Variability Analysis of the Gaia data

Laurent Eyer
Nami Mowlavi, Dafydd W. Evans, Berry Holl, Lorenzo Rimoldini, Alessandro Lanzafame, Leanne Guy,
Shay Zucker, Brandon Tingley, Isabelle Lecoer-Taibi, Maroussia Roelens, Jan Cuypers,
Joris De Ridder, Sara Regibo, Manuel López, Jonas Debosscher, Maria Süveges, Luis Sarro,
Gisella Clementini, Silvio Leccia, Vincenzo Ripepi, Fabio Barblan, André Moitinho,
Diego Ordoñez-Blanco, Krzysztof Nienartowicz, Jonathan Charnas, Grégory Jéwardat de Fombelle

1 Department of Astronomy, University of Geneva, Versoix, Switzerland
2 Institute of Astronomy, University of Cambridge, Cambridge, United Kingdom
3 Dipartimento di Fisica e Astronomia, Universita di Catania, Catania, Italy
4 Department of Geosciences, Tel Aviv University, Tel Aviv, Israel
5 Institut for Fysik og Astronomi, Aarhus Universitet, Aarhus, Denmark
6 Royal Observatory of Belgium, Brussels, Belgium
7 Instituut voor Sterrenkunde, KU Leuven, Leuven, Belgium
8 Centro de Astrobiologia, Departamento de Astrofisica, Villanueva de la Canada, Spain
9 Departamento de Inteligencia Artificial, UNED, Madrid, Spain
10 INAF Osservatorio Astronomico di Bologna, Bologna, Italy
11 INAF-Osservatorio Astronomico di Capodimonte, Napoli, Italy
12 Faculdade de Ciencias de Universidade de Lisboa, Lisboa, Portugal

Gaia Mission Status and First Data Processing
Division A, IAU
Hawaii, USA
Friday August 7, 2015
Gaia Variability Processing and Analysis

1 billion sources observed by Gaia
790,000 REAL Gaia data from Ecliptic Poles
Time series of 70 (40-250) measurements over 5 years
Time series up to 170 measurements over 28 days

Calibrated photometry (CU5)

General Variability Detection (GVD)  Special Variability Detection (SVD)

Characterization

Classification

Specific Object Studies (SOS)
Selection of sources observed by Gaia from Ecliptic Pole Scanning Law (790,000 sources)

~28 days

(Equatorial coordinates, deg)

Warning: Sampling regular, different from Nominal Scanning law
South Ecliptic Pole region (part of Large Magellanic Cloud): Gaia and other surveys
To get the data flavour
Comparison with OGLE

Image of the Week (March 05, 2015):
RR Lyrae stars

Take-home message:
Gaia G band photometry is good!
Gaia Variability Processing and Analysis

790,000 REAL Gaia data from Ecliptic Poles

Time series up to 170 measurements over 28 days

Calibrated photometry (CU5)

General Variability Detection (GVD)  Special Variability Detection (SVD)

Characterization

Classification

Specific Object Studies (SOS)
Two fundamental quantities to estimate:
- Completeness
- Contamination

Detection was done with a classifier (Random Forest) attributes were computed a training set was defined (based on OGLE)

**Classifier result: The confusion matrix**

<table>
<thead>
<tr>
<th></th>
<th>VARIABLE</th>
<th>CONSTANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>376 VARIABLE</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>546 CONSTANT</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>Contamination</td>
<td>8</td>
<td>13</td>
</tr>
</tbody>
</table>
Special Variability Detection:

Statistical Parameters

Special Variability Detection
Special Variability Detection: short time scale

Implementation of variogram: “variance” for all the paired magnitude difference separated by a certain time lag

One example of per-ccd data:

Probed Time-scales: ~10 seconds
Special Variability Detection: exo-planet transits

Two algorithms:

1) Box-Least Square

2) Outlier Probability, Tingley (A&A 2011)

Box Least Square algorithm gives about 200 candidates

We **do not** claim any detection! yet!

Conclusion Box-Least Square is functioning well
Gaia Variability Processing and Analysis

790,000 REAL Gaia data from Ecliptic Poles

Time series up to 170 measurements over 28 days

Calibrated photometry (CU5)

General Variability Detection (GVD)  Special Variability Detection (SVD)

Characterization

Classification

Specific Object Studies (SOS)

Astrometric char (CU3+CU4)
Spectroscopic char (CU6)
Astrophysical param (CU8)
Calibrated photometry (CU5)
Radial velocities (CU6)
Variables catalogue (CU7)
Characterisation

Time series per object:

\[ \text{Time}_{(i)}, G-, BP-, RP- \text{ mag}_{(i)} [ \text{ or radial velocity}_{(i)} ] \]

\[ i=1, \ldots, \text{number of measurements} \]

Goal: To define attributes

- statistical parameters

- Modelling
  - Period search
  - Fourier Series and polynomial fit
Characterisation: few examples of modelling
Gaia Variability Processing and Analysis

790,000 REAL Gaia data from Ecliptic Poles

Time series up to 170 measurements over 28 days

Calibrated photometry (CU5)

General Variability Detection (GVD)  Special Variability Detection (SVD)

Characterization

Classification

Specific Object Studies (SOS)
Classification

Supervised classification (several methods):

- Multistage tree: Bayesian networks
- Multistage tree: Gaussian mixture
- Random Forest

Tree for Gaussian Mixture:

1. Eclipsing binaries: EA, EB, EW
2. GDOR
3. RRAB, RRC
4. 5. Cepheids: DCEP, DCEPS, CWA, CWB, CEP(B)
5. SPB
6. DSCTC, DSCT+SXPHE

Furnish training set built from Crossmatched data

Classification Error Rate against Number of Attributes Selected:

- Backward Elimination
- Forward Selection
### Classification

Confusion matrix of Random Forest using cross-matched data (OGLE, Hipparcos, AAVSO, Milliquas)

<table>
<thead>
<tr>
<th>Class</th>
<th>CONSTANT</th>
<th>QSO</th>
<th>ECL</th>
<th>OTHER</th>
<th>RRLYR</th>
<th>LPV</th>
<th>ELL</th>
<th>DSCT</th>
<th>DCEP</th>
<th>UG</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>79</td>
<td>8</td>
<td>5</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QSO</td>
<td>7</td>
<td>88</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECL</td>
<td>4</td>
<td>3</td>
<td>84</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER</td>
<td>13</td>
<td>4</td>
<td>12</td>
<td>69</td>
<td></td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRLYR</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Contamination:
- LPV: 30, 19, 7, 30, 15
- ELL: 25, 12, 62
- DSCT: 14, 29, 29, 29
- DCEP: 50, 50
- UG: 100

Total:
- LPV: 34, 20, 19, 34, 5, 43, 100, _, _, _
- ELL: _, _, _, _, _, _, _
- DSCT: _, _, _, _, _, _, _
- DCEP: _, _, _, _, _, _, _
- UG: _, _, _, _, _, _, _
Gaia Variability Processing and Analysis

790,000 REAL Gaia data from Ecliptic Poles

Time series up to 170 measurements over 28 days

Calibrated photometry (CU5)

- General Variability Detection (GVD)
- Special Variability Detection (SVD)

Characterization

Classification

Specific Object Studies (SOS)
Specific Object studies: a list of variability types

<table>
<thead>
<tr>
<th>Eclipsing Binaries</th>
<th>Cepeids &amp; RRLyrae stars</th>
<th>Long Period Variables</th>
<th>Exoplanet transits</th>
<th>Be stars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short time scales</td>
<td>Pre-main sequence</td>
<td>Flaring Stars</td>
<td>Rotational Modulation</td>
<td>μ-lensing events</td>
</tr>
</tbody>
</table>

- Cataclysmic Variables
- AGN
- Rapid-phases
Specific Object Studies: Eclipsing binaries

Eclipsing binaries go to a dedicate treatment (Université Libre de Bruxelles) for a full modelling
Here, simple modelling are made
The solutions enable a ranking

Highest rank

Phase

G mag
Specific Object Studies: Eclipsing binaries

Lowest rank

- $P = 0.089681373$, $\text{rChi2} = 4.42915$, rank = 0.2232, s/n = 0.24179, 2 pts in Eclipse, Abbe = 0.773
- $P = 0.1017395$, $\text{rChi2} = 4.97379$, rank = 0.2231, s/n = 0.25739, 2 pts in Eclipse, Abbe = 0.691
- $P = 0.16715213$, $\text{rChi2} = 2.79541$, rank = 0.2225, s/n = 0.18585, 7 pts in Eclipse, Abbe = 0.711
- $P = 0.12502477$, $\text{rChi2} = 4.71783$, rank = 0.2213, s/n = 0.24358, 3 pts in Eclipse, Abbe = 0.399
- $P = 0.37760976$, $\text{rChi2} = 9.25721$, rank = 0.2208, s/n = 0.35019, 10 pts in Eclipse, Abbe = 0.803
- $P = 0.40026117$, $\text{rChi2} = 9.83707$, rank = 0.2148, s/n = 0.33139, 4 pts in Eclipse, Abbe = 0.884
Gaia Variability Processing and Analysis

790,000 REAL Gaia data from Ecliptic Poles

Time series up to 170 measurements over 28 days

Calibrated photometry (CU5)

General Variability Detection (GVD)
Special Variability Detection (SVD)

Characterization

Classification

Specific Object Studies (SOS)

Global processing

Per source processing
Global Variability studies:

Comparison of distribution functions of RR Lyrae stars

BN 3.2  RF 4.2  OGLE IV GSEP

Log (frequency [1/day])

Log(Amplitude [G & I mag])

Courtesy of L. Sarro (modified by L.Eyer)
Gaia Variability Processing and Analysis

First take-home message: 1 billion sources observed by Gaia
Time series of 70 (40-250) measurements over 5 years

Second take-home message: Gaia Variability Analysis is in good "shape"

General Variability Detection (GVD)
Characterization
Classification
Specific Object Studies (SOS)

Calibrated photometry (CU5)
Astrometric char (CU3+CU4)
Spectroscopic char (CU6)
Astrophysical param (CU8)

Global Variability Studies (GVS)
Variables catalogue (CU7)
# Release scenario

<table>
<thead>
<tr>
<th>Year</th>
<th>Release</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>1</td>
<td>- (\alpha) and (\delta), mean G-magnitude&lt;br&gt;- Commissioning data&lt;br&gt;- 100K proper motion stars (Hipparcos+Gaia)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>- 5-parameter astrometric solutions for single star (parallax)&lt;br&gt;- Integrated BP/RP + Astrophysical parameters&lt;br&gt;- Mean (V_{\text{rad}}) (for non variable)</td>
</tr>
<tr>
<td>2015</td>
<td>2</td>
<td>&lt;br&gt;- + (?) Groups of variables: RR Lyrae stars + Cepheids + Short times scales</td>
</tr>
<tr>
<td>2016</td>
<td>3</td>
<td>- Mean (V_{\text{rad}})&lt;br&gt;- 5-par astrometry&lt;br&gt;- Object classifications and Astrophysical Parameters&lt;br&gt;- Orbital solution of binaries&lt;br&gt;- Mean RVS spectra</td>
</tr>
<tr>
<td>2017</td>
<td>4</td>
<td>- Variable stars classification&lt;br&gt;- Non-single star catalogue&lt;br&gt;- Solar system objects</td>
</tr>
<tr>
<td>2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Final release:**
  - everything !

**Courtesy of B.Holl**
Conclusions

Gaia is unique, also for variability analysis, because:

• surveys the entire sky with one set of instruments
• includes the “bright sky”
• performs nearly simultaneous measurements with different instruments
• furnishes exquisite parallaxes
• has a very distinctive sampling (complementary with respect to ground based surveys)

Final take-home message:

Gaia variable census will be exceptional!
Thank you for your attention