

---

## 132. Variability, DR3, and Citizen Science

---

I DESCRIBED the discovery and characterisation of variable stars with Gaia DR2 in essay #61, and variability across the Hertzsprung–Russell diagram in essay #62.

Gaia DR2 (April 2018) was based on observations between July 2014 and May 2016, and some basic statistics are worth recalling. Before Gaia, the number of known variable stars was some 150 000. As a result of the global variability analysis (Holl et al., 2018), which I outlined in essay #61, Gaia DR2 contained 550 737 sources classified as variable. These included 228 904 RR Lyrae and 11 438 Cepheids, 151 761 long-period variables, 147 535 stars showing rotation modulation, 8882  $\delta$  Scu and SX Phe stars, and 3018 short-timescale variables.

G AIA DR3 (June 2022), included all satellite observations between July 2014 and May 2017 (34 months of data). The updated details of the variability processing and analysis are given by Eyer et al. (2022).

Of the 1.8 billion sources in DR3, 10 509 536 were identified as variable (9.5 million stars and one million galaxies/quasars), with associated time series in the photometric bands  $G$ ,  $G_{BP}$ , and  $G_{RP}$  and, in some cases, with the radial velocity time series. The classified variables lie mostly in the range  $G = 12 - 21$  mag, with the majority around  $G = 16 - 20$  mag, and amplitudes ranging from a few milli-magnitudes to several magnitudes.

After period search and time-series modelling, supervised classification assigned the variables to 24 classes (Rimoldini et al., 2022). These include  $\beta$  Cephei, Cepheids, cataclysmic variables, eclipsing binaries, ellipsoidal variables, exoplanets, long-period variables, RR Lyrae, RS Can Ven, sub-dwarf B, supernovae, solar-like stars, slowly pulsating B-stars, symbiotic stars, pulsating white dwarfs, and young stellar objects.

Even compared with DR2, the numbers are transformational. There are 297 778 RR Lyrae, 16 141 Cepheids, 1475  $\beta$  Cepheid stars, 2352 775 long-period-variables, 474 026 showing rotational modulation, 7 480 058  $\delta$  Scu and SX Phe stars, 1 934 844 solar-like, 65 300 ellipsoidal variables, 7306 cataclysmic variables, 910 white dwarfs, 254 microlensing events, and 214 exoplanet transits.

LET ME NOW FOCUS on the ‘citizen science’ contributions to these Gaia variability studies. Broadly, ‘citizen science’ is scientific work undertaken by the general public, usually under the direction of professional scientists. It is being used in a wide range of studies, most prominently in the areas of biology and conservation.

Typically, volunteers help to collect or classify data, thereby assisting in various aspects of the scientific process, while helping to inform and educate the wider public. Depending on the project, participants may collect and enter data in dedicated databases (e.g. species of birds or plants observed), helping to classify the data, or assist in training machine-learning algorithms.

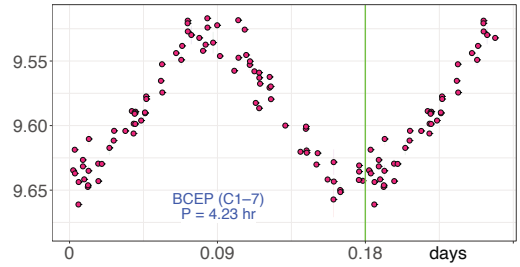
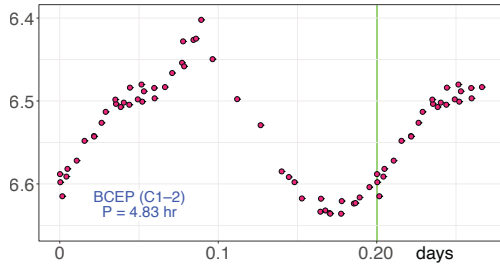
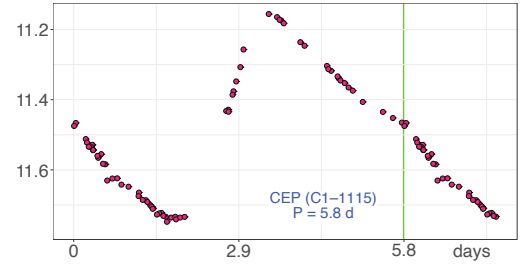
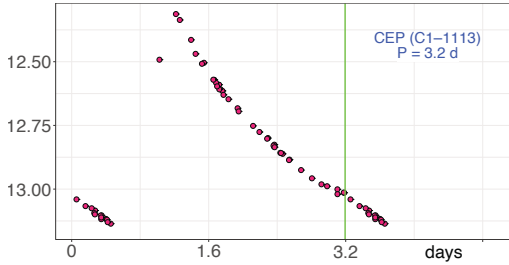
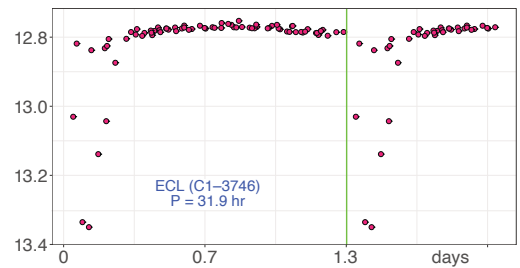
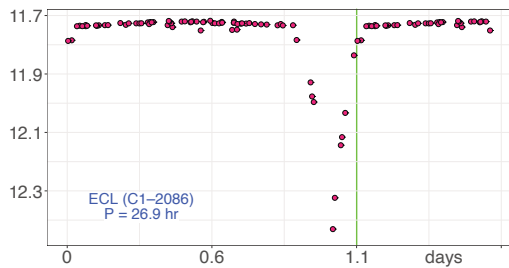
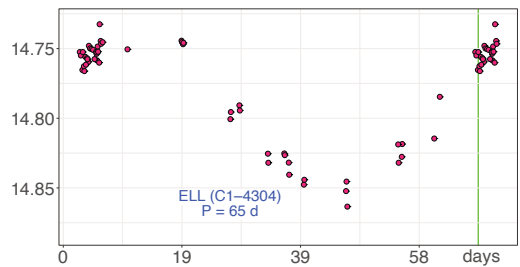
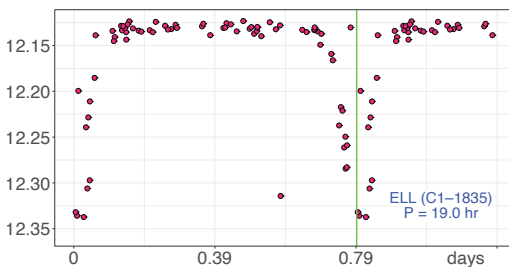
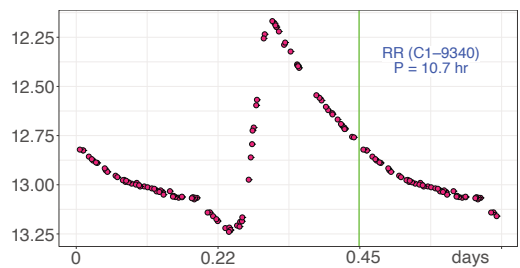
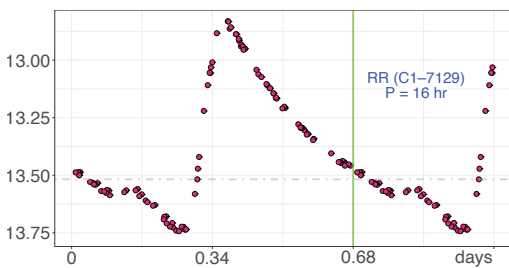
Zooniverse is one such web portal, with around two million registered volunteers, and 25 space-related projects, amongst them Galaxy Zoo, Gravity Spy, Planet Hunters, Solar Jet Hunter, and Cloudspotting on Mars.

WITHIN ZOONIVERSE, *Gaia-Vari* was launched in June 2022. Led by Krzysztof Nienartowicz and Elsa Matias, it is the first citizen science project related to Gaia. Light curves of more than 10 000 DR3 sources, including eclipsing binaries, Cepheids, RR Lyrae, and long-period variables, are made available, folded at the period from Gaia’s automated analysis system.

Volunteers, as of today more than 600, are then invited to decide, from visual inspection, whether the period is (say) wrong by a factor 2, which may occur for eclipsing binaries, or whether the variability type has been incorrectly assigned. By improving the fidelity of the training set, this work is expected to significantly improve the content of the next Gaia data release, DR4.

I HAVE SELECTED 10 of these folded light curves by way of illustration, and these are shown on the following page: two examples each of just five variability classes.

The vertical axis is the Gaia  $G$  magnitude, the horizontal axis is the observation time (in days) folded at the inferred period (shown as the vertical green line). The C1–xxx numbers are the identification numbers assigned within the Gaia–Vari project.

 *$\beta$  Cephei variables [BCEP]**Cepheid variables [CEP]**Eclipsing binaries [ECL]**Ellipsoidal variables [ELL]**RR Lyrae variables [RR]*