*Highlights on Spanish Astrophysics IX, Proceedings of the XII Scientific Meeting of the Spanish Astronomical Society held on July 18 – 22, 2016, in Bilbao, Spain. S. Arribas, A. Alonso-Herrero, F. Figueras, C. Hernández-Monteagudo, A. Sánchez-Lavega, S. Pérez-Hoyos (eds.)*

# Treatment of spurious detections in Gaia

**N. Garralda, C. Fabricius, J. Castañeda, J. Portell, M. Clotet, J.J. González-Vidal, and J. Torra.**

Institut de Ciències del Cosmos, Universitat de Barcelona (IEEC-UB), C/Martí Franquès 1, E08028 Barcelona, Spain.

## Abstract

Gaia is a space instrument that observes some 50 million point like sources per day in an autonomous way, using a rapid on-board detection software. The detection uses two sets of CCD, with a first detection followed by an explicit confirmation in order to remove detections due to e.g. cosmic rays. Nevertheless, false positives, detections of non-existing sources, or spurious detections, are frequent.

A detection classifier algorithm is used to identify the spurious observations to prevent them from entering processing stages, where they would be added as new sources in the Gaia source list. Different types of spurious detections have being identified, e.g. detections located at the diffraction spikes of a bright source, false detections due to background noise or due to cosmic rays. A big effort has been put into identifying the different types and develop a model for the detection classification.

## 1 Introduction

The Gaia satellite [5] is composed by two optical telescopes and one focal plane (with 106 CCDs) in which images from both telescopes are combined. The Gaia satellite rotates around its axis at a spin rate of 4 revolutions per day, observing all stars crossing one of the 2 telescopes. Gaia has a slow precession motion, producing a continuous drift on the sky scanned, surveying the full sky in six months, and obtaining in average around 70 different measurements of each star in the 5 years mission.

The on-board detection/confirmation/observation algorithms constitute a critical process, because Gaia observes in an autonomous way, not using an input tracking catalogue. The on-board processing power is limited and the number of operations per second is high, thus, this process has to be simple, and false positives frequently happen producing detections of non-real stars.

The on-board software was originally configured to observe up to magnitude 20, but during the mission this faint limit has been extended up to magnitude 20.7. With this higher magnitude limit, also the frequency of spurious detections has increased and for the first two years of mission it lies around 20%. The on-board configuration (rejection coefficients) has been recently updated in order to decrease to 4% the percentage of spurious detections.

The necessity of having an on-ground system dedicated to cleaning the spurious detections in the daily on ground processing was envisaged since the beginning of Gaia operations, and with this purpose, the DetectionClassifier algorithm has been developed. A simple version of this algorithm runs on a daily basis avoiding that spurious detections enter the Cross-Match step [3] where they would lead to false new entries in the preliminary Gaia Catalogue. A more sophisticated version of this algorithm runs once every processing cycle [1], after all detections are accumulated, with the same purpose but for the final Gaia Catalogue.

## 2 Detection process

The Gaia spacecraft carries three scientific instruments, the Astrometric Instrument to determine positions of stars, the Photometric Instrument to determine colour and chemical composition of stars by taking measurements of two low-resolution spectra in the red and blue ranges, and the Radial Velocity Spectrometer to determine the radial velocity of the stars by taking high-resolution spectra. The three instruments share the focal plane [4]. The 106 CCDs compose a matrix of seven rows and each one is commanded by a Video-processing unit (VPU). The VPUs are in charge of detecting, confirming, and observing objects in the nominal magnitude range (magnitude 5 to 20.7). Video-Processing algorithms were designed to detect point like sources, but also moving asteroids.

The CCDs operate in Time-delay integration (TDI) mode, reading out at the same velocity as Gaia rotates (the rotation of the satellite adapts to the TDI step), thus, a star crosses the focal plane with a regular motion. Each row contains a set of CCDs for detecting the stellar objects and a set of CCDs for taking the astro/photo/radial velocity measurements.

The detection process is done by the Sky Mapper CCDs, SM1 detects objects coming from telescope 1, and SM2 from telescope 2. The detection is based on computing the centroid location and magnitude of all visible objects, and in case of available resources, assign a window to track each detected object. The VPU uses the data gathered on the first CCD of the astrometric field (AF1) to confirm an object detection in order to discard false detections. Once a detection is confirmed, that object is followed through the entire focal plane following the CCD windowing strategy, collecting small images from each CCD.

The sampling strategy depends on the estimated on-board magnitude, two dimensional (2D) windows are acquired for bright stars up to magnitude 13, and 1D windows (collapsed in the across-scan direction) for the faint objects up to magnitude 20.7. The confirmation is by default for the 2D windows, which often contain saturated samples.

## 3 Spurious detections

Different types of spurious detections have being identified [3] since the beginning of Gaia operations:

1. Spurious detections around bright sources (magnitude < 16), located along the diffraction spikes. For Very Bright Stars (VBS), they appear not only along the diffraction spikes but also in more complicated patterns near the source. See Fig. 1 showing two examples of spurious detections around a star of magnitude 10.25 and a star of magnitude 0.14 (Rigel).

2. Spurious detections around a Solar System Object (SSO). See Fig. 2 showing spurious detections around a transit of Jupiter.

3. Spurious detections due to unexpected light paths and internal reflections within the spacecraft when a VBS or an SSO (with very bright apparent magnitude) is close to the focal plane. See Fig. 2 showing an example of spurious detections when Venus is close to the focal plane.

4. Spurious detections due to cosmic rays. This only happen for 2D windows, as the confirmation on-board is by default for that magnitude range. These are relatively harmless because they happen randomly across the sky. See Fig. 3.

5. Spurious detections due to on-board detection problems, or what we call "phantom" detections. See Fig. 3.

The dominating spuriuos type is the spurious from the diffraction spikes around bright detections, this group represents about 85-90%.

## 4 Detection classifier algorithm

The DetectionClassifier algorithm's main goal is to classify the spurious detections in order to filter them from the Cross-Match process, where each detection is linked to a source from the Gaia Catalogue, and if no suitable sources are found it will create a new entry in the Catalogue [3]. The observations are processed normally in every other respect, and downstream users may re-classify some blacklisted detections. More information about the Cross-Match in [2].

The Detection Classifier algorithm follows two different strategies:

1. Combined detections analysis: It is based on analyzing a group of detections located in the surroundings of a bright star (magnitude < 16), by comparing their location w.r.t the bright detection, along (AL) and across (AC) scan distances, and the magnitude differences. The model is a function of the magnitude, it defines a set of distances AL and AC for parametrizing the maximum longitude of each spike, a minimum density of detections along one spike, and a magnitude decay between consecutive detections within one spike. Figure 4 shows two spurious patterns for different magnitudes.

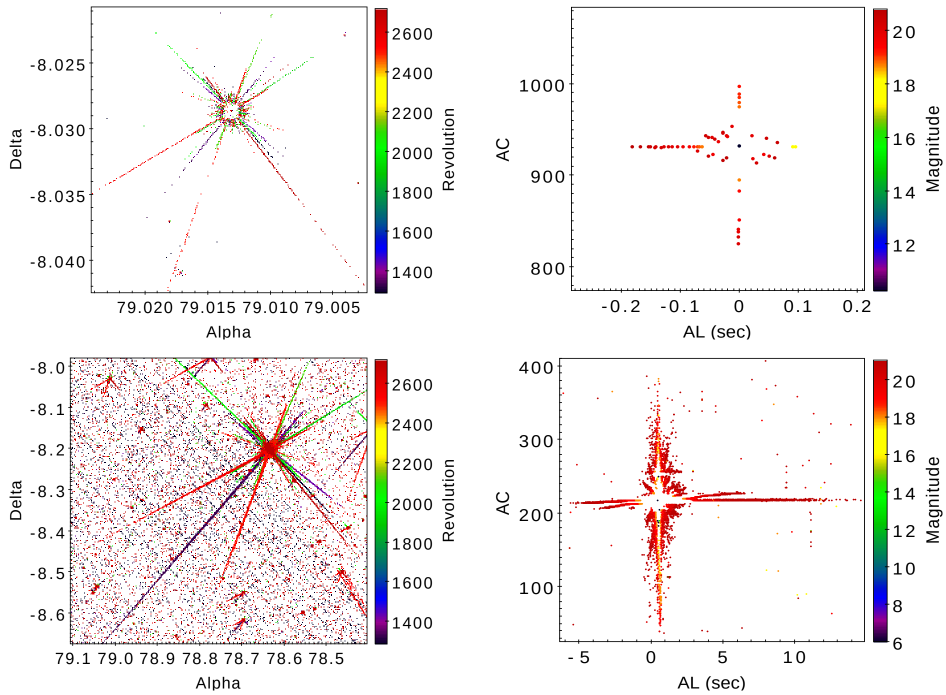


Figure 1: Spurious detections along the diffraction spikes for two different magnitudes *top*: 10.25 mag, *bottom*: 0.14 mag. *Left*: the detections from several scans with the colour scale representing the time (in revolutions), and *right*: the detections from a single scan with the colour scale representing the magnitude.

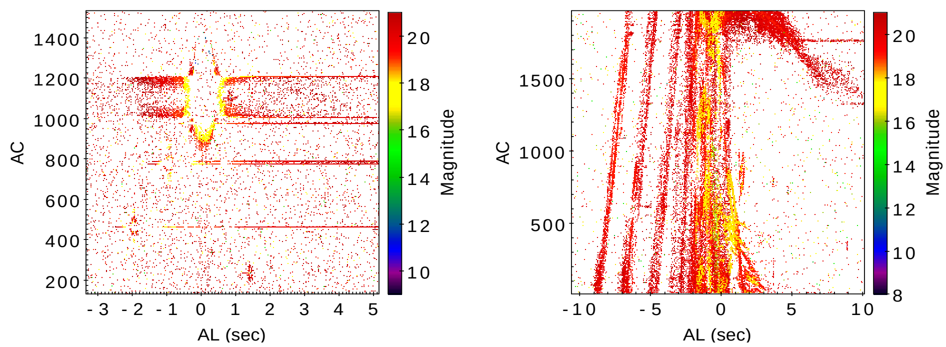


Figure 2: Spurious detections around (*left*) a transit of Jupiter, (*right*) a transit of Venus close to the focal plane.

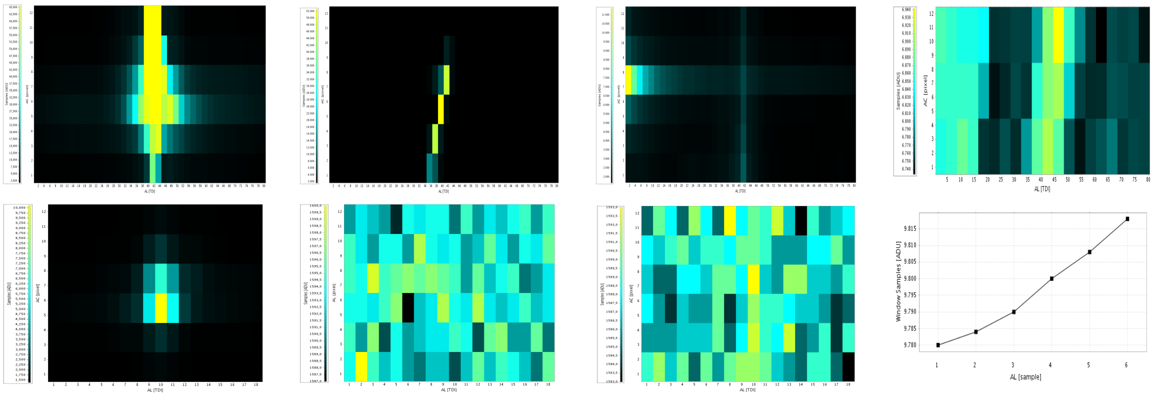

Figure 3: Windows from four transits, *top*: Sky mapper windows, *bottom*: AF1 windows. First colum is a bright real detection, the second column cosmic spurious, the third column phantom spurious and the fourth column spurious from non expected ligth paths.

Furthermore, for the classification of the spurious around a non-detected VBS, the algorithm can use the prediction of those transits as an input to trigger the processing.

For classifying spurious from SSO transits, this strategy is also followed but using a simpler model (parametrized for planet). The model defines an area (AL and AC distance) around the predicted transit of the planet and a magnitude threshold.

2. Individual raw sample analysis: It is based on analyzing the raw samples from the SM and AF1 windows from each received detection. Detections are classified as cosmics, phantoms, and detections from unexpected light paths when:

    - Cosmics: detections brighter than 14 mag in which the AF1 window is almost empty. AF1 raw flux is computed and compared with the assigned on-board magnitude.
    - Phantom: detections brighter than 10 mag without saturated samples in SM.
    - Detections from unexpected light paths: detections with unexpected across (AC) and along (AL) window samples profiles. The raw samples from SM and AF windows are collapsed in AC and AL, and correlated with an expected signal in order to detect windows without the flux accumulated in the window centre.

## 5 Conclusions

The DetectionClassification alorithm is designed to be flexible enough to cope with the different spurious patterns seen before and after the on-board detection parameter configuration change. The implemented software runs on a daily basis providing reliable results, and the process is repeated later in the mission, during the cyclic processing to reach higher reliability after accumulating all the data received until that moment. For the first Gaia Data Release (DR1), around 20% of the detections have been classified as spurious detections.

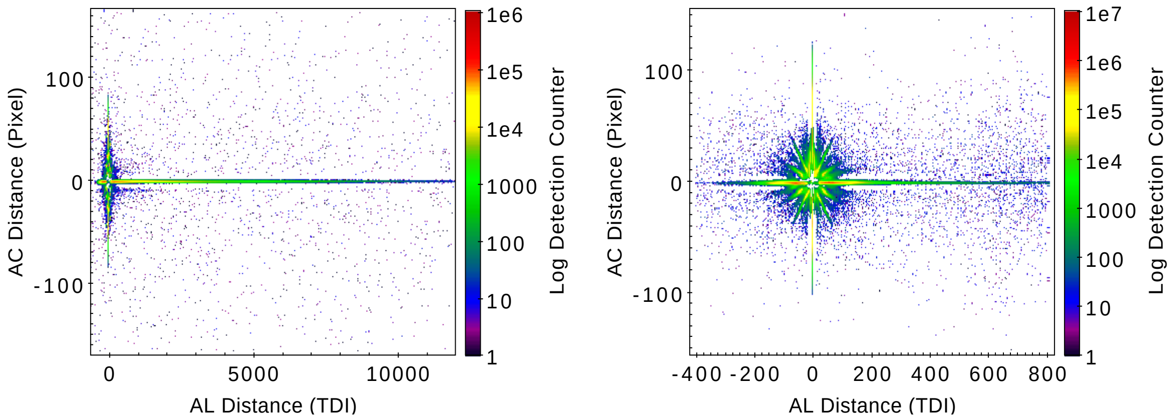


Figure 4: Maps of the accumulated faint detections around brighter detections. *Left*: bright range 8 to 9 mag., *rigth*: bright range 11 to 12 mag.

## Acknowledgments

This work was supported by the MINECO (Spanish Ministry of Economy) - FEDER through grant ESP2014-55996-C2-1-R and MDM-2014-0369 of ICCUB (Unidad de Excelencia *María de Maeztu*)

## References

[1] Castañeda, J., Fabricius, C., Portell, J., Garralda, N., Gonzalez-Vidal, J.J., Clotet, M. and Torra, J. 2016, in Highlights of Spanish Astrophysics IX (this volume).

[2] Clotet, M., Gonzalez-Vidal, J.J., Castañeda, J., Garralda, N., Portell, J., Fabricius, C. and Torra, J. 2016, in Highlights of Spanish Astrophysics IX (this volume).

[3] Fabricius, C., Bastian, U. Portell, J., et al. 2016, A&A, in press (Gaia SI).

[4] Gaia Collaboration (Brown et al) 2016, A&A, in press (Gaia SI).

[5] Gaia Collaboration (Prusti et al.) 2016, A&A, in press (Gaia SI).