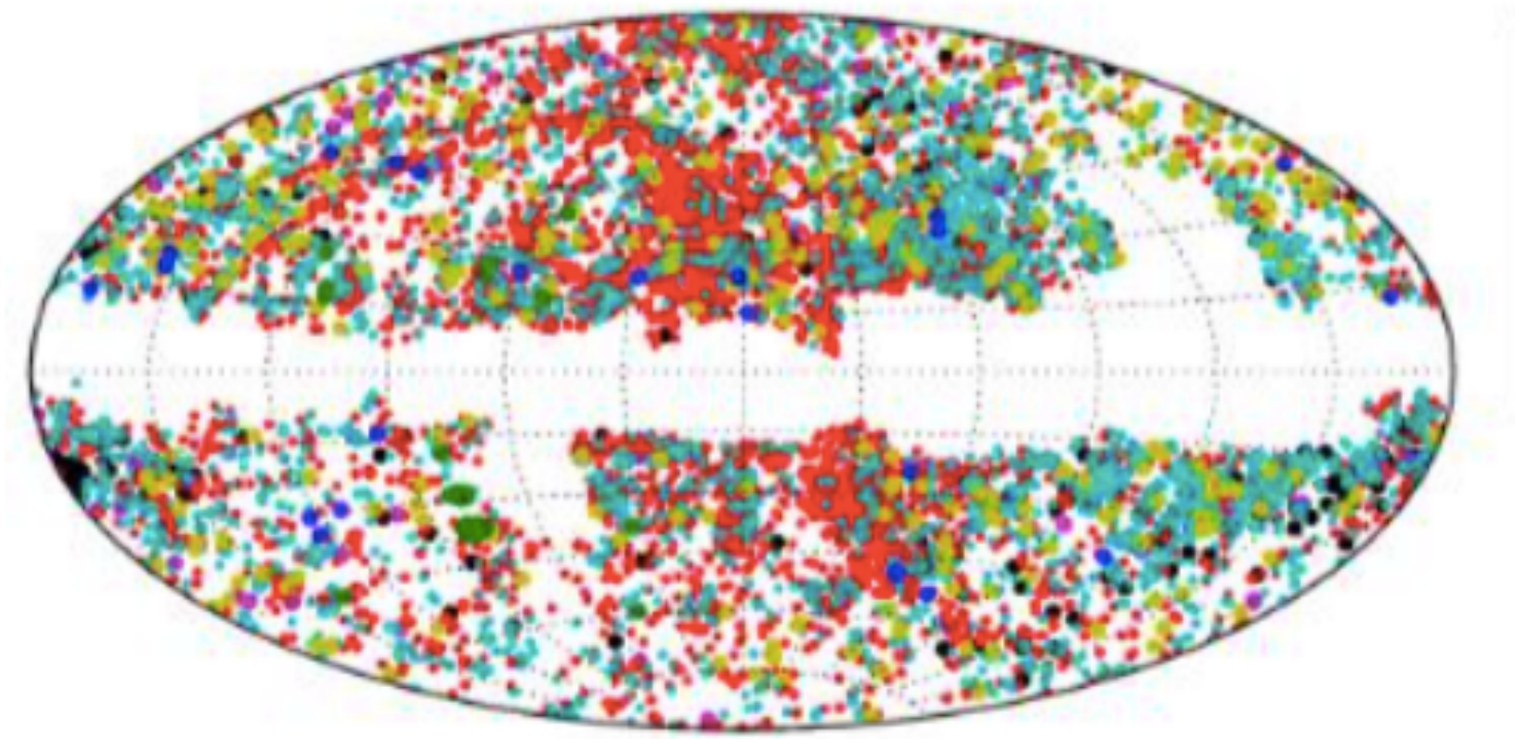
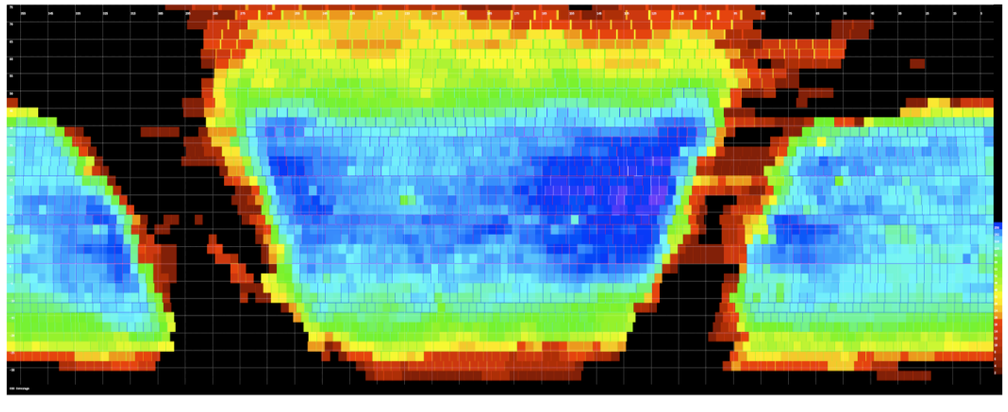
The Big Picture from the Bottom-Up



Ashish Mahabal
aam at astro.caltech.edu
Center for Data-Driven Discovery, Caltech
Co-chair, LSST TransientsVariables Collaboration
EWASS, AstroInformatics, 2017-06-30

Time-domain tentacles



- **Sky Surveys:**
  - CRTS, [PZ]TF, LSST, Pan-STARSS, ASAS-SN, Gaia ++
- **Exoplanets - detection and characterization**
- **Gravitational Waves - various aspects (e.g. EM)**
- **Math/stats approaches (e.g. SAMSI's ASTRO)**

LSST Science Drivers

all relevant to transients + variable Universe!

- Dark energy and dark matter (via measurements of strong and weak lensing, large-scale structure, clusters of galaxies, and ***supernovae***)

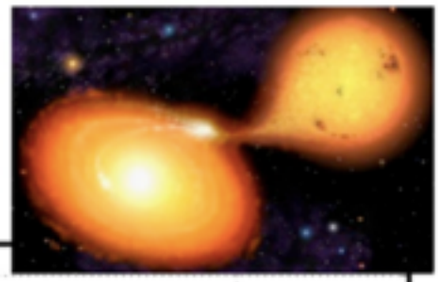
- Exploring the ***transient and variable universe***

- Studying the structure of the Milky Way galaxy and its neighbors via ***resolved stellar populations***

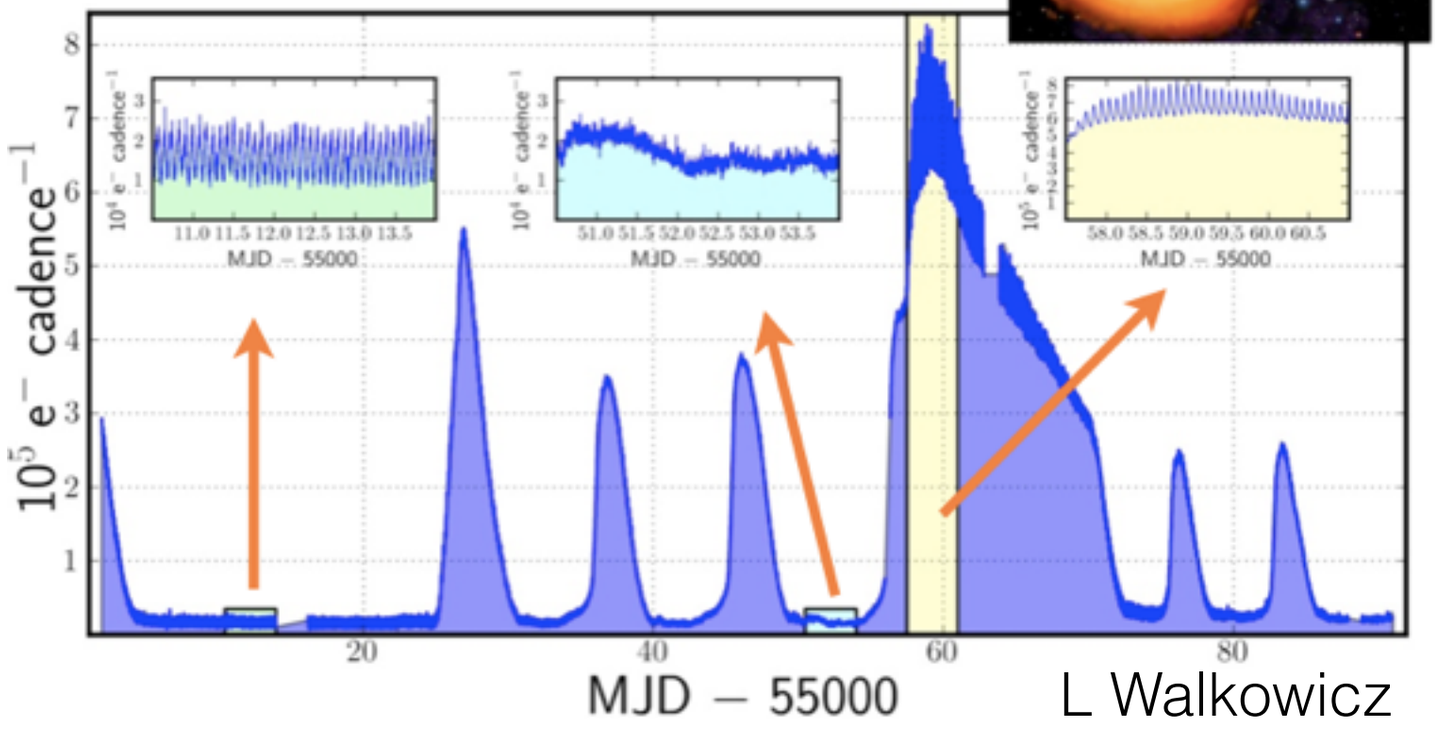
- An inventory of the Solar System, including Near Earth Asteroids and Potential Hazardous Objects, Main Belt Asteroids, and Kuiper Belt Objects ***moving objects***

Quick recap of time-domain roadblocks

- Large number of surveys
  - different apertures,
  - different filters,
  - different cadence
    - irregular
    - large gaps
    - varying error bars



Dwarf nova in the Kepler field



What do survey's do?

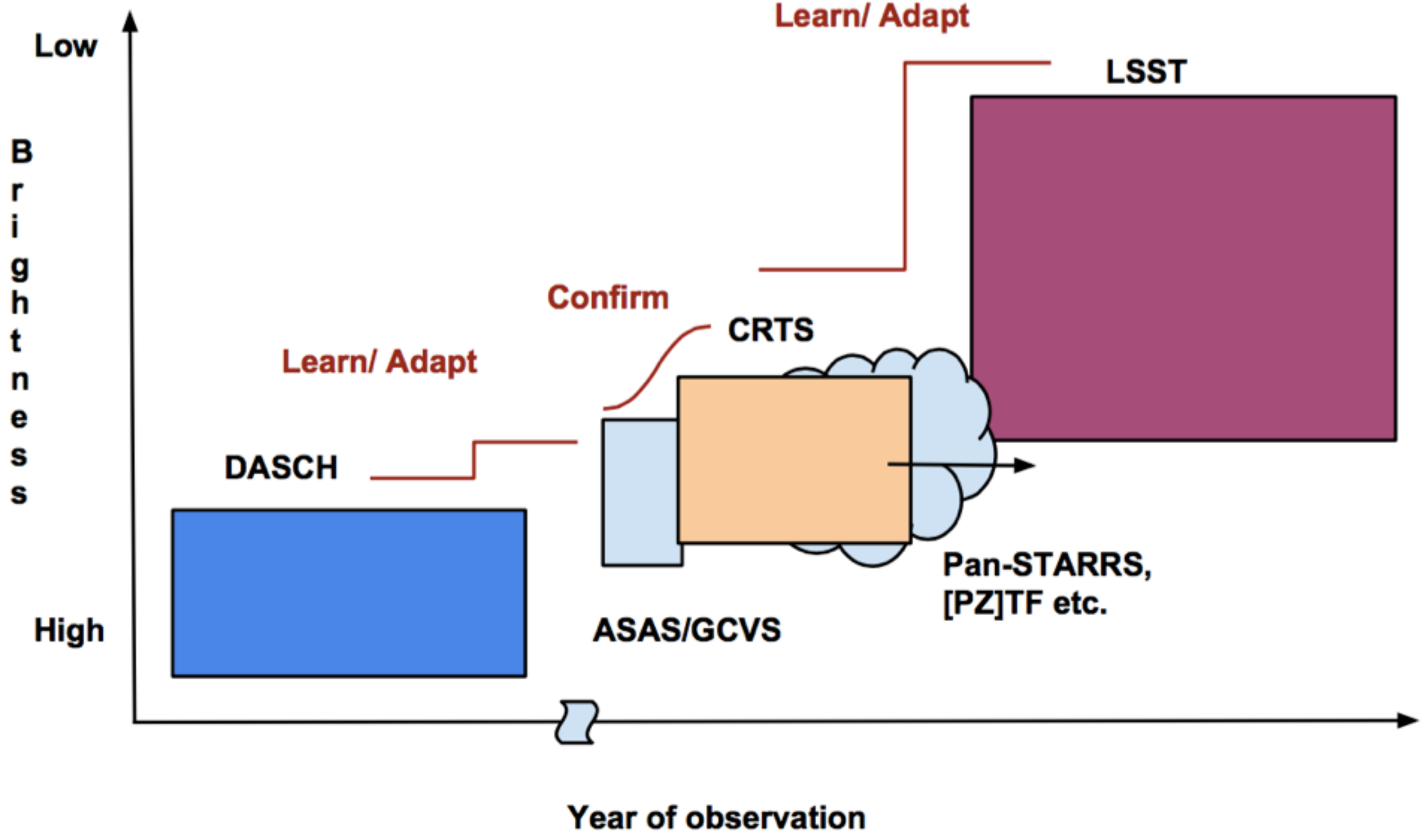

- Pick low-hanging fruit
  - select best objects, easy science
  - get spectroscopy
- That does push the envelope
  - but also leaves gaps

1 year = pi * 10^7 seconds

↓

↑ Exaggerated coverage for individual objects by ZTF, the fastest large survey (to be)

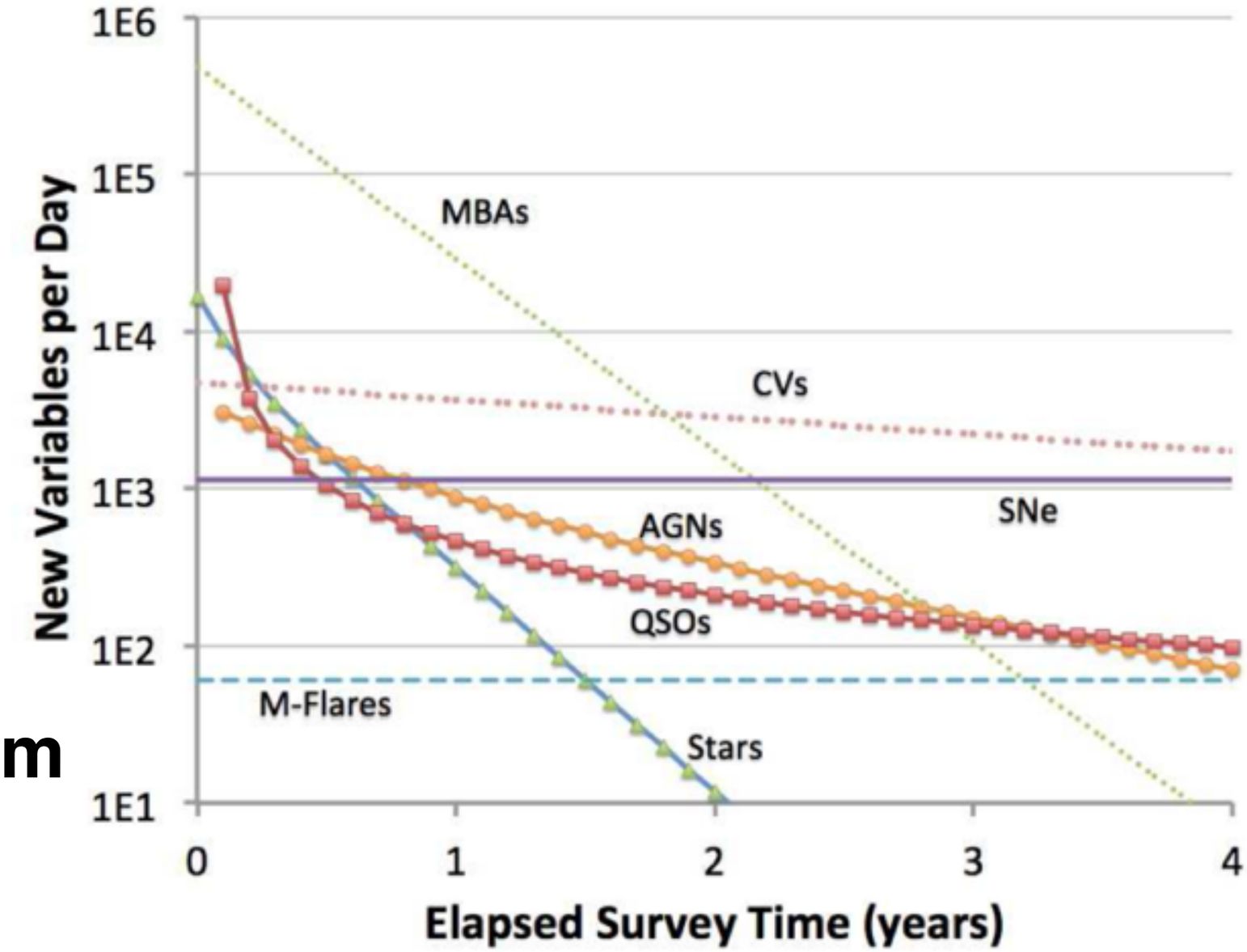
We need to be able to look at all surveys holistically



This shows just two dimensions

Number of transients will soon explode

CRTS II
ZTF
LSST



10M per night
6M solar system
20B galaxies
17B stars

1M/night with ZTF

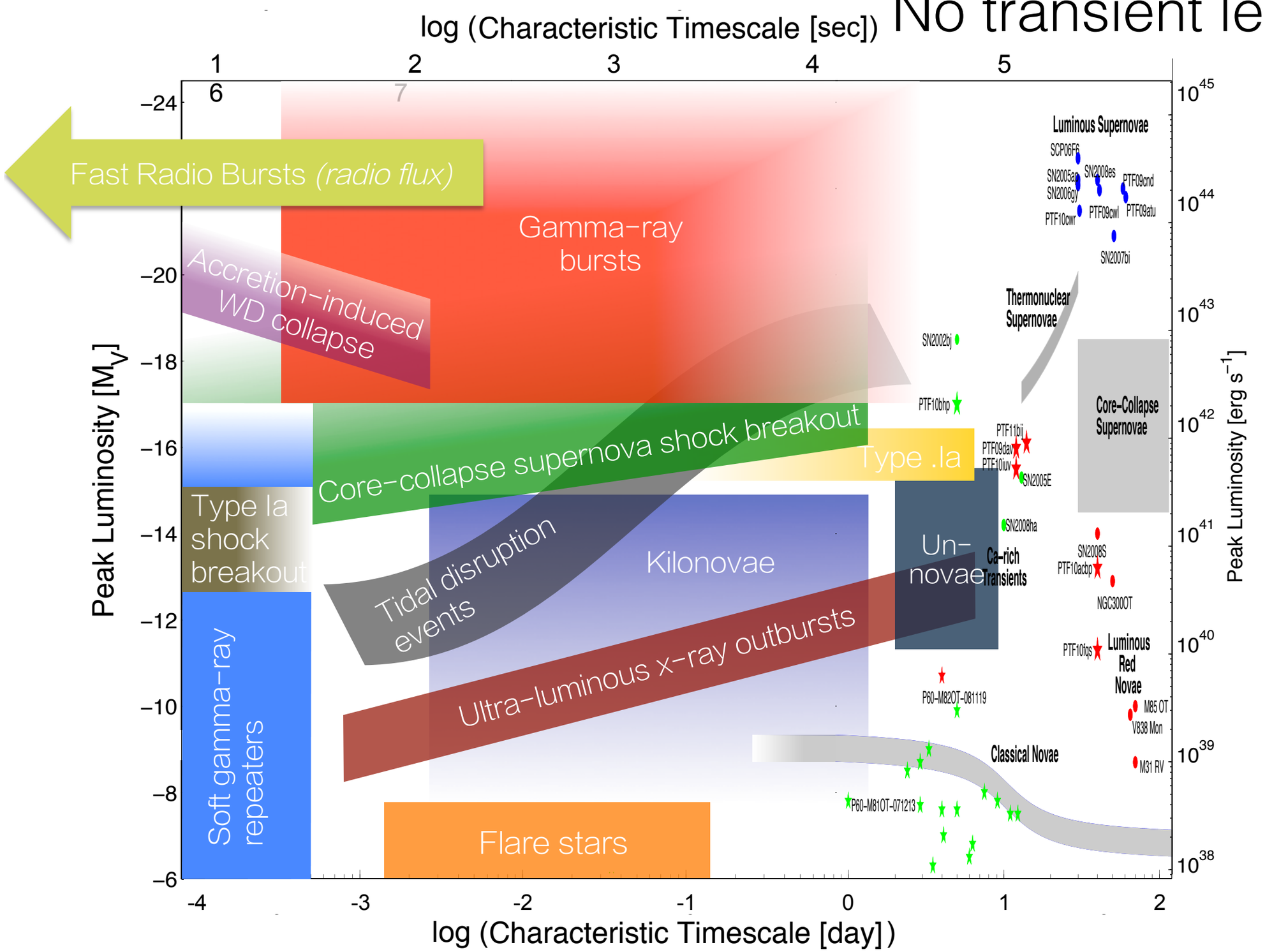
transient|variable

Ridgeway et al.,

arXiv: 1409.3265

Challenge: Characterize/Classify as much with as little data as possible

No transient left behind



J Cooke

Despite the heterogeneity, gaps, heteroskedasticity

Variable Stars

Periodic Variable Type	Examples of target science	Amplitude	Timescale
RR Lyrae	Galactic structure, distance ladder, RR Lyrae properties	large	day
Cepheids	Distance ladder, cepheid properties	large	day
Long Period Variables	Distance ladder, LPV properties	large	weeks
Short period pulsators	Instability strip, white dwarf interior properties, evolution	small	min
Periodic binaries	Eclipses, physical properties of stars, distances, ages, evolution, apsidal precession, mass transfer induced period changes, Applegate effect	small	hr-day
Rotational Modulation	Gyrochronology, stellar activity	small	days
Young stellar populations	Star and planet formation, accretion physics	small	min-days

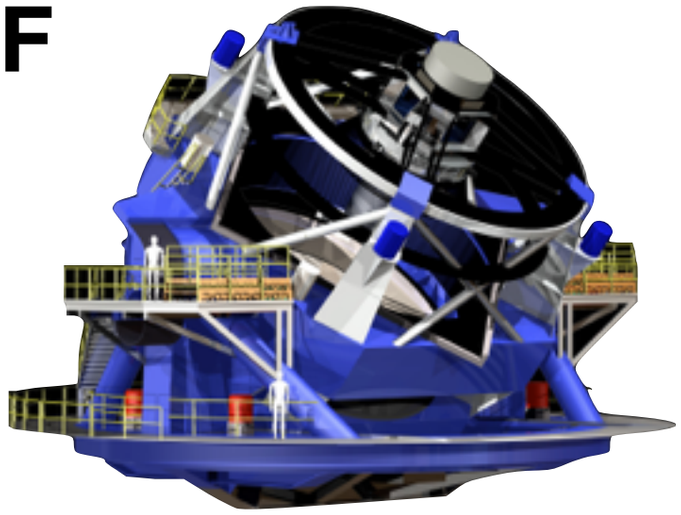
Planetary transits	Planet formation, life in the Universe	small	weeks-years

Transients

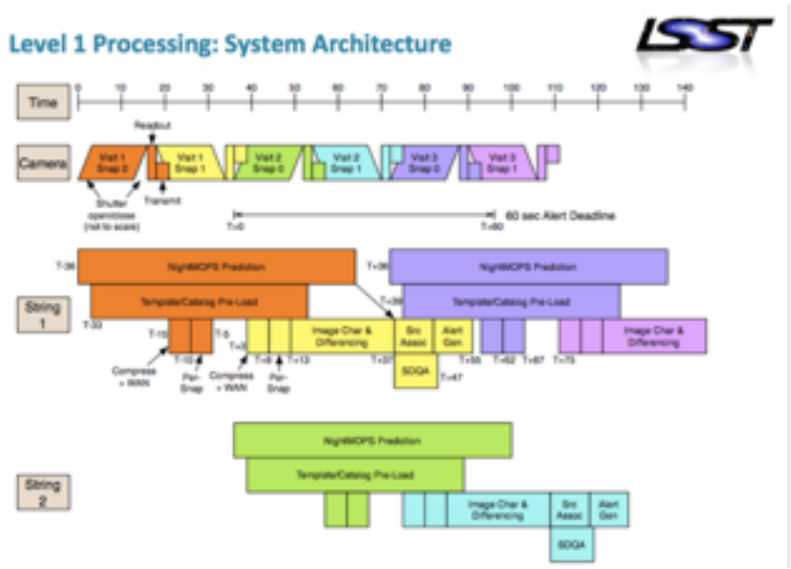
Transient Type	Science drivers	Amplitude	Time Scale
Flare stars	Flare frequency, energy, stellar age	large	min
X-ray Novae	Interacting binaries, stellar evolution, SN progenitors, nuclear physics	large	weeks
Cataclysmic variables	Interacting binaries, stellar evolution, compact objects	large	min - days
LBV variability	Late stages stellar evolution, Mass loss, SN progenitors	large	weeks-years
Massive star eruptions	Late stages stellar evolution, Mass loss, SN progenitors	extreme	weeks-years
Super Novae	stellar evolution, feedback, chemical enrichment, cosmology	extreme	days - months
Gamma Ray Bursts	SN connection, stellar evolution, cosmology	extreme	min - days
Tidal Disruption Events	Massive BH demographic	large	months
LIGO detections	EM characterization	unknown	unknown
Unknown	Discovery	unknown	unknown

Workflow

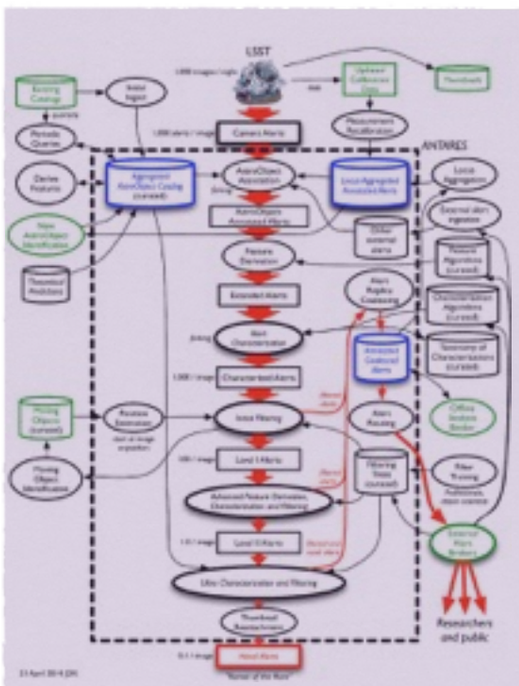
CRTS-II
ZTF
…



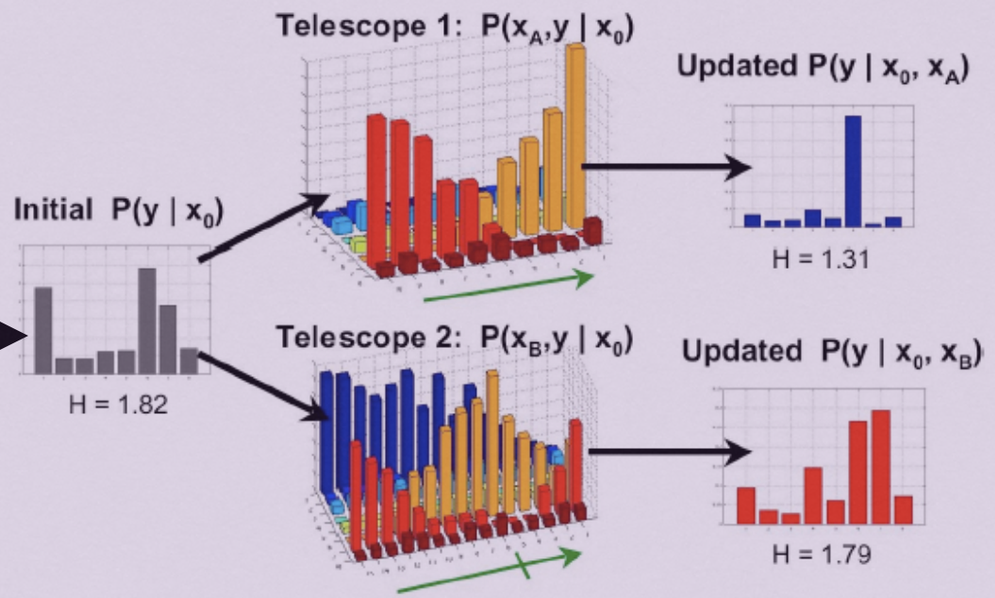
Observe



Alert



Classify



Follow-up Optimization

SAMSI - ASTRO 2016-7 program

WG2 subgroups

1. **Data Challenge**
2. **Designer Features**
3. **Scheduling Obs**
4. Interpolating Lightcurves
5. Incorporating Non-Structured Ancillary Info
6. Outlier Detection
7. **Domain Adaptation**
8. Lightcurve Decomposition
9. **Period Finding**

Overall WG2 leaders:
Ashish Mahabal
Jogesh Babu

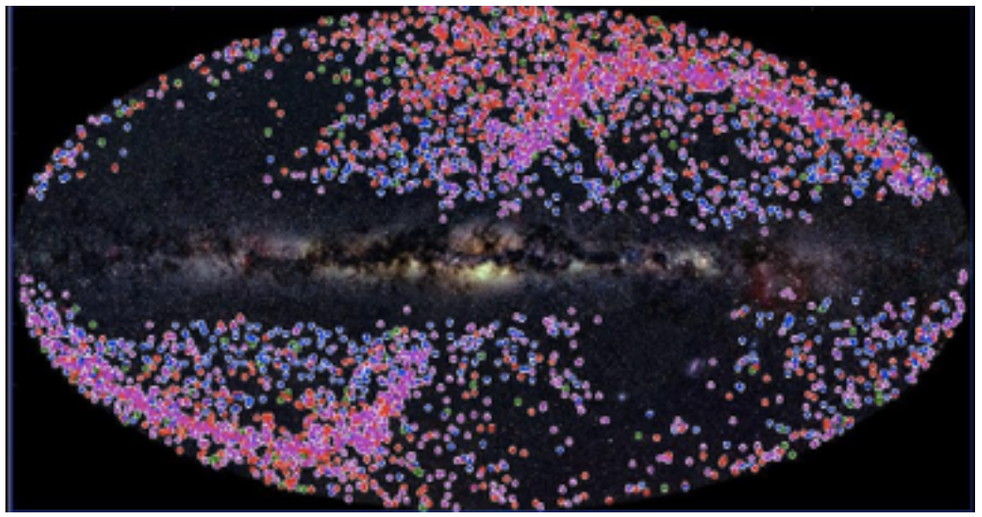
~25 members
Opening Workshop
biweekly telecons
Follow-up meetings
Connection to LSST "community"

Interconnectivity of the subgroups

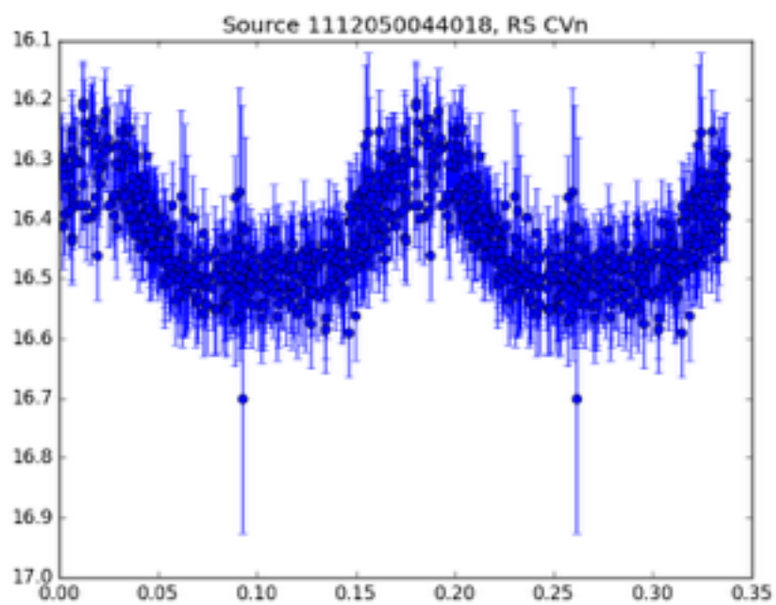
Meeting at ICTS in India (GW + TD)

Intricacies of a data challenge

- SNe data challenge (Kessler et al.)
  - full light-curves
  - first six data points

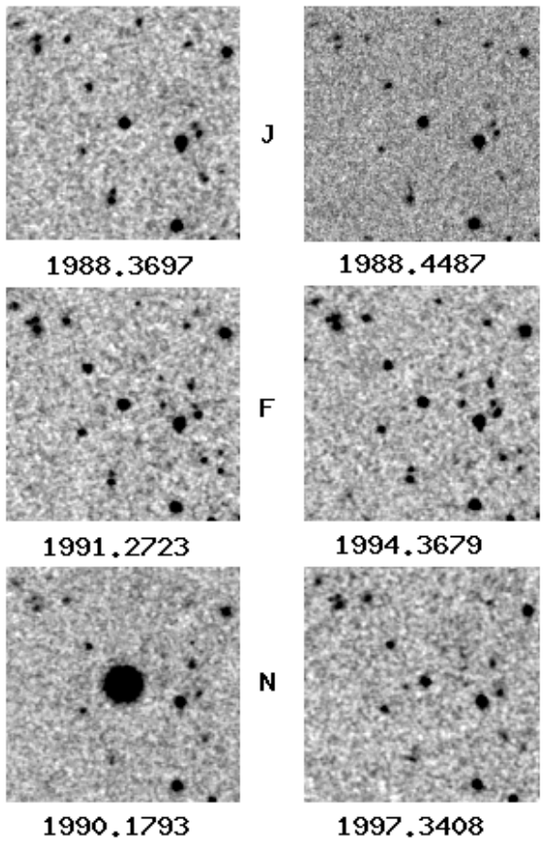


- Great3 challenge (Cosmology)
- Kaggle (Widely popular platform)
- Our plans: new challenge(s)



The data challenge challenge

Collect datasets.
And priors!
diversity in:
aperture
filters
cadence …



Attract non-experts.
Requires stripping domain knowledge
Without dumbing down the challenge

Challenges:
Not enough labels
Not enough private labels
Size of the challenge
Number of challenges

(SAMSI) Data challenge details

- Possible Datasets:
  - Catalina Real-Time Transient Survey
  - MACHOs survey
  - OGLE
  - Pan-STARRS
  - PTF
  - SDSS STRIPE82

Simulations
Theory

Rafael Martinez-Galarza
Peter Freeman
Matthew Graham
Shashi Kanbur
Vivek Kohar
James Long
Ashish Mahabal
Wenlong Yuan

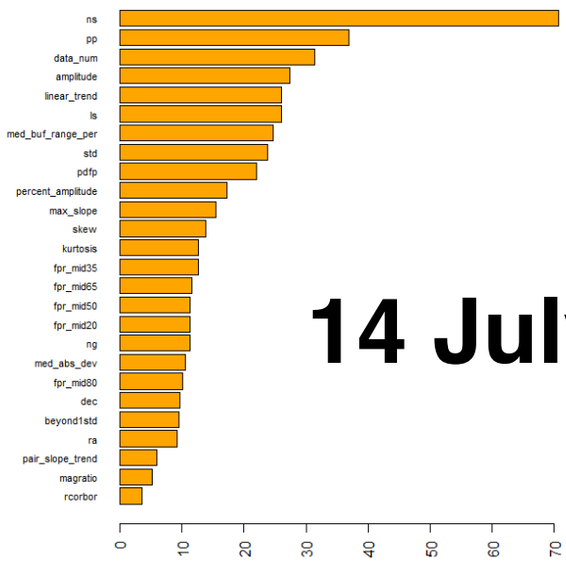
https://community.lsst.org/t/data-challenge-to-characterize-transient-and-variable-objects/1061/14

LSSTC funded meeting

The 2017ClassificationChallenge:
An LSST Photometric Classification Challenge

Ren´ee Hloˇzek, Tina Peters, Rick Kessler, Dan Scolnic, Saurabh Jha, Ashish Mahabal, Federica Bianco, Hiranya Peiris, Michelle Lochner, Jason McEwen, Robert Schuhmann, Rob Firth, Mark Sullivan, Alex Malz, Llu´ıs Galbany, Emille Ishida, Rahul Biswas, Bob Nichol, Elizabeth Swann, Mi Dai, Philippe Gris, Johanna Pasquet, Dominique Fouchez, Chad Schafer.

Updated SN templates (2010+)
LSST Simulations
Robustness to non-representative training sets
Incorporating contextual information
Evaluating competing classifiers
Update challenge
map to different survey strategies



14 July 2017 (Manhattan, following DESC meeting)
Hlozek, Kessler

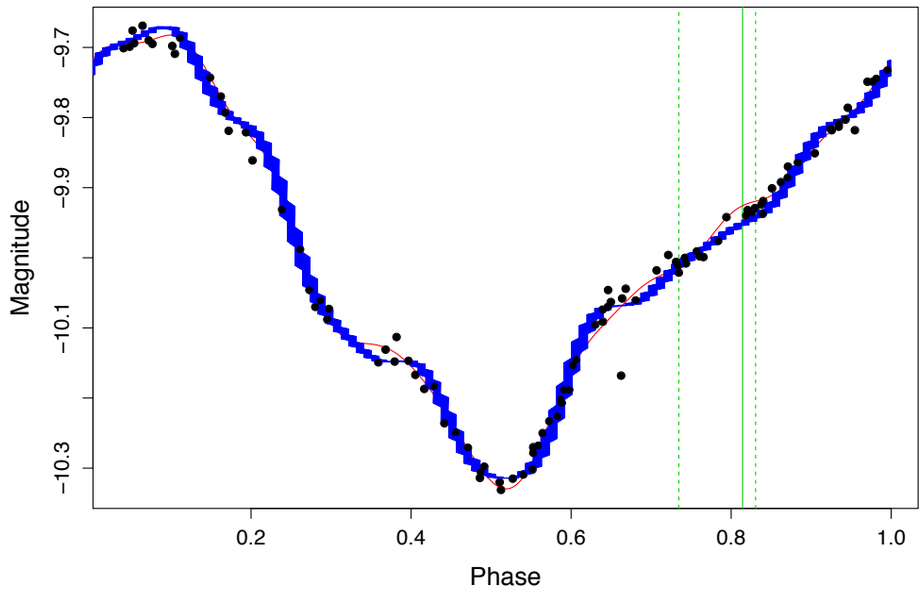
12-14 March 2018 (UK)

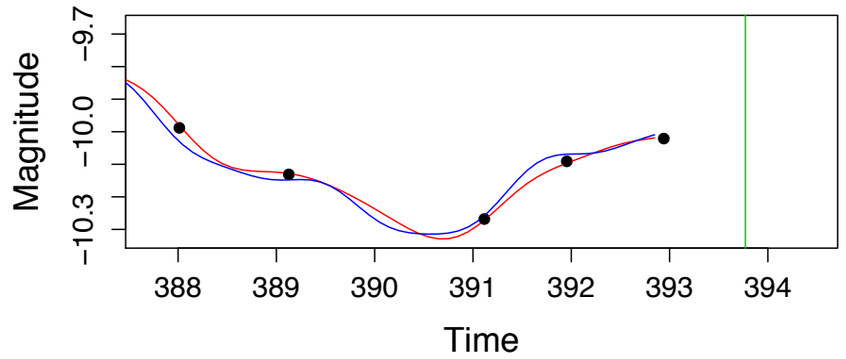
Scheduling observations

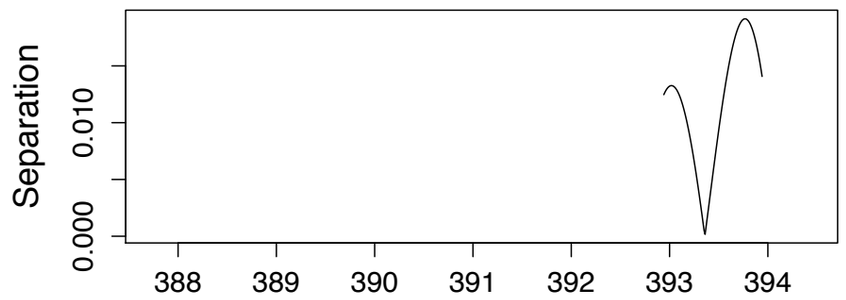
Lead: David Jones

Sujit Ghosh, James Long,
Zhenfeng Lin, Ashish Mahabal
Ana-Maria Staicu, Jogesh Babu

Toy Cepheid example







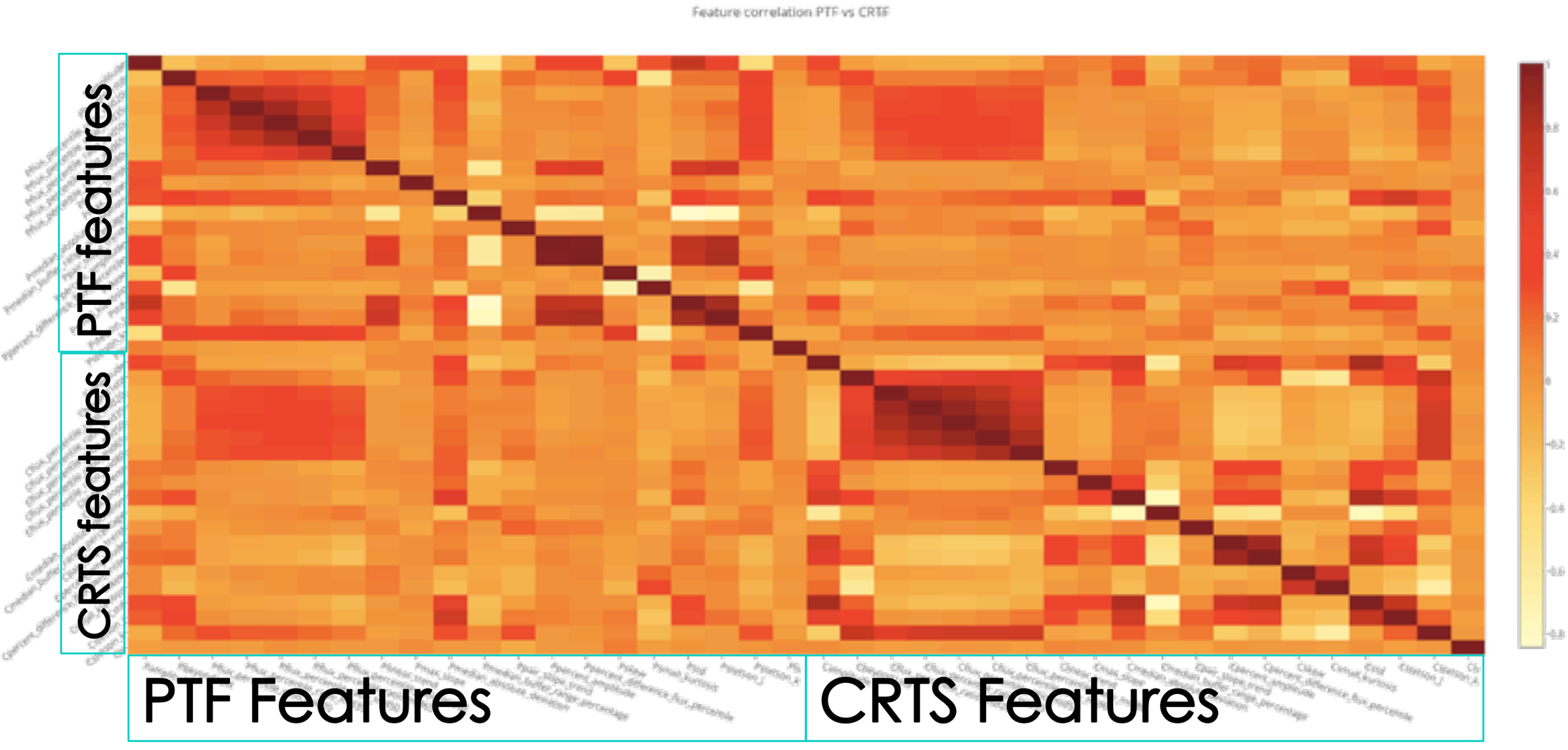
Class / Model 1: basis model with correct period
Class / Model 2: basis model with slightly misspecified period

Left: solid green line shows the optimal (posterior mean) time for a new observation in a one day interval indicated by vertical dashed lines. Red and blue curves show current posterior mean fits for models 1 and 2.

Right: top shows the optimal observation time with the two model means plotted for a single posterior draw of the parameters. Bottom shows the corresponding posterior draw of the separation between the model means

Using basis functions

Domain Adaptation with CRTS/PTF



A Mahabal, J Li, S Vaijanapurkar

~20 features
6 classes
~50K objects

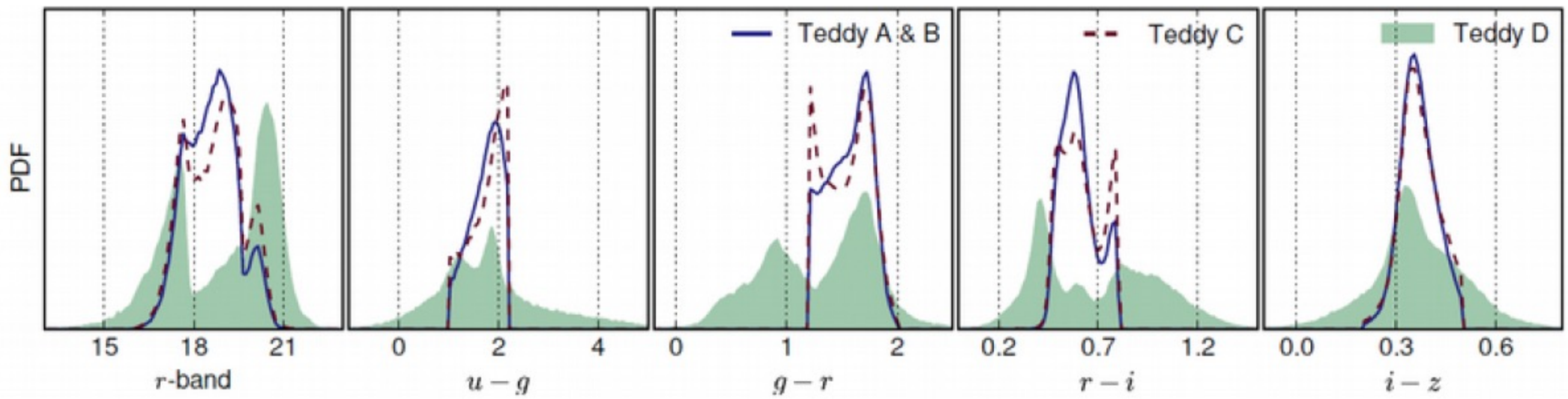
General Goal

Estimate galaxy redshifts exploiting two domains: spectroscopic and photometric data.

Problem: Distributions are different.

Strategy: Use domain adaptation and active learning to learn from both domains.

Ricardo Vilalta
Jogesh Babu, Ashish Mahabal
Ji Meng, E. Ishida, R. de Souza,
E. Feigelson, S. Lahiri, R. Beck,
C. A. Lin



Designer features

Matthew Graham
Ashish Mahabal

- Supernova from just archival information
- R Cor Bor plateaus
- Role of ancillary data (e.g. archival radio source)

Also based on lightcurve decomposition

SN v. non-SN

normalized

based on peaks

$$\left(\frac{1}{t_{span}}\left(\frac{1}{N}\Sigma_i w_i (p_i - p_m)^2\right)^{1/2}\right.$$

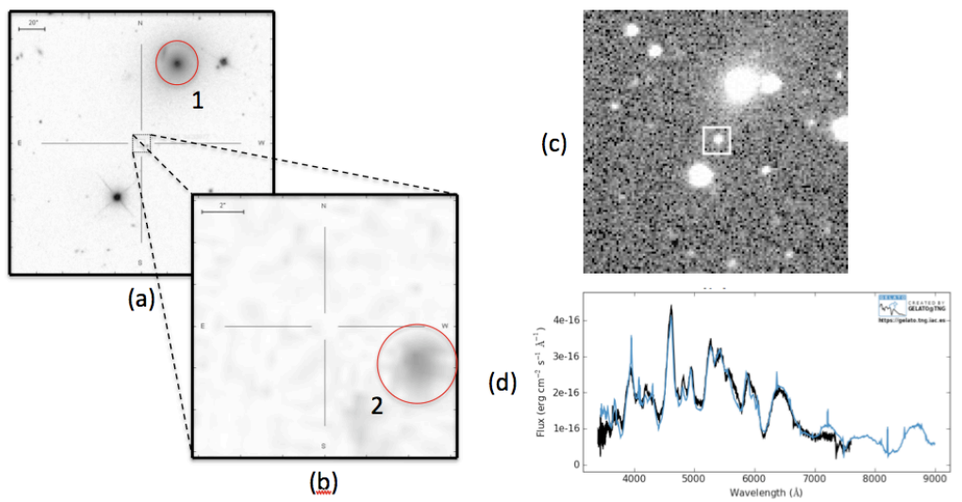
Incorporating ancillary info

Lead: James Lang

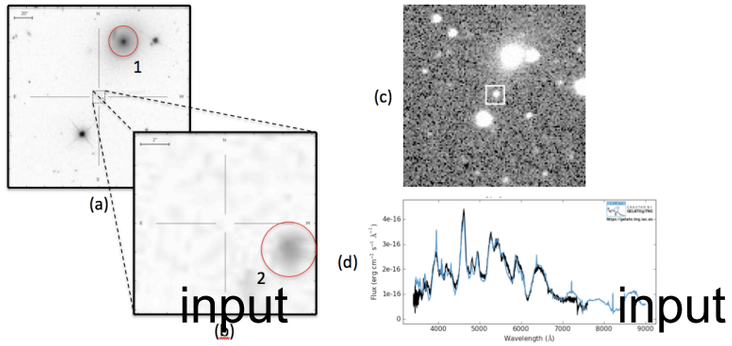
David Jones
Ashish Mahabal
Virisha Timmaraju

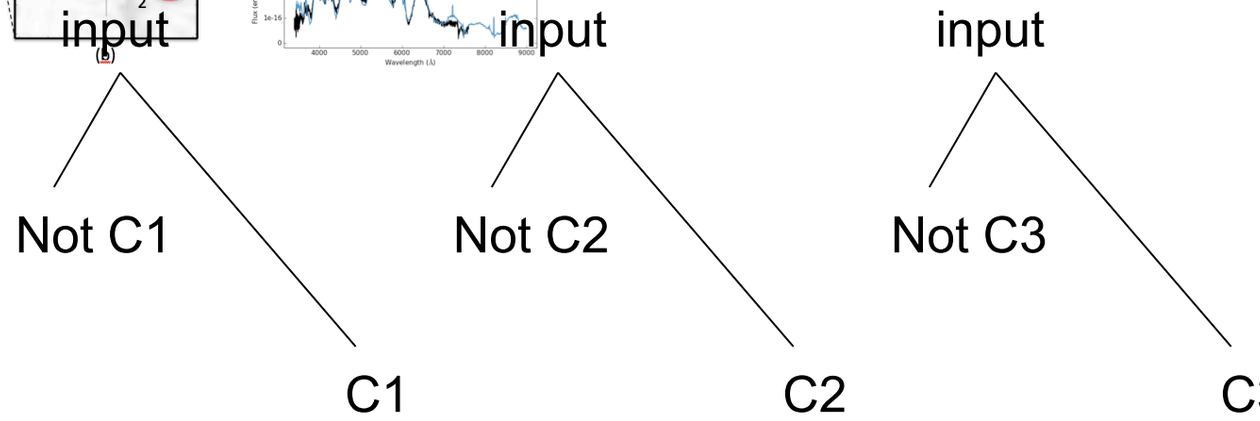
- Parameters like
  - Galactic latitude (Galactic versus extra-galactic)
  - Nearest galaxy (Supernova versus non-)
  - Nearest radio source (blazar or not)

Natural language
Best guesses



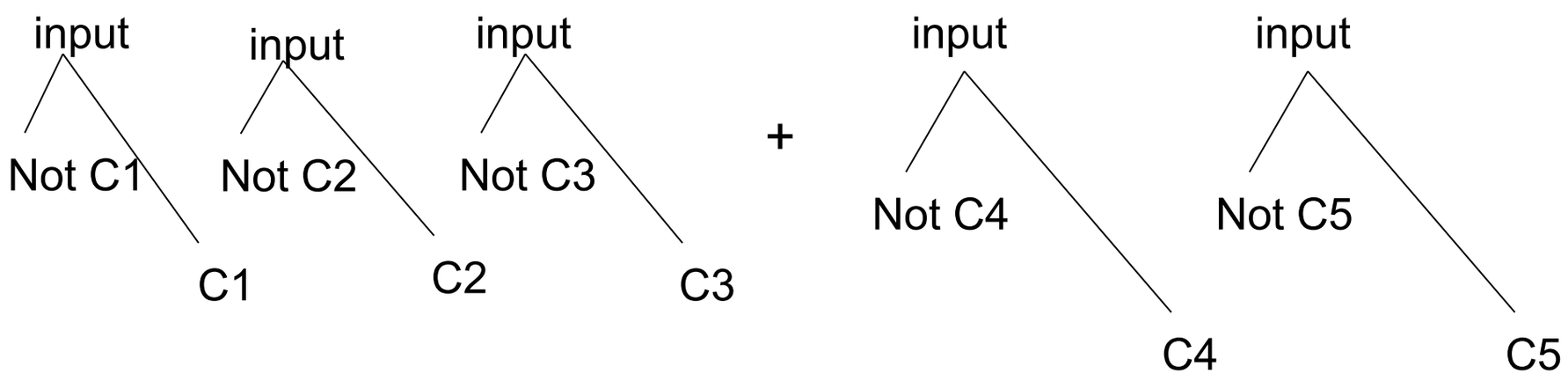
Binary Brokers





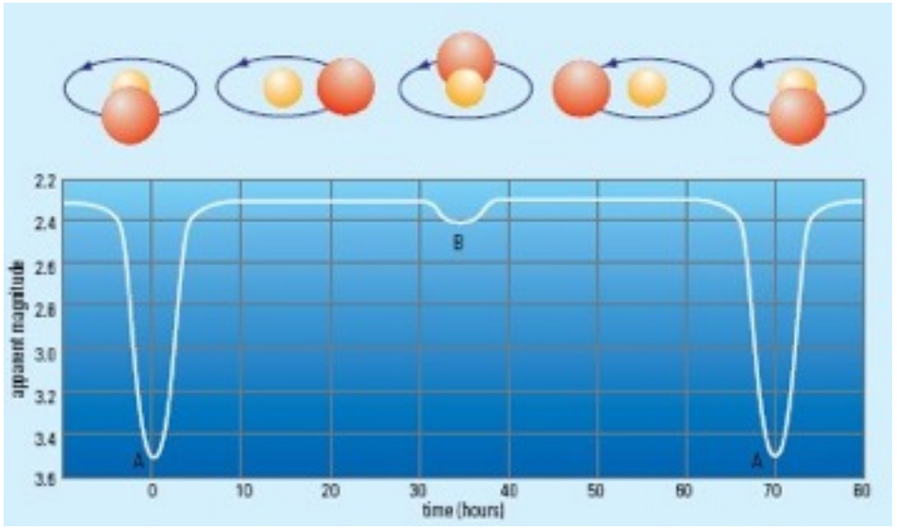
Inputs:
Light-curves
Nearby objects
Archival catalogs
...

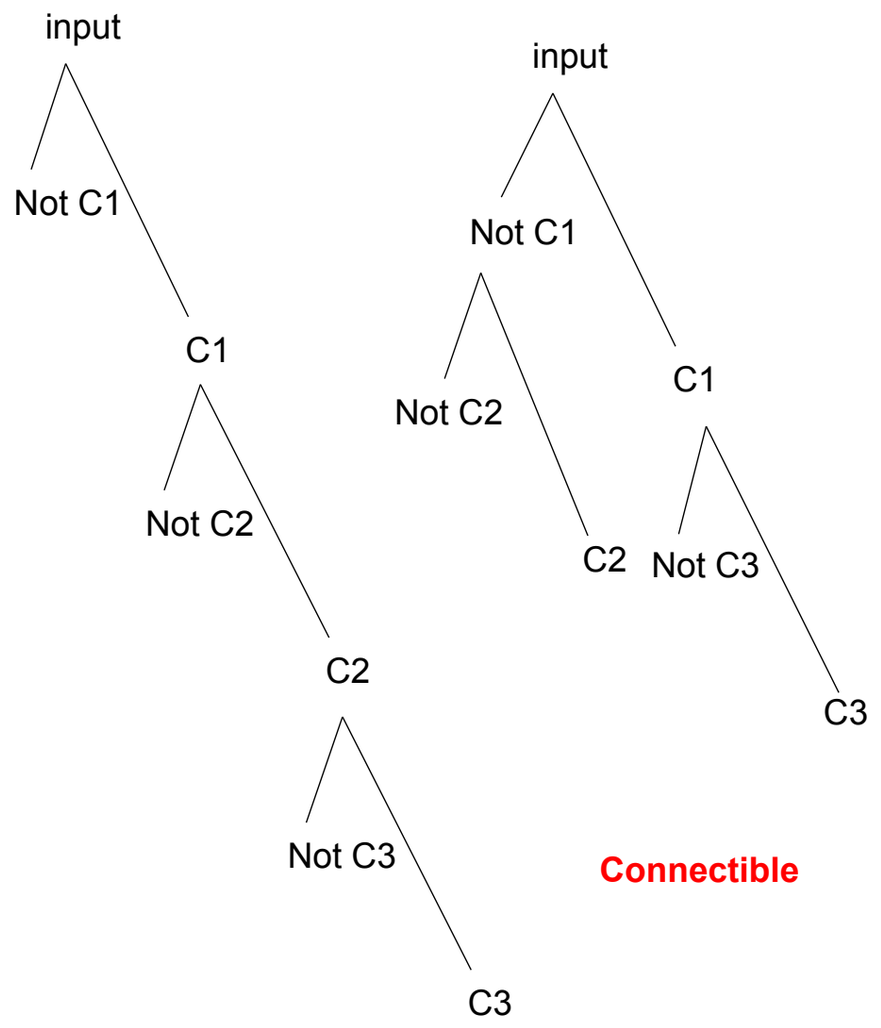
Modular



Extendible

Periodic Binaries

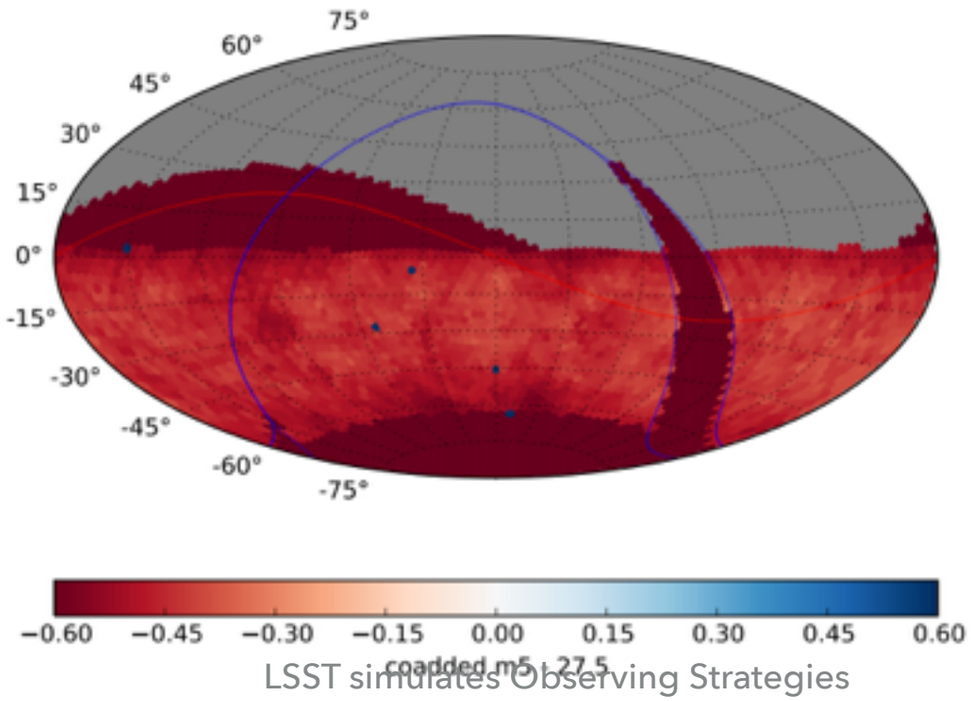




Supernova?
Blazar?
Periodic variable?
Periodic variable of type eclipsing binary?
Eclipsing binary with low metallicity?

OpSim

LSST developed operation simulations (A. Connoly)



LSST simulates Observing Strategies

MAF API

Metric Analysis Framework (Peter Yoachim, Lynne Jones)

Getting Help in MAF

This notebook is a collection of snippets of how to get help on the various bits of the **MAF** ecosystem. It shows some of the **MAF** provided help functions. It also uses the `help` function. The `help` function used below is a Python standard library function. It can be used on any module, class or function. Using `help` should give clarity to the parameters used in associated functions. It will also list functions associated with modules and classes. The notebook also uses the `dir` command which is another Python standard library function. This is useful for getting a list of names from the target object (module/class/function).

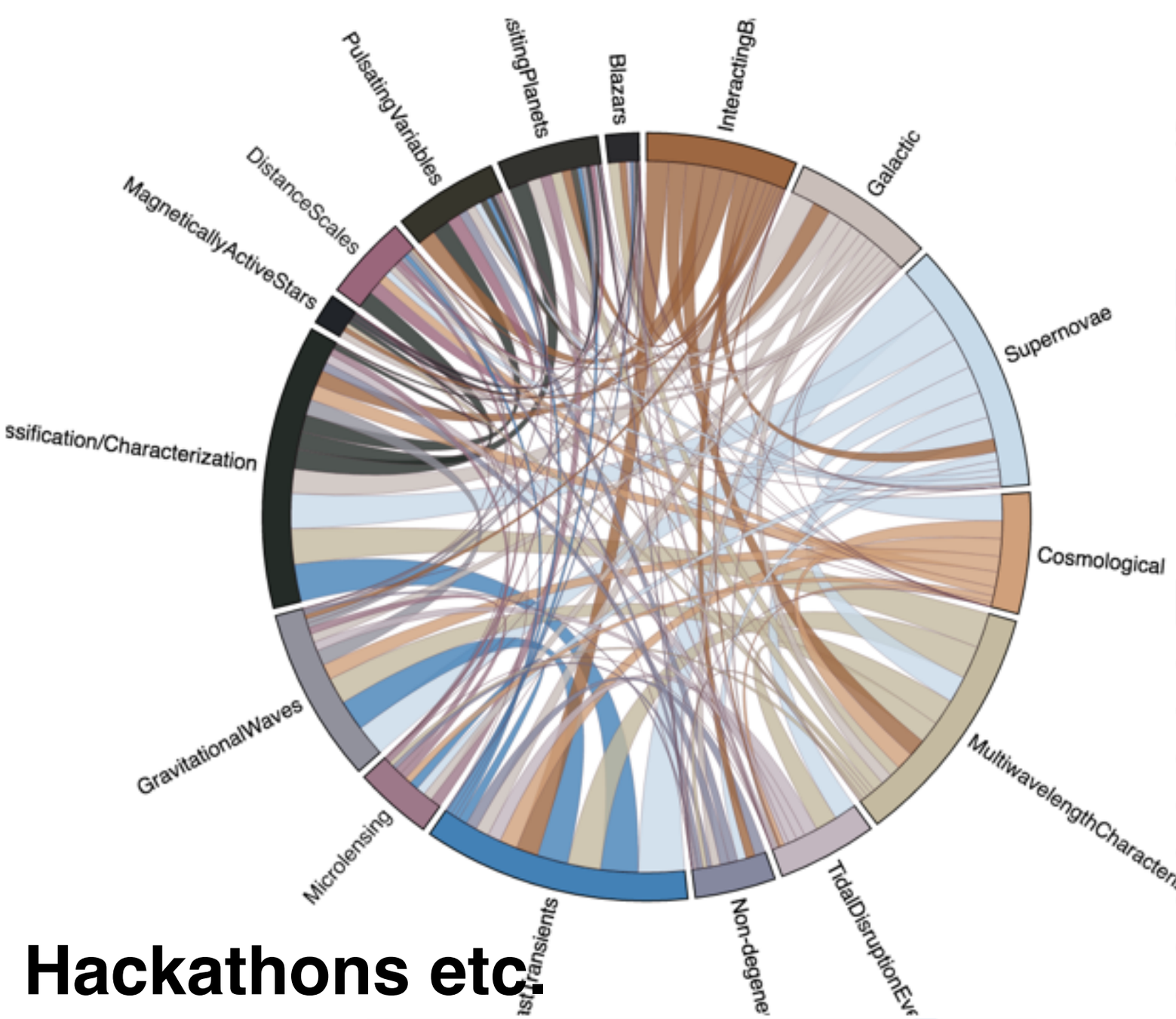
In [1]:
```
# Need to import everything before getting help!
import lsst.sims.maf
import lsst.sims.maf.metrics as metrics
import lsst.sims.maf.slicers as slicers
import lsst.sims.maf.stackers as stackers
import lsst.sims.maf.plots as plots
```

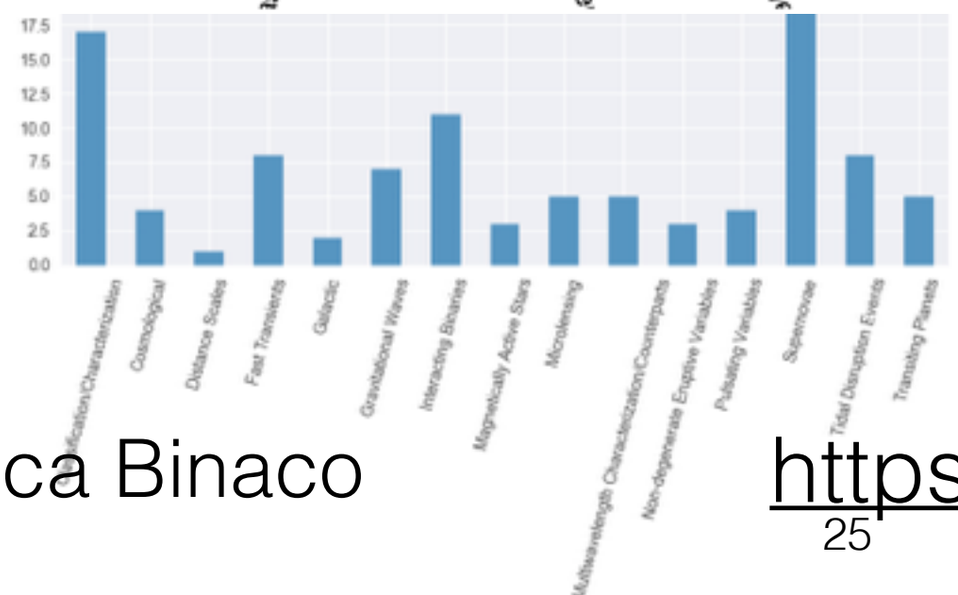
In [2]:
```
# Show the list of metrics with a little bit of documentation
metrics.BaseMetric.list(doc=True)
```

```
----  AveSlewFracMetric  ----
None
----  BinaryMetric  ----
Return 1 if there is data.
----  Coadd5Metric  ----
Calculate the coadded m5 value at this gridpoint.
----  CompletenessMetric  ----
Compute the completeness and joint completeness
```

https://github.com/LSST-nonproject/

Tools for cadence diplomacy





LSST TVS Science Collaboration

me | Projects | Subgroups | Documents | Apply

Co-chairs: Ashish Mahabal, Federica Bianco, Rachel Street

Home » Subgroups

Subgroups

View | Edit | Revisions

- Cosmological
- Classification/Characterization
- Distance Scale
- Fast Transients
- Galactic
- Gravitational Waves
- Interacting Binaries
- Magnetically Active Stars
- Microlensing Subgroup
- Multiwavelength Characterization/Counterparts
- Non-degenerate Eruptive Variables
- Pulsating Variables
- Supernovae Subgroup
- Tidal Disruption Events
- Transiting Planets

Hackathons etc.

Federica Binaco

https://tvs.science.lsst.org/home

Summary

Please join the fun!

- Interconnectedness of different groups
  - Physics of individual types
  - Classification as an over-arching themes
  - Nature of light-curves: filling gaps, decomposing them, features to separate classes, subspaces to match cadences, determining outliers, incorporating ancillary information and determine best times to classify the sources
  - That is the grand (data-)challenge

Ashish Mahabal
aam at astro.caltech.edu

Today
5:30 PM