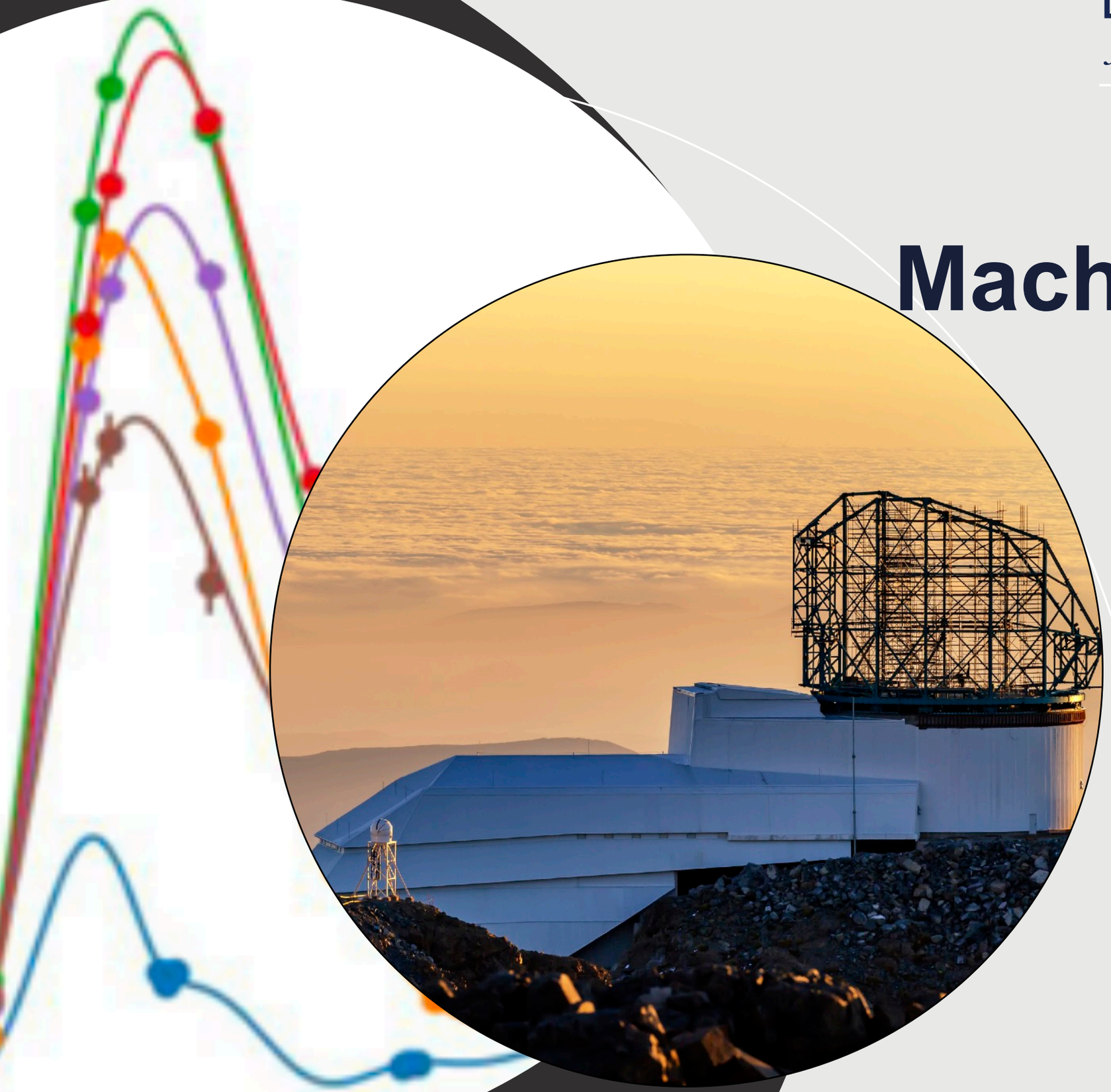
DUNLAP INSTITUTE
for ASTRONOMY & ASTROPHYSICS

Learn from Machine Learning:

PLAsTiCC



Tina Peters
Postdoctoral Fellow

Dunlap Institute for
Astronomy & Astrophysics
UNIVERSITY OF TORONTO

August 2018 - From LSST.org - Credit Gianluca Lombardi

Featured Prediction Competition

PLAsTiCC Astronomical Classification

Can you help make sense of the Universe?

kaggle

$25,000
Prize Money

LSST Project · 1,094 teams · a month ago

Overview | Data | Kernels | Discussion | Leaderboard | Rules | Team | Host | My Submissions | **Late Submission**

Overview — Edit

- **Description**
- **Evaluation**
- **Prizes**
- **Timeline**
- **PLAsTiCC's Team**

+ Add Page

Help some of the world's leading astronomers grasp the deepest properties of the universe.

The human eye has been the arbiter for the classification of astronomical sources in the night sky for hundreds of years. But a new facility -- the Large Synoptic Survey Telescope (LSST) -- is about to revolutionize the field, discovering 10 to 100 times more astronomical sources that vary in the night sky than we've ever known. Some of these sources will be

Slide Credit: Renée Hložek

NSF AURA U.S. DEPARTMENT OF ENERGY SLAC CHARLES AND LISA SIMONYI FUND FOR ARTS AND SCIENCES LSST CORPORATION LSST TVS Science Collaboration LSST DESC Dark Energy Science Collaboration

Photometric **L**SST **As**tronomical **Ti**me-series **C**lassification **C**hallenge

The PLAsTiCC Team

Tarek Allam Jr., Anita Bahmanyar, Rahul Biswas, Alexandre Boucaud, Mi Dai,
Lluís Galbany, Emille E. O. Ishida, Saurabh W. Jha, David O. Jones, Richard Kessler,
Michelle Lochner, Ashish A. Mahabal, Alex I. Malz,
Kaisey S. Mandel, Juan Rafael Martínez-Galarza, Jason D. McEwen,
Daniel Muthukrishna, Gautham Narayan, Hiranya Peiris, Christina M. Peters,
Kara Ponder, Christian N. Setzer,
The LSST Dark Energy Science Collaboration,
The LSST Transients, Variable Stars Science Collaboration

More than 1 million new SEDs across several new models (from the community):

15 **classes** in final test set, one class not represented in training test

~ 3.5 million **objects** in test set w/ < 8000 objects for training

~ 450 million **observations** (LSST WFD + DDF) in 6 bands ~ 18.5 GB of data

Slide Credit: Renée Hložek

Model Flux

SNIa (SALT2)

SNIa-91bg

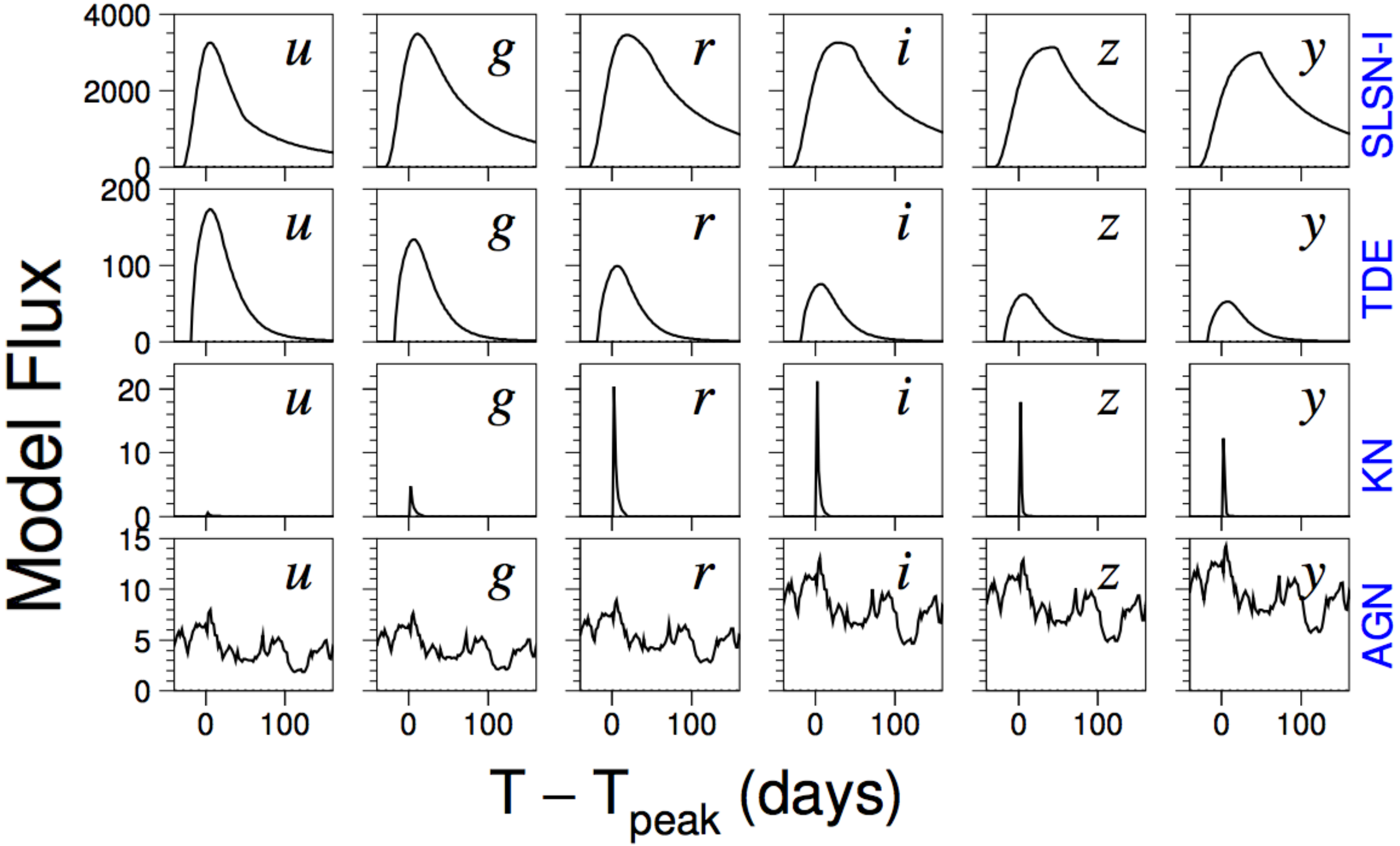
SNIax

SNII (NMF)

SNIbc (MOSFIT)

$T - T_{peak}$ (days)

Kessler et al. (2019)



Kessler et al. (2019)

time (days)

Model Flux

u g r i z y

RRL

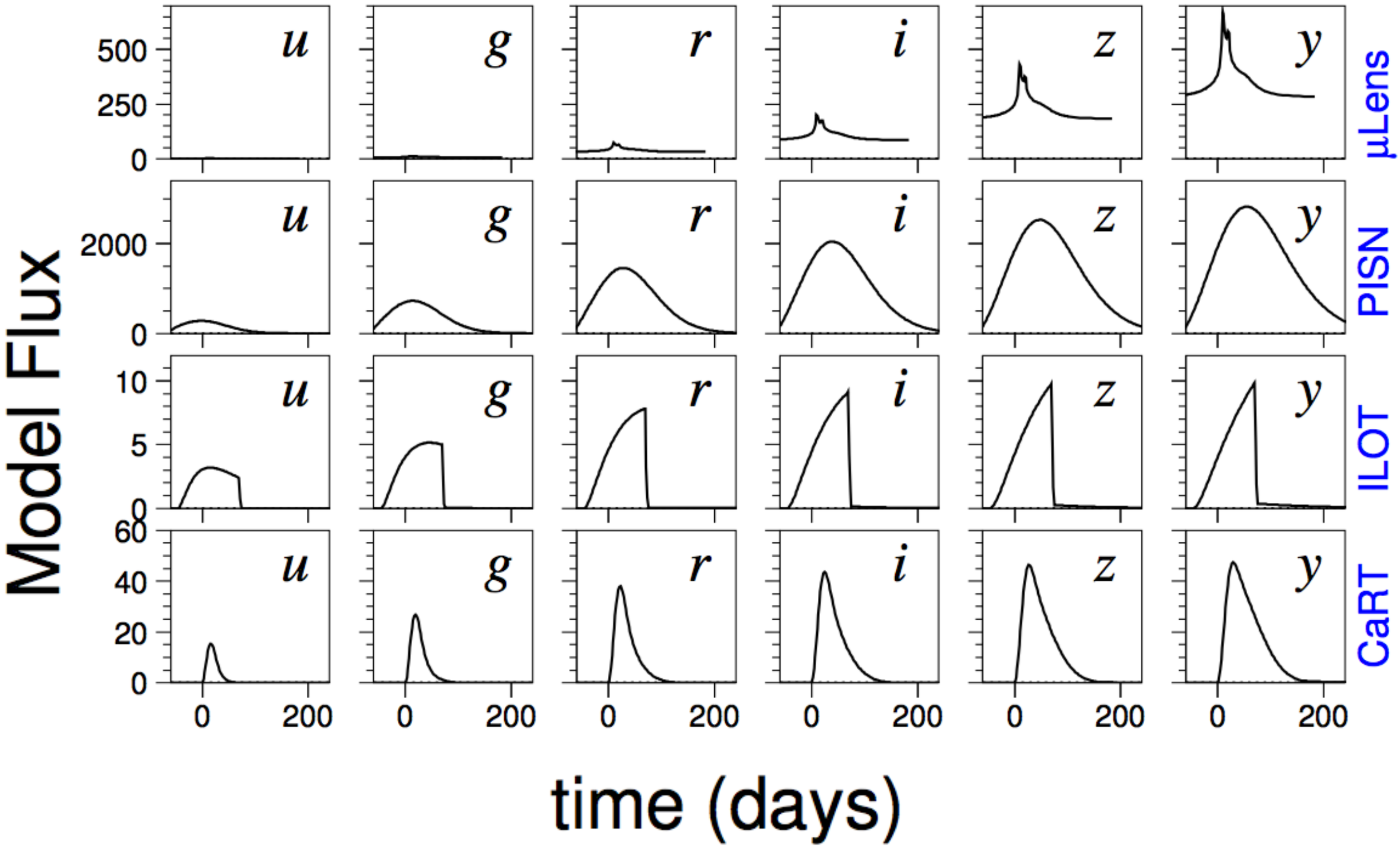
M-dwarf

EB

MIRA

µLens Single

Kessler et al. (2019)



Kessler et al. (2019)

Summary of Models used in PLAsTiCC

(Published in PASP)

model class num[a]: name	model description	contributor(s)[b]	Nevent Gen[c]	Nevent train[d]	Nevent test[e]	redshift range[f]
90: SNIa	WD detonation, Type Ia SN	RK	16,353,270	2,313	1,659,831	< 1.6
67: SNIa-91bg	Peculiar type Ia: 91bg	SG,LG	1,329,510	208	40,193	< 0.9
52: SNIax	Peculiar SNIax	SJ,MD	8,660,920	183	63,664	< 1.3
42: SNII	Core Collapse, Type II SN	SG,LG:RK,JRP:VAV	59,198,660	1,193	1,000,150	< 2.0
62: SNIbc	Core Collapse, Type Ibc SN	VAV:RK,JRP	22,599,840	484	175,094	< 1.3
95: SLSN-I	Super-Lum. SN (magnetar)	VAV	90,640	175	35,782	< 3.4
15: TDE	Tidal Disruption Event	VAV	58,550	495	13,555	< 2.6
64: KN	Kilonova (NS-NS merger)	DK,GN	43,150	100	131	< 0.3
88: AGN	Active Galactic Nuclei	SD	175,500	370	101,424	< 3.4
92: RRL	RR lyrae	SD	200,200	239	197,155	0
65: M-dwarf	M-dwarf stellar flare	SD	800,800	981	93,494	0
16: EB	Eclipsing Binary stars	AP	220,200	924	96,572	0
53: Mira	Pulsating variable stars	RH	1,490	30	1,453	0
6: μLens-Single	μ-lens from single lens	RD,AA:EB,GN	2,820	151	1,303	0
991: μLens-Binary	μ-lens from binary lens	RD,AA	1,010	0	533	0
992: ILOT	Intermed. Lum. Optical Trans.	VAV	4,521,970	0	1,702	< 0.4
993: CaRT	Calcium Rich Transient	VAV	2,834,500	0	9,680	< 0.9
994: PISN	Pair Instability SN	VAV	5,650	0	1,172	< 1.9
995: μLens-String	μ-lens from cosmic strings	DC	30,020	0	0	0
TOTAL	Sum of all models		117,128,700	7,846	3,492,888	—

[a]num>990 were all in unknown class 99 during the competition. An extra digit is added here to distinguish each model.
[b]Co-author initials. Colon separates independent methods.
[c]Number of generated events, corresponding to the true population without observational selection bias.
[d]Labeled subset from spectroscopic classification. 0 → predicted from theory, not convincingly observed, or very few observations.
[e]Unlabeled sample. `PLAsTiCC` goal is to label this sample.
[f]Redshift> 0 for extragalactic models; Redshift= 0 for Galactic models.

Model Contributors:

AA: Arturo Avelino (Harvard U.)
EB: Etienne Bachelet (LCO)
DC: David Chernoff (Cornell U.)
MD: Mi Dai (Rutgers U.)
SD: Scott Daniel (U.Washington)
RD: Rosanne Di Stefano (Harvard U.)
LG: Lluís Galbany (U.Pitt)
SG: Santiago González-Gaitán (U.Lisbon)
RH: Renée Hlozek (U.Toronto)
SJ: Saurabh Jha (Rutgers U.)
DK: Dan Kasen (U.C. Berkeley)
RK: Rick Kessler (U.Chicago)
GN: Gautham Narayan (STScI)
JRP: Justin Pierel (U. South Carolina)
AP: Andrej Prsa (Villanova U.)
VAV: Ashley Villar (Harvard U.)

Unblinded Data Files: http://doi.org/10.5281/zenodo.2539456

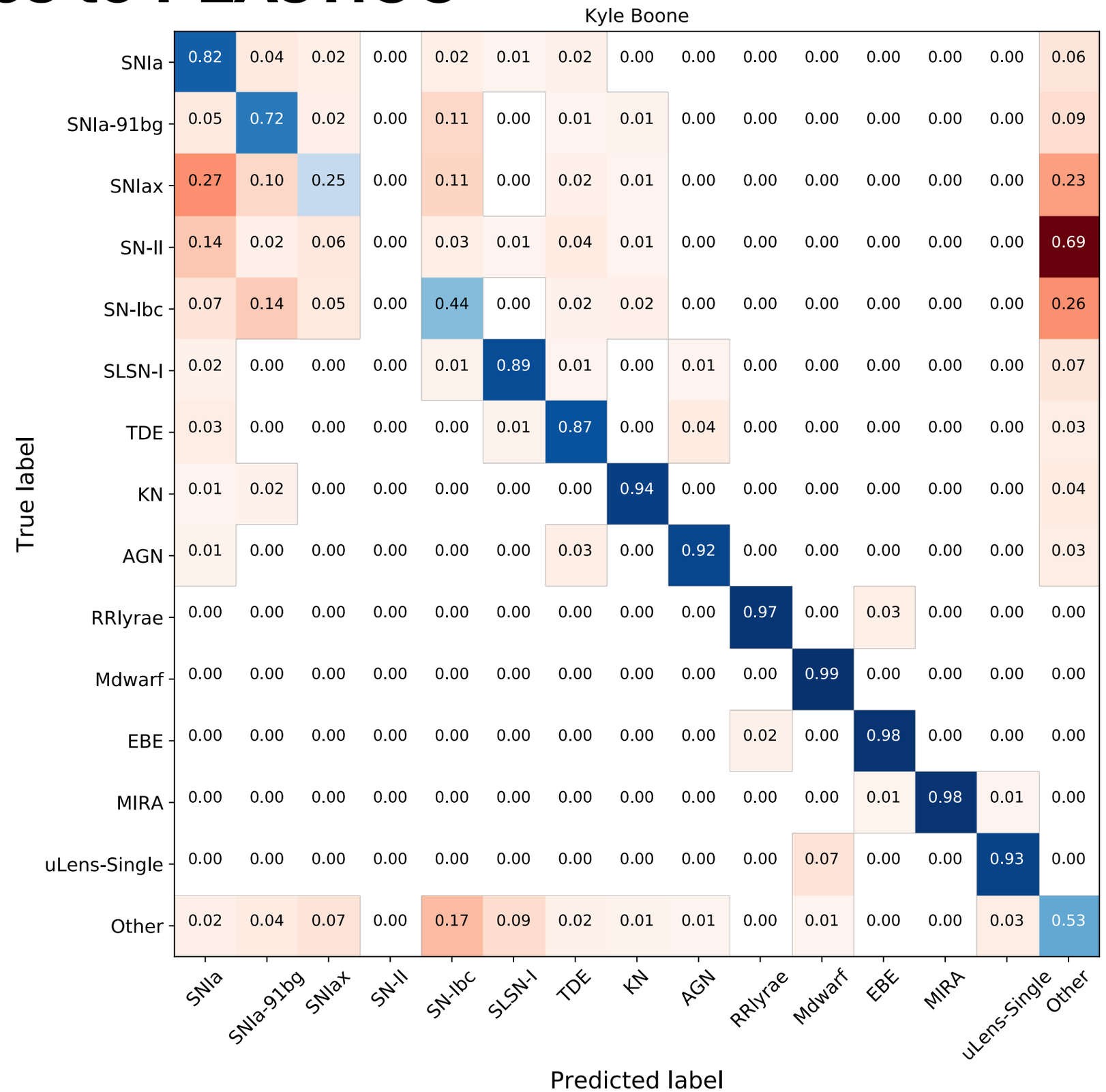
Simulation Source code: http://snana.uchicago.edu
(model libraries will be released with article)

Slide credit: Rick Kessler

Evaluating Entries to PLAsTiCC

PLAsTiCC was designed for more than one science case, but Kaggle requires a single winner.

Need to define a metric that requires a good performance overall.

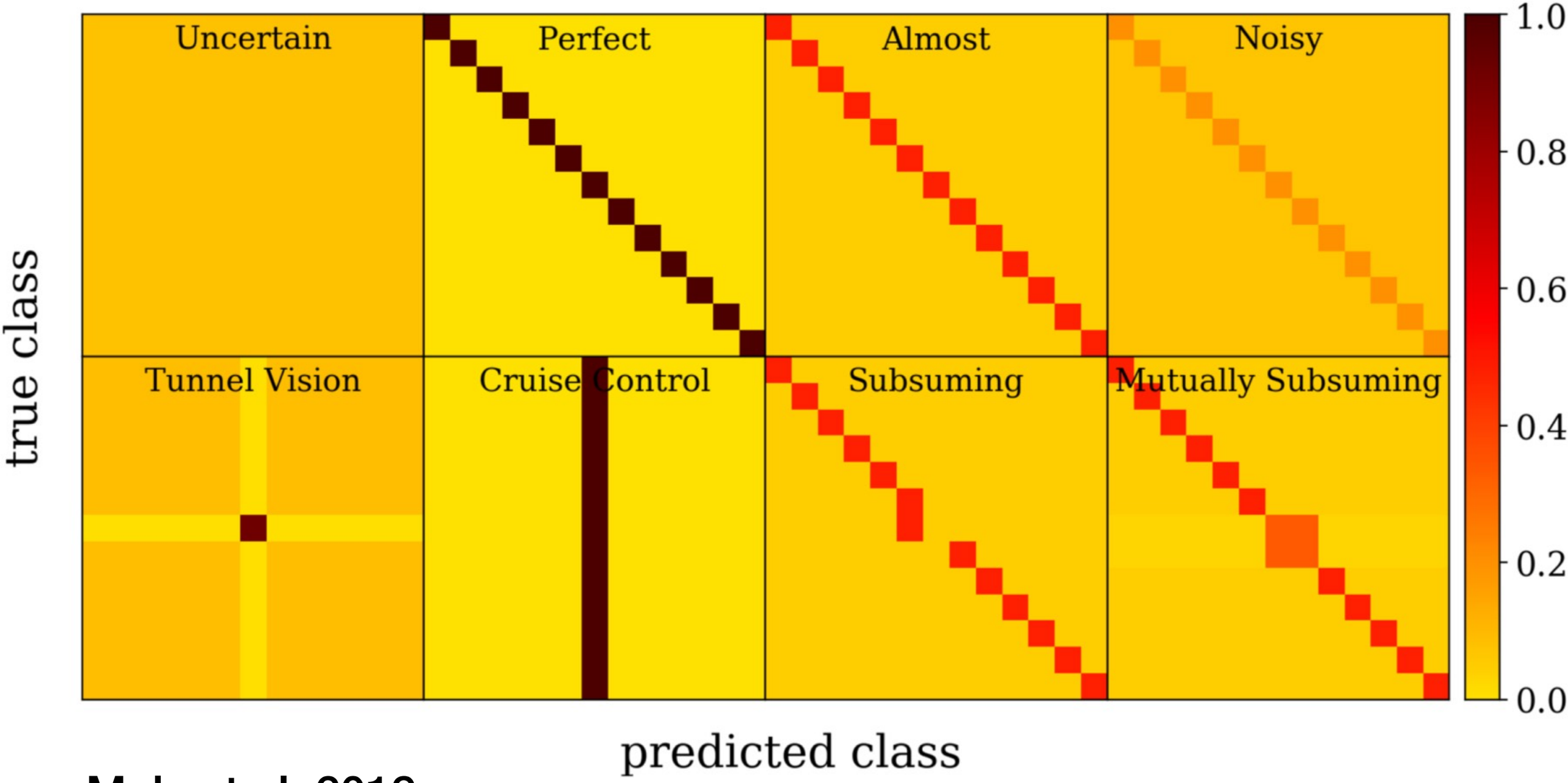


Hložek et al. *in prep*

DUNLAP INSTITUTE
for ASTRONOMY & ASTROPHYSICS

Dunlap Institute for
Astronomy & Astrophysics
UNIVERSITY OF TORONTO

Evaluating Entries to PLAsTiCC



Malz et al. 2019

The leaderboard

Public Leaderboard | Private Leaderboard

This leaderboard is calculated with approximately 33% of the test data.

The final results will be based on the other 67%, so the final standings may be different.

Raw Data | Refresh

In the money | Gold | Silver | Bronze

#	△1w	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲1	**Kyle Boone**			0.67056	104	16d
2	▲2	**Major Tom**			0.68042	366	16d
3	▲2	**AhmetErdem**			0.69132	233	16d
4	▼3	**Mike & Silogram**			0.69365	176	16d
5	▼2	**SKZ Lost in Translation**			0.73968	343	16d
6	▲2	**Three Musketeers**			0.79207	313	16d
7	▲14	**rapids.ai**			0.79222	133	16d
8	▲5	**Stefan Stefanov**			0.79334	28	16d

Slide Credit: Renée Hložek

The leaderboard

Public Leaderboard | Private Leaderboard

The private leaderboard is calculated with approximately 67% of the test data.

This competition has completed. This leaderboard reflects the final standings.

Refresh

In the money | Gold | Silver | Bronze

#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	—	**Kyle Boone**			0.68503	104	16d
2	▲ 2	**Mike & Silogram**			0.69933	176	16d
3	▼ 1	**Major Tom**			0.70016	366	16d
4	▼ 1	**AhmetErdem**			0.70423	233	16d
5	—	**SKZ Lost in Translation**			0.75229	343	16d
6	▲ 2	**Stefan Stefanov**			0.80173	28	16d
7	▲ 3	**hklee**			0.80836	63	21d
8	▼ 1	**rapids.ai**			0.80905	133	16d
		eteers			0.81312	313	16d

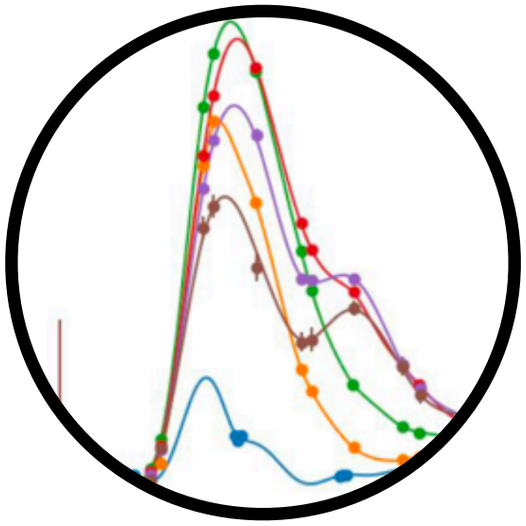
Slide Credit: Renée Hložek

Photometric LSST Astronomical Time Series Classification

DUNLAP INSTITUTE
for ASTRONOMY & ASTROPHYSICS

Dunlap Institute for
Astronomy & Astrophysics
UNIVERSITY OF TORONTO

Photometric LSST Astronomical Time Series Classification

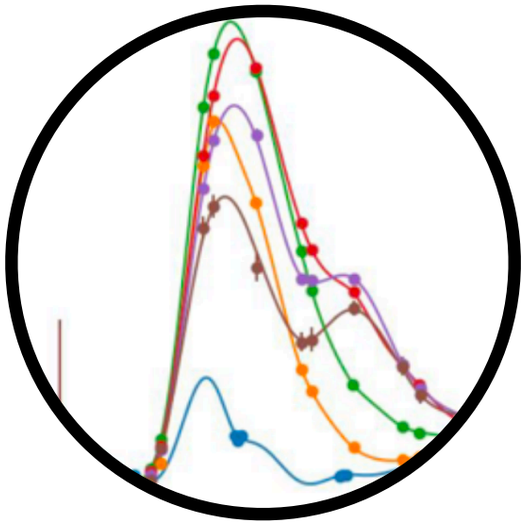


kaggle

DUNLAP INSTITUTE
for ASTRONOMY & ASTROPHYSICS

Dunlap Institute for
Astronomy & Astrophysics
UNIVERSITY OF TORONTO

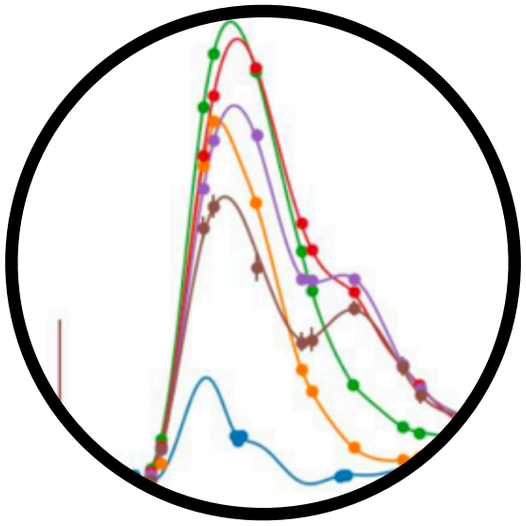
Photometric LSST Astronomical Time Series Classification



DUNLAP INSTITUTE
for ASTRONOMY & ASTROPHYSICS

Dunlap Institute for
Astronomy & Astrophysics
UNIVERSITY OF TORONTO

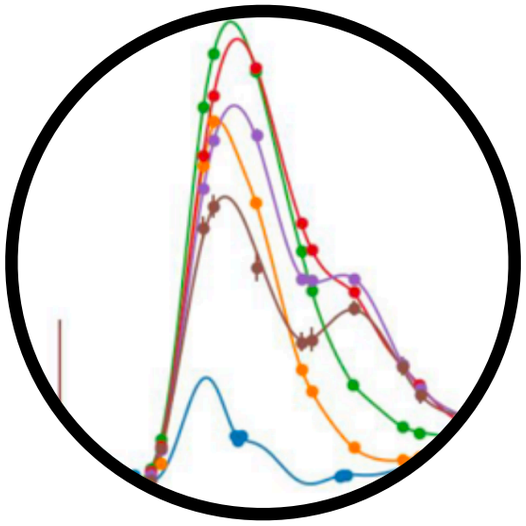
Photometric LSST Astronomical Time Series Classification



DUNLAP INSTITUTE
for ASTRONOMY & ASTROPHYSICS

Dunlap Institute for
Astronomy & Astrophysics
UNIVERSITY OF TORONTO

Photometric LSST Astronomical Time Series Classification



kaggle

Challenge Results and Implications Thereof, in prep

DUNLAP INSTITUTE
for ASTRONOMY & ASTROPHYSICS

Dunlap Institute for
Astronomy & Astrophysics
UNIVERSITY OF TORONTO