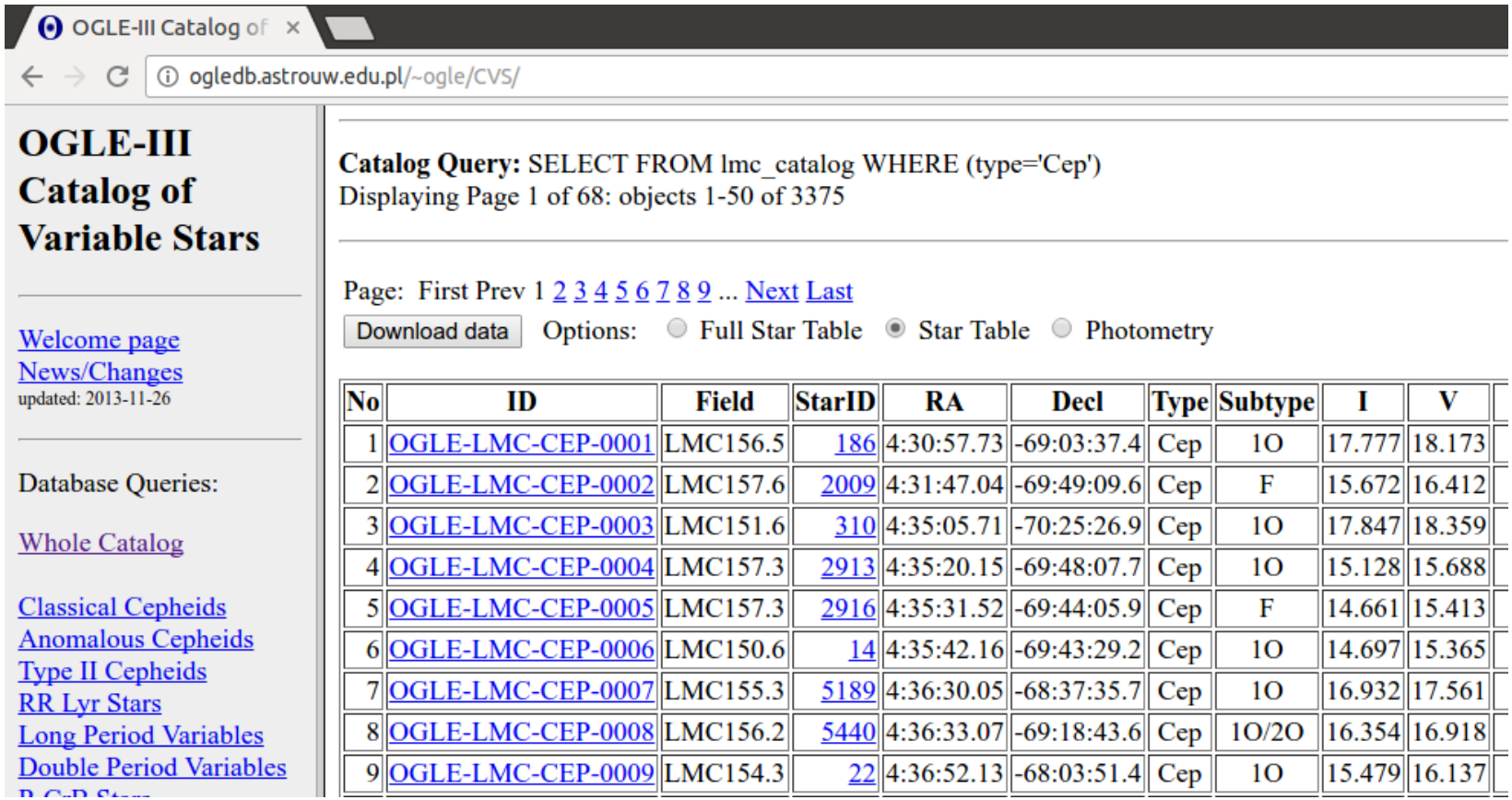
Discussion of Overview session: Science of transients

James Long

March 20, 2017

1. Collaborative Classification

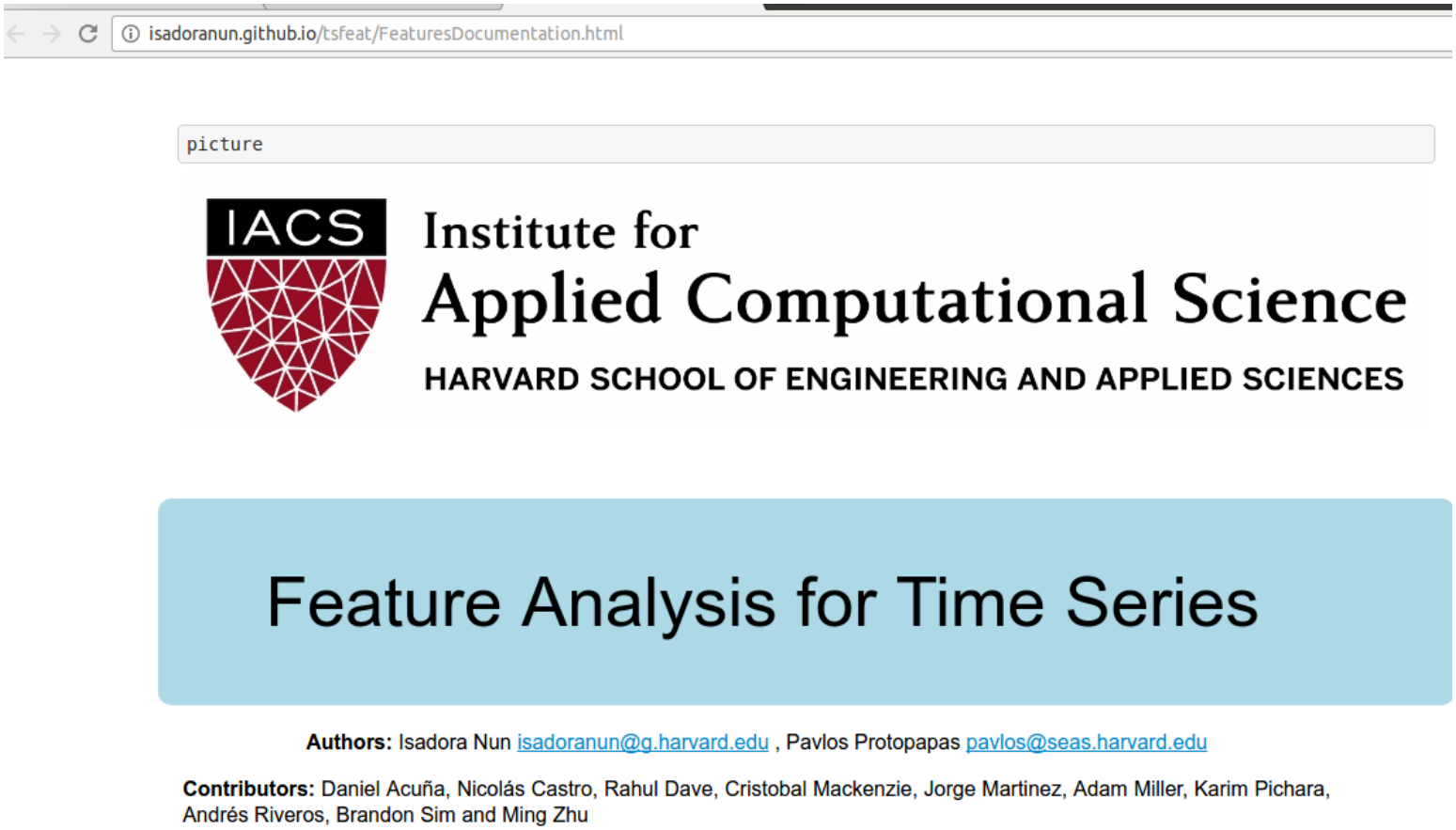
Data Sets



No	ID	Field	StarID	RA	Decl	Type	Subtype	I	V
1	OGLE-LMC-CEP-0001	LMC156.5	186	4:30:57.73	-69:03:37.4	Cep	1O	17.777	18.173
2	OGLE-LMC-CEP-0002	LMC157.6	2009	4:31:47.04	-69:49:09.6	Cep	F	15.672	16.412
3	OGLE-LMC-CEP-0003	LMC151.6	310	4:35:05.71	-70:25:26.9	Cep	1O	17.847	18.359
4	OGLE-LMC-CEP-0004	LMC157.3	2913	4:35:20.15	-69:48:07.7	Cep	1O	15.128	15.688
5	OGLE-LMC-CEP-0005	LMC157.3	2916	4:35:31.52	-69:44:05.9	Cep	F	14.661	15.413
6	OGLE-LMC-CEP-0006	LMC150.6	14	4:35:42.16	-69:43:29.2	Cep	1O	14.697	15.365
7	OGLE-LMC-CEP-0007	LMC155.3	5189	4:36:30.05	-68:37:35.7	Cep	1O	16.932	17.561
8	OGLE-LMC-CEP-0008	LMC156.2	5440	4:36:33.07	-69:18:43.6	Cep	1O/2O	16.354	16.918
9	OGLE-LMC-CEP-0009	LMC154.3	22	4:36:52.13	-68:03:51.4	Cep	1O	15.479	16.137

- 400,000 + classified variable star light curves
- `http://ogledb.astrouw.edu.pl/~ogle/CVS/`
- many others out there

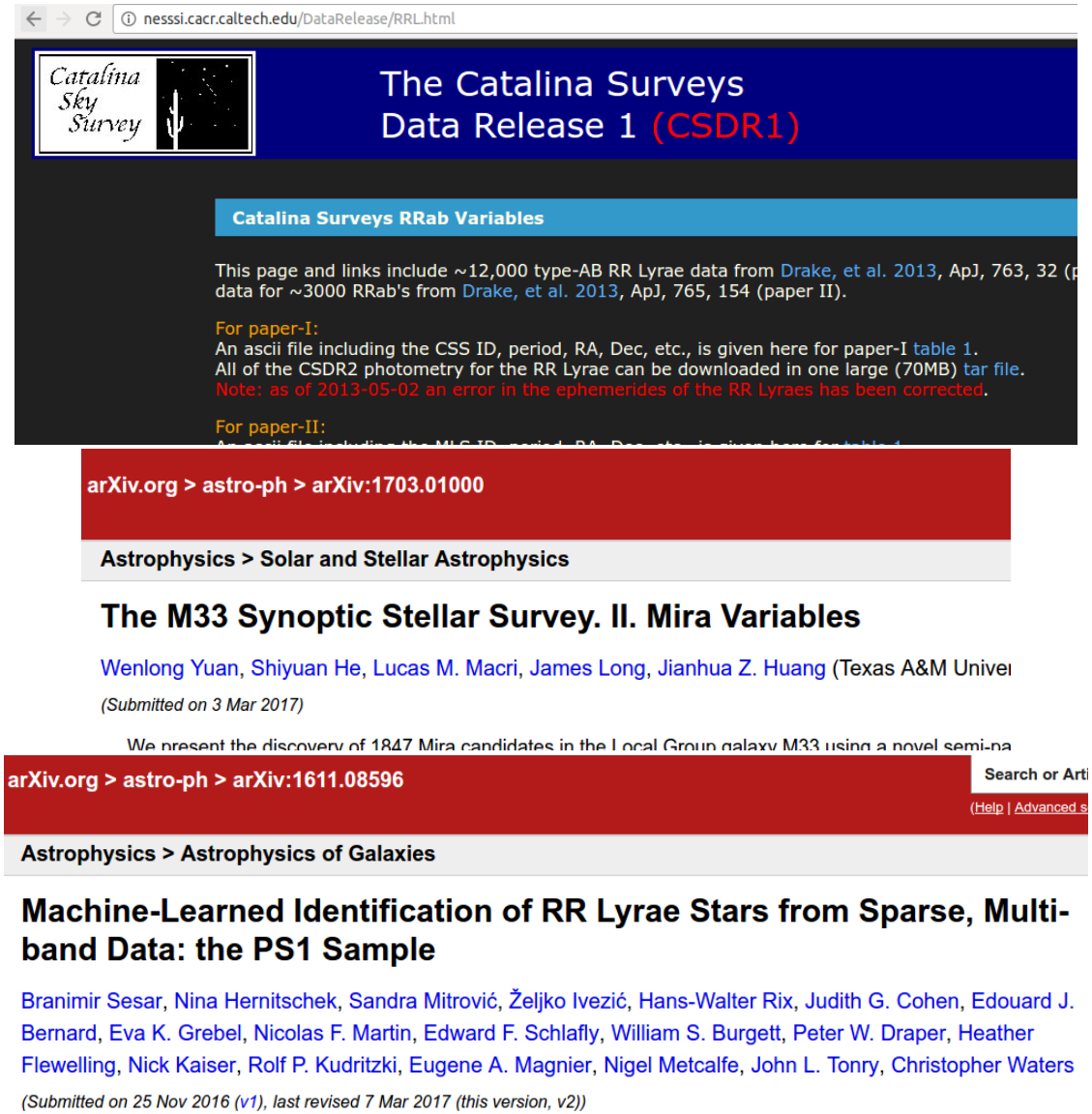
Feature Extractors



- quickly derive many useful features
- `http://isadoranun.github.io/tsfeat/FeaturesDocumentation.html`
- several others, eg vartools: `http://www.astro.princeton.edu/~jhartman/vartools.html`

Output from Classifiers

Three of many:



nesssi.cacr.caltech.edu/DataRelease/RRL.html

Catalina Sky Survey

The Catalina Surveys
Data Release 1 (CSDR1)

Catalina Surveys RRab Variables

This page and links include ~12,000 type-AB RR Lyrae data from Drake, et al. 2013, ApJ, 763, 32 (p
data for ~3000 RRab's from Drake, et al. 2013, ApJ, 765, 154 (paper II).

For paper-I:
An ascii file including the CSS ID, period, RA, Dec, etc., is given here for paper-I table 1.
All of the CSDR2 photometry for the RR Lyrae can be downloaded in one large (70MB) tar file.
Note: as of 2013-05-02 an error in the ephemerides of the RR Lyraes has been corrected.

For paper-II:

arXiv.org > astro-ph > arXiv:1703.01000

Astrophysics > Solar and Stellar Astrophysics

The M33 Synoptic Stellar Survey. II. Mira Variables

Wenlong Yuan, Shiyuan He, Lucas M. Macri, James Long, Jianhua Z. Huang (Texas A&M Univer

(Submitted on 3 Mar 2017)

We present the discovery of 1847 Mira candidates in the Local Group galaxy M33 using a novel semi-pa

arXiv.org > astro-ph > arXiv:1611.08596

Search or Arti

(Help | Advanced s

Astrophysics > Astrophysics of Galaxies

Machine-Learned Identification of RR Lyrae Stars from Sparse, Multi-band Data: the PS1 Sample

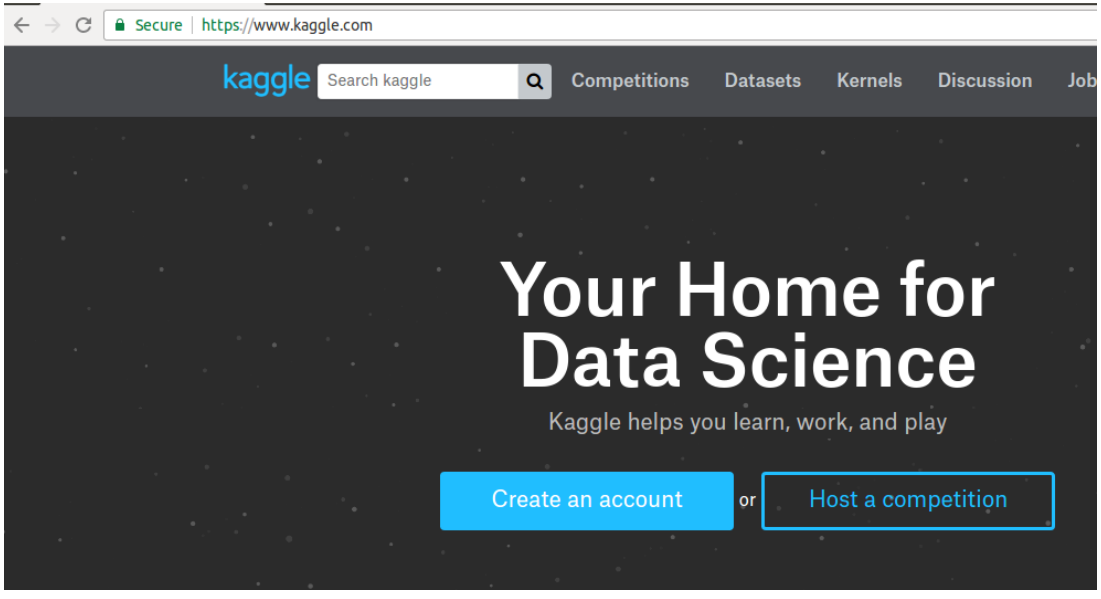
Branimir Sesar, Nina Hernitschek, Sandra Mitrović, Željko Ivezić, Hans-Walter Rix, Judith G. Cohen, Edouard J. Bernard, Eva K. Grebel, Nicolas F. Martin, Edward F. Schlafly, William S. Burgett, Peter W. Draper, Heather Flewelling, Nick Kaiser, Rolf P. Kudritzki, Eugene A. Magnier, Nigel Metcalfe, John L. Tonry, Christopher Waters

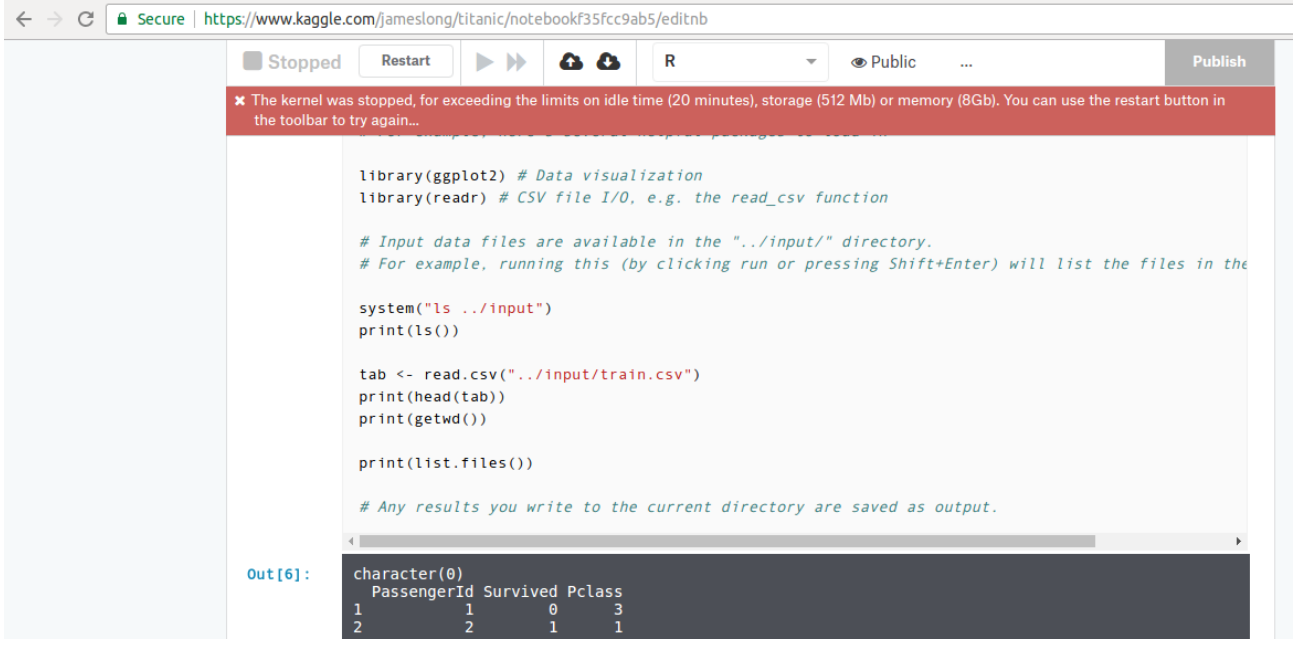
(Submitted on 25 Nov 2016 (v1), last revised 7 Mar 2017 (this version, v2))

Can we integrate data sets, features, and classifications at a single location?

- accelerate development and testing of new features
- encourage multidisciplinary collaboration
- merge labeled data sets to improve results
- easy comparison of methodologies and results

Kaggle Has Done This With Data Science

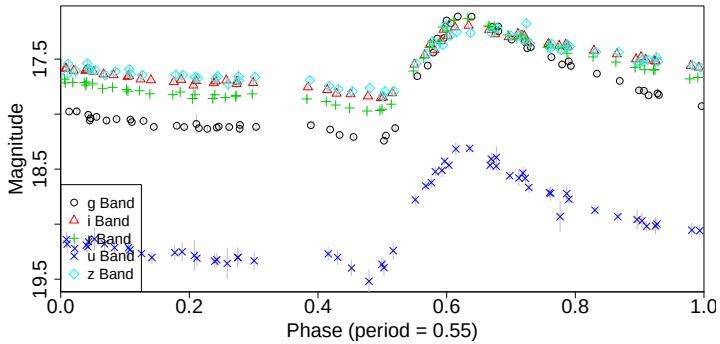




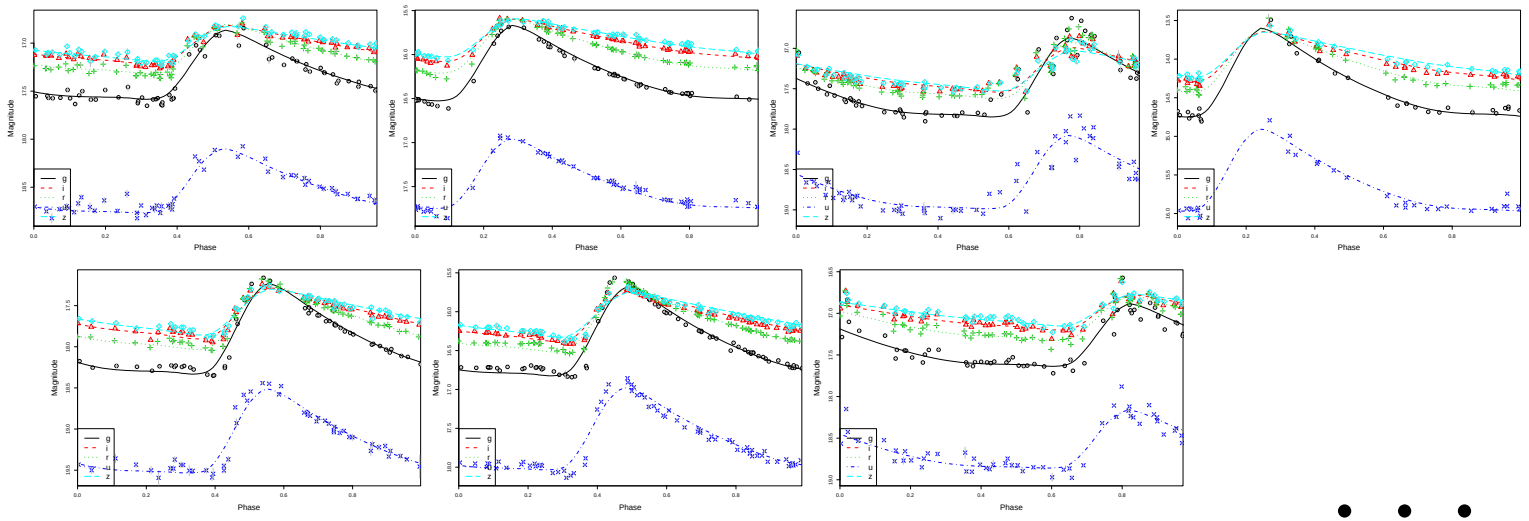
2. Object Classification to Population Modeling

Object Classification to Population Modeling

Object Classification: I found an RR Lyrae.



Population Modeling: I found 1000 RR Lyrae, what is their period, luminosity, metallicity relationship?



• • •

Uncertainty and Biases in Parameter Estimates

- Classifier will incorrectly label some objects, providing an incomplete understanding of the object class.
- Incompleteness may systematically exclude objects with certain characteristics (eg dimmer variable stars, GW at certain frequencies).

How much overlap should there be between classification work (e.g. SAMSI WG2 **Synoptic Time Domain Surveys**) and Population Modeling work (e.g. SAMSI WG4 **Astrophysical populations**) to address potential biases?