Variable Stars and how to find them

Filip Novotný
Miroslav Lžičař
Tomáš Bárta

With help of Dr. Martin Jelínek – Czech academy of sciences

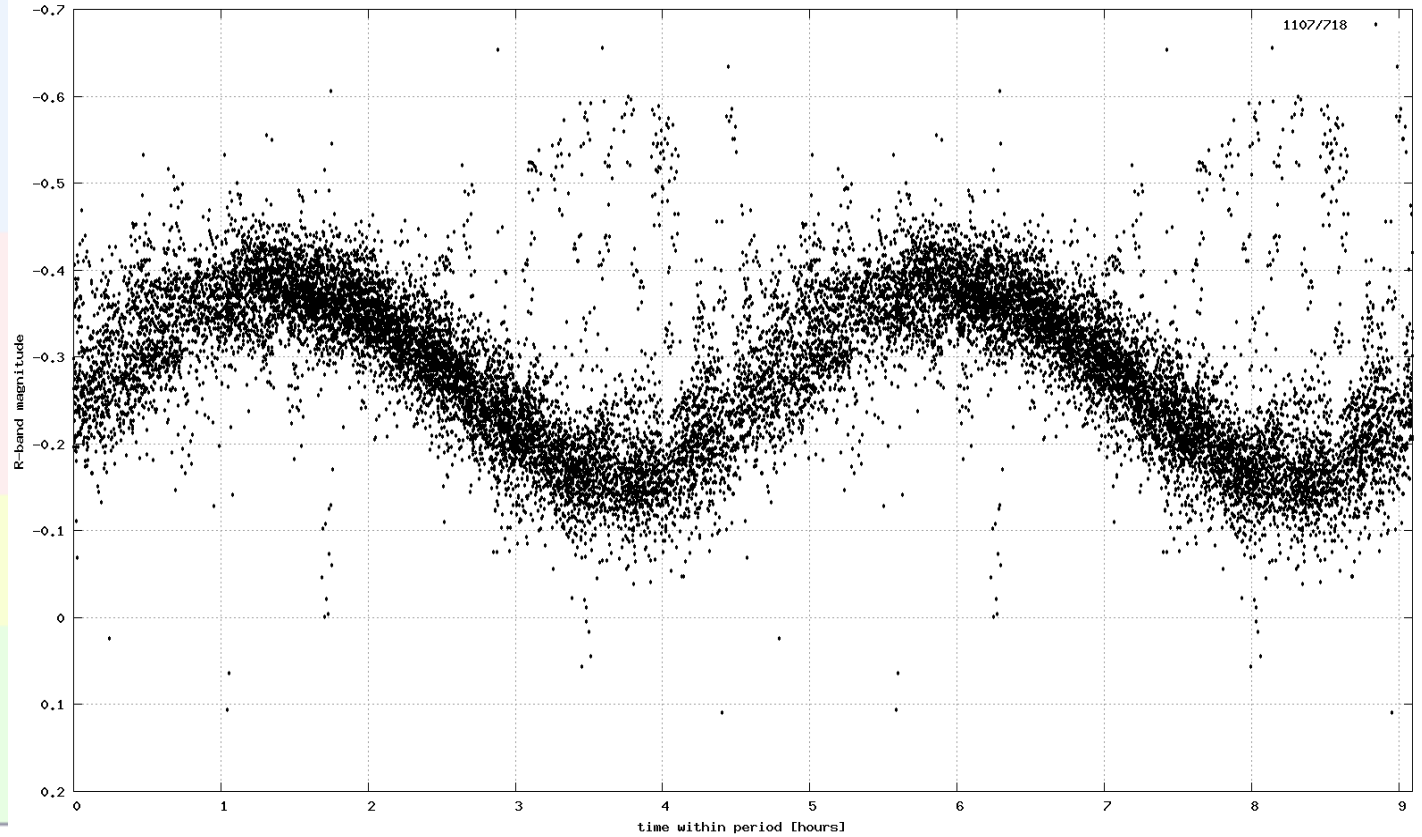
Discovering variable stars

Independently on 3rd-party catalogues

Unique approach

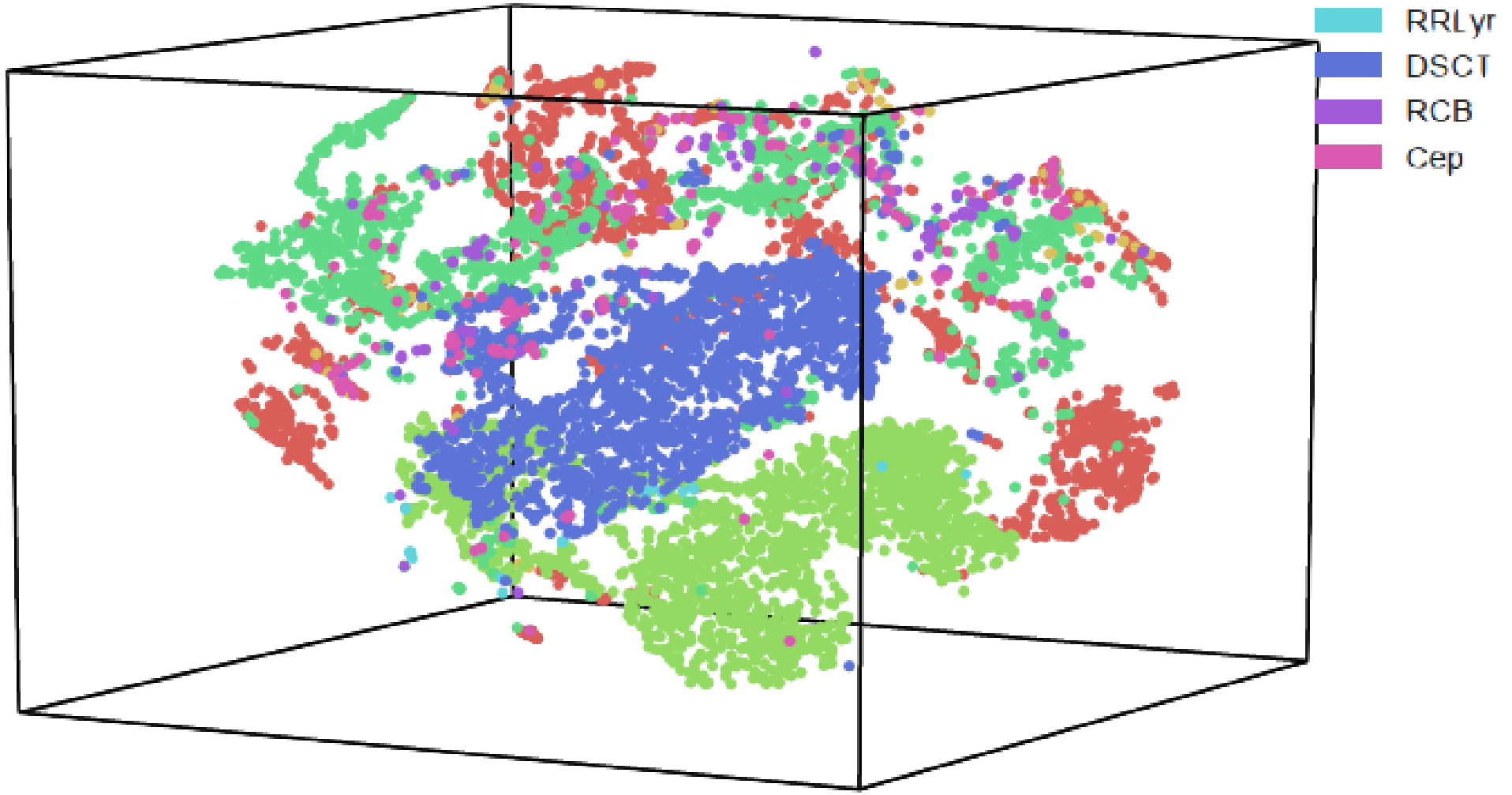
CzeV 1174 Aql

Název:	
Souřadnice:	RA: 19 10 58.92 DE: +08 30 05.06
Cross-identifikace:	UCAC4 453-089706
Typ:	DSCT
Max:	14.345
Min (prim):	14.592
Obor:	R
Perioda:	0.189456 +/- 0.000093 d
Epocha:	2457634.5625 +/- 0.00694
Rok objevu:	2017
Objevitel:	Filip Novotny, Martin Jelinek



Classification

Using artificial intelligence and novel dimensionality reduction method

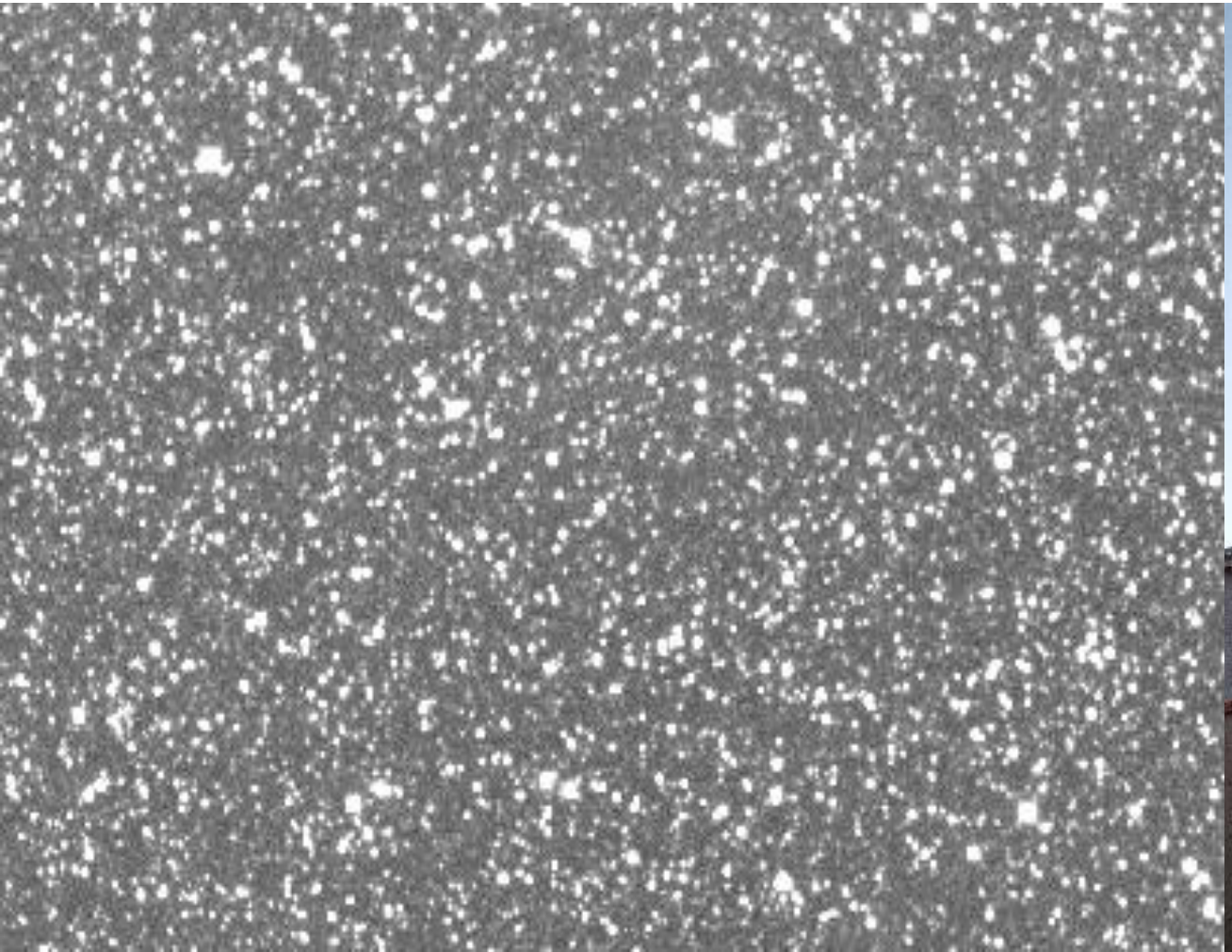


Finding stars cookbook

- Capturing astronomical images
- Calibrating images
- Extracting data
- Magnitudes synchronization
- Finding varying light curves (to spot variable stars)
- Classifying the ,,most variable“ stars

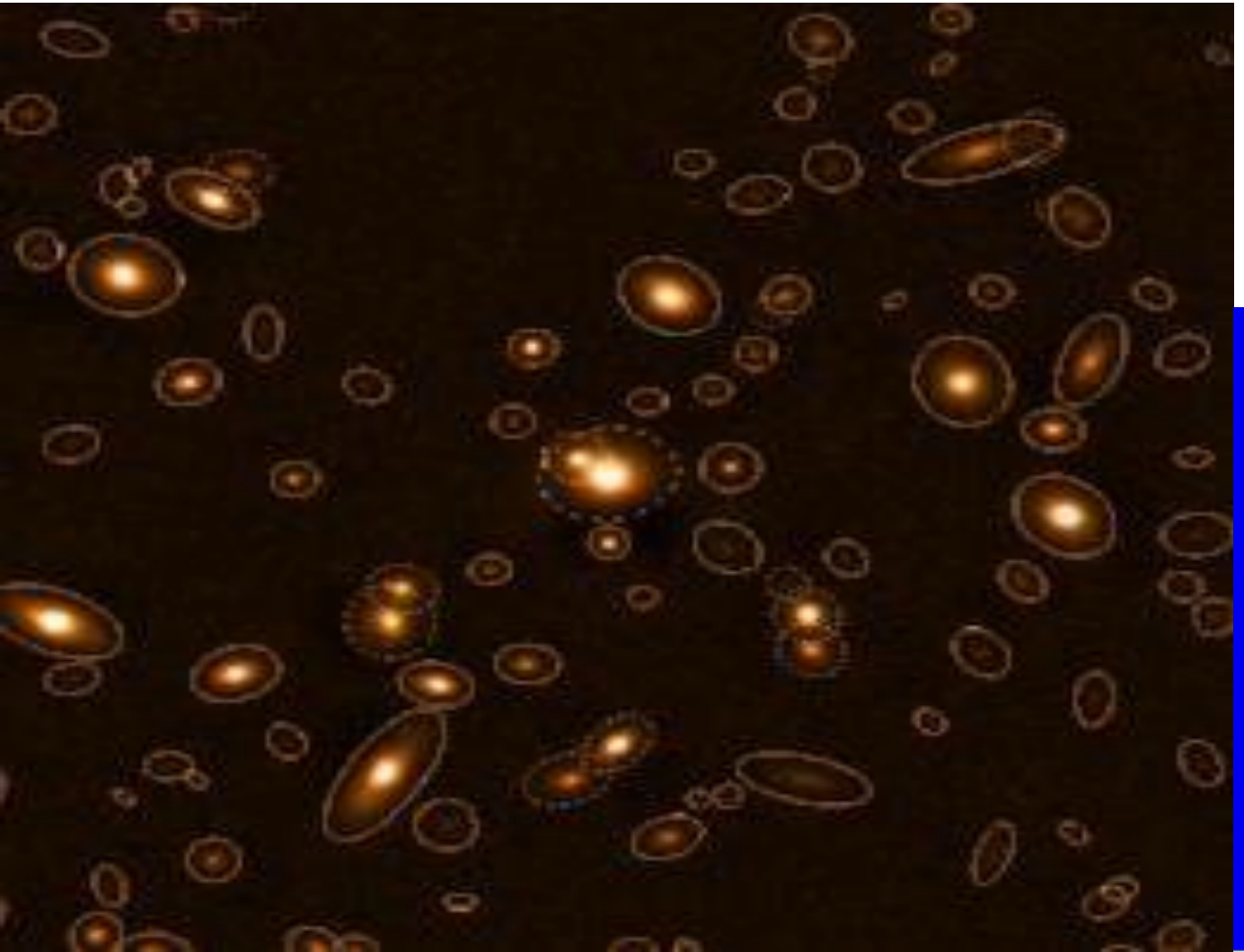
Observing the sky and extracting data

Obtaining raw data



Observing from AsU AV D50

Extracting from .FITS



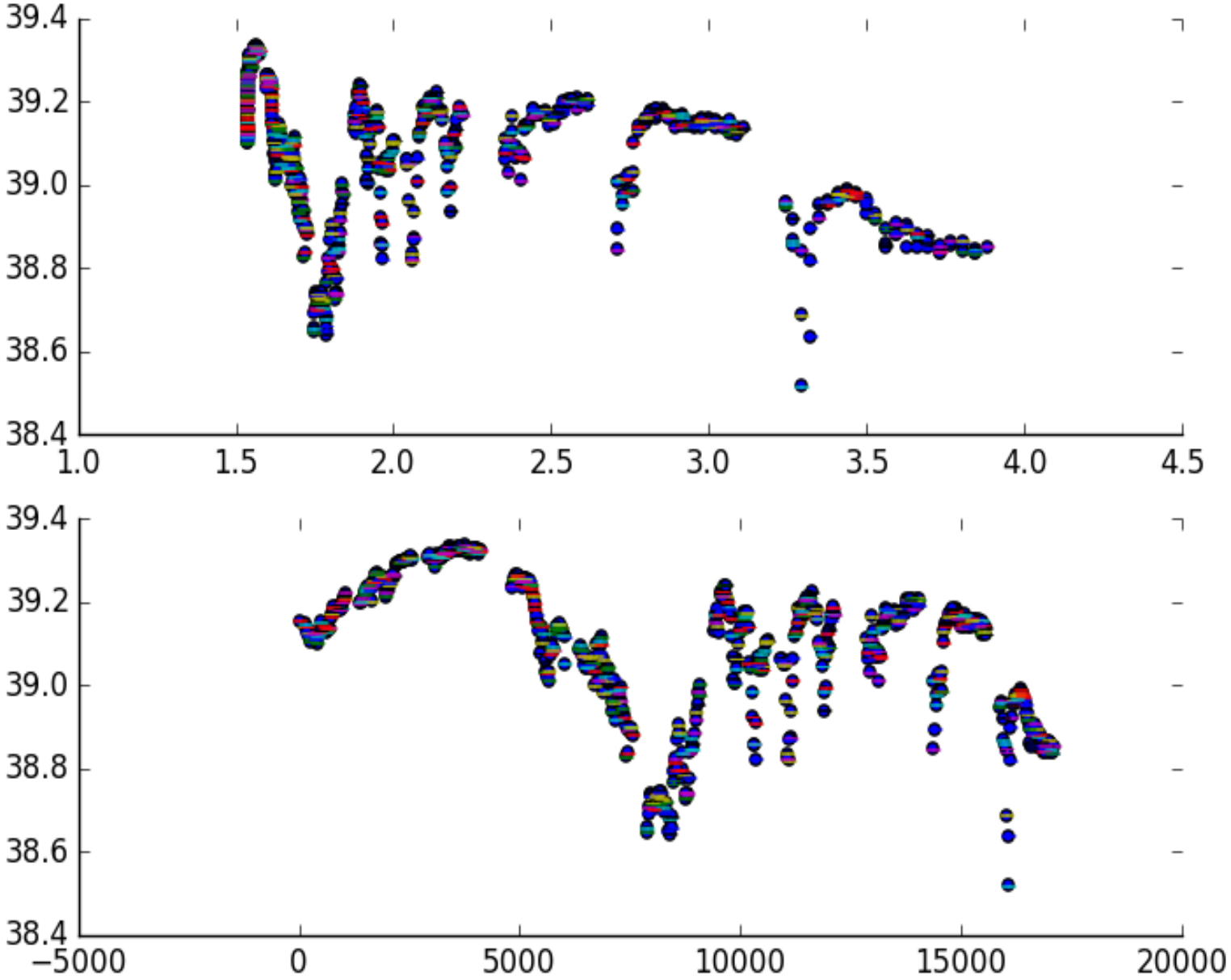
SExtractor
*v*2.13

Finding bright objects

1	295.1971713	+30.0031453	-7.3117	0.1583	152.2118	3.5481	5.6038515148e-02	8.0748162933e-03	4.63	0.617	0
2	295.1968326	+30.1572604	-7.2837	0.1429	622.3002	4.3888	3.4911047098e-02	8.1294988886e-03	4.07	0.530	0
3	295.1967809	+30.2832202	-7.1404	0.1499	1006.5091	4.4812	4.4968610232e-02	1.2343739974e-02	3.73	0.575	2
4	295.1970732	+30.0338559	-7.9799	0.1075	245.8869	3.7965	1.7914501118e-02	4.7472152531e-03	5.87	0.477	16
5	295.1966086	+30.2019611	-7.6987	0.1222	758.6502	4.9634	1.8836641197e-02	1.0024078060e-02	3.33	0.226	16
6	295.1966522	+30.0990849	-7.4655	0.1355	444.8540	4.8846	1.6675150607e-02	1.6522986674e-02	4.74	0.058	19
7	295.1962885	+30.0271287	-7.3594	0.1626	225.3760	5.8698	2.8736620874e-02	1.5870400558e-02	4.10	0.285	0
8	295.1970447	+29.9932900	-7.8168	0.1671	122.1525	3.8855	2.1787314117e-02	2.9101179172e-03	7.43	0.628	16
9	295.1965756	+30.2433412	99.0000	99.0000	884.8699	5.0358	2.8675677860e-02	2.4620995875e-02	5.78	0.546	16
10	295.1967633	+30.2770971	99.0000	99.0000	987.8323	4.5298	2.9552607279e-02	1.2464680118e-02	5.37	0.492	18
11	295.1966396	+30.1689868	-7.7803	0.1482	658.0706	4.8933	9.4562064324e-03	5.5645573322e-03	4.96	0.310	16
12	295.1963121	+30.1608214	-6.6449	0.4147	633.1676	5.7592	1.4634340238e-02	1.2863633004e-02	4.30	0.297	16
13	295.1933424	+30.2019939	-6.8967	0.2769	758.7848	13.5677	3.7167990796e-02	1.6161286061e-02	3.41	0.377	0
14	295.1958734	+30.0974163	-8.3909	0.0985	439.7727	6.9390	1.9938420968e-02	1.4589173439e-02	5.17	0.269	19
15	295.1970058	+30.1932734	-8.1634	0.1191	732.1466	3.9199	2.7272877008e-03	1.5714505514e-03	5.12	0.372	19
16	295.1967512	+30.0534379	-7.2588	0.3038	305.6198	4.6395	4.5340377969e-03	2.3211787091e-03	5.51	0.262	16
17	295.1966669	+30.2732383	99.0000	99.0000	976.0630	4.7848	3.6274258538e-03	2.1094420840e-03	8.60	0.339	16
18	295.1946018	+30.1831722	-7.8720	0.2055	701.3611	10.2577	3.1777833314e-02	2.7482587937e-02	6.12	0.201	0
19	295.1962575	+30.1201772	-7.7407	0.1970	509.1945	5.9179	6.7575540050e-03	5.0070735380e-03	4.99	0.191	16
20	295.1969540	+29.9738149	-9.4832	0.0431	62.7499	4.1317	2.2394744935e-03	1.1218965836e-03	5.50	0.286	16
21	295.1946981	+30.1707623	-8.4962	0.1308	663.5070	10.0090	1.8368003136e-02	1.8852382419e-02	4.45	0.222	3
22	295.1937553	+30.0704660	-8.1478	0.1257	357.5909	12.5365	1.8404892462e-02	1.0040137869e-02	4.77	0.232	1
23	295.1968465	+30.2560742	-8.7314	0.0941	923.7061	4.3179	3.4296881528e-03	1.1371372024e-03	6.60	0.446	19
24	295.1965681	+30.0477105	-7.2376	0.3131	288.1519	5.1244	3.8423802331e-03	3.0252742328e-03	4.75	0.165	16
25	295.1967047	+30.1906300	99.0000	99.0000	724.0866	4.7142	3.3132160993e-02	2.4567026325e-03	10.25	0.726	19
26	295.1965732	+30.2187758	-9.0941	0.0636	809.9396	5.0507	1.9184187777e-03	1.3959960541e-03	10.07	0.220	16
27	295.1952616	+30.0804873	-8.7518	0.0747	388.1421	8.5591	1.2702467130e-02	8.8029901715e-03	5.42	0.409	16
28	295.1967008	+30.1728314	99.0000	99.0000	669.7968	4.7308	6.4294111515e-03	2.7463557586e-03	9.03	0.430	19
29	295.1914361	+30.2174040	-6.8347	0.2811	805.8098	18.5822	4.0432807639e-02	1.7423745204e-02	3.44	0.410	0
30	295.1961544	+30.1259589	-8.2274	0.1242	526.8308	6.1876	5.0495201179e-03	4.6182956855e-03	4.56	0.198	16
31	295.1962852	+30.2515601	99.0000	99.0000	909.9429	5.7975	7.9477444771e-03	3.6475933508e-03	7.74	0.339	16
32	295.1949286	+30.0900306	-7.8469	0.1252	417.2549	9.4335	1.4182079940e-02	2.6561230730e-02	6.44	0.273	0

Fail

Data are not cross-calibrated



Reference stars method

- Finding reference stars in images
- Comparing to web catalog
- Calculating zero-point $m = m_m - Z_{img}$

Disadvantages

- Relyance on web catalogue data (contains error)
- Complicated search of reference stars (and not always possible)

Self referencing

Lets say most of the stars are constant

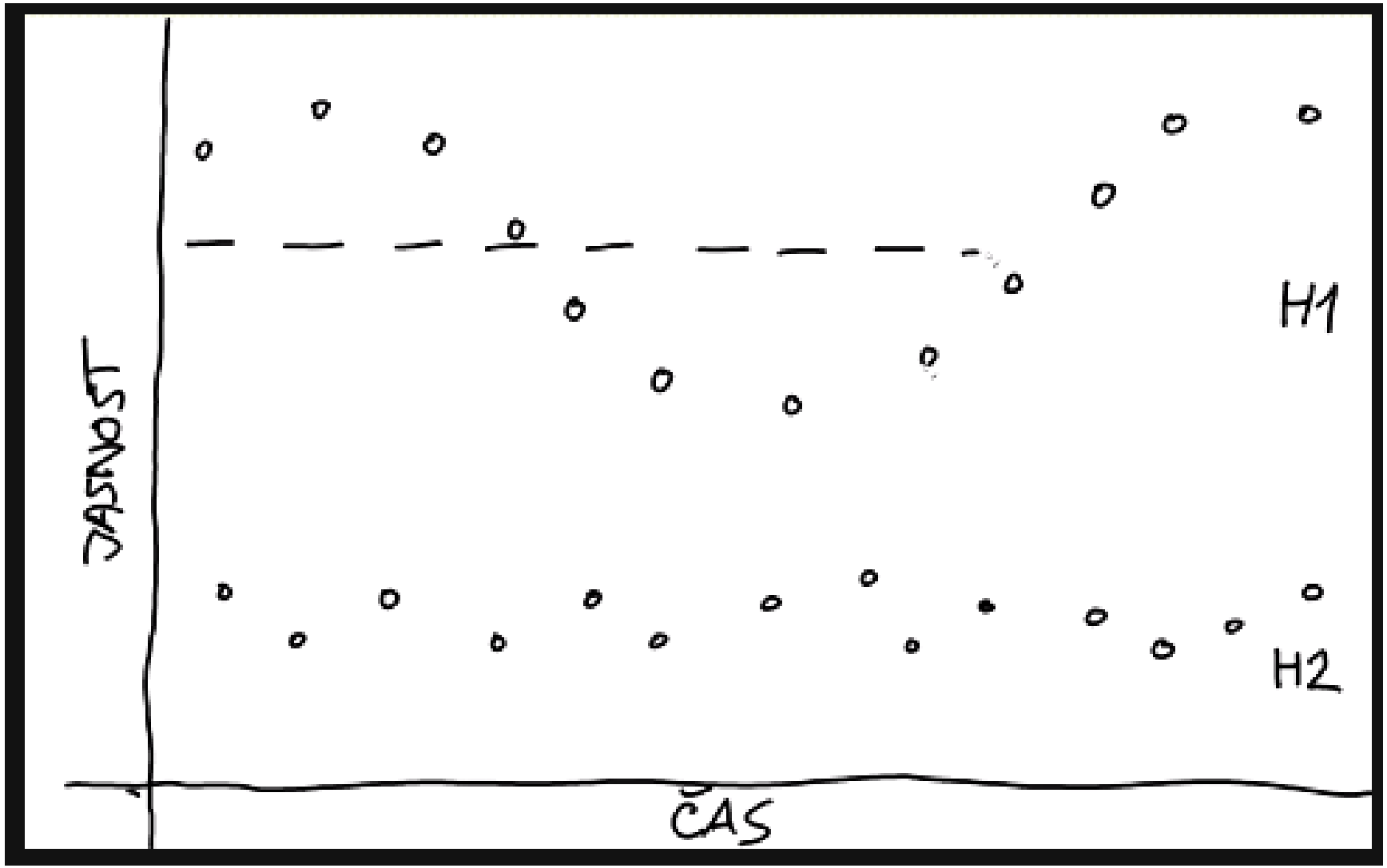
Most stars have a flat magnitude curvature

And we are finding only the varying data

Magnitudes are relative, they only make sense within this data source

$$s^2 = \frac{\sum(x - \bar{x})^2}{N - 1}$$

Schyzo parameter

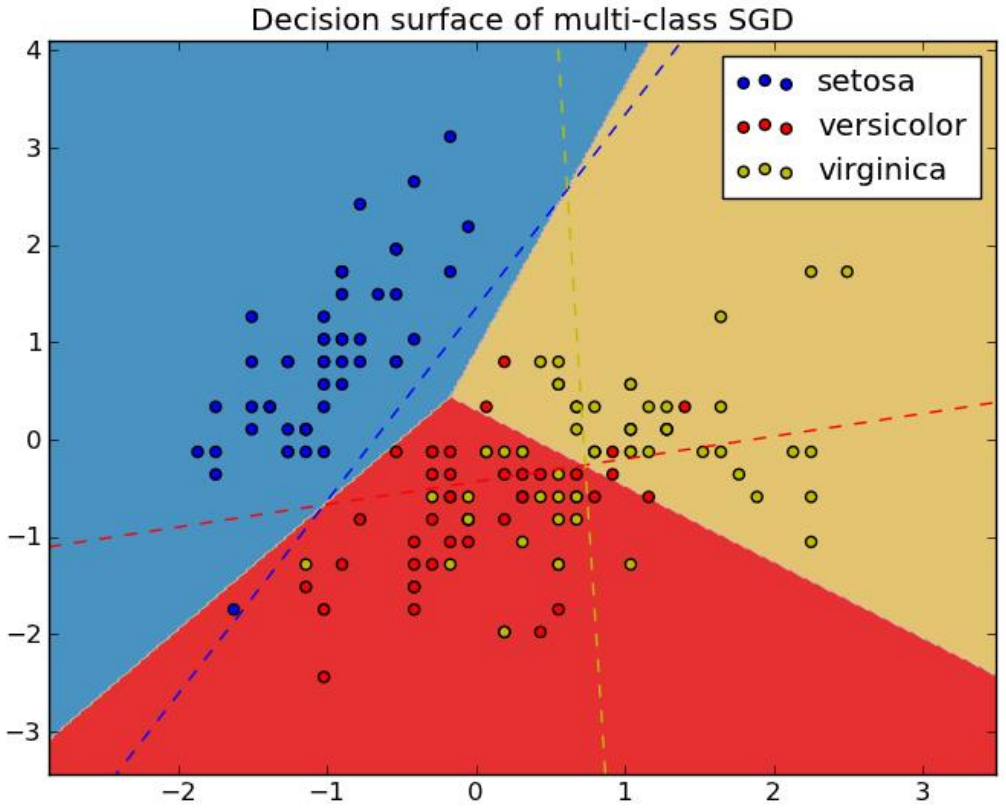


Variability index
=
Variance/schyzo

Catalogue comparison

Light curves classification

- Obtaining training data from web catalogues
- Feature engineering
- Training classifiers
- Selecting best models

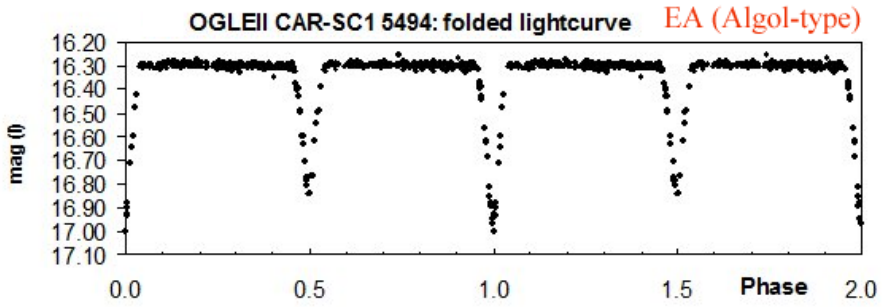


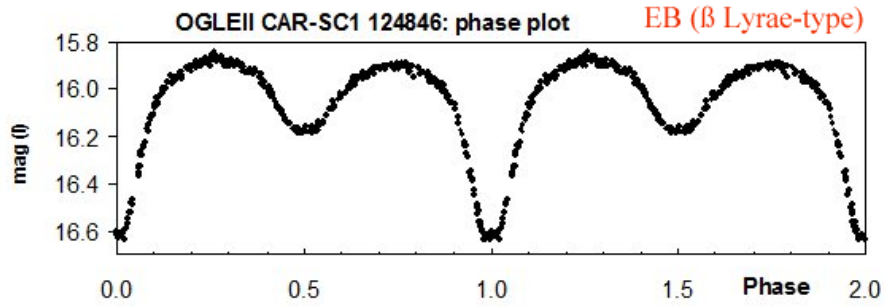
Obtaining training data

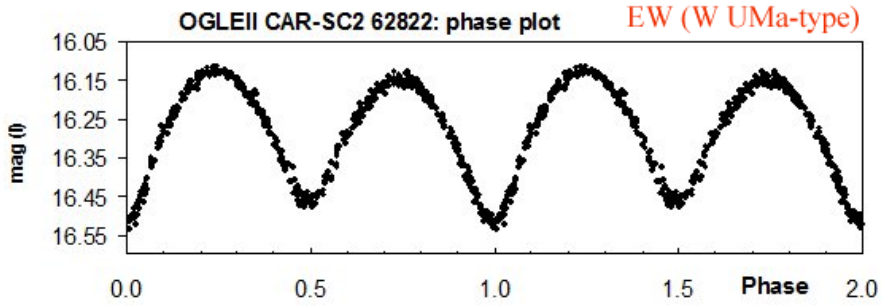
- Representative set of light curves from OGLE survey (hard to make them of simmilar distribution)

2000 Classical Cepheids
82 Anomalous Cepheids
203 Type II Cepheids
2000 RRLyr Stars
2000 Long Period Variables
137 Double Period Variables
23 R CrB Stars
2000 δ Sct Stars

Different stars – different light curves



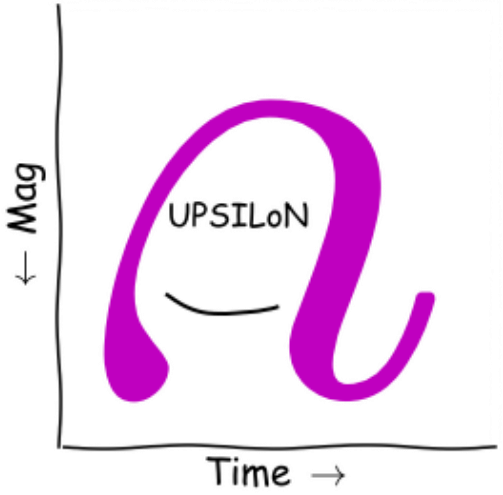




Descripting graphs by features

Period, amplitude, curvature, variance skewness...

24 features inspired by UPSILoN paper

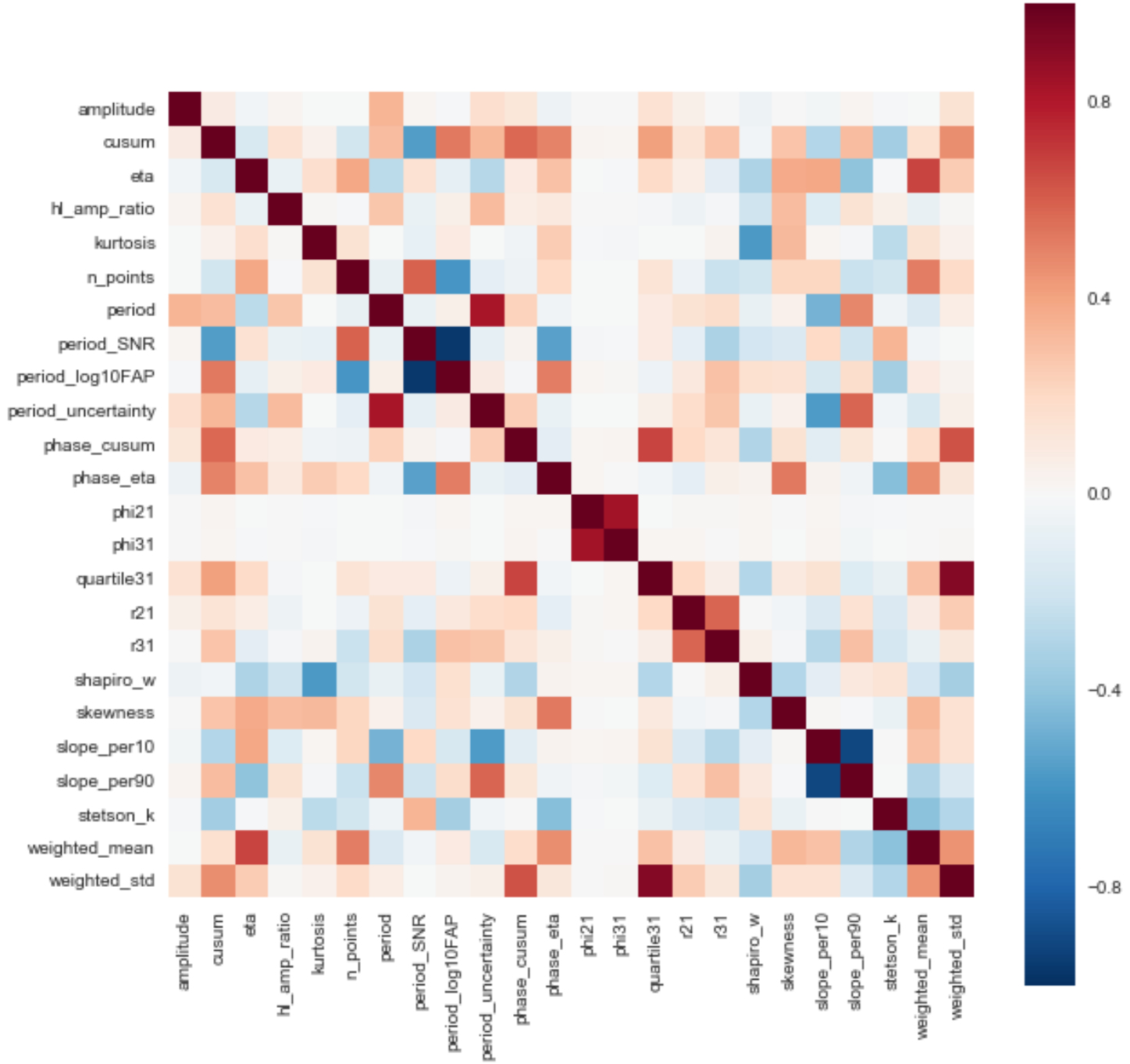


Key	Description
amplitude (+)	Amplitude from the Fourier decomposition
hl_amp_ratio (+)	Ratio of higher and lower magnitudes than the average
kurtosis (+)	Kurtosis
period (+)	Period
phase_cusum (+)	Cumulative sum index over a phase-foled ligit curve
phase_eta (+)	Eta over a phase-foled ligit curve
phi21 (+)	2nd and 1st phase difference from the Fourier decomposition

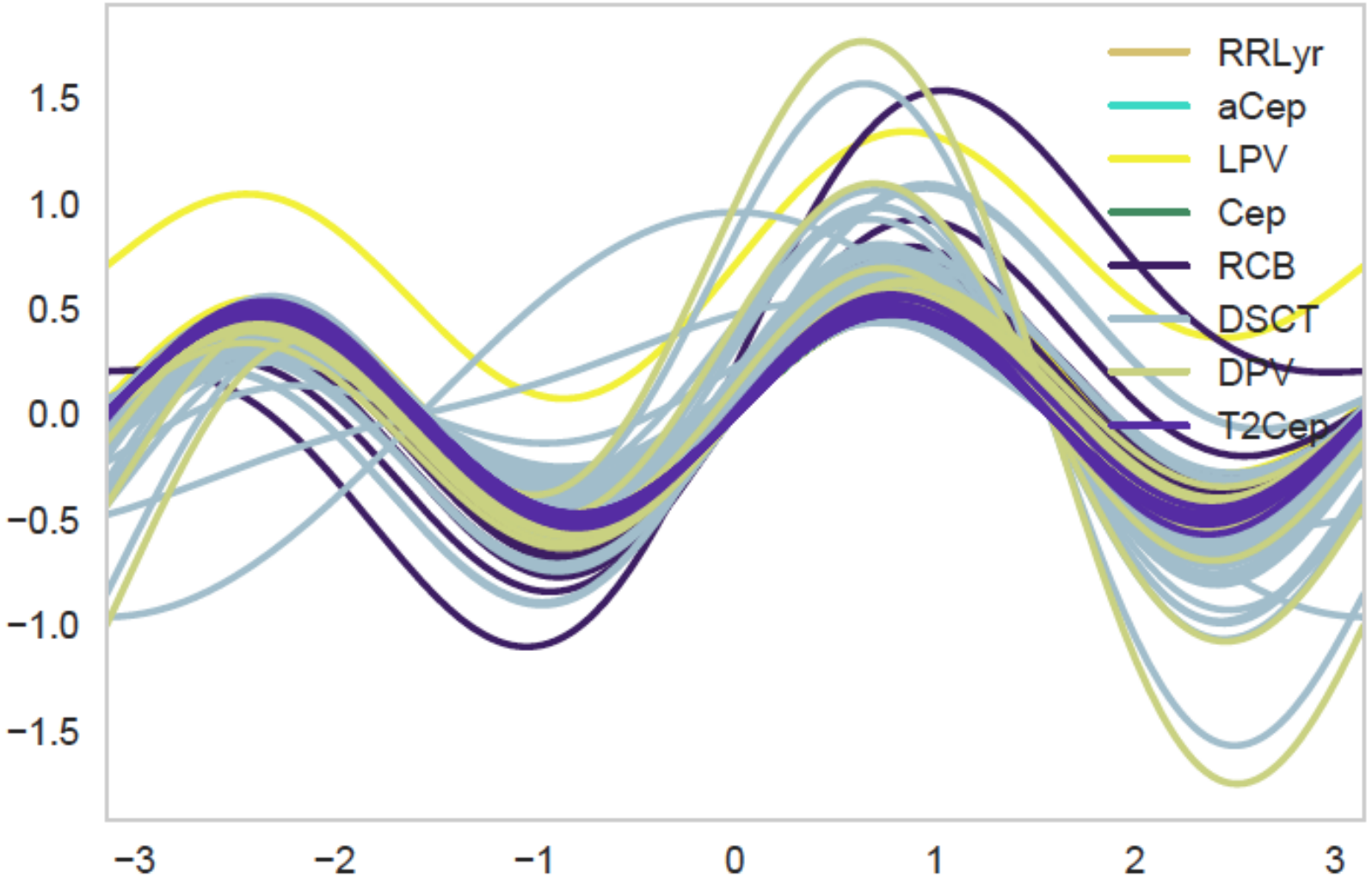
	type	amplitude	cusum	eta	hl_amp_ratio	kurtosis	n_points	period	period_SNR	period_log10FAP	...	quartile31
0	aCep	0.173048	0.126777	0.915260	0.689657	-1.110821	380	0.976699	96.164451	-49.621281	...	0.26025
1	aCep	0.221470	0.166455	1.451716	0.402886	-1.183087	725	0.381785	133.016293	-119.716574	...	0.32300
2	aCep	0.235030	0.151673	1.965137	1.035610	-1.178892	361	1.862021	99.042597	-52.283372	...	0.33700
3	aCep	0.106398	0.070888	1.999927	0.617563	-1.077602	370	0.932103	105.255041	-55.526561	...	0.14550
4	aCep	0.145742	0.084975	2.632110	0.933617	-1.075075	370	0.849591	105.562812	-56.930509	...	0.20075

Feature correlation

Lower correlation (white squares) means better differentiantion



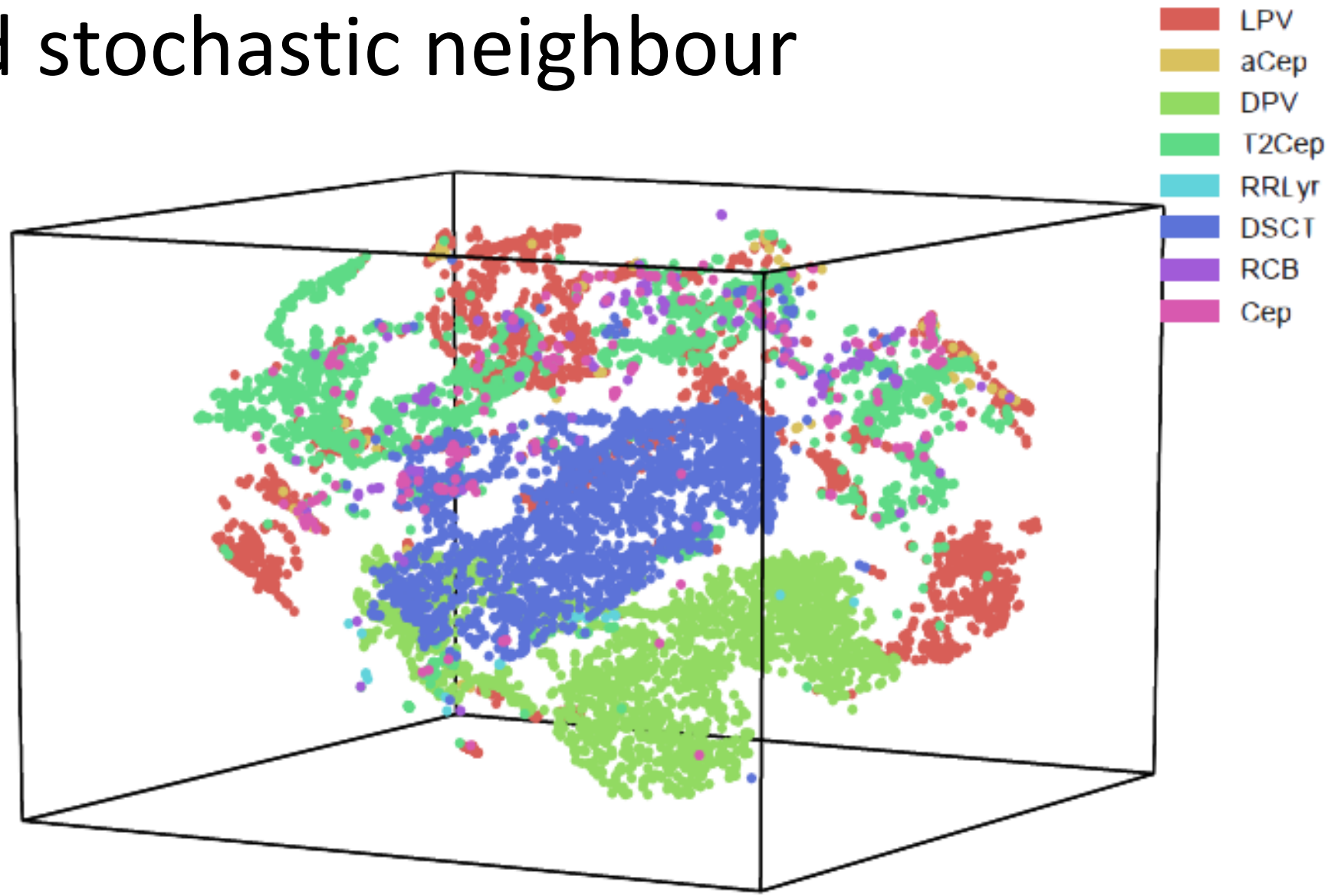
Dealing with high dimensionality

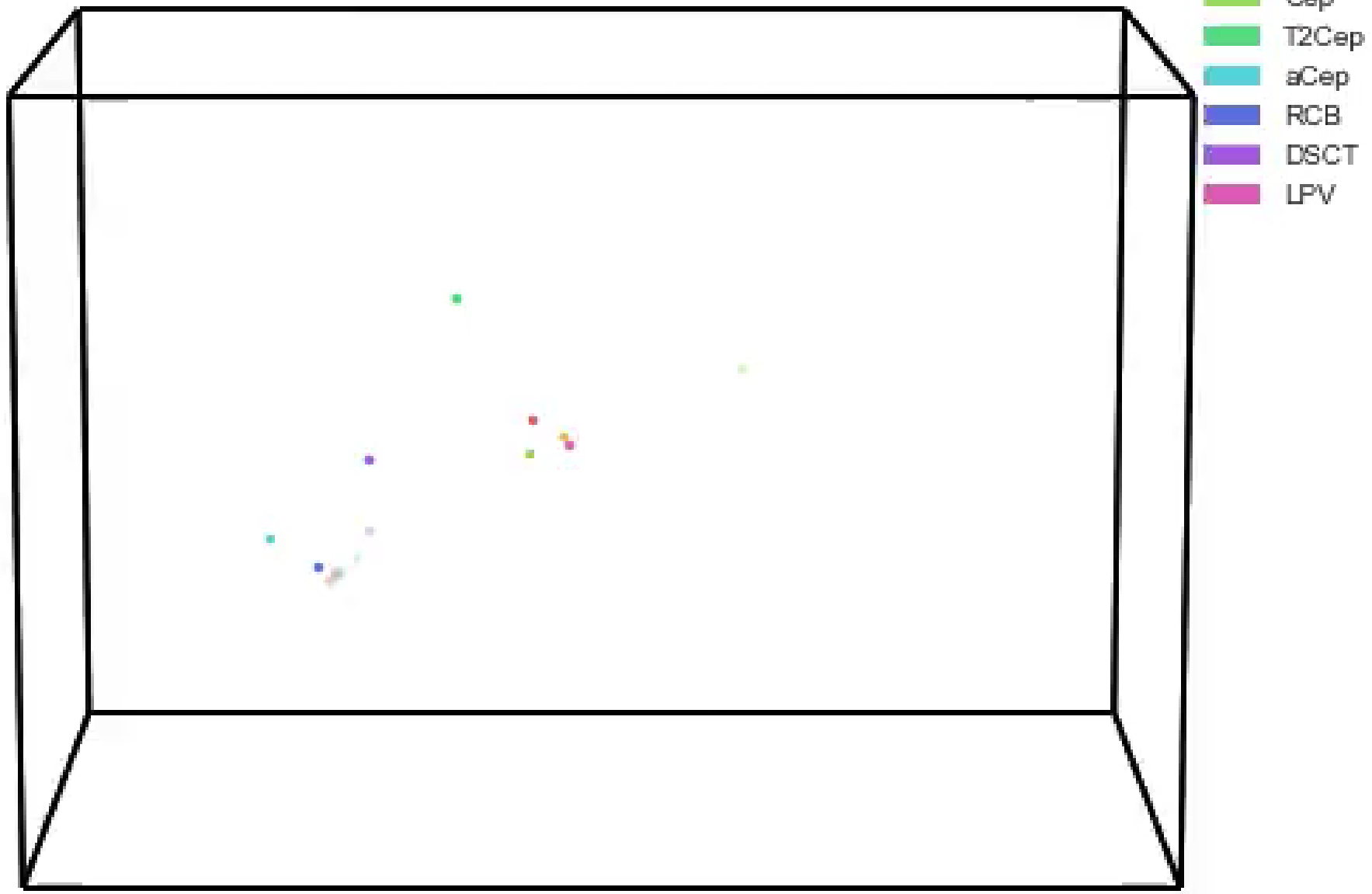


Andrew`s curves mapping of features on amplitude, curvature, period and skewness

f.e. on these features T2Cep looks more differentiated than LPV

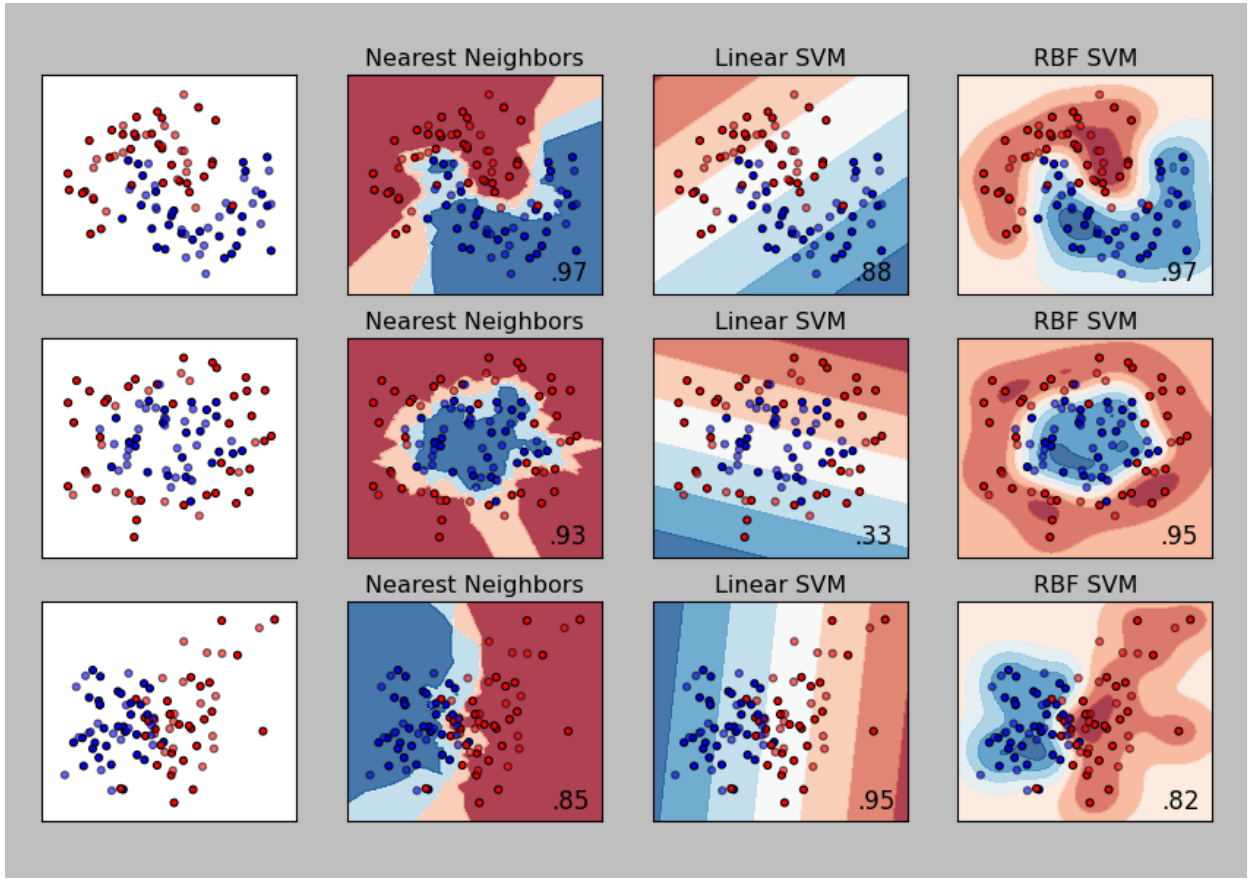
t-distributed stochastic neighbour embedding

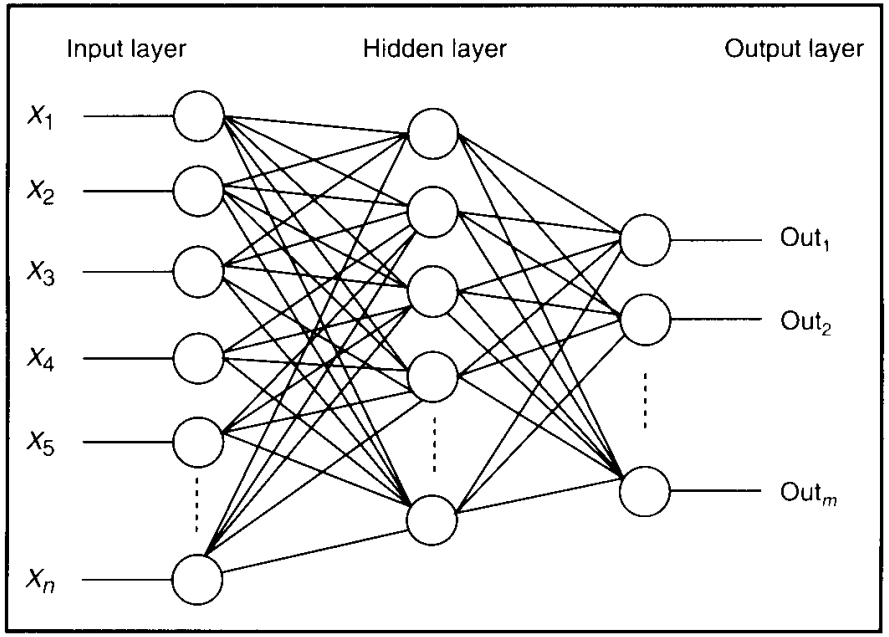



RRLyr
DPV
Cep
T2Cep
aCep
RCB
DSCT
LPV

Training classifiers

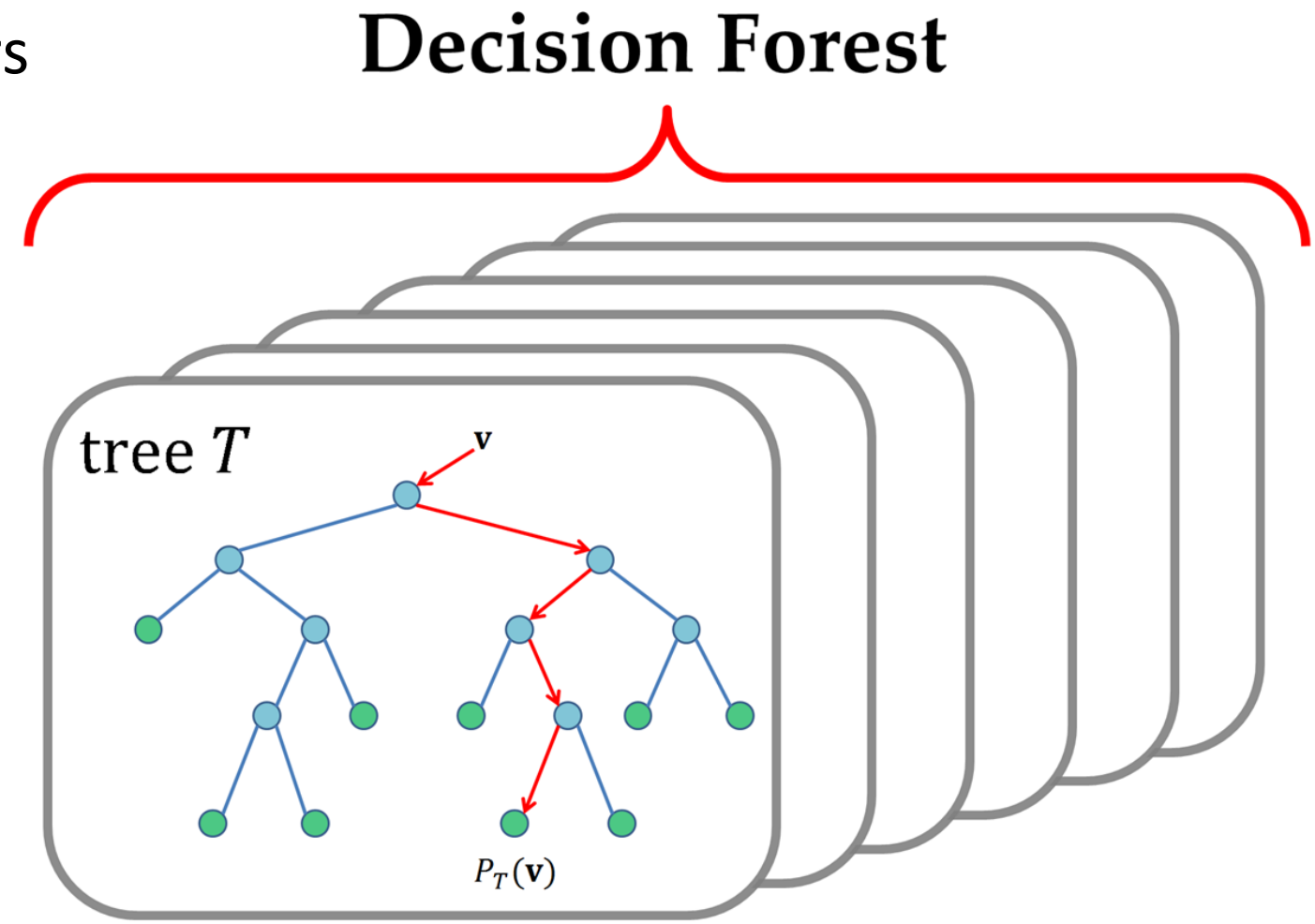
- Trying neural networks, support vector machines, logistic regression...





Random forest

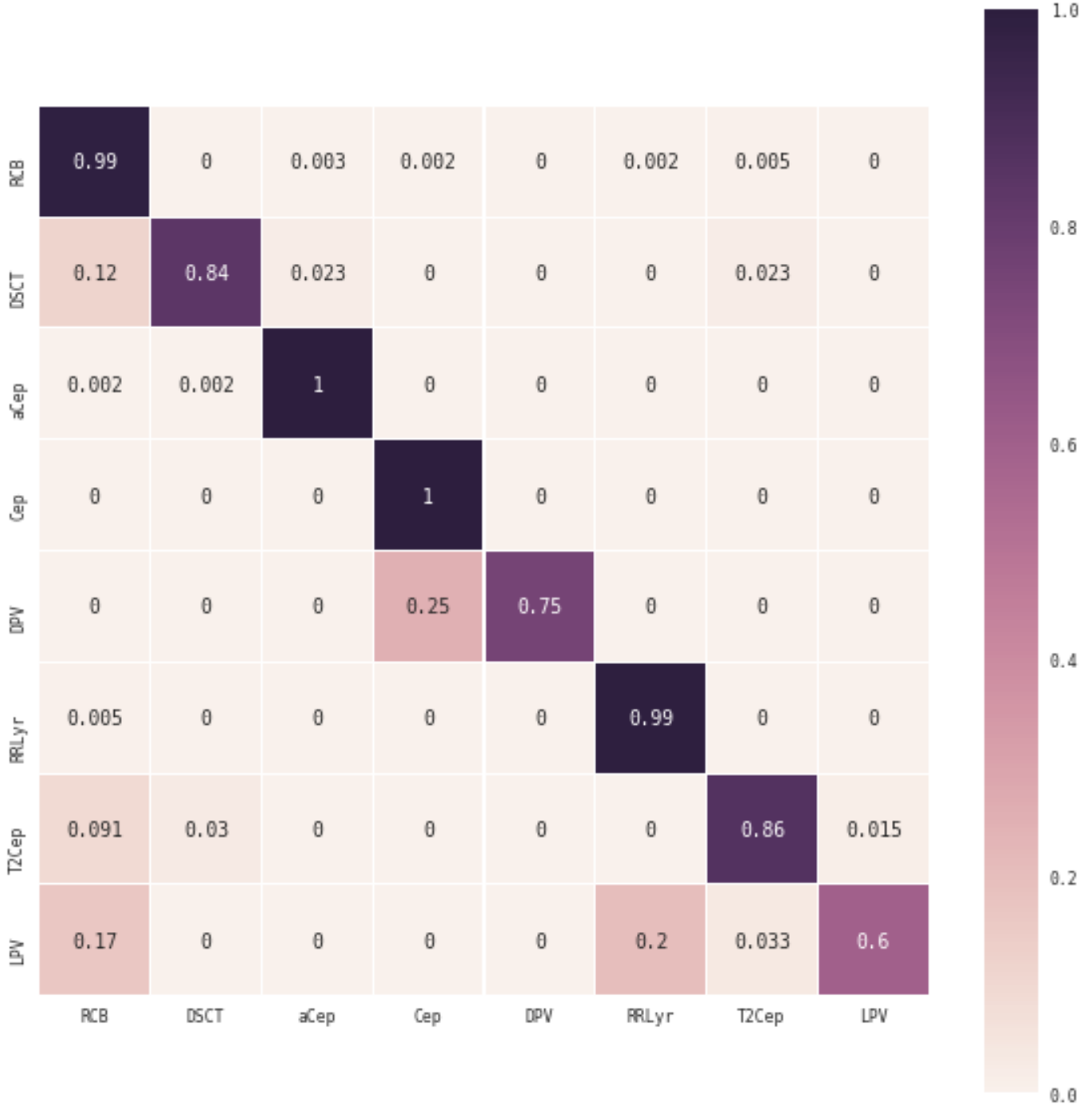
- Overperformed all others
- Able to generalize most



Results

Trained on 2/3 of training data and evaluated on remaining 1/3 (that the algortihm never seen before)

To ensure the algorithm can generalise



Command line interface

time	mag	mage
int	float	float
1472849577	-1.04948111e+01	2.28000004e-02
1472844716	-1.05995226e+01	1.97000001e-02
1472847239	-1.05626822e+01	1.82000007e-02
1472852102	-1.06074657e+01	2.67999992e-02
1472859341	-1.08553429e+01	2.48000007e-02
1472859490	-1.08047218e+01	2.26000007e-02
1472850216	-1.05295258e+01	2.40000002e-02
1472858499	-1.07993689e+01	2.00999994e-02
1472851974	-1.06263018e+01	3.04000005e-02
1472856272	-1.07689209e+01	2.15000007e-02
1472844759	-1.06238003e+01	1.93000007e-02
1472854342	-1.06142759e+01	2.50000004e-02
1472845780	-1.06167355e+01	1.82000007e-02
1472857124	-1.07746162e+01	1.89999994e-02
1472845035	-1.06064148e+01	1.77999996e-02
1472858286	-1.07329960e+01	2.03000009e-02
1472852902	-1.06398029e+01	1.93000007e-02
1472856123	-1.07224150e+01	2.06000004e-02
1472846133	-1.05883675e+01	1.79999992e-02
1472854938	-1.07115955e+01	1.86999999e-02

"718.tbl" 628L, 30772C

IN

OUT

```
The star is most probably RRLyr with probability 28.999999999999996%

___All probabilities___
LPV 9.0 %
aCep 16.33 %
DPV 12.33 %
T2Cep 17.67 %
RRLyr 28.67 %
DSCT 1.0 %
RCB 15.0 %
Cep 0.0 %
```

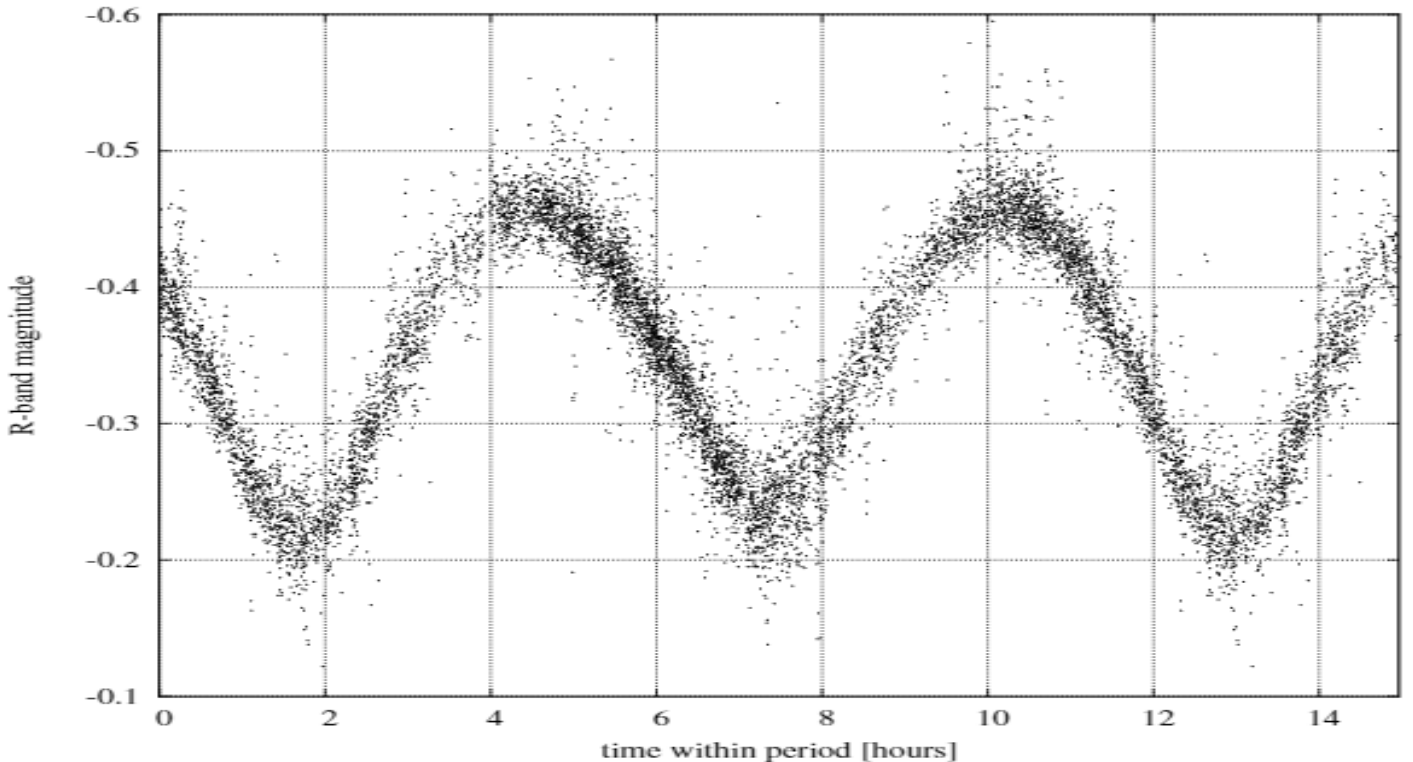
Results

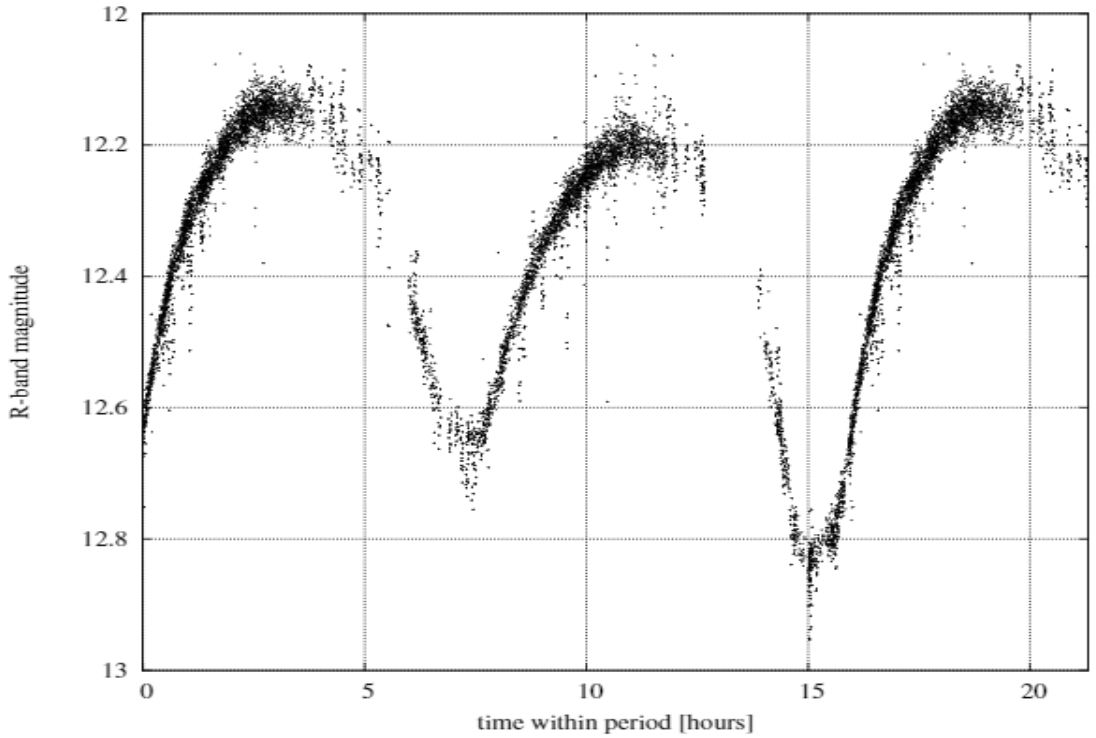
1 Officialy registrated

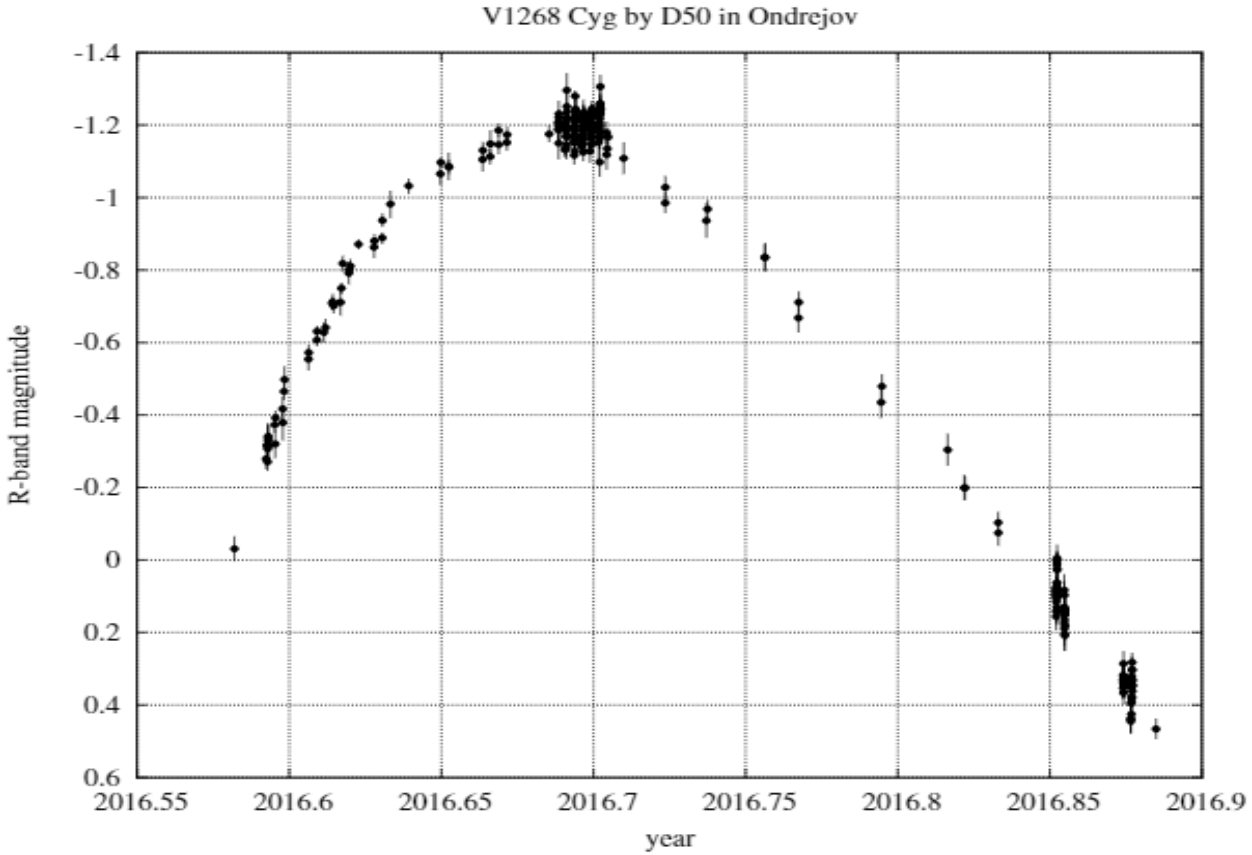
and many other pending...

CzeV 1174 Aql

Název:	
Souřadnice:	RA: 19 10 58.92 DE: +08 30 05.06
Cross-identifikace:	UCAC4 453-089706
Typ:	DSCT
Max:	14.345
Min (prim):	14.592
Obor:	R
Perioda:	0.189456 +/- 0.000093 d
Epocha:	2457634.5625 +/- 0.00694
Rok objevu:	2017
Objevitel:	Filip Novotny, Martin Jelinek

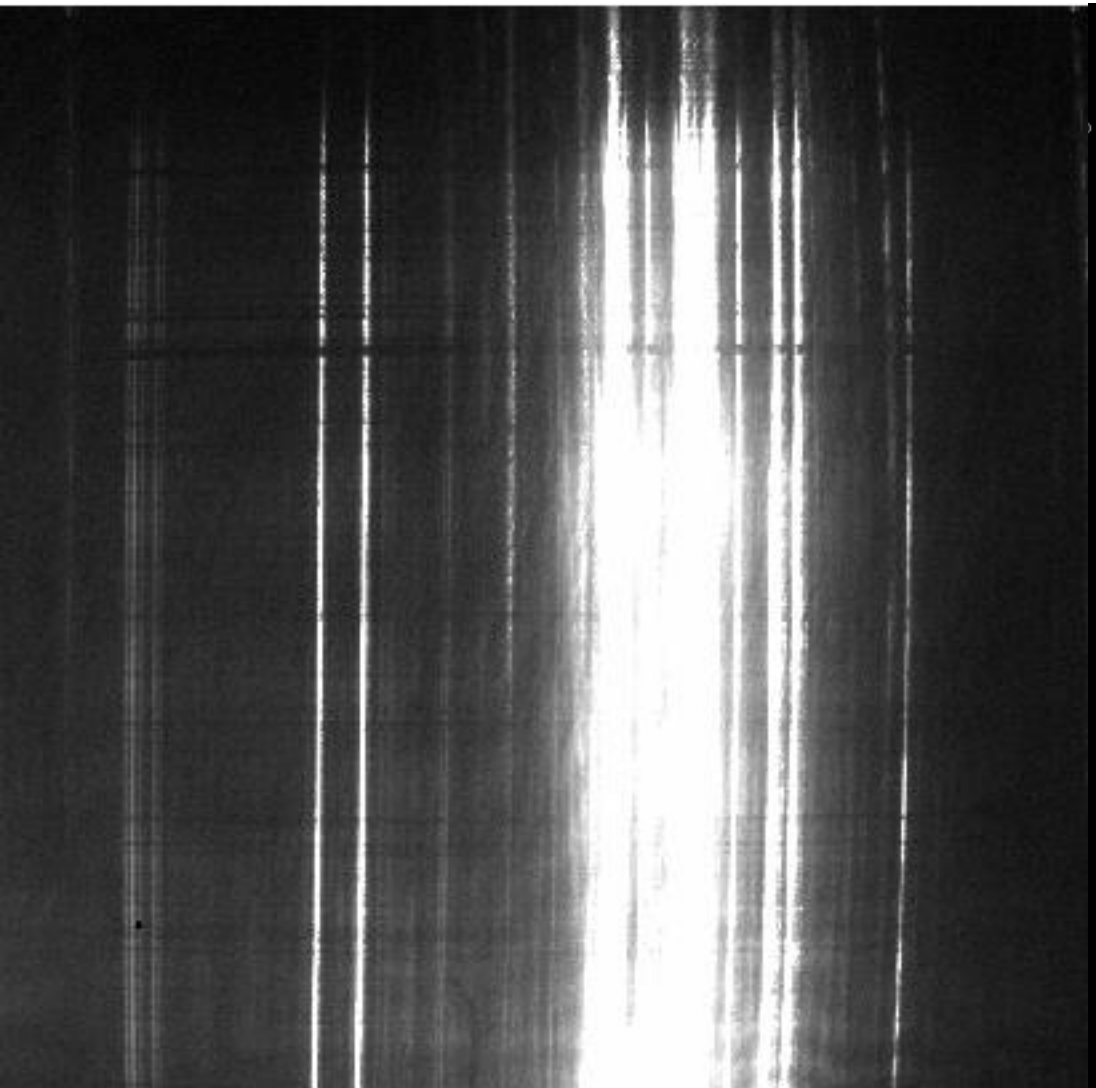


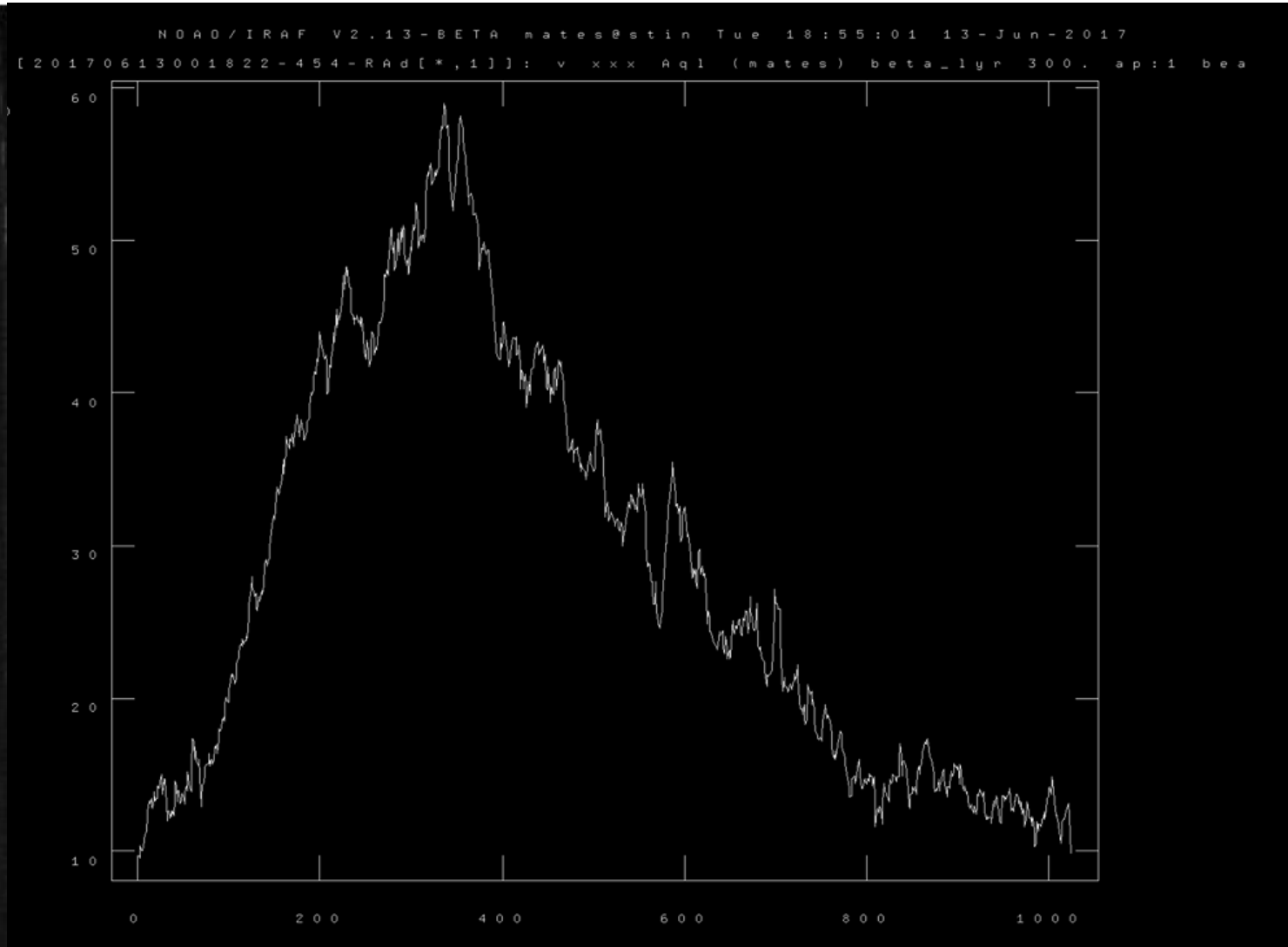




Id	*dec*	*ra*	*period*	*type*
822/01111	294.8821186	30.1943058	300d	Mira Ceti type
369/11107	287.83333329 19h 11min 20.0s	[illegible]	15.98h	W Uma type
718/11107	287.74555020 19h 10 min 58.93s	0.5016745999 0h 02min 0.40s	4.5h	RRC or delta sc
664/11107	287.7233886 19h 10min 53.61s	0.7057814999 0 hod 2min 49.39s	11.22h	not sure

Trying spectroscopy…





Impacts and social relevance

- New methods and findings
  - photometry
  - classification
- Secondary science: get most from reusing older data
  - (even students can find their own)
- Zero-cost (no expenses on new images)

Resources

Bakis, V., Burgaz, U., Butterley, T., et al. 2016, The Astronomer's Telegram, 9376

E. Bertin, S. A. 1996, Astronomy and astrophysics supplement series, 393

Nekola, M., Hudec, R., Jelnek, M., et al. 2010, Experimental Astronomy, 28, 79

Terlouw, J. P. & Vogelaar, M. G. R. 2015, Kapteyn Package, version 2.3, Kapteyn Astronomical Institute, Groningen, available from http://www.astro.rug.nl/software/kapteyn/

van Rossum, G. 1998, Python: a computer language. Version 1.5.1., Amsterdam, sticting Mathematisch Centrum.

Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." *Journal of Machine Learning Research* 12.Oct (2011): 2825-2830.

Kim & Bailer-Jones 2015 A Package for the Automated Classification of Periodic Variable Stars A&A 587, A18 (2016)

Maaten, Laurens van der, and Geoffrey Hinton. "Visualizing data using t-SNE." *Journal of Machine Learning Research* 9.Nov (2008): 2579-2605.

Udalski, Szymanski, Soszynski and Poleski, 2008, *Acta Astron.*, 58, 69 (OGLE-III photometry)

Thank you for your attention