹⁰, A. J. Falcão⁹¹, M. Farràs Casas¹³, L. Federici³¹, G. Fedorets⁵, P. Fernique⁴⁵, F. Figueras¹³, F. Filippi³³, K. Findeisen¹⁸, A. Fonti³³, E. Fraile⁸⁸, M. Fraser^{9,92}, M. Gai³, S. Galleti³¹, D. Garabato⁵⁴, F. García-Sedano⁶⁰, A. Garofalo^{93,31}, N. Garralda¹³, A. Gavel⁴⁹, P. Gavras^{18,37,86}, J. Gerssen⁸⁷, R. Geyer²⁴, P. Giacobbe³, G. Gilmore⁹, S. Girona⁹⁴, G. Giuffrida^{57,56}, F. Glass²², M. Gomes⁵⁰, A. Gueguen^{18,95}, A. Guerrier⁶², R. Gutiérrez-Sánchez²⁸, R. Haigron¹⁸, D. Hatzidimitriou^{86,37}, M. Hauser^{21,20}, M. Haywood¹⁸, U. Heiter⁴⁹, A. Helmi⁷⁶, J. Heu¹⁸, T. Hilger²⁴, D. Hobbs²⁵, W. Hofmann²¹, G. Holland⁹, H. E. Huckle³⁰, A. Hypki^{15,96}, V. Icardi³³, K. Janßen⁸⁷, G. Jevardat de Fombelle³⁴, P.G. Jonker^{79,63}, Á. L. Juhász^{97,98}, F. Julbe¹³, A. Karamelas^{86,99}, A. Kewley⁹, J. Klar⁸⁷, A. Kochoska^{100,101}, R. Kohley¹², K. Kolenberg^{102,39,74}, M. Kontizas⁸⁶, E. Kontizas³⁷, S. E. Kposov^{9,103}, G. Kordopatis¹, Z. Kostrzewa-Rutkowska^{79,63}, P. Koubsky¹⁰⁴,

* Corresponding author: F. Spoto, e-mail: fspoto@oca.eu

S. Lambert³⁶, A. F. Lanza⁵², Y. Lasne⁶², J.-B. Lavigne⁶², Y. Le Fustec¹⁰⁵, C. Le Poncin-Lafitte³⁶, Y. Lebreton^{18,106}, S. Leccia¹⁰⁷, N. Leclerc¹⁸, I. Lecoeur-Taibi³⁴, H. Lenhardt²¹, F. Leroux⁶², S. Liao^{3,108,109}, E. Licata⁷³, H. E. P. Lindström^{110,111}, T.A. Lister¹¹², E. Livanou⁸⁶, A. Lobel⁷, M. López⁶⁹, S. Managau⁶², R. G. Mann³⁸, G. Mantelet²¹, O. Marchal¹⁸, J. M. Marchant¹¹³, M. Marconi¹⁰⁷, S. Marinoni^{56,57}, G. Marschalkó^{97,114}, D. J. Marshall¹¹⁵, M. Martino³³, G. Marton⁹⁷, N. Mary⁶², D. Massari⁷⁶, G. Matijević⁸⁷, T. Mazeh⁹⁰, P. J. McMillan²⁵, S. Messina⁵², D. Michalik²⁵, N. R. Millar⁹, D. Molina¹³, R. Molinaro¹⁰⁷, L. Molnár⁹⁷, P. Montegriffo³¹, R. Mor¹³, R. Morbidelli³, T. Morel⁴⁰, D. Morris³⁸, A. F. Mulone³³, T. Muraveva³¹, I. Musella¹⁰⁷, G. Nelemans^{63,39}, L. Nicastro³¹, L. Noval⁶², W. O'Mullane^{12,44}, D. Ordóñez-Blanco³⁴, P. Osborne⁹, C. Pagani⁷⁰, I. Pagano⁵², F. Pailler¹⁰, H. Palacin⁶², L. Palaversa^{9,22}, A. Panahi⁹⁰, M. Pawlak^{116,117}, A. M. Piersimoni⁸⁴, F.-X. Pineau⁴⁵, E. Plachy⁹⁷, G. Plum¹⁸, E. Poggio^{68,3}, A. Prša¹⁰¹, L. Pulone⁵⁶, E. Racero³², S. Ragaini³¹, N. Rambaux², M. Ramos-Lerate¹¹⁸, S. Regibo³⁹, C. Reylé⁸, F. Riclet¹⁰, V. Ripepi¹⁰⁷, A. Riva³, A. Rivard⁶², G. Rixon⁹, T. Roegiers¹¹⁹, M. Roelens²², M. Romero-Gómez¹³, N. Rowell³⁸, F. Royer¹⁸, L. Ruiz-Dern¹⁸, G. Sadowski²⁶, T. Sagristà Sellés²¹, J. Sahlmann^{12,120}, J. Salgado¹²¹, E. Salguero⁴¹, N. Sanna⁴, T. Santana-Ros⁹⁶, M. Sarasso³, H. Savietto¹²², M. Schultheis¹, E. Sciacca⁵², M. Segol¹²³, J. C. Segovia³², D. Ségransan²², I.-C. Shih¹⁸, L. Siltala^{5,124}, A. F. Silva⁵⁰, R. L. Smart³, K. W. Smith²⁰, E. Solano^{69,125}, F. Solitro³³, R. Sordo¹⁶, S. Soria Nieto¹³, J. Souchay³⁶, A. Spagna³, U. Stampa²¹, I. A. Steele¹¹³, H. Steidelmüller²⁴, C. A. Stephenson²⁸, H. Stoev¹²⁶, F.F. Suess⁹, J. Surdej⁴⁰, L. Szabados⁹⁷, E. Szegedi-Elek⁹⁷, D. Tapiador^{127,128}, F. Taris³⁶, G. Tauran⁶², M. B. Taylor¹²⁹, R. Teixeira⁸⁵, D. Terrett⁵⁹, P. Teyssandier³⁶, A. Titarenko¹, F. Torra Clotet¹³⁰, C. Turon¹⁸, A. Ulla¹³¹, E. Utrilla⁴⁷, S. Uzzi³³, M. Vaillant⁶², G. Valentini⁸⁴, V. Valette¹⁰, A. van Elteren¹⁵, E. Van Hemelryck⁷, M. van Leeuwen⁹, M. Vaschetto³³, A. Vecchiato³, J. Veljanoski⁷⁶, Y. Viala¹⁸, D. Vicente⁹⁴, S. Vogt¹¹⁹, C. von Essen¹³², H. Voss¹³, V. Votruba¹⁰⁴, S. Voutsinas³⁸, M. Weiler¹³, O. Wertz¹³³, T. Wevers^{9,63}, Ł. Wyrzykowski^{9,116}, A. Yoldas⁹, M. Žerjal^{100,134}, H. Ziaepour⁸, J. Zorec¹³⁵, S. Zschocke²⁴, S. Zucker¹³⁶, C. Zurbach⁴⁸, T. Zwitter¹⁰⁰

(Affiliations can be found after the references)

Received 24 February 2018 / Accepted 10 April 2018

ABSTRACT

Context. The *Gaia* spacecraft of the European Space Agency (ESA) has been securing observations of solar system objects (SSOs) since the beginning of its operations. Data Release 2 (DR2) contains the observations of a selected sample of 14,099 SSOs. These asteroids have been already identified and have been numbered by the Minor Planet Center repository. Positions are provided for each *Gaia* observation at CCD level. As additional information, complementary to astrometry, the apparent brightness of SSOs in the unfiltered *G* band is also provided for selected observations.

Aims. We explain the processing of SSO data, and describe the criteria we used to select the sample published in *Gaia* DR2. We then explore the data set to assess its quality.

Methods. To exploit the main data product for the solar system in *Gaia* DR2, which is the epoch astrometry of asteroids, it is necessary to take into account the unusual properties of the uncertainty, as the position information is nearly one-dimensional. When this aspect is handled appropriately, an orbit fit can be obtained with post-fit residuals that are overall consistent with the a-priori error model that was used to define individual values of the astrometric uncertainty. The role of both random and systematic errors is described. The distribution of residuals allowed us to identify possible contaminants in the data set (such as stars). Photometry in the *G* band was compared to computed values from reference asteroid shapes and to the flux registered at the corresponding epochs by the red and blue photometers (RP and BP).

Results. The overall astrometric performance is close to the expectations, with an optimal range of brightness $G \sim 12 - 17$. In this range, the typical transit-level accuracy is well below 1 mas. For fainter asteroids, the growing photon noise deteriorates the performance. Asteroids brighter than $G \sim 12$ are affected by a lower performance of the processing of their signals. The dramatic improvement brought by *Gaia* DR2 astrometry of SSOs is demonstrated by comparisons to the archive data and by preliminary tests on the detection of subtle non-gravitational effects.

Key words. astrometry – minor planets, asteroids: general – methods: data analysis – space vehicles: instruments

1. Introduction

The ESA *Gaia* mission ([Gaia Collaboration 2016](#)) is observing the sky since December 2013 with a continuous and pre-determined scanning law. While the large majority of the observations concern the stellar population of the Milky Way, *Gaia* also collects data of extragalactic sources and solar system objects (SSOs). A subset of the latter population of celestial bodies is the topic of this work.

Gaia has been designed to map astrophysical sources of very small or negligible angular extension. Extended sources, like the major planets, that do not present a narrow brightness peak are indeed discarded by the onboard detection algorithm. This mission is therefore a wonderful facility for the study of the population of SSOs, including small bodies, such as asteroids, Jupiter trojans, Centaurs, and some trans-Neptunian objects (TNO) and planetary satellites, but not the major planets.

The SSO population is currently poorly characterised, because basic physical properties including mass, bulk density, spin properties, shape, and albedo are not known for the vast majority of them.

The astrometric data are continuously updated by ground-based surveys, and they are sufficient for a general dynamical classification. Only in rare specific situations, however, their accuracy allows us to measure subtle effects such as non-gravitational perturbations and/or to estimate the masses. In this respect, *Gaia* represents a major step forward.

Gaia is the first global survey to provide a homogeneous data set of positions, magnitudes, and visible spectra of SSOs, with extreme performances on the astrometric accuracy (Mignard et al. 2007; Cellino et al. 2007; Tanga et al. 2008, 2012; Hestroffer et al. 2010; Delbo' et al. 2012; Tanga & Mignard 2012; Spoto et al. 2017). *Gaia* astrometry, for ~350 000 SSOs by the end of the mission, is expected to produce a real revolution. The additional physical data (low-resolution reflection spectra, accurate photometry) will at the same time provide a much needed physical characterisation of SSOs.

Within this population, the *Gaia* DR2 contains a sample of 14 099 SSOs (asteroids, Jupiter trojans, and a few TNOs) for a total of 1 977 702 different observations, collected during 22 months since the start of the nominal operations in July 2014. A general description of *Gaia* DR2 is provided in Gaia Collaboration (2018).

The main goal of releasing SSO observations in *Gaia* DR2 is to demonstrate the capabilities of *Gaia* in the domain of SSO astrometry and to also allow the community to familiarise itself with *Gaia* SSO data and perform initial scientific studies. For this reason, the following fundamental properties of the release are recalled first.

- Only a sub-sample of well-known SSOs was selected among those observed by *Gaia*. Moreover, this choice is not intended to be complete with respect to any criterion based on dynamics of physical categories.
- The most relevant dynamical classes are represented, including near-Earth and main-belt objects, Jupiter trojans, and a few TNOs.
- For each of the selected objects, all the observations obtained over the time frame covered by the *Gaia* DR2 are included, with the exception of those that did not pass the quality tests described later in this article.
- Photometric data are provided for only a fraction of the observations as a reference, but they should be considered as preliminary values that will be refined in future data releases.

The goals of this paper are to illustrate the main steps of the data processing that allowed us to obtain the SSO positions from *Gaia* observations and to analyse the results in order to derive the overall accuracy of the sample, as well as to illustrate the selection criteria that were applied to identify and eliminate the outliers.

The core of our approach is based on an accurate orbital fitting procedure, which was applied on the *Gaia* data alone, for each SSO. The data published in the DR2 contain all the quantities needed to reproduce the same computations. The post-fit orbit residuals generated during the preparation of this study are made available as an auxiliary data set on the ESA Archive¹. Its object is to serve as a reference to evaluate the performance of independent orbital fitting procedures that could be performed by the archive users.

More technical details on the data properties and their organisation, which are beyond the scope of this article, are illustrated in the *Gaia* DR2 documentation accessible through the ESA archive.

This article is organised as follows. Section 2 illustrates the main properties of the sample selected for DR2 and recalls the features of *Gaia* that affect SSO observations. For a more comprehensive description of *Gaia* operations, we refer to Gaia Collaboration (2016). The data reduction procedure is outlined in Sect. 3, while Sect. 4 illustrates the properties of the photometric data that complement the astrometry. Section 5 is devoted to the orbital fitting procedure, whose residuals are then used to explore the data quality. This is described in Sects. 6 and 7.

2. Data used

We recall here some basic properties of the *Gaia* focal plane that directly affect the observations. As the *Gaia* satellite rotates at a constant rate, the images of the sources on the focal plane drift continuously (in the along-scan direction, AL) across the different CCD strips. A total of nine CCD strips exists, and the source in the astrometric field (AF, numbered from one to nine, AF1, AF2, . . . AF9) can cross up to these nine strips.

Thus each transit published in the *Gaia* DR2 consists at most of nine observations (AF instrument). Each CCD operates in time-delay integration (TDI) mode, at a rate corresponding to the drift induced by the satellite rotation. All observations of SSOs published in the *Gaia* DR2, both for astrometry and photometry, are based on measurements obtained by single CCDs.

The TDI rate is an instrumental constant, and the spacecraft spin rate is calibrated on the stars. The exposure time is determined by the crossing time of a single CCD, that is, 4.4 s. Shorter exposure times are obtained when needed to avoid saturation, by intermediate electric barriers (the so-called gates) that swallow all collected electrons. Their positioning on the CCD in the AL direction is chosen in such a way that the distance travelled by the source on the CCD itself is reduced, thus reducing the exposure time.

To drastically reduce the data volume processed on board and transmitted to the ground, only small patches around each source (windows) are read out from each CCDs. The window is assigned after the source has been detected in a first strip of CCD, the sky mapper (SM), and confirmed in AF1. For the vast majority of the detected sources ($G > 16$), the window has a size of 12×12 pixels, but the pixels are binned in the direction perpendicular to the scanning direction, called across-scan (AC). Only 1D information in the AL direction is thus available, with the exception of the brightest sources ($G < 13$), for which a full 2D window is transmitted. Sources of intermediate brightness are given a slightly larger window (18×12 pixels), but AC binning is always present.

As the TDI rate corresponds to the nominal drift velocity of stars on the focal plane, the image of an SSO that has an apparent sky motion is slightly spread in the direction of motion. Its AL position also moves with respect to the window centre during the transit. The signal is thus increasingly truncated by the window edge. For instance, the signal of an SSO with an apparent motion (in the AL direction) of 13.6 mas s^{-1} moves by one pixel during a single CCD crossing, with corresponding image smearing.

We can assume that the image is centred in the window at the beginning of the transit, when it is detected first by the SM, and its position is used to define the window coordinates. Then, while drifting on the focal plane and crossing the AF CCDs, due to its motion relative to the stars, the SSO will leave the window

¹ <https://gea.esac.esa.int/archive/>

center. When the AF5 strip is reached, about half of the flux will be lost.

In practice, the uncertainty in determining the position of the source within the window is a function of its centring and can vary over the transit due to the image drift described above. This contribution to the error budget is computed for each position and published in *Gaia* DR2.

2.1. Selection of the sample

For *Gaia* DR2, the solar system pipeline worked on a pre-determined list of transits in the field of view (FOV) of *Gaia*. To build it, a list of accurate predictions was first created by cross-matching the evolving position of each asteroid to the sky path of the *Gaia* FOVs. This provides a set of predictions of SSO transits that were then matched to the observed transits. At this level, the information on the SSO transits comes from the output of the daily processing (Fabricius et al. 2016) and in particular from the initial data treatment (IDT). IDT proceeds by an approximate, daily solution of the astrometry to derive source positions with a typical uncertainty of the order of ~ 70 – 100 mas. There was typically one SSO transit in this list for every 100 000 stellar transits.

SSO targets for the *Gaia* DR2 were selected following the basic idea of assembling a satisfactory sample for the first mass processing of sources, despite the relatively short time span covered by the observations (22 months). The selection of the sample was based on some simple rules:

- The goal was to include a significant number of SSOs, between 10 000 and 15 000.
- The sample had to cover all classes of SSOs: near-Earth asteroids (NEAs), main-belt asteroids (MBAs), Jupiter trojans, and TNOs.
- Each selected object was requested to have at least 12 transits in the 22 months covered by the *Gaia* DR2 data.

The final input selection contains 14 125 SSOs, with a total of 318 290 transits. Not all these bodies are included in *Gaia* DR2: 26 objects were filtered out for different reasons (see Sects. 3.2 and 5). The coverage in orbital semi-major axes is represented in Fig. 1.

2.2. Time coverage

The *Gaia* DR2 contains observations of SSOs from 5 August 2014 to 23 May, 2016². During the first two weeks of the period covered by the observations, a special scanning mode was adopted to obtain a dense coverage of the ecliptic poles (Gaia Collaboration 2016, the ecliptic pole scanning law, EPSL). Due to the peculiar geometry of the EPSL, the scan plane crosses the ecliptic in the perpendicular direction with a gradual drift of the node longitude at the speed of the Earth orbiting the Sun.

A smooth transition then occurred towards the nominal scanning law (NSL) between 22 August and 25 September 2014 that was maintained constant afterwards. In this configuration, the spin axis of *Gaia* precesses on a cone centred in the direction of the Sun, with a semi-aperture of 45° and period of 62.97 days (Fig. 4). As a result, the scan plane orientation changes continuously with respect to the ecliptic with inclinations between 90° and 45° . The nodal direction has a solar elongation between 45° and 135° .

² As a rule, *Gaia* DR2 data start on 25 July 2014, but for SSOs and for technical reasons, no transits have been retained before August 5.

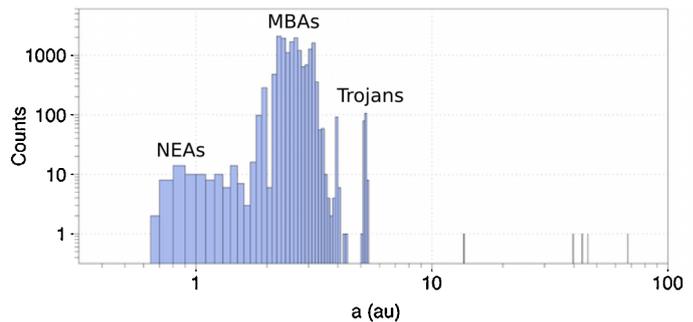


Fig. 1. Distribution of the semi-major axes of the 14 125 SSOs contained in the final input selection. Not all the bodies shown in this figure are included in *Gaia* DR2: 26 objects were filtered out for different reasons (see Sects. 3.2 and 5).

The general distribution of the observations is rather homogeneous in time, with very rare gaps, in general shorter than a few hours; these are due to maintenance operations (orbital maneuvers, telescope refocusing, micrometeoroid hits, and other events; Fig. 2).

A more detailed view of the distribution with a resolution of several minutes (Fig. 3) reveals a general pattern that repeats at each rotation of the satellite (6 h) and is dominated by a sequence of peaks that correspond to the crossing of the ecliptic region by the two FOVs, at intervals of ~ 106 min (FOV 1 to FOV 2) and ~ 254 min (FOV 2 to FOV 1). The peaks are strongly modulated in amplitude by the evolving geometry of the scan plane with respect to the ecliptic.

The observation dates are given in barycentric coordinate time (TCB) *Gaia*-centric³, which is the primary timescale for *Gaia*, and also in coordinated universal time (UTC) *Gaia*-centric. Timings correspond to mid exposure, which is the instant of crossing of the fiducial line on the CCD by the photocentre of the SSO image.

The accuracy of timing is granted by a time-synchronisation procedure between the atomic master clock onboard *Gaia* (providing onboard time, OBT) and OBMT, the onboard mission timeline (Gaia Collaboration 2016). OBMT can then be converted into TCB at *Gaia*. The absolute timing accuracy requirements for the science of *Gaia* is $2 \mu\text{s}$. In practice, this requirement is achieved throughout the mission, with a significant margin.

2.3. Geometry of detection

The solar elongation is the most important geometric feature in *Gaia* observations of SSOs. By considering the intersection of the scan plane with the ecliptic, as shown in Fig. 5, it is clear that SSOs are always observed at solar elongations between 45° and 135° .

This peculiar geometry has important consequences on solar system observations. The SSOs are not only observed at non-negligible phase angles (Fig. 12), in any case never close to the opposition, but also in a variety of configurations (high/low proper motion, smaller or larger distance, etc.), which have some influence on many scientific applications and can affect the detection capabilities of *Gaia* and the measurement accuracy.

The mean geometrical solar elongation of the scan plane on the ecliptic is at quadrature. In this situation, the scan plane is

³ Difference between the barycentric JD time in TCB and 2455197.5.

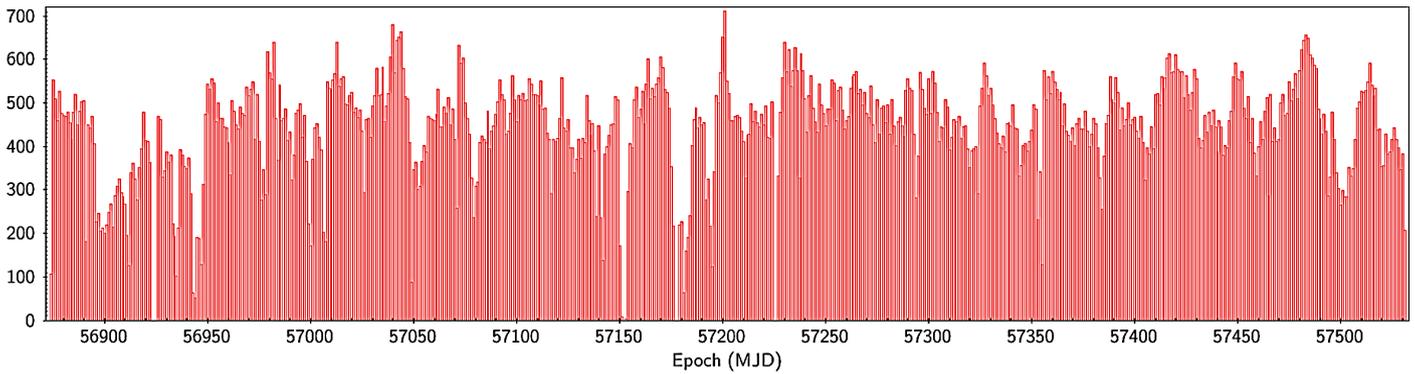


Fig. 2. Distribution in time of the SSO observations published in DR2. The bin size is one day.

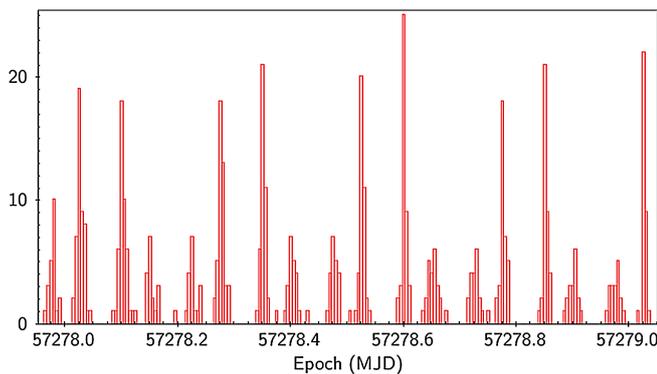


Fig. 3. Detail over a short time interval of the distribution shown in Fig. 2.

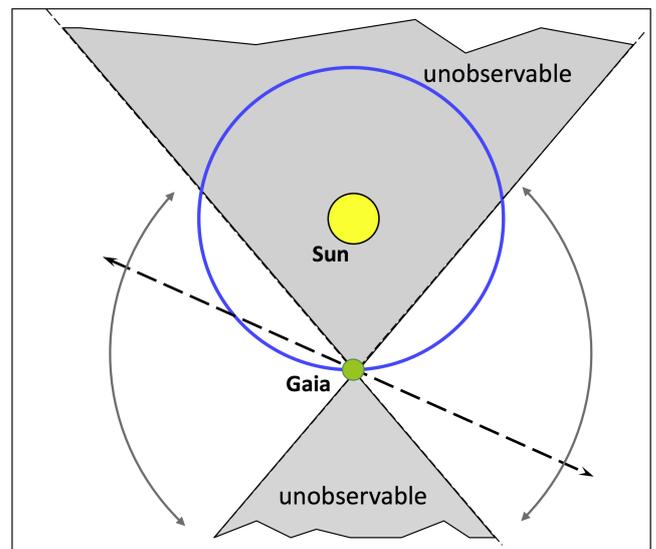


Fig. 5. By drawing the intersection of the possible scan plane orientations with the ecliptic, in the reference rotating around the Sun with the *Gaia* spacecraft, the two avoidance regions corresponding to the cones of Fig. 4 emerge in the direction of the Sun and around opposition. The dashed line represents the intersection of the scanning plane and the ecliptic at an arbitrary epoch. During a single rotation of the satellite, the FOVs of *Gaia* cross the ecliptic in two opposite directions. The intersection continuously scans the allowed sectors, as indicated by the curved arrows.

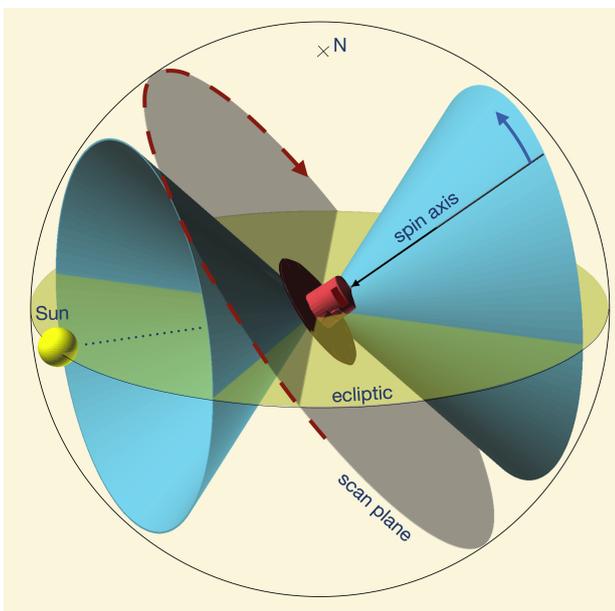


Fig. 4. Geometry of the *Gaia* NSL on the celestial sphere, with ecliptic north at the top. The scanning motion of *Gaia* is represented by the red dashed line. The precession of the spin axis describes the two cones, aligned on the solar–anti-solar direction, with an aperture of 90° . As a consequence, the scan plane, here represented at a generic epoch, is at any time tangent to the cones. When the spin axis is on the ecliptic plane, *Gaia* scans the ecliptic perpendicularly, at extreme solar elongations. The volume inside the cones is never explored by the scan motion.

inclined by 45° with respect to the ecliptic. During the precession cycle, the scan plane reaches the extreme inclination of 90° on the ecliptic. In this geometry, the SSOs with low-inclination orbits move mainly in the AC direction when they are observed by *Gaia*. As the AC pixel size and window are approximately times larger than AL, the sensitivity to the motion in terms of flux loss, image shift, and smearing will thus be correspondingly lower.

These variations of the orientation and the distribution of the SSO orbit inclinations translate into a wide range of possible orientations of the velocity vector on the (AL, AC) plane. Even for a single object, a large variety of velocities and scan directions is covered over time.

2.4. Errors and correlations

The SSO apparent displacement at the epoch of each observation is clearly a major factor affecting the performance of *Gaia*,

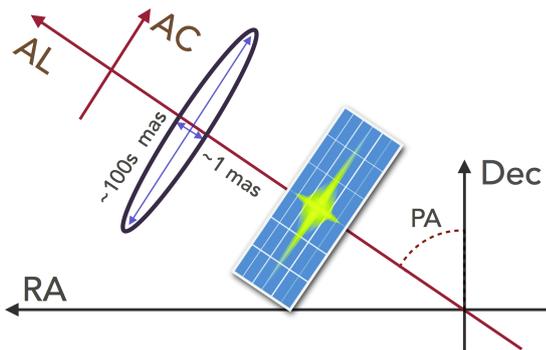


Fig. 6. Approximate sketch illustrating the effects of the strong difference between the astrometry precision in AL (reaching sub-mas level) and in AC (several 100 s mas). The approximate uncertainty ellipse (not to be interpreted as a 2D Gaussian distribution) is extremely stretched in the AC direction. The position angle (PA) is the angle between the declination and the AC direction.

even within a single transit. Other general effects acting on single CCD observations exist, such as local CCD defects, local point spread function (PSF) deviations, cosmic rays, and background sources. For all these reasons, the exploitation of the single data points must rely on a careful analysis that takes both the geometric conditions of the observations and appropriate error models into account.

A direct consequence of the observation strategy employed by *Gaia* is the very peculiar error distribution for the single astrometric observation.

Because of the AC binning, most accurate astrometry in the astrometric field for most observations is only available in the AL direction. This is a natural consequence of the design principle of *Gaia*, which is based on converting an accurate measurement of time (the epoch when a source image crosses a reference line on the focal plane) into a position. In practical terms, the difference between AC and AL accuracy is so large that we can say that the astrometric information is essentially one-dimensional.

As illustrated in Fig. 6, the resulting uncertainty on the position can be represented by an ellipse that is extremely stretched in the AC direction. When this position is converted into another coordinate frame (such as the equatorial reference α , δ), a very strong correlation appears between the related uncertainties σ_α , σ_δ . Therefore it is of the utmost importance that the users take these correlations into account in their analysis. The values are provided in the ESA Archive and must be used to exploit the full accuracy of the *Gaia* astrometry and to avoid serious misuse of the *Gaia* data.

3. Outline of the data reduction process

The solar system pipeline (Fig. 7) collects all the data needed to process the identified transits (epoch of transit on each CCD, flux, AC window coordinates, and many auxiliary pieces of information).

A first module, labelled “Identification” in the scheme, computes the auxiliary data for each object, and assigns the identifying correct identification label to each object. Focal plane coordinates are then converted into sky coordinates by using the transformations provided by AGIS, the astrometric global iterative solution, and the corresponding calibrations (astrometric reduction module). This is the procedure described below in Sect. 3.1. We note that this approach adopts the same principle as absolute stellar astrometry (Lindegren et al. 2018): a local

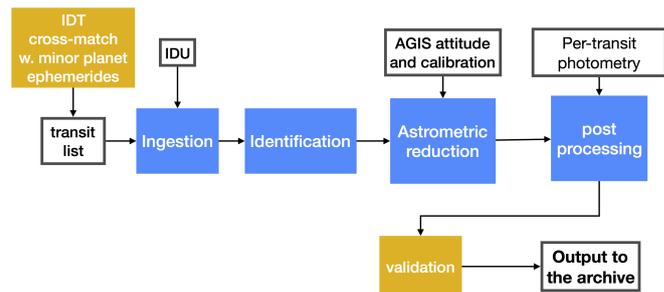


Fig. 7. Main step of the solar system pipeline that collects all the data needed to process identified transits.

information equivalent to the usual small-field astrometry (i.e. position relative to nearby stars) is never used.

Many anomalous data are also rejected by the same module. The post-processing appends the calibrated photometry to the data of each observation (determined by an independent pipeline, see Sect. 4) and groups all the observations of a same target. Eventually, a “Validation” task rejects anomalous data.

The origin of the anomalies are multiple: for instance, data can be corrupted for technical reasons, or a mismatch with a nearby star on the sky plane can enter the input list. Identifying truly anomalous data from peculiarities of potential scientific interest is a delicate task. Most of this article is devoted to the results obtained on the general investigation of the overall data properties, and draws attention to the approaches needed to exploit the accuracy of *Gaia* and prepare a detailed scientific exploitation.

3.1. Astrometric processing

We now describe the main steps of the astrometric processing. A more comprehensive presentation is available in the *Gaia* DR2 documentation and Lindegren et al. (2016, 2018). The basic processing of the astrometric reduction for SSOs consists of three consecutive coordinate transformations.

The first step in the processing of the astrometry is the computation of the epoch of observations, which is the reconstructed timing of crossing of the central line of the exposure on the CCD. The first coordinate transformation is the conversion from the Window Reference System (WRS) to the Scanning Reference System (SRS). The former consist of pixel coordinates of the SSO inside the transmitted window along with time tagging from the On Board Mission Timeline (OBMT), the internal time scale of *Gaia* (Lindegren et al. 2016). The origin of the WRS is the reference pixel of the transmitted window. The SRS coordinates are expressed as two angles in directions parallel and perpendicular to the scanning direction of *Gaia*, and the origin is a conventional and fixed point near the centre of the focal plane of *Gaia*.

The second conversion is from SRS to the centre-of-mass reference system (CoMRS), a non-rotating coordinate system with origin in the centre of mass of *Gaia*.

The CoMRS coordinates are then transformed into the barycentric reference system (BCRS), with the origin in the barycentre of the solar system. The latter conversion provides the instantaneous direction of the unit vector from *Gaia* to the asteroid at the epoch of the observation after removal of the annual light aberration (i.e., as if *Gaia* were stationary relatively to the solar system barycenter). These positions, expressed in right ascension (α) and declination (δ), are provided in DR2. They are similar to astrometric positions in classical ground-based astrometry.

A caveat applies to SSO positions concerning the relativistic bending of the light in the solar system gravity field. In *Gaia* DR2, this effect is over-corrected by assuming that the target is at infinite distance (i.e. a star). In the case of SSOs at finite distance, this assumption introduces a small discrepancy (always <2 mas) that must be corrected for to exploit the ultimate accuracy level.

3.2. Filtering and internal validation

An SSO transit initially includes at most nine positions, each corresponding to one AF CCD detection (see Sect. 2). However, in many cases, fewer than nine observations in a transit are available in the end. The actual success of the astrometric reduction depends on the quality of the recorded data: CCD observations of too low quality are quickly rejected; the same holds true if an observation occurs in the close vicinity of a star or within too short a time from a cosmic ray event, the software fails to produce a good position.

These problems represent only a small part of all the possible instances encountered in the astrometric processing, which has required an efficient filtering. Observations have been carefully analysed inside the pipeline to ensure that positions that probably do not come from an SSO are rejected, as well as positions that do not meet high quality standards. We applied the filtering both at the level of individual positions and at the level of complete transits. We list the main causes of rejection below.

- Problematic transit data. The positions were rejected when some transit data were too difficult to treat or if they gave rise to positions with uncertain precision.
- Error-magnitude relation. Positions with reported uncertainties that were too large or too small for a given magnitude are presumably not real SSO detections, and they were discarded.
- No linear motion. At a solar elongation of more than 45° , an SSO should show a linear motion in the sky during a single transit, where linear means that both space coordinates are linear functions of time. We considered all those positions to be false detections that did not fit the regression line to within the estimated uncertainties.
- Minimum number of positions in a transit. The final check was to assess how many positions were left in a transit. For *Gaia* DR2, we set the limit to two because we relied on an a priori list of transits to be processed (see Sect. 2.1). SSOs have also gone through a further quality check and filtering according to internal processing requirements established to take into account some expected peculiarities of SSO signals.

Three control levels were implemented:

- Standard window checking. Only centroids/fluxes from windows with standard characteristics were accepted and transmitted to the following step of the processing pipeline.
- Checking of the quality codes in the input data, resulting from the signal centroiding. Only data that successfully passed the centroid determination were accepted.
- A filtering depending on the magnitude and apparent motion of the source and the location of its centroid inside the window in order to reject observations with centroids close to the window limits, where the interplay between the distortion of the PSF due to motion and the signal truncation would introduce biases in centroid and flux measurements.

3.3. Error model for astrometry

Between CCD positions within a transit, the errors are not entirely independent, since in addition to the uncorrelated random noise, there are some systematics, like the attitude error, that have a coherence time longer than the few seconds interval between two successive CCDs. This induces complex correlations between the errors in the different CCDs from the same transit that are practically impossible to account for rigorously. Hence, we adopted a simplified approach separating the error into a systematic and a random part. Systematic errors are the same for all positions of the same transit, while random errors are statistically independent from one CCD to another. One of the main error sources is the error from the centroiding. It is propagated in the pipeline down from the signal processing in pixels in the coordinate system (AL, AC), and it is eventually converted into right ascension and declination. The errors in AL and AC are usually uncorrelated, but the rotation from the system (AL, AC) to the system $(\alpha \cos \delta, \delta)$ makes them highly correlated.

Along-scan uncertainties are very small (of the order of 1 mas), and they show the extreme precision of *Gaia*. The error on the centroiding represents the main contribution to the random errors for SSOs fainter than magnitude 16. For SSOs fainter than magnitude 13, all pixels are binned in AC to a single window, and the only information we have is that the object is inside the window. Therefore the position is given as the centre of the window, and the uncertainty is given as the dispersion of a rectangular distribution over the window. The errors in AC are thus very large (of the order of 600 mas) and highly non-Gaussian. For SSOs brighter than magnitude 13, the uncertainty in AC is smaller. In these cases, a 2D centroid fitting is possible, but the error in AC is generally still more than three times larger than in AL direction, essentially because of the shape of the *Gaia* pixels.

An important consequence is that uncertainties given in the $(\alpha \cos \delta, \delta)$ coordinate system may appear to be large as a result of the large uncertainties in AC, which contributes to the uncertainty in both right ascension and declination after the coordinate transformation.

Other errors also affect the total budget, such as the error from the satellite attitude and the modelling errors that are due to some corrections that are not yet fully calibrated or implemented. They contribute to both the random and the systematic error and are of the order of a few milliarcseconds.

4. Asteroid photometry in *Gaia* DR2

The *Gaia* Archive provides asteroid magnitudes in *Gaia* DR2 in the *G* band (measured in the AF *white band*), for 52% of the observations. This fraction is a result of a severe selection that is described below.

Asteroids, due to their orbital motion, move compared to stellar sources on the focal plane of *Gaia*. Hence, it is possible that they can drift out of the window during the observations of the AFs. This drift can be partial or total, resulting in potential loss of flux during the AF_1, \dots, AF_x with $x > 1$ observations. Asteroid photometry at this stage is processed with the same approach as is used for stellar photometry (Carrasco et al. 2016; Riello et al. 2018) and no specific optimisation is currently in place to account for flux loss in moving sources. This situation is expected to improve significantly in the future *Gaia* releases.

The photometry of *Gaia* DR2 is provided at transit level: the brightness values (magnitude, flux, and flux error) repeat identically for each entry of the *Gaia* archive that is associated with the same transit. The transit flux is derived from the

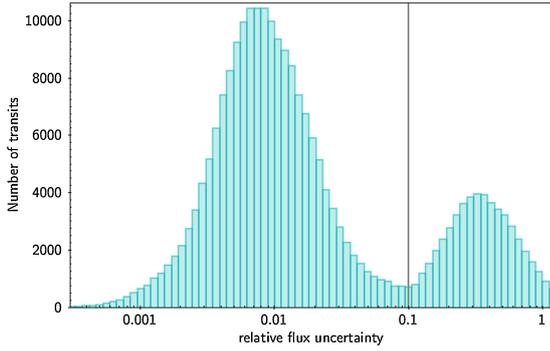


Fig. 8. Relative error in magnitude σ_G for the whole sample of transit-level G values. The vertical line at $\sigma_G \sim 0.1$ represents the cut chosen to discard the data with low reliability.

average of the calibrated fluxes recorded in each CCD strip of the AF, weighted by the inverse variance computed using the single CCD flux uncertainties. This choice minimises effects that are related, for instance, to windows that are off-centred with respect to the central flux peak of the signal. However, when the de-centring becomes extreme during the transit of a moving object, or worse, when the signal core leaves the allocated window, significant biases propagate to the value of the transit average and increase its associated error. This happens in particular for asteroids whose apparent motion with respect to stars is non-negligible over the transit duration. A main-belt asteroid with a typical motion of 5 mas s^{-1} drifts with respect to the computed window by several pixels during the $\approx 40\text{s}$ of the transit in the *Gaia* FOV.

As provided by the photometric processing, a total of 234 123 transits of SSOs have an associated, fully calibrated magnitude (81% of the total). Figure 8 shows the distribution of the relative error per transit σ_G of the whole dataset before filtering. We found out that the sharp bi-modality in the distribution correlates positively with transits of fast moving objects. For this reason, we decided to discard all transits that fell in the secondary peak of large estimated errors $\sigma_G > 10\%$ as they almost certainly correspond to fluxes with a large random error and might be affected by some (unknown) bias.

A second rejection was implemented on the basis of a set of colour indices, estimated by using the red and blue photometer (RP and BP), the two low-resolution slitless spectrophotometers. Again due to asteroid motion, the wavelength calibration of RP/BP can be severely affected, and this in turn can affect the colour index that is used to calibrate the photometry in AF. In future processing cycles, when the accurate information on the position of asteroids, produced by the SSO processing system, will become available to the photometric processing, we expect to have a significant improvement in the calibration of the low-resolution spectra and photometric data for these objects. After checking the distribution of the observations of SSOs on a space defined by three colour indices (BP-RP, RP-G, and G-BP), we decided to discard the photometric data falling outside a reasonable range of colour indices, corresponding to the interval (0.0, 1.0) for both RB-G and G-RP.

The two criteria above, based on the computed uncertainty and on the colour, are not independent. Most transits that were rejected due to poor photometry in the G band also showed colour problems, which proves that the two issues are related.

Both filtering procedures together result in the rejection of a rather large sample of 48% of the initial brightness measurements available. In the end, 52% of the the transits

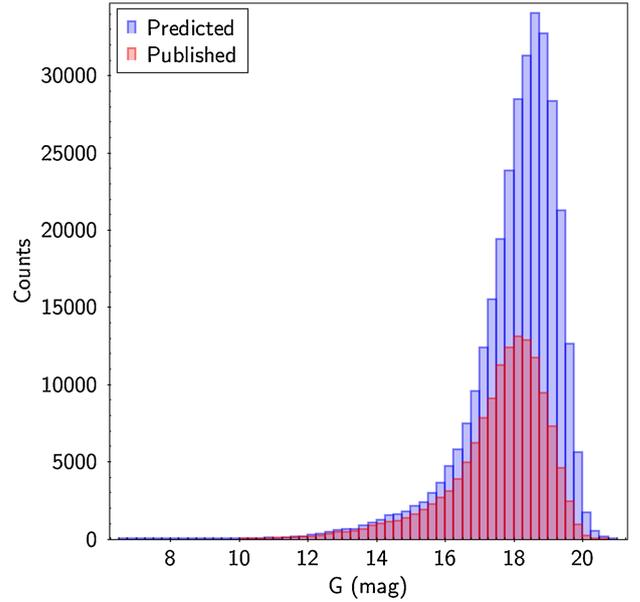


Fig. 9. Distribution of the apparent magnitude of the SSOs in *Gaia* DR2 at the transit epochs. For the whole sample the brightness derived from ephemerides (adopting the (H, G) photometric system) is provided (label: "predicted"). The sub-sample contains the magnitude values that are published in *Gaia* DR2. The shift of the peak towards brighter values indicates a larger fraction of ejected values among faint objects.

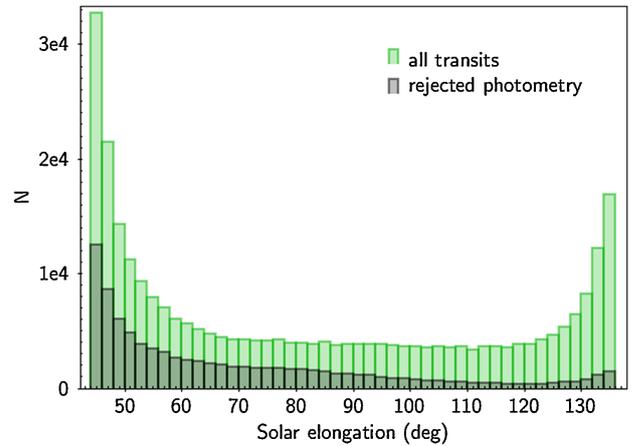


Fig. 10. Distribution of the asteroid sample in *Gaia* DR2 as a function of solar elongation. The whole sample is compared to the sub-sample of asteroids with rejected photometric results (histogram of lower amplitude).

of SSOs in *Gaia* DR2 have an associated G -band photometry (Fig. 9).

Figure 10 shows the difference in distribution of solar elongation angles, between the entire *Gaia* DR2 transit sample and the transits for which the magnitude is rejected. Figure 11 shows the same comparison on the AL velocity distribution. The majority of rejections occurs at low elongations, where their average apparent velocity is higher.

The resulting distribution of phase angles and reduced magnitudes (G_{red} , at 1 au distance from *Gaia* and the Sun) for the transits in *Gaia* DR2 is plotted in Fig. 12. In addition to the core of the distribution represented by MBAs, a small sample of NEAs reaching high phase angles is visible, as well as some transits associated with large TNOs at the smallest phase angles.

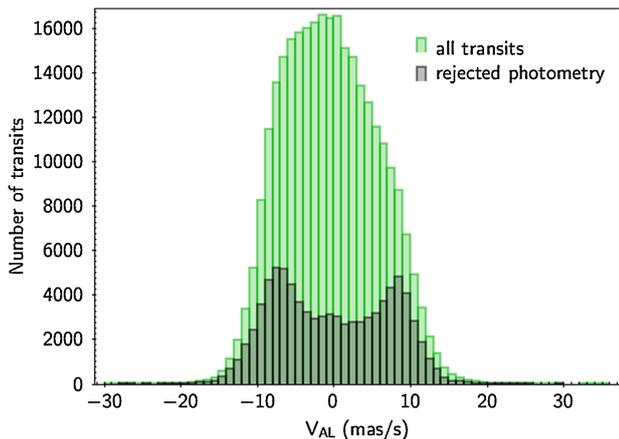


Fig. 11. Distribution of the asteroid sample in *Gaia* DR2 as a function of AL velocity. The whole sample is compared to the sub-sample of asteroids with rejected photometric results (histogram of lower amplitude).

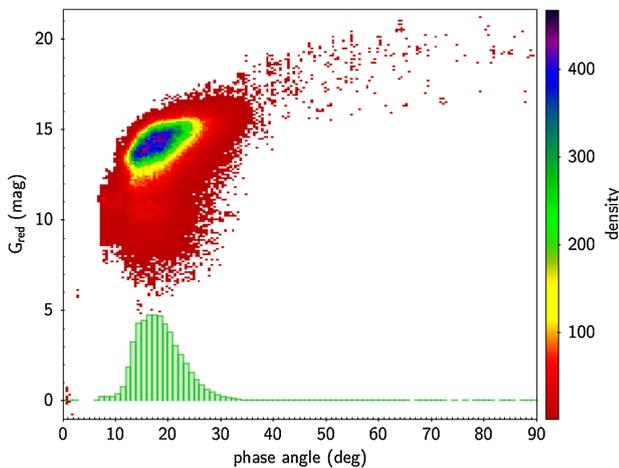


Fig. 12. Reduced asteroid magnitude as a function of phase angle. The histogram of phase angles is superposed on the bottom part (arbitrary vertical scale).

Despite the severe rejection of outliers, assessing the reliability of the published photometry at the expected accuracy of *Gaia*, specifically for solar system bodies, is not straightforward. The intrinsic variability of the asteroids due to their changing viewing and illumination geometry and to their complex shapes makes the comparison of observed fluxes with theoretical ones very challenging. Sunlight scattering effects from the asteroid surfaces also play a role and must be modelled to reproduce the observed brightness.

We attempted to model the observed brightness following two different approaches, on a small sample of asteroids. First, we used a genetic inversion algorithm derived from a full inversion algorithm developed by Cellino et al. (2009) and massively tested by Santana-Ros et al. (2015) to derive for a few selected objects the best-fitting three-axial ellipsoid (axis ratios) from *Gaia* observations alone. The procedure assumes known values of the spin period and spin-axis direction (“asteroid pole”) available in the literature for objects that have been extensively observed from the ground, and takes into account a linear phase-magnitude dependence. The procedure is extensively explained in the *Gaia* DR2 documentation.

Independently, we exploited the detailed shape models available for the two asteroids (21) Lutetia and (2867) Šteins derived by combining ground-based data with those obtained during the ESA Rosetta flybys to reproduce their observed *Gaia* brightness. Both attempts, of course, concern modelling the flux variations relative to a given observation in the sample, not its absolute value.

The results from the sparse photometry inversion are presented in Figs. 13–15. They are obtained by assuming a Lommel-Seeliger scattering law, a realistic choice when a more detailed mapping of the scattering properties across the surface is not available (Muinonen & Lumme 2015; Muinonen et al. 2015).

Despite the very simplified shape model, the residuals (observations minus computed) O-C are always within ± 0.05 magnitudes, and the typical scatter can be estimated around 2–3%. Using the shape models of (21) Lutetia (Carry et al. 2010) and (2867) Šteins (Jorda et al. 2012), we tried to assess the photometric accuracy limit of *Gaia* on asteroids. In the case of (21) Lutetia, it was found that *Gaia* data are in very good agreement with expectations based upon the best available shape model of this asteroid, derived from disk-resolved imaging by Rosetta (which only imaged one hemisphere of the object) and a lower-resolution model based on disk-integrated, ground-based photometry. The high-resolution shape model reproduces the *Gaia* photometry with a small RMS value of 0.025 mag, corresponding to 2.3% RMS in flux. This strongly suggests that *Gaia* photometry is probably better than 2% RMS, within the limitations imposed by the shape model accuracy and the assumptions on the scattering model. Moreover, *Gaia* data seem to offer an opportunity to improve the currently accepted shape solution for Lutetia, which is based partly upon ground-based data.

The results obtained for (2867) Šteins, for which a high-resolution shape model is also available, strongly support the conclusion that the photometry is indeed very accurate. For (2867) Šteins two pole solutions exist, essentially differing only by the value of the origin of the rotational phase. By directly using the shape model to reproduce *Gaia* data, resampled at 5° resolution, with a Lommel-Seeliger scattering corresponding to E-type asteroid phase functions, the RMS value of the O-C is 1.64% and 1.51% for the two pole solutions, a very good result. Changing the resolution to 3° does not improve the fit further. The remaining limitations in the case of (2867) Šteins are still related to details of the shape, and to the assumptions made (and/or scattering properties) when it was derived from Rosetta images.

In conclusion, our validation appears to show that *Gaia* epoch photometry, appropriately filtered to eliminate the outliers, probably has an accuracy below 1–2% up to the magnitude of (2867) Šteins, in the range G 17–19. However, given the current limitations on the calibration and processing, we cannot exclude that the sample published in *Gaia* DR2 still contains a non-negligible fraction of anomalous data. For this reason, we recommend detailed analysis and careful checks for any applications based on *Gaia* DR2 photometry of asteroids.

5. Validation of the astrometry

The processing of the solar system data described above has eventually produced a list with 14 124 objects (all numbered SSOs), 290 704 transits, and 2 005 683 CCD observations. The sky distribution is shown in Fig. 16 in a density plot in equatorial coordinates. As expected, most SSOs are found in a

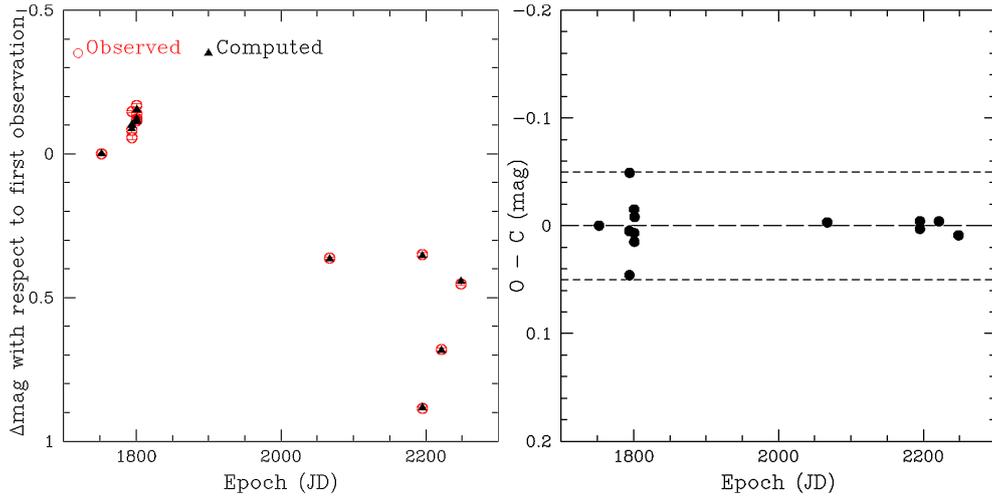


Fig. 13. Observed and computed magnitude from the best fit of *Gaia* observations of an ellipsoidal model for the asteroid (39) Laetitia. In the *right panel*, we show the corresponding residuals. The origin of the time axis is J2010.0.

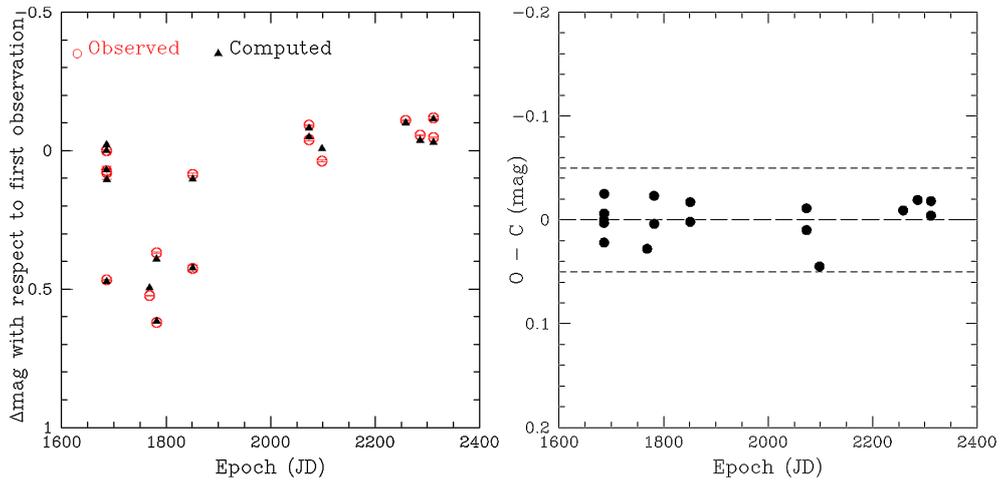


Fig. 14. As in Fig. 13 for the asteroid (283) Emma.

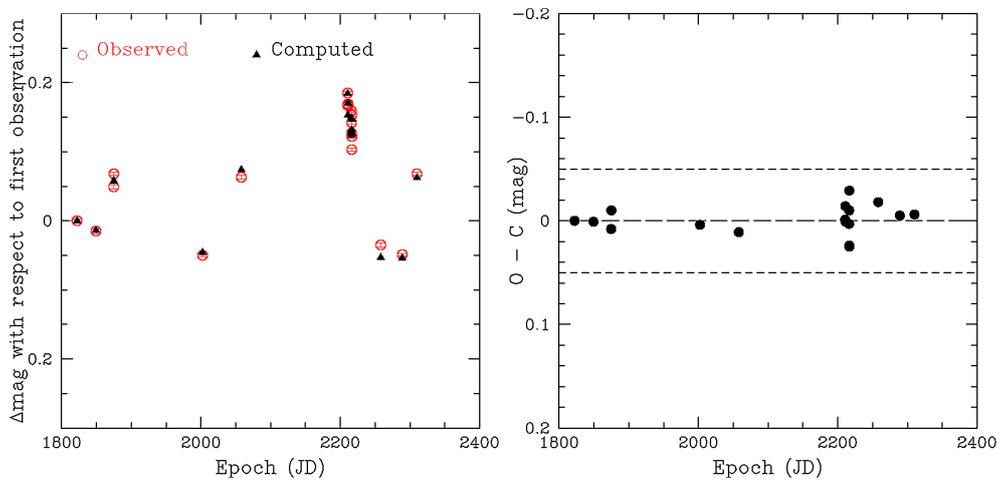


Fig. 15. As in Fig. 13 for the asteroid (704) Interamnia.

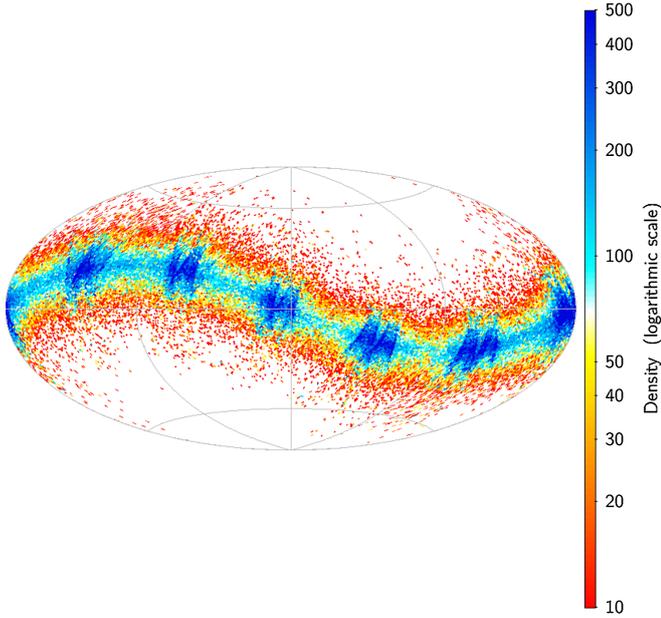


Fig. 16. Sky distribution (equatorial coordinates) of the 2005 683 observations for the 14 124 asteroid in the validation sample. This sky map uses an Aitoff projection in equatorial (ICRS) coordinates with $\alpha = \delta = 0$ at the centre, north up, and α increasing from right to left. The observation density is higher in blue areas. The pattern in ecliptic longitude is a consequence of the *Gaia* scanning law over a small fraction of the five-year nominal mission.

limited range of ecliptic latitudes. The distribution in longitude is not uniform because over a relatively short duration of 22 months, the *Gaia* scanning returned to the same regions of the sky, only in a limited number of areas.

Assessing the quality of the astrometry is challenging, and it needs an ad hoc treatment. Various filters have been applied during the activity of the astrometric reduction processing. The filtering process ensures the rejection of a maximum number of bad detections, while keeping the number of good positions that are rejected as small as possible (for more details, see the *Gaia* DR2 documentation). To prove that *Gaia* is already close to the performances expected at the end of the mission, we designed an ad hoc procedure for the external validation of the results. To this end, we fitted an orbit (initialising the fit with the best existing orbit) using only the available 22 months of *Gaia* observations, and we examined the residuals in right ascension and declination, and also in AL and AC (see Sect. 5.1). The main differences between *Gaia* and ground-based observations (or any other satellite observations) can be summarised as follows:

- *Gaia* observations are given in TCB, which is the primary timescale for *Gaia*.
- Positions (right ascension and declination) are given in the BCRS as the direction of the unit vector from the centre of mass of *Gaia* to the SSOs.
- The observation accuracies are up to the order of few $\sim 10^{-9}$ radians (sub-mas level) in the AL direction.
- The error model contains the correlations in $\alpha \cos \delta$ and δ because of the rotation from the (AL, AC) plane to the $(\alpha \cos \delta, \delta)$ plane (Sect. 3.3).

5.1. Orbit determination process

The orbit determination process usually consists of a set of mathematical methods for computing the orbit of objects such as

planets or spacecraft, starting from their observations. For our validation purpose, we considered only the list of numbered asteroids for which the orbits were already well-known from ground-based (optical or radar)/satellite observations. We used the least-squares method and the differential correction algorithm (see Milani & Gronchi 2010) to fit orbits on 22 months of *Gaia* observations, using as initial guess the known orbits of these objects. To be consistent with the high quality of the data, we employed a high-precision dynamical model, which includes the Newtonian pull of the Sun, eight planets, the Moon, and Pluto based on JPL DE431 Planetary ephemerides⁴. We also added the contribution of 16 massive main-belt asteroids (see Appendix A). We used a relativistic force model including the contribution of the Sun, the planets, and the Moon, namely the Einstein-Infeld-Hoffman approximation (Moyer 2003 or Will 1993). As a result of the orbit determination process, we obtained for every object a corrected orbit fitted on *Gaia* data only together with the post-fit residuals.

The core of the least-squares procedure is to minimise the target function (Milani & Gronchi 2010),

$$Q = \frac{1}{m} \xi^T W \xi, \quad (1)$$

where m is the number of observations, ξ are the residuals (observed positions minus computed positions), and W is the weight matrix. The solution is given by the normal equations,

$$C = B^T W B; \quad D = -B^T W \xi \quad \left(B = \frac{\delta \xi}{\delta \mathbf{x}} \right), \quad (2)$$

where \mathbf{x} is the vector of the parameters to be solved for. The differential corrections produce the adjustments $\Delta \mathbf{x}$ to be applied to the orbit:

$$\Delta \mathbf{x} = C^{-1} D.$$

It is clear from Eqs. (1) and (2) that the weight matrix plays a fundamental role in the orbit determination. It is usually the inverse of a diagonal matrix (Γ) that contains on the diagonal the square of the uncertainties in right ascension and declination for each observation, according to the existing debiasing and error models (as in Farnocchia et al. 2015). Each *Gaia* observation comes with its uncertainties on both coordinates and the correlation, which are key quantities in the orbit determination process. Therefore the weight matrix in our case is $W = \Gamma^{-1}$, where

$$\Gamma = \begin{bmatrix} \sigma_{\alpha_1}^2 & \text{cov}(\alpha_1, \delta_1) & 0 & \cdots & 0 \\ \text{cov}(\alpha_1, \delta_1) & \sigma_{\delta_1}^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \cdots & \sigma_{\alpha_m}^2 & \text{cov}(\alpha_m, \delta_m) \\ 0 & 0 & \cdots & \text{cov}(\alpha_m, \delta_m) & \sigma_{\delta_m}^2 \end{bmatrix}.$$

The uncertainties used to build the W matrix are given by the random component of the error model, but we also take into account the systematic contribution when this is needed, as explained in the following section.

⁴ We also performed the orbit determination process using INPOP13c (Fienga et al. 2014) ephemerides and did not find significant differences in the results.

5.2. Outlier rejection procedure

The rejection of the outliers is a fundamental step in the orbit determination procedure. Since we assumed that the residuals are distributed as normal variables, the rejection was based on the post-fit χ^2 value for each observation, computed as in [Carpino et al. \(2003\)](#):

$$\chi_i^2 = \xi_i \gamma_{\xi_i}^{-1} \xi_i^T \quad i = 1, \dots, m,$$

where m is the total number of observations, ξ_i is the vector of the residuals for each observation, and γ_{ξ_i} is the expected covariance of the residuals. Each χ_i^2 has a distribution of a χ^2 variable with two degrees of freedom. We call outlier each observation whose χ^2 value is greater than 25. The choice of 25 as a threshold was driven by the fact that we wished to keep as many good observations as possible and wished to discard only the observations (or the transits) that are very far from the expected *Gaia* performances. During this procedure, we took random and systematic errors into account.

Firstly, we rejected all the observations whose χ^2 value was greater than 25. Then, when the systematic part was larger than the random part, we performed a second step in the outlier rejection, described as follows:

- We computed the mean of the residuals for each transit.
- We checked if the value of the mean is lower than the systematic error for the transit.
- If the value was higher than the systematic error, we discarded the entire transit.
- If the value was lower than the systematic error, we computed for each observation the difference between the residual and the mean value.
- We checked whether the difference was smaller than the random error. When that was the case, we kept the observation, otherwise we discarded it.

This approach is consistent with the uncertainties produced for *Gaia* DR2. Its underlying hypothesis is that the error correlations over a single transit can be completely represented by just one quantity, which is the value of the systematic component alone. Although this is an approximation, we currently do not have the impression that a more complex correlation model is required.

5.3. Results

We fitted the orbits of the 14 124 asteroids contained in the validation sample using an updated version of the *OrbFit* software⁵, developed to handle *Gaia* observations and *Gaia* error model (Sects. 5.1 and 5.2).

We were unable to fit the observations to the existing orbit for only three asteroids because the time span covered by the available transits was too short. They were removed from the final output of *Gaia* DR2. We also removed 22 bright objects (transits with $G < 10$) whose residuals were substantially larger than the uncertainties we expected, and we considered these solutions as not reliable.

Figure 17 shows the distribution in semi-major axis and eccentricity of the 14 099 SSOs published in *Gaia* DR2. We can distinguish the NEA population ($q < 1.3$) from the MBAs (including Jupiter trojans in this class) and the TNOs ($q > 28$).

The total number of fitted observations is 2 005 683, which corresponds to 290 704 transits. During the outlier rejection procedure (Sect. 5.2), we discarded 27 981 observations ($\sim 1\%$ of the total). Figure 18 shows all the residuals in the (AL, AC) plane in

⁵ <http://adams.dm.unipi.it/orbfit/>

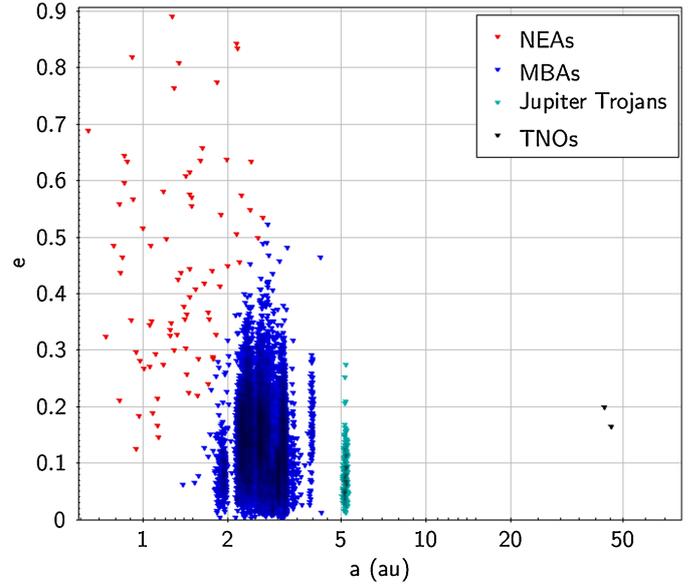


Fig. 17. Distribution of the 14 099 asteroids published in *Gaia* DR2 in semi-major axis a (au) and eccentricity e . The sample shows that all the broad categories of SSOs are represented (NEAs, MBAs, Jupiter trojans, and TNOs).

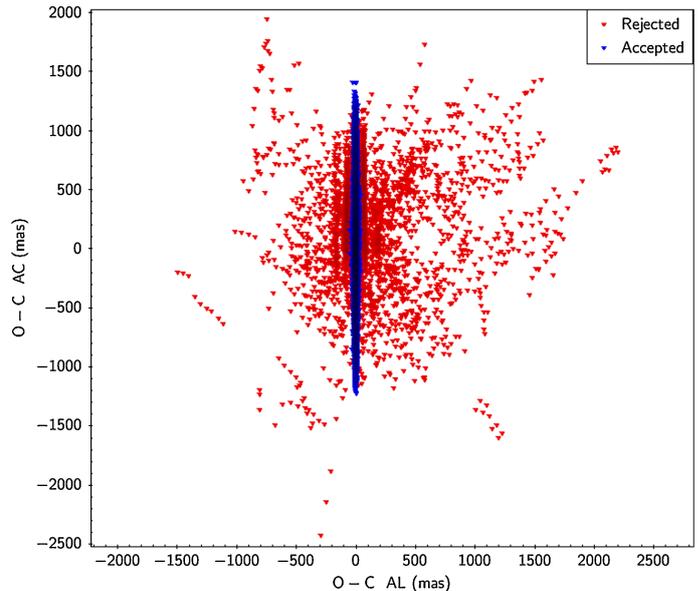


Fig. 18. Residuals in the (AL, AC) plane in milliarcseconds. Outliers are marked in red, while the blue thick line in the middle contains all the residuals for the accepted observations. The total number of fitted observations is 2 005 683, and there are 27 981 outliers ($\sim 1\%$ of the total).

mas. Outliers are the red points, while accepted observations are all contained in the blue thick line in the centre of the figure.

After the filtering and the outlier rejection, *Gaia* DR2 contains 1 977 702 observations, corresponding to 14 099 SSOs and 287 904 transits. Figure 19 represents a density plot of the residuals at CCD level in the (AL, AC) plane, for all the observations published in *Gaia* DR2. This plot, together with the plots of the residuals (Figs. 20 and 21), shows the epoch-making change brought about by *Gaia* astrometry: 96% of the AL residuals fall in the interval $[-5, 5]$ mas and 52% are at sub-milliarcsecond

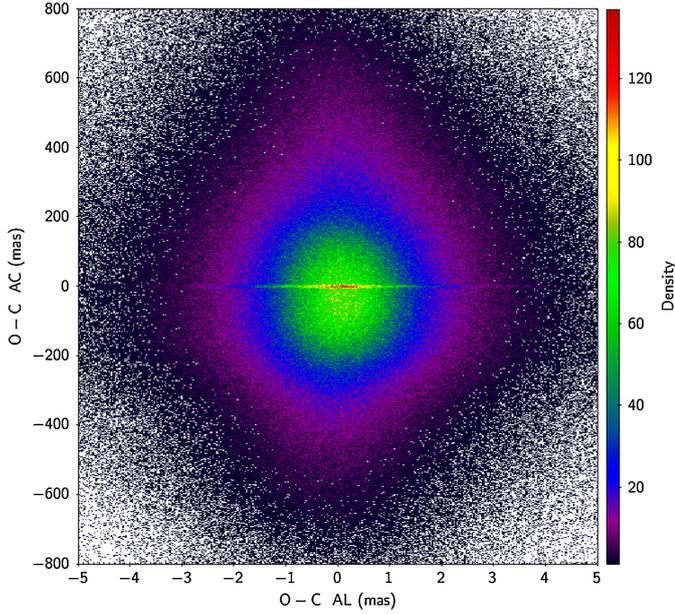


Fig. 19. Zoom-in of the density plot of the residuals in the (AL, AC) plane expressed in milliarcsecond for all the observations published in *Gaia* DR2. 96% of the AL residuals fall in the interval $[-5, 5]$, and 52% are at sub-milliarcsecond level. Almost all the residuals in AC (98%) fall in the interval $[-800, 800]$.

level. The behaviour of the residuals in AC is markedly different as a result of the geometry of the spacecraft observations. The AC residuals as a rule are much larger than AL, for the reasons detailed in Sect. 3.3. Figures 20 and 21 highlight the performances of *Gaia* even better. They show the residuals in AL and AC with respect to the G magnitude of the asteroids at the epoch of observation. The right panels of both Figs. 20 and 21 display the histogram of the residuals in AL and AC, respectively. The clear peak around 0 is well visible in the histogram of the residuals in AL (Fig. 20), and it corresponds to the core of the density plot in Fig. 19.

Even the histogram of the residuals in AC (Fig. 21) shows a peak around 0, which is strictly related to the distribution of the residuals as a function of the G magnitude. For objects brighter than $G = 13$, the full 2D window is transmitted (see Sect. 2), thus the accuracies in AC and AL are similar (although still slightly larger in the across direction because of the pixel size), while for objects fainter than $G = 13$, the errors in AC are much larger.

6. Interpretation of asteroid residuals

The residuals in AL, as in Figs. 19 and 20, show the quality and extreme precision of *Gaia* observations. We now examine the residuals in more detail, using first as an example a main-belt asteroid, and then showing the properties at transit-level.

6.1. CCD-level residuals

We chose (367) Amicitia as a test case to analyse the residuals in right ascension and declination and to explain their relationship (and main differences) with the residuals in AL and AC. Residuals in right ascension and declination are important because they are the direct output of the fit and are probably more easily understood than the residuals in AL and AC, which are closely related to the *Gaia* scanning mode. The residuals in right ascension and declination are by-products of the orbit correction and

computed as the differences between the observed and computed positions. Afterwards, given the position angle shown in Fig. 6, we can rotate the residual vector in $\alpha \cos \delta$ and δ to compute the residuals in AL and AC (Eq. (3)),

$$\begin{pmatrix} \Delta AL \\ \Delta AC \end{pmatrix} = \begin{bmatrix} \sin(PA) & \cos(PA) \\ -\cos(PA) & \sin(PA) \end{bmatrix} \begin{pmatrix} \Delta \alpha \cos \delta \\ \Delta \delta \end{pmatrix}. \quad (3)$$

For the sake of simplicity, we call $(\Delta \alpha \cos \delta, \Delta \delta)$ the vector of the residuals in $\alpha \cos \delta$ and δ and $(\Delta AL, \Delta AC)$ the vector of the residuals in AL and AC.

Asteroid (367) Amicitia is a typical object among SSOs observed by *Gaia*: its average G magnitude is 14, which means that it is not one of the brightest objects for which we have the 2D window. It has 22 *Gaia* transits, which corresponds to 132 CCD observations. Figure 22 shows all the residuals in right ascension and declination. Each transit is represented by a different symbol.

Figure 23 shows the residuals rotated in the (AL, AC) plane. They are clearly very small in AL and considerably larger in AC (as expected for an asteroid of that magnitude with almost no across-scan data). The majority of the residuals in AL are at sub-milliarcsecond level (Fig. 23, right panel), which means that almost all are inside 1σ (dashed vertical lines).

We have already mentioned the particular features of the residuals in the (AL, AC) plane (Sects. 3.3 and 5.3). We now focus on the residuals in $(\alpha \cos \delta, \delta)$. They are a linear combination of the residuals in AL and AC as in Eq. (3). Thus they are in general larger in both coordinates (Fig. 22) and highly correlated. Even if the residuals in the (AL, AC) plane are uncorrelated, the large difference in their standard deviations gives rise to a very strong correlation between the residuals expressed as $(\Delta \alpha \cos \delta, \Delta \delta)$. This strong correlation between $\Delta \alpha \cos \delta$ and $\Delta \delta$ is expressed visually by the fact that the residuals for each transit lie roughly on a straight line.

The error ellipses associated with each observation, and thus with each residual, are shown in Fig. 24. They are very elongated as a result of the large errors in AC, and they are not parallel to the axes because of the orientation of the *Gaia* scanning law, which changes from one transit to the next.

To explain the distance of each point from the centre (0,0) of the residuals for each transit, we illustrate the detailed properties of a single transit (Fig. 25). To this end, we consider a single transit. We chose one of the transits for which none of the nine CCD observations (AF1–AF9) was rejected during the processing (Fig. 25). In this case, the asteroid motion was not fast enough to move it outside the window during the transit.

The main striking feature is the alignment of the residuals. The direction of maximum spread corresponds to the motion across scan, as oriented during the transit. The AC motion of the SSOs is due both to the slow precession of the spin axis of *Gaia* as imposed by the NSL and to the proper motion of the asteroid.

Another important feature is the order of the AF CCDs, provided by their numbering, in the sequence of residuals. While on average the drift proceeds in the figure from AF1 at the bottom right towards AF9 at top left, the sequence is sometimes inverted in direction. This is an effect of the quantisation of the window position AC at integer steps of one full sample size.

The compact clustering of the residuals shows that the in the AL direction, the determination of the position during the transit has a very low spread, as expected thanks to the much higher accuracy.

The last point we need to analyse is the trend of the residuals in the negative side of the Cartesian plane. The black vectors