Machine Learning Variability Classification in the OGLE Project

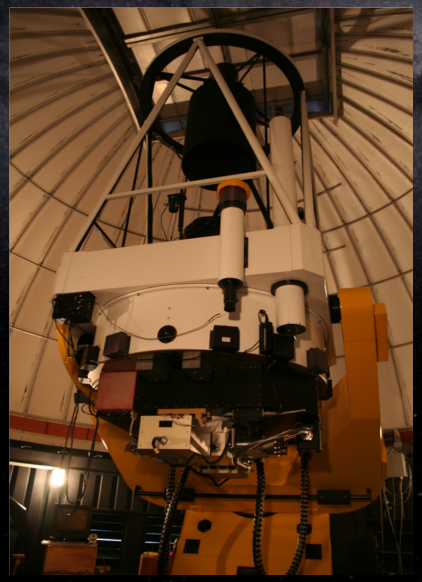
ADASS XXVI, Trieste, 19.10.2016

Michał Pawlak

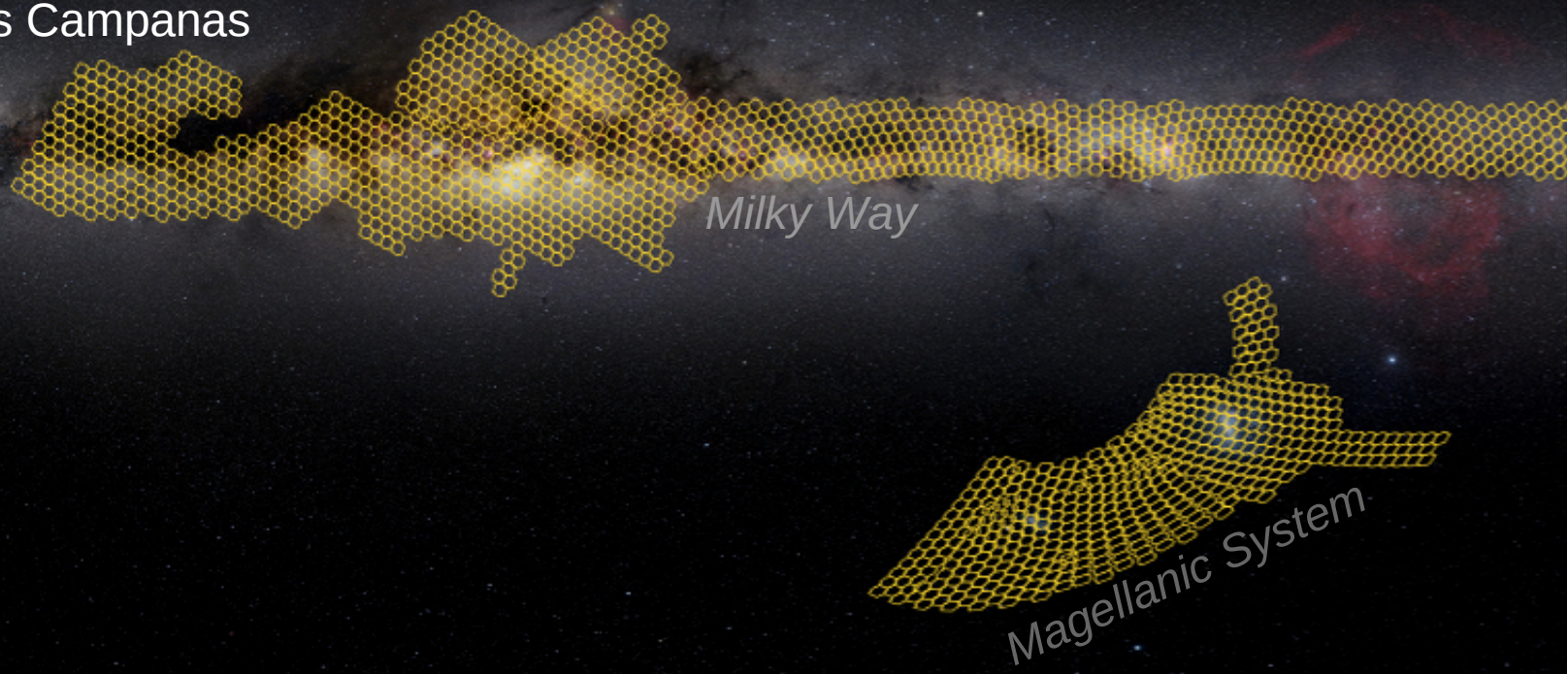
Warsaw University Astronomical Observatory

OGLE – a Large Variability Sky Survey

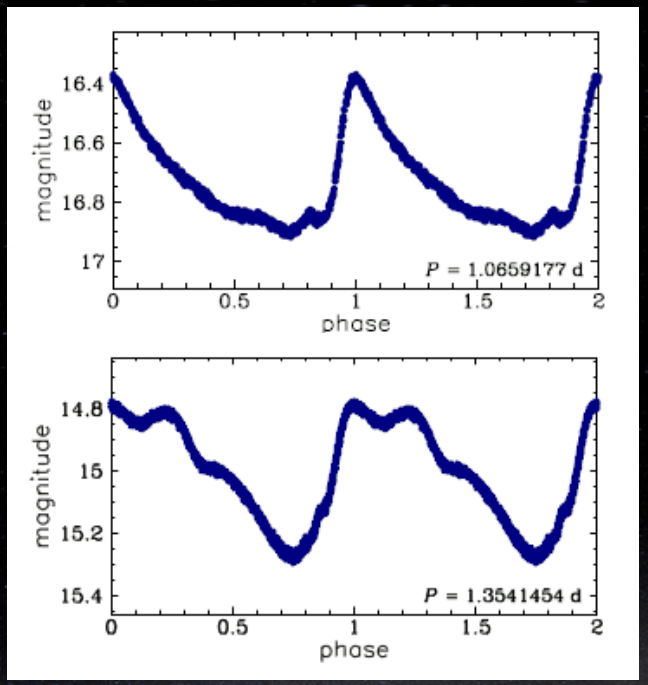
Warsaw 1.3-m @ Las Campanas



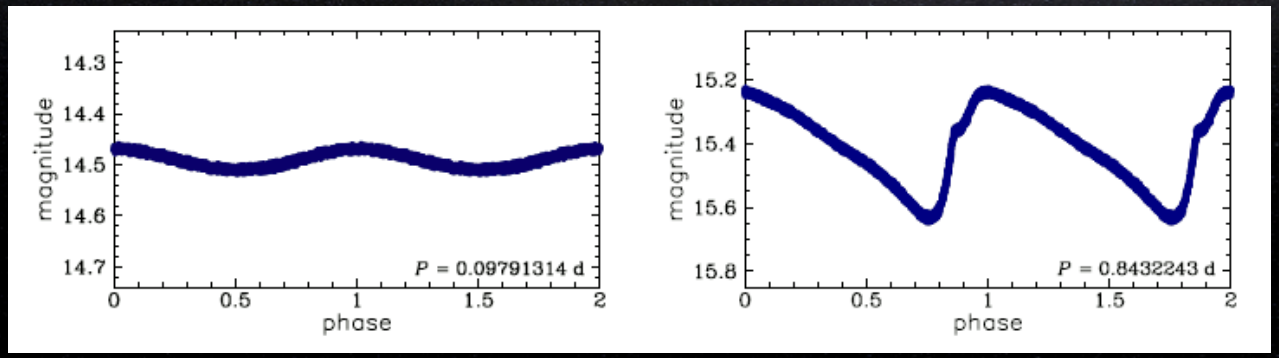
- in operation since 1992
- since 2010 as OGLE-IV (Udalski *et al.* 2015)
- 3500 deg^2 sky coverage
- 1.3 billion sources monitored every night
- 10^{12} photometric measurements by 2016
- over 17,000 microlensing detections
- more than 50 extrasolar planets discovered
- 500,000 new variable stars

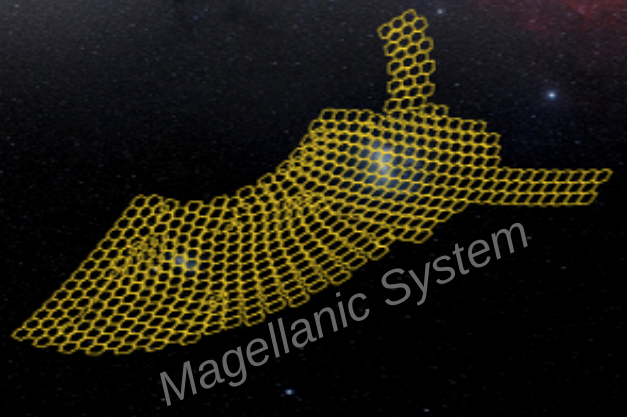


The OGLE Collection of Variable Stars



- about 500,000 new variable stars in
  Galactic Bulge, Galactic Disk and Magellanic Clouds
- pulsating stars: Cepheids, RR Lyraes, LPVs:
  Soszyński *et al.* 2004, ..., 2015
- eclipsing and ellipsoidal binaries:
  Graczyk *et al.* 2011, Petrukowicz *et al.* 2013
  Pawlak *et al.* 2013
-





Machine Learning in the OGLE Project - overview

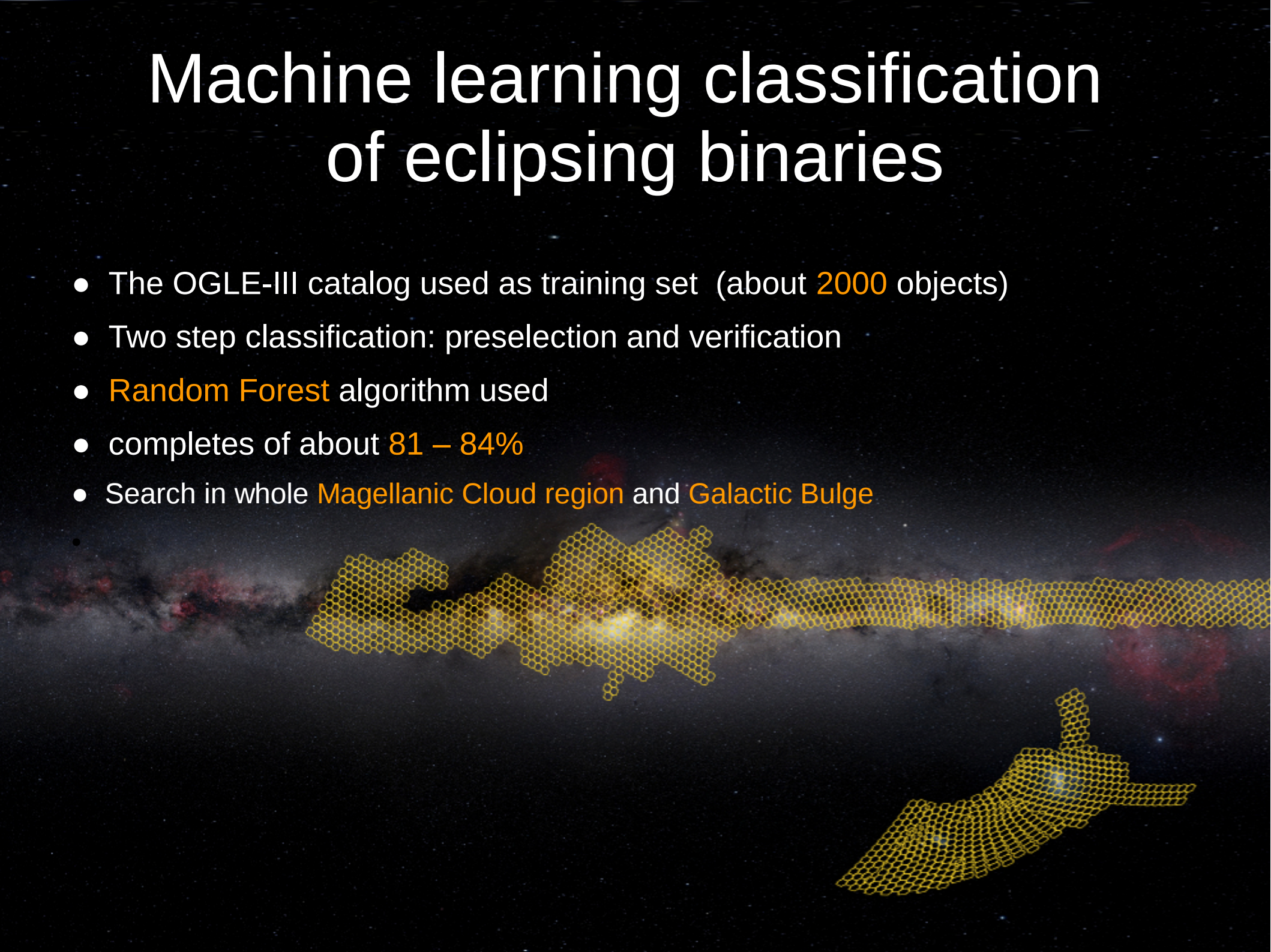
- search for eclipsing binaries with Artificial Neural Networks: (Wyrzykowski *et al.* 2003, 2004)
- search for eclipsing binaries with Decision Trees: (Pawlak *et al.* (2013)
- search for microlensing events with Random Forest: (Wyrzykowski *et al.* 2015)
- real-time transient detection with Self-Organizing Maps: (Klencki & Wyrzykowski 2016)

In progress:

- search for eclipsing binaries with Random Forest: (Pawlak *et al.* 2016)

Machine learning classification of eclipsing binaries

- The OGLE-III catalog used as training set (about 2000 objects)
- Two step classification: preselection and verification
- Random Forest algorithm used
- completes of about 81 – 84%
- Search in whole Magellanic Cloud region and Galactic Bulge



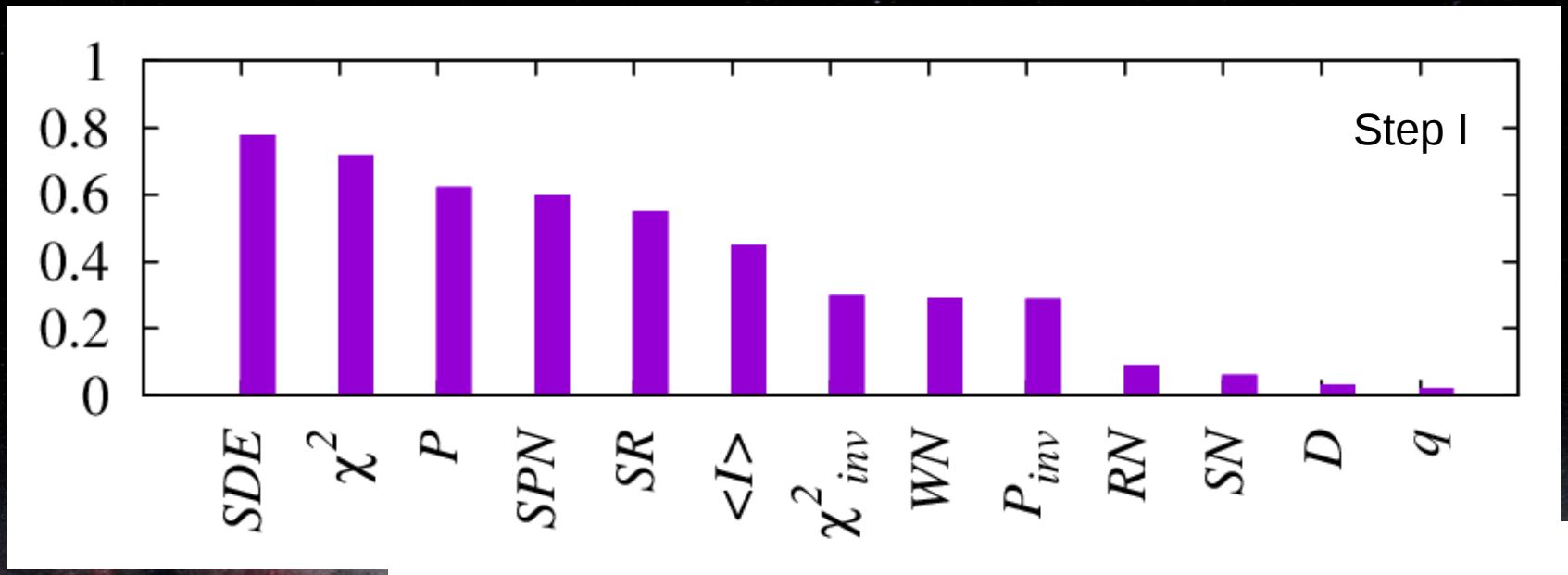
Machine learning classification of eclipsing binaries - parameters

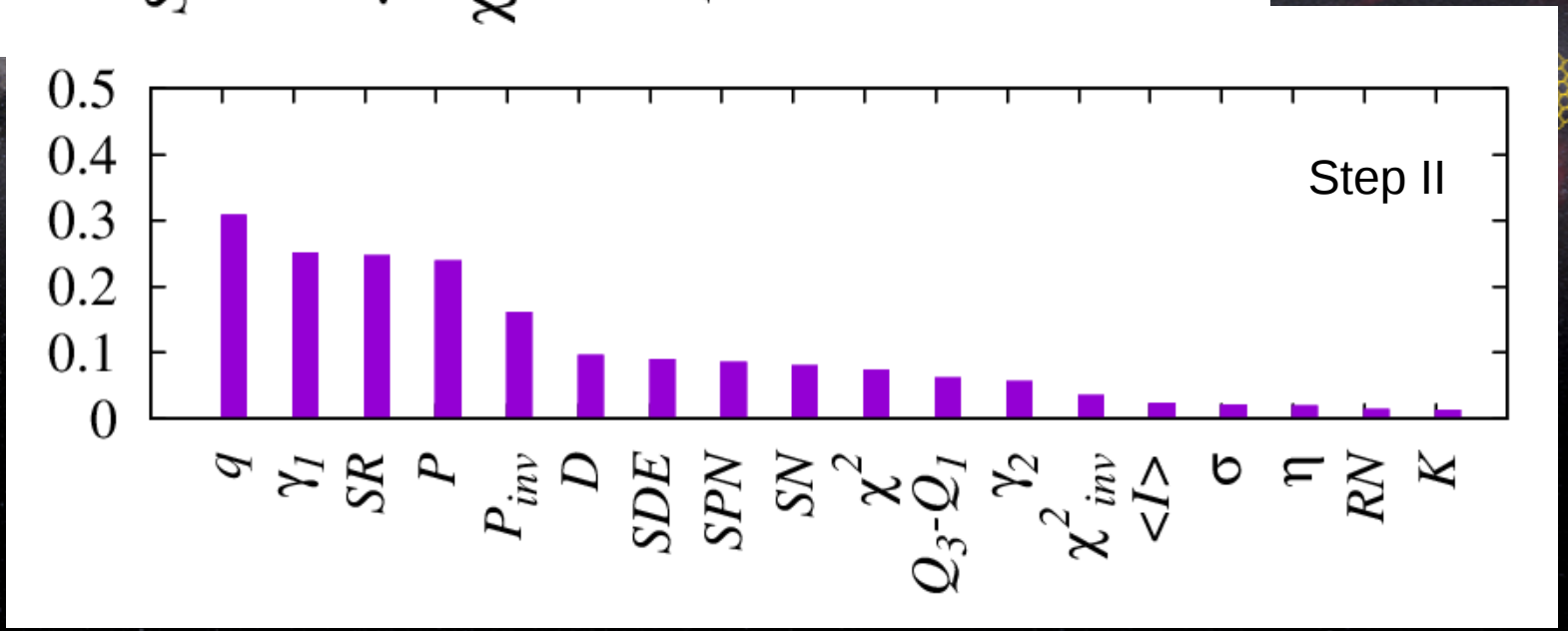
Parameter	Description
P	Period
SR	Singnal Residue
SN	Signal to Noise Ratio
SDE	Signal Detection Efficiency
D	Depth of transit
q	Fraction of the phase in ransit
χ^2	χ^2 of the transit model fit
RN	Red Noise
WN	White Noise
SPN	Signal to Pink Noise Ratio
$P_{\rm inv}$	Period of inverse transit
$\chi^2_{\rm inv}$	χ^2 of the inverted transit model fit
I	mean I-band magnitude

BLS, Kovacs *et al.* (2002)

Parameter	Description
σ	standard deviation
γ_1	skewness
γ_2	kurtosis
$Q_3 - Q_1$	difference between third and first quartile
η	Von Neumann index
K	Stetson K index

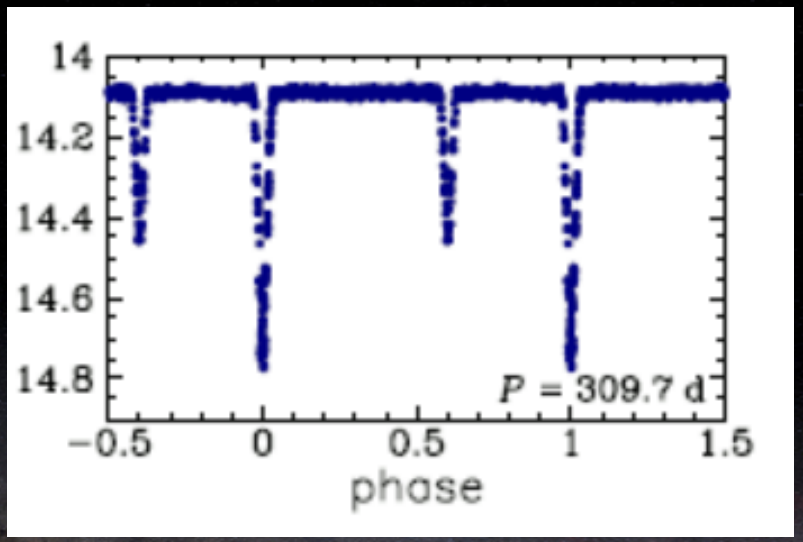
Machine learning classification of eclipsing binaries – information gain

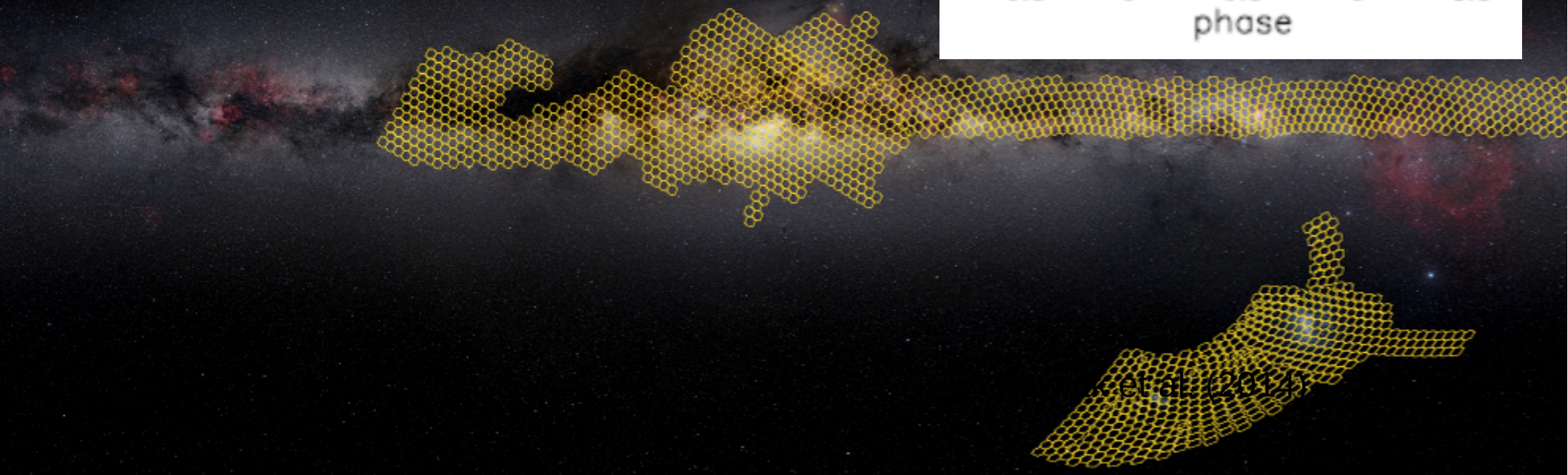




OGLE-IV Catalogs of Eclipsing Binaries

- Magellanic Clouds: over 48 000 objects
  (Pawlak *et al.* 2016)
- Galactic Bulge: about 400 000 objects
  (Soszyński *et al.* 2016)
- Coming soon!





References

Graczyk D. *et al.* 2011, Acta Astronomica, 61, 103
Klencki, J. & Wyrzykowski, Ł., arXiv:1601.06151
Kovács, G., Zucker, S., Mazeh, T. 2002, A&A, 391, 369
Pawlak, M. *et al.* 2013, Acta Astronomica, 63, 323
Pawlak, M *et al.* 2016, in preparation
Pietrukowicz, P. *et al.* 2013, Acta Astronomica, 63, 115
Soszyński, I. *et al.* 2002, Acta Astronomica, 52, 369
Soszyński, I. *et al.* 2015, Acta Astronomica, 65, 297
Soszyński, I. *et al.* 2016, in preparation
Udalski, A., *et al.* 2015, Acta Astronomica, 65, 1
Wyrzykowski, Ł. *et al.* 2003, Acta Astronomica, 53, 1
Wyrzykowski, Ł. *et al.* 2004, Acta Astronomica, 54, 1
Wyrzykowski, Ł. *et al.* 2015, ApJS, 216, 17

br0br0ken!!

Thank you for your attention!

Acknowledgment
This work has been supported by Polish National Science Center grant no. 2016/20/T/ST9/00170