Highlights of Spanish Astrophysics VIII, Proceedings of the XI Scientific Meeting of the Spanish Astronomical Society held on September 8–12, 2014, in Teruel, Spain. A. J. Cenarro, F. Figueras, C. Hernández-Monteagudo, J. Trujillo Bueno, and L. Valdivielso (eds.)

Overview and stellar statistics of the expected Gaia Catalogue using the Gaia Object Generator

M. Palmer [1], X. Luri [1], F. Arenou [2], E. Masana [1], J. de Bruijne [3], E. Antiche [1], C. Babusiaux [2], R. Borrachero [1], P. Sartoretti [2], F. Julbe [1], Y. Isasi [1], O. Martinez [1], A.C. Robin [4], C. Reylé [4], C. Jordi [1], and J.M. Carrasco [1]

[1] Dept. d'Astronomia i Meteorologia, Institut de Ciències del Cosmos, Universitat de Barcelona (IEEC-UB), Martí Franquès 1, E08028 Barcelona, Spain
[2] GEPI, Observatoire de Paris, CNRS, Université Paris Diderot, 5 place Jules Janssen, 92190, Meudon, France
[3] Scientific Support Office of the Directorate of Science and Robotic Exploration of the European Space Agency, European Space Research and Technology Centre, Keplerlaan 1, 2201 AZ, Noordwijk, The Netherlands
[4] Institut UTINAM, CNRS UMR 6213, Observatoire des Sciences de l'Univers THETA Franche-Comté-Bourgogne, Université de Franche Comté, Observatoire de Besançon, BP 1615, 25010 Besançon Cedex, France

Abstract

An effort has been made to simulate the expected Gaia Catalogue, including the effect of observational errors. We statistically analyse this simulated Gaia data to better understand what can be obtained from the Gaia astrometric mission. This catalogue is used to investigate the potential yield in astrometric, photometric, and spectroscopic information and the extent and effect of observational errors on the true Gaia Catalogue. The simulated catalogue of one billion objects is presented, with detailed information on the 523 million individual single stars it contains.

1 Introduction

During its five years of data collection, Gaia is expected to transmit some 150 terabytes of raw data to Earth, leading to production of a catalogue of 10^9 individual objects. After on-ground processing, the full database is expected to be in the range of one to two petabytes of data. Preparation for acquiring this huge amount of data is essential. Work has begun to model the expected output of Gaia in order to predict the content of the Gaia Catalogue, to facilitate the production of tools required to effectively validate the real data before publication, and to analyse the real data set at the end of the mission.

To this end, the Gaia Data Processing and Analysis Consortium (DPAC) has been preparing a set of simulators, including a simulator called the Gaia Object Generator (GOG), which simulates the end-of-mission catalogue, including observational errors. Here a full description of GOG is provided, including the models assumed for the performance of the Gaia satellite and an overview of its simulated end-of-mission catalogue. A selection of statistics from this catalogue is provided to give an idea of the performance and output of Gaia.

2 The simulator

One basic component of the Gaia Simulator is its Universe Model (UM), which is used to create object catalogues down to a particular limiting magnitude (in our case $G = 20$ mag for Gaia). For stellar sources, the UM is based on the Besançon galaxy model [3]. This model simulates the stellar content of the Galaxy, including stellar distribution and a number of object properties. It produces stellar objects based on the four main stellar populations (thin disk, thick disk, halo, and bulge), each population with its own star formation history and stellar evolutionary models. Additionally, a number of object-specific properties are also assigned to each object, dependent on its type. Possible objects are stars (single and multiple), stellar clusters, diffuse light, planets, asteroids, comets, resolved galaxies, unresolved extended galaxies, quasars, AGN, and supernovae. Therefore, the UM is capable of simulating almost every object type that Gaia can potentially observe. It can therefore construct simulated object catalogues down to Gaia's limiting magnitude.

Building on this, the UM creates for any time, over any section of the sky (or the whole sky), a set of objects with positions and assigns each a set of observational properties [2]. These properties include distances, apparent magnitudes, spectral characteristics, and kinematics. The GOG is capable of transforming this UM catalogue into Gaia's simulated intermediate and final catalogue data. This is achieved through the use of analytical and numerical error models (see: [1] to create realistic observational errors in astrometric, photometric, and spectroscopic parameters. In this way, GOG transforms 'true' object properties from the UM into 'observed' quantities that have an associated error that depends on the object's properties, Gaia's instrument capabilities, and the type and number of observations made.

3 Mock catalogue

The GOG simulator has been used to generate the simulated final mission catalogue for Gaia, down to magnitude $G = 20$. The simulation was performed on the MareNostrum super computer[1] at the Barcelona Supercomputing Centre (Centre Nacional de Supercomputació),

[1]The authors thankfully acknowledge the computer resources, technical expertise and assistance provided by the Red Española de Supercomputación.

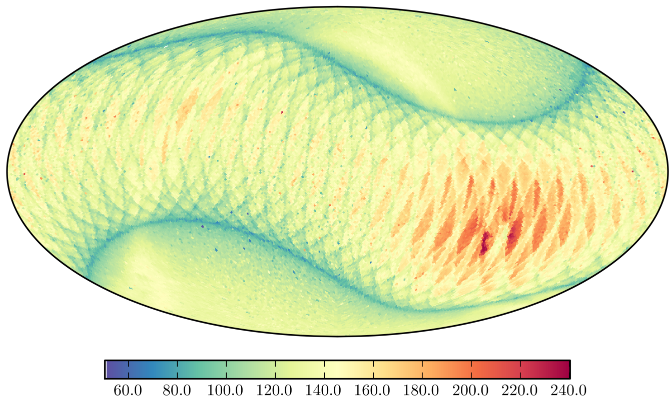

Figure 1: Sky map (healpix) of mean parallax error for all single stars in equatorial coordinates. Colour scale is mean parallax error in μas. The red area is the location of the bulge.

and it took 400 thousand CPU hours. An extensive set of validations and statistics have been produced to validate performance of the simulator. Below we include some examples of these statistics to give an overview of the expected Gaia Catalogue (further statistics can be found in [1]).

In total, GOG has produced a catalogue of about one billion objects, consisting of 523 million individual single stars and 484 million binary or multiple systems. The total number of stars, including the components of binary and multiple systems is 1.6 billion.

The error in parallax measurements for Gaia depends on the magnitude of the source, the number of observations made, and the true value of the parallax. Figure 1 shows the mean parallax error over the sky. Its shape clearly follows that of the Gaia scanning law. The red area corresponding to the region of worst precision is due to the bulge population, which suffers from high levels of reddening. The faint ring around the centre of the figure corresponds to the disk of the Galaxy, remembering that the plot is given in equatorial coordinates. The blue areas corresponding to regions of improved mean precision are areas with a higher number of observations. The characteristic shape of this plot is due to the Gaia scanning law.

In Fig. 2, the mean error for parallax, position, proper motion are given as a function of G magnitude. The sharp jumps in the mean error in astrometric parameters between 8 and 12 mag are due to the activation of gates for the brighter sources in an attempt to prevent CCD saturation.

Gaia will produce low-resolution spectra, in addition to measuring the magnitude of each source in the Gaia bands G, G_{BP}, G_{RP}, and G_{RVS}. Figure 3 shows the distribution in the error of each photometric measurement. As can be seen in this figure, the error in G is much lower than for the other instruments, and for all stars it is less than 8 mmag.

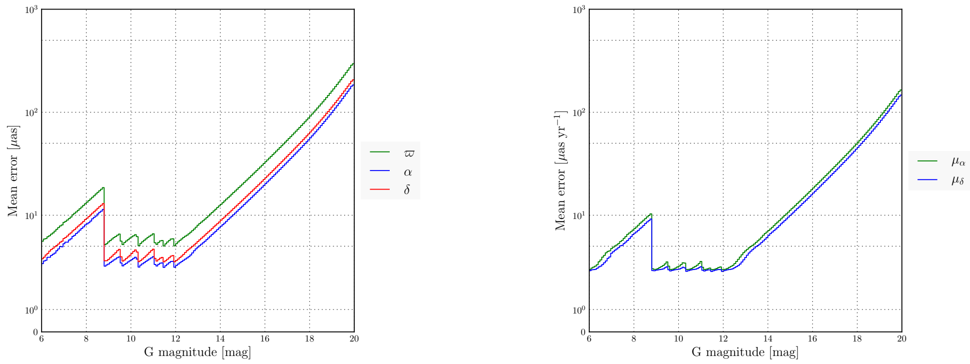

Figure 2: Mean end-of-mission error as a function of G magnitude for parallax and position (*left*), and proper motion (*right*).

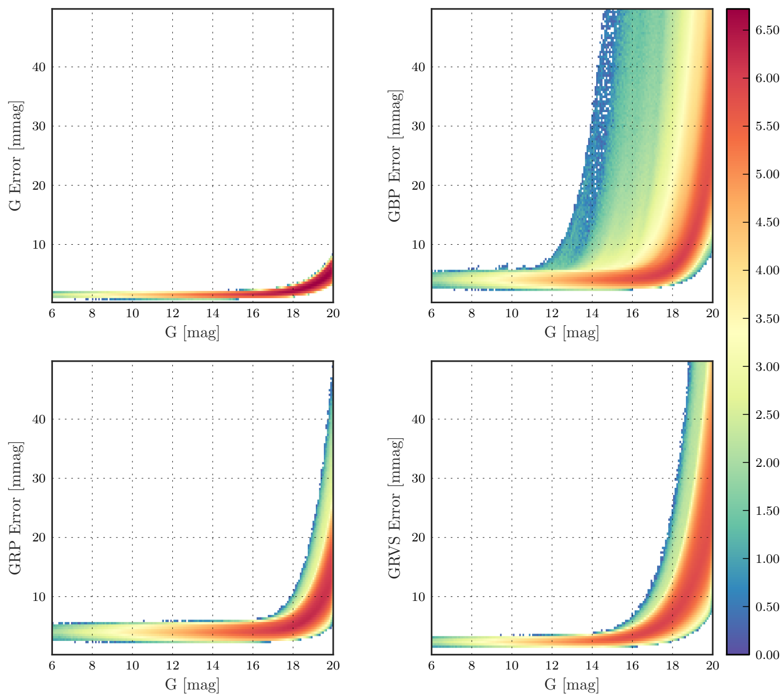

Figure 3: End-of-mission errors in photometry as a function of G magnitude. The colour scale represents log density of objects in a bin size of 80 mmag by 0.4 mmag. Top left, G magnitude; top right, G_{BP}; bottom left, G_{RP}; bottom right, G_{RVS}. White area represents zero stars.

4 Conclusions

The Gaia Object Generator provides the most complete picture to date of what can be expected from the Gaia astrometric mission. Its simulated catalogue provides useful insight into how various types of objects will be observed and how each of their observables will appear after including observational errors and instrument effects. The simulated catalogue includes directly observed quantities, such as sky position and parallax, as well as derived quantities, such as interstellar extinction and metallicity.

Additionally, the full sky simulation described here is useful for gaining an idea of the size and format of the eventual Gaia Catalogue, for preparing tools and hardware for hosting and distribution of the data, and for becoming familiar with working with such a large and rich dataset.

Here we have focussed on the simulated catalogue from the inbuilt Gaia Universe Model, based on the Besançon Galaxy model. However, GOG can alternatively be supplied with an input catalogue generated by the user. This way, simulated data from any other model can be processed with GOG to obtain simulated Gaia observations of specific interest to the individual user. The input can be either synthetic data on a specific star or catalogue, or an entire simulated survey such as those generated using Galaxia [4], provided a minimum of input information is supplied (e.g. position, distance, apparent magnitude, and colour).

With GOG, the capabilities of the instrument can be explored, and it is possible to gain insight into the expected performance for specific types of objects. While only a very small subset of the available statistics have been reproduced here, more are available in [1], and it is possible to obtain the full set of available statistics at request.

Acknowledgments

GOG is the product of many years of work from a number of people involved in DPAC and specifically CU2. The authors would like the thank the various CUs for contributing predicted error models for Gaia. The processing of the GOG data made significant use of the Barcelona Supercomputing Center (MareNostrum), and the authors would specifically like to thank Javi Castañeda, Marcial Clotet, and Aidan Fries for handling our computation and data handling needs. Additionally, thanks go to Sergi Blanco-Cuaresma for his help with Matplotlib. I thankfully acknowledge the computer resources, technical expertise and assistance provided by the Red Española de Supercomputación. This work was supported by the MINECO (Spanish Ministry of Economy) - FEDER through grant AYA2012-39551-C02-01 and ESP2013-48318-C2-1-R and the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement GREAT-ITN FP7 264895.

References

[1] Luri, X., Palmer, M., Arenou, F., et al. 2014, A&A, 566, A119

[2] Robin, A. C., Luri, X., Reyle, C., et al. 2012, A&A, 543, A100

[3] Robin, A. C., Reyle, C., Derriere, S., & Picaud, S. 2003, A&A, 409, 523

[4] Sharma, S., Bland-Hawthorn, J., Johnston, K. V., & Binney, J. 2011, ApJ, 730, 3