GOPREAUX

Open-source Code and Data to Model Multi-wavelength Emission of Extragalactic Transients using Gaussian Processes

Tyler Pritchard
University of Maryland / NASA Goddard Space Flight Center

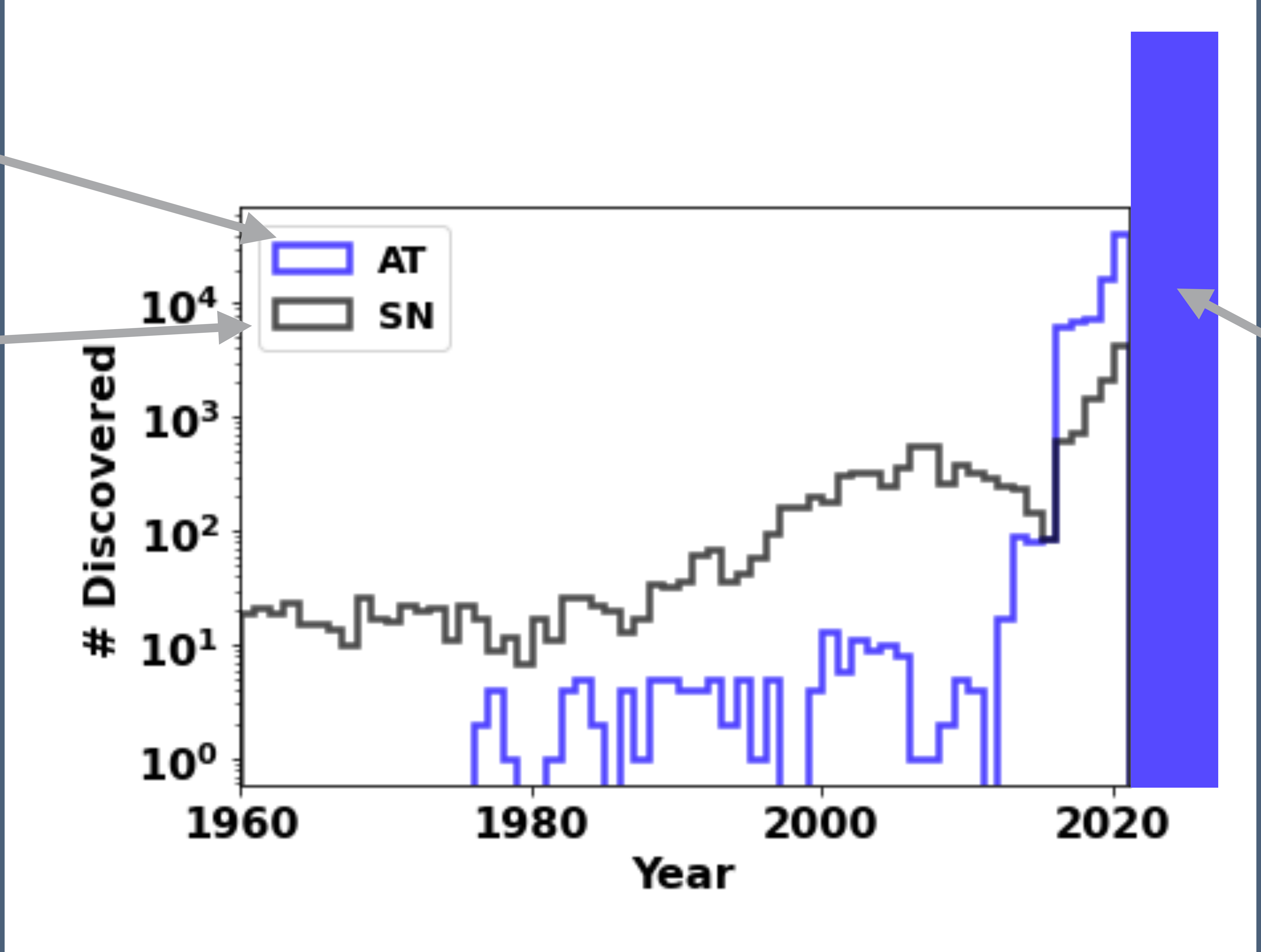
Craig Pellegrino (NASA GSFC), Maryam Modjaz (UVA), Somayeh Khakpash (Lehigh)Adrian Crawford (UVA), Federica Bianco (UDel)

Motivation

The number of transient events discovered is *exploding*

Spectroscopically Classified

Photometric detection



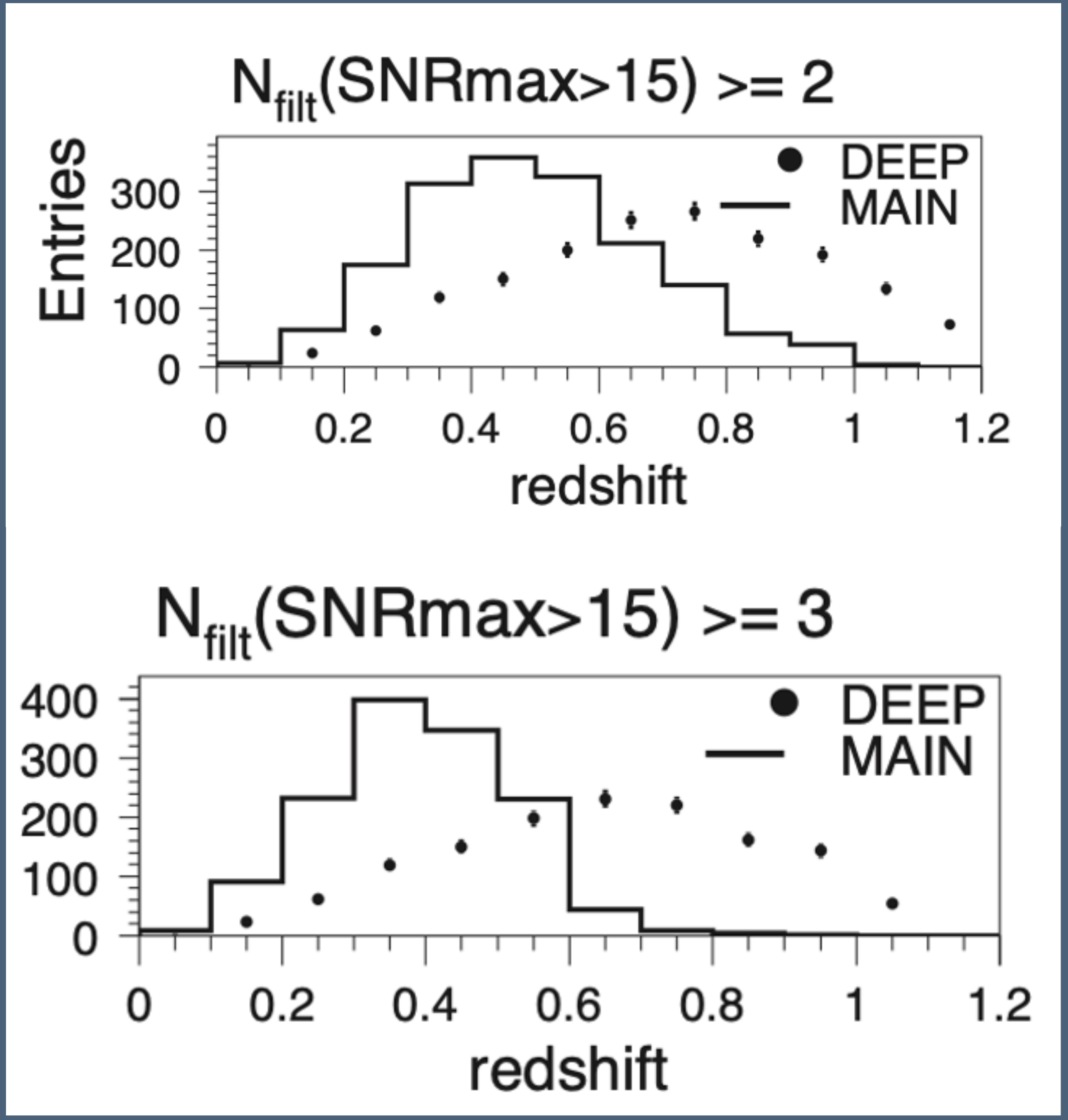
Rubin and Roman are going to blow the axes here

Motivation

... and many of these discoveries are happening at increasingly high redshifts

Today, most supernovae are discovered in the local universe - z~ 0.02 or less

Rubin & Roman Observatories will be push *routine* supernovae discoveries out to z~1+



Vera C. Rubin LSST Science Book

Motivation

… and that poses a challenge to our detection and interpretation of events

We'll be sampling significantly different wavelengths due to redshift (uv-rest frame -> opt/IR observed)

We won't have spectra for most* events

Our observing cadence may be longer as follow-up resources are more limited for objects at fainter magnitudes with fewer filter repeats

How do we detect, classify, and compare these future events against our local understanding?

gopreaux

Archive, tools and models for Rubin, Roman and beyond

The Archive

A multi-wavelength, UV-focussed dataset that scrapes public sources for as much complementary data as possible about specific events

The Tools

Combine multi-wavelength photometry from heterogeneous sources to model an object as a low-resolution SED

Create a data-driven model across wavelength, time using information from the object and others deemed similar

The Models

3D spectral energy "surfaces" to predict light curves, photometry, and infer physical properties

... Can be used to create data-driven training sets, apples-to-apples comparisons of light curves, sample property comparisons, and more!

The Archive

Selection Criteria

- Publicly reported transients observed by *Swift* Through August 2023 with at least 3 detections in two filters
- Objects with publicly Available NIR Data through August 2023

- Any publicly available data obtained the above objects
- ZTF, ATLAS, ASAS-SN, ATLAS, Open SNe Catalog

The Sample

The Archive

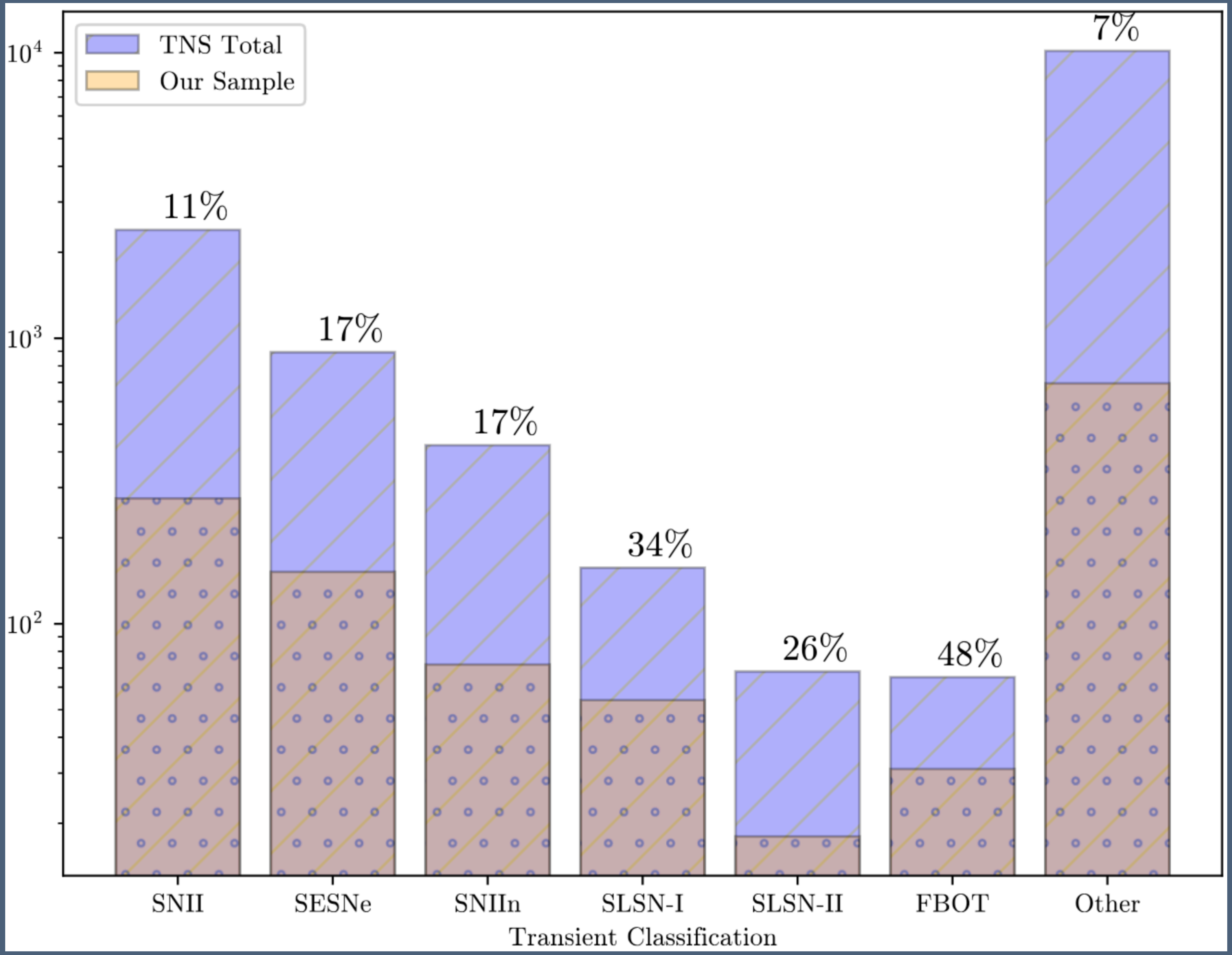
Type	# Objects	#Detections	#Nondetections
SN II	275	44,585	13,781
SESNe	172	16,817	5049
SN II	72	10,044	2997
SLSN-I	54	10,380	3443
SLSN-II	18	3322	760
FBOT	31	2625	3297
Other*	695	58,301	17,016

* combination of Type Ia SNe, unclassified, and/or non-SNe, and/or non-transient objects

The Sample II

Compared to TNS

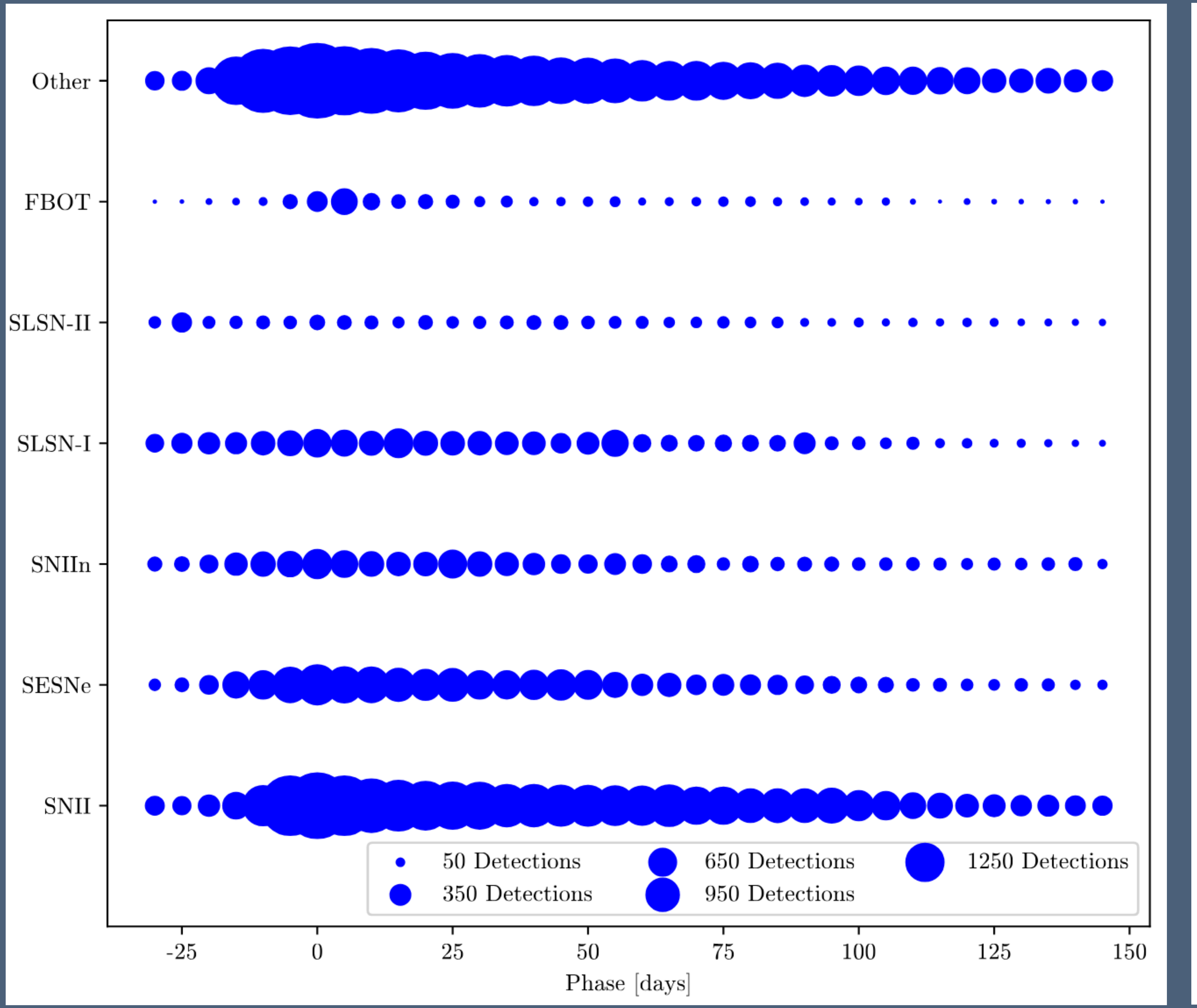
1300+ objects with more than 146,000 data points - one of the largest uniform samples to date!

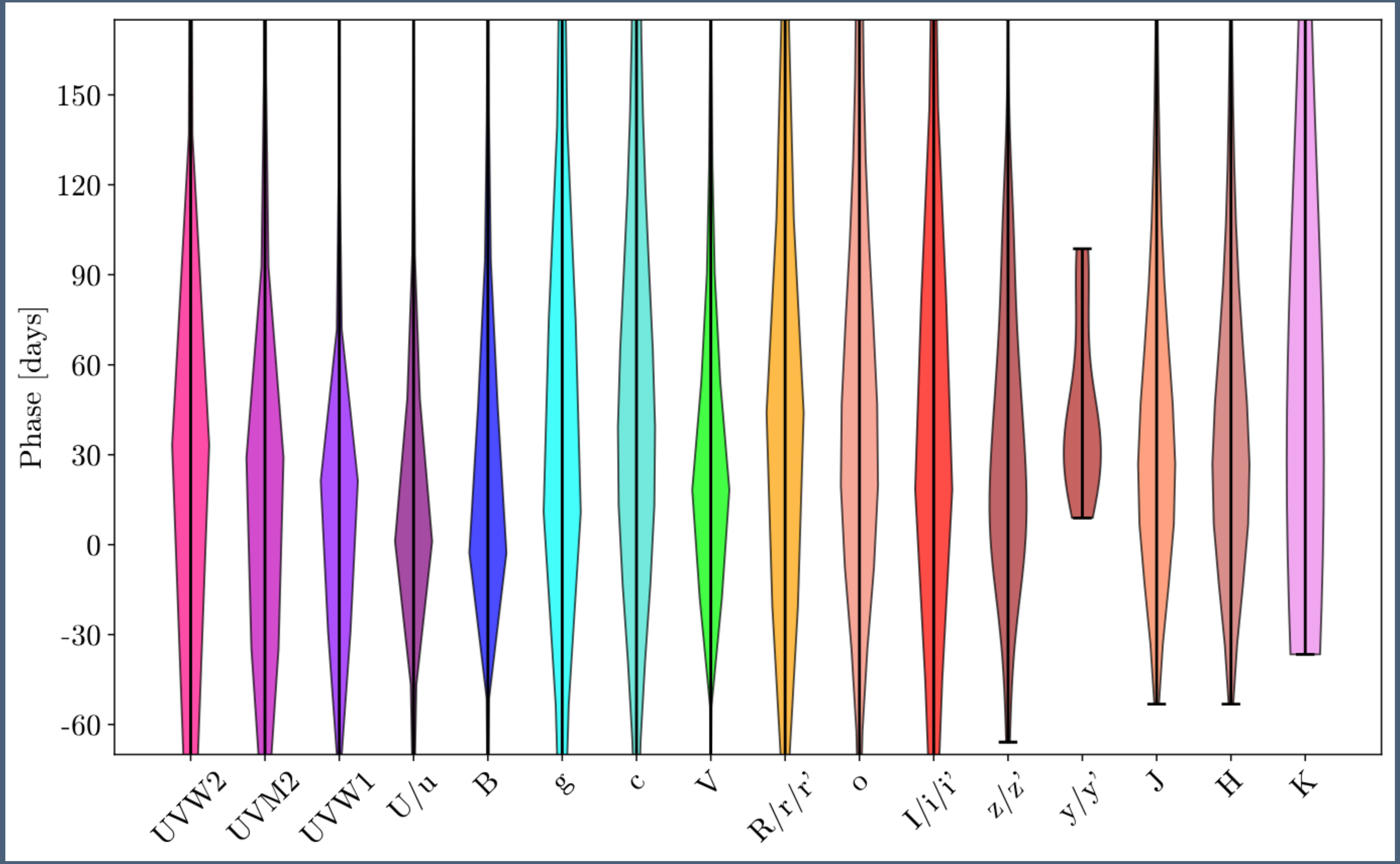


We focus on core collapse supernovae in the paper, but we do have lots of Ia's in the archive. They are under "other" here.

The Sample III

A collection of fairly well observed objects





Pellegrino et al. 2026

The Archive

All publicly available now on GitHub

SN2021gno	SESNe	CaST	0.006211	183.042875	13.249181	59305.56318551056	17.49515786162694	g
SN2022hgk	SESNe	SNIb	0.033	212.598667	44.233681	59719.49550269742	18.5037680925394	g
SN2020hvp	SESNe	SNIb	0.005247	245.439125	-2.289269	58976.11913072734	15.616503994803644	g
SN2006dn	SESNe	SNIb	0.053	127.2227503	1.1107514			
SN2017bgu	SESNe	SNIb	0.008	253.997852235	42.5600231429	57813.685925949496	15.765782804233334	V
SN2016bau	SESNe	SNIb	0.003856	170.245917	53.173778	57483.97462290909	15.927507846387066	V

- SN2020hvp_asassn_raw.dat
- SN2020hvp_atlas_fp.json
- SN2020hvp_atlas_fp_raw.dat
- SN2020hvp_datacube_mangled.csv
- SN2020hvp_opensn.json
- SN2020hvp_uvotB15.1.dat
- SN2020hvp_ztf_fp.json
- batchfp_req0001527637_lc.txt

Table outlining what objects we have data on

Photometry Files from the relevant sources are included without additional processing

Uniformly processed *Swift* UV Data

The Tools

Goals

A suite of data-driven tools to:

Describe a single object, or a class of objects

To describe an object if it was observed a little differently

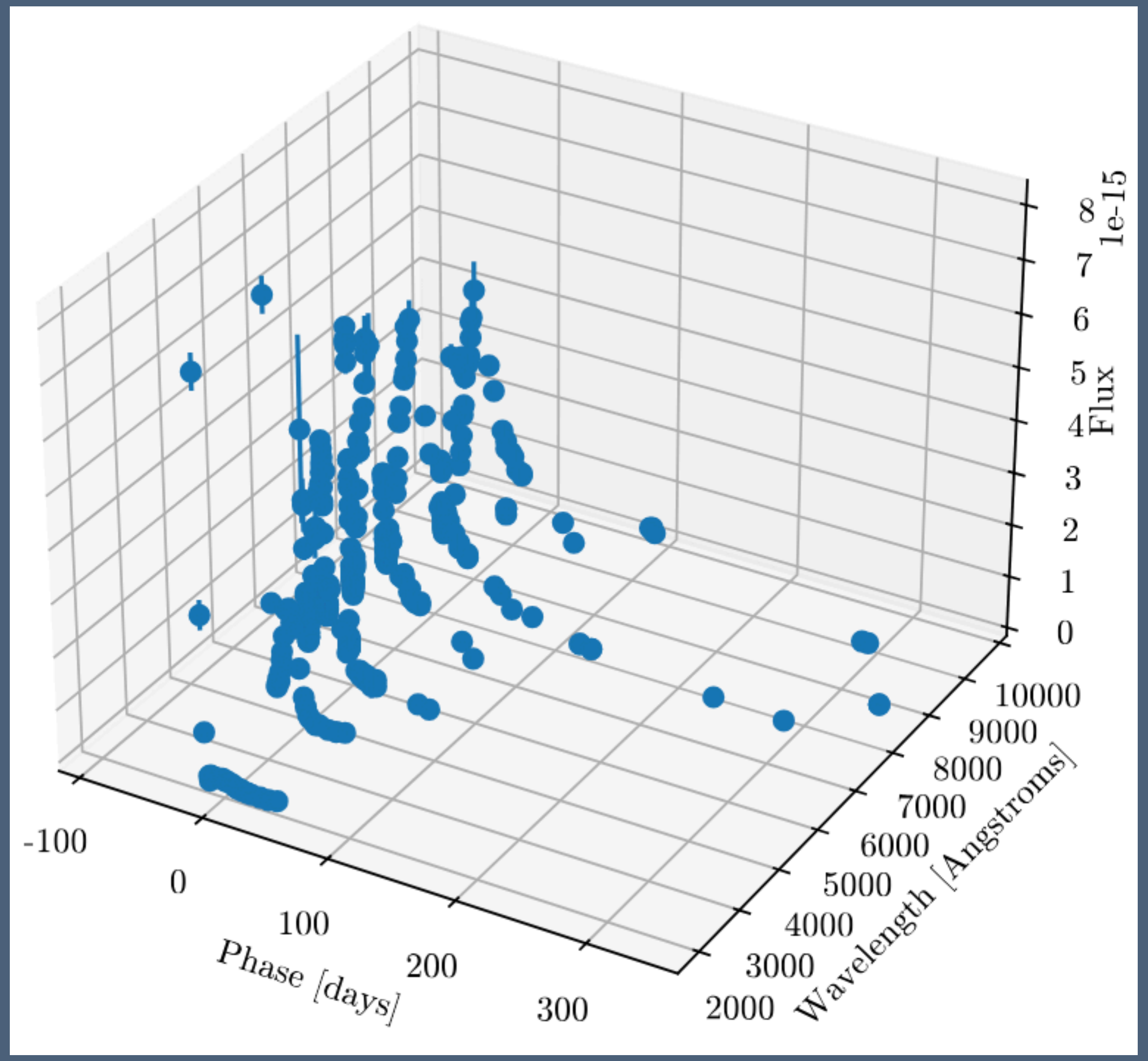
To help answer questions like:

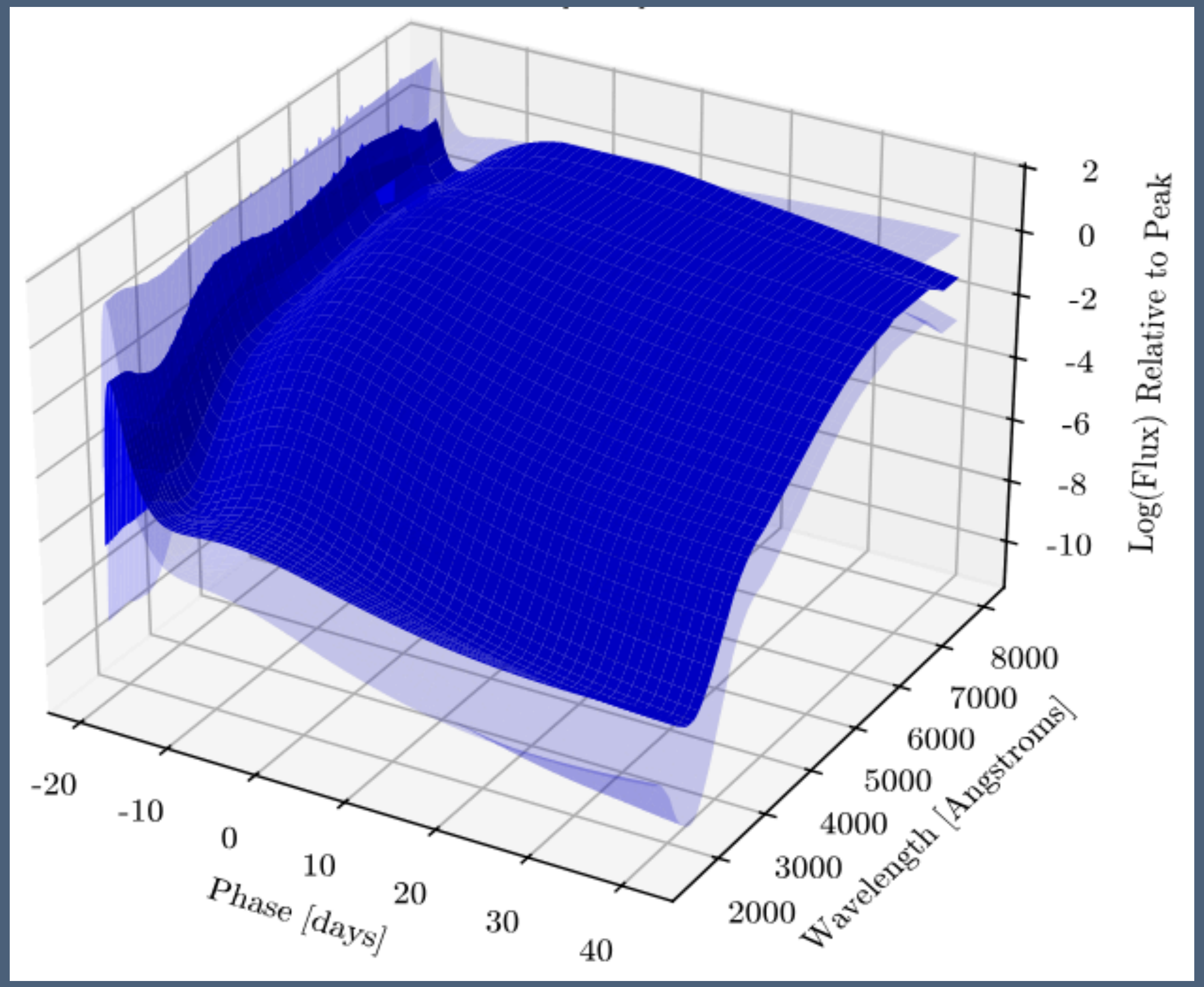
What might this object look like in a different filter? At a different redshift? With different extinction?

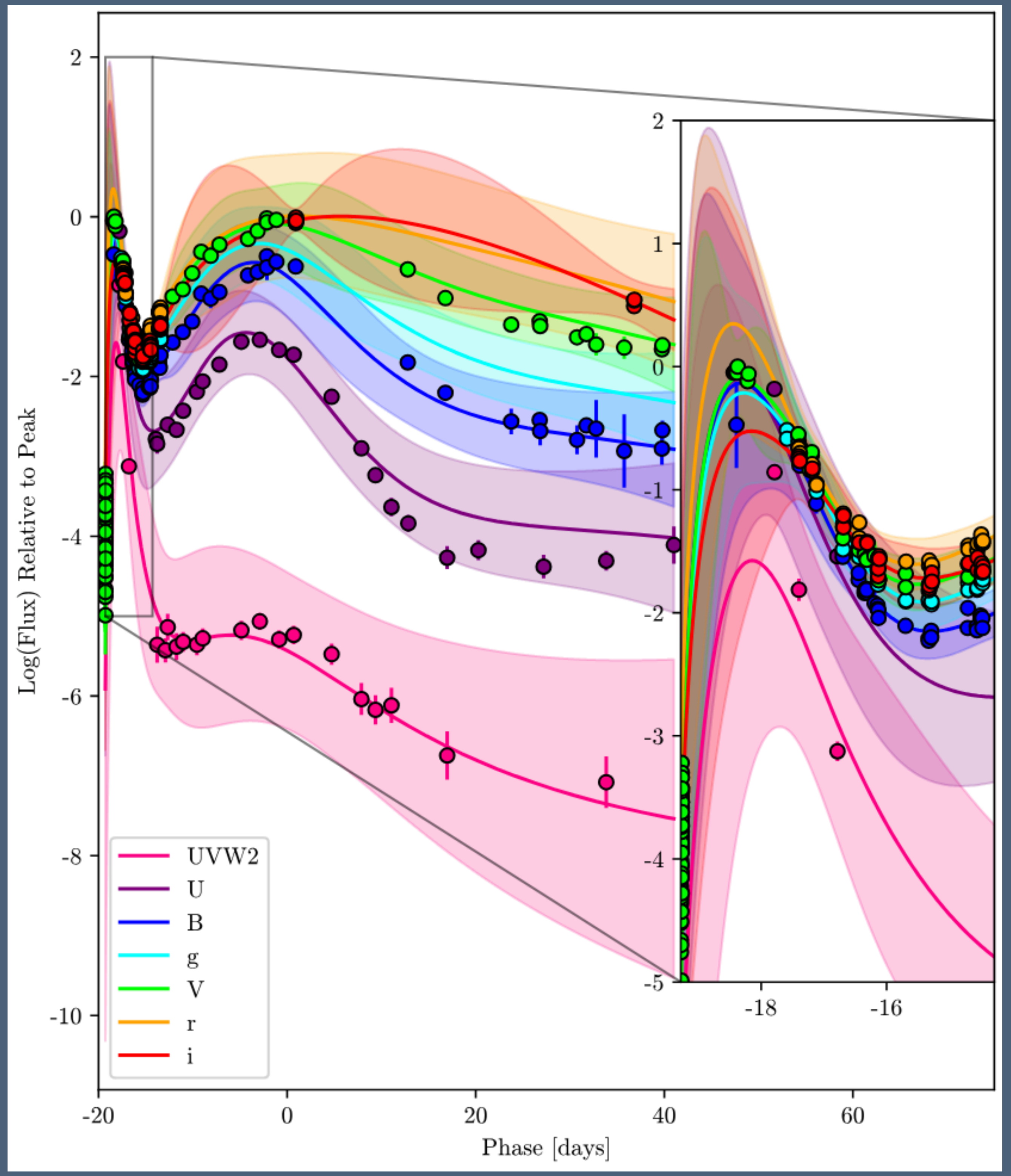
How does my favourite niche group of outlier objects behave in aggregate?

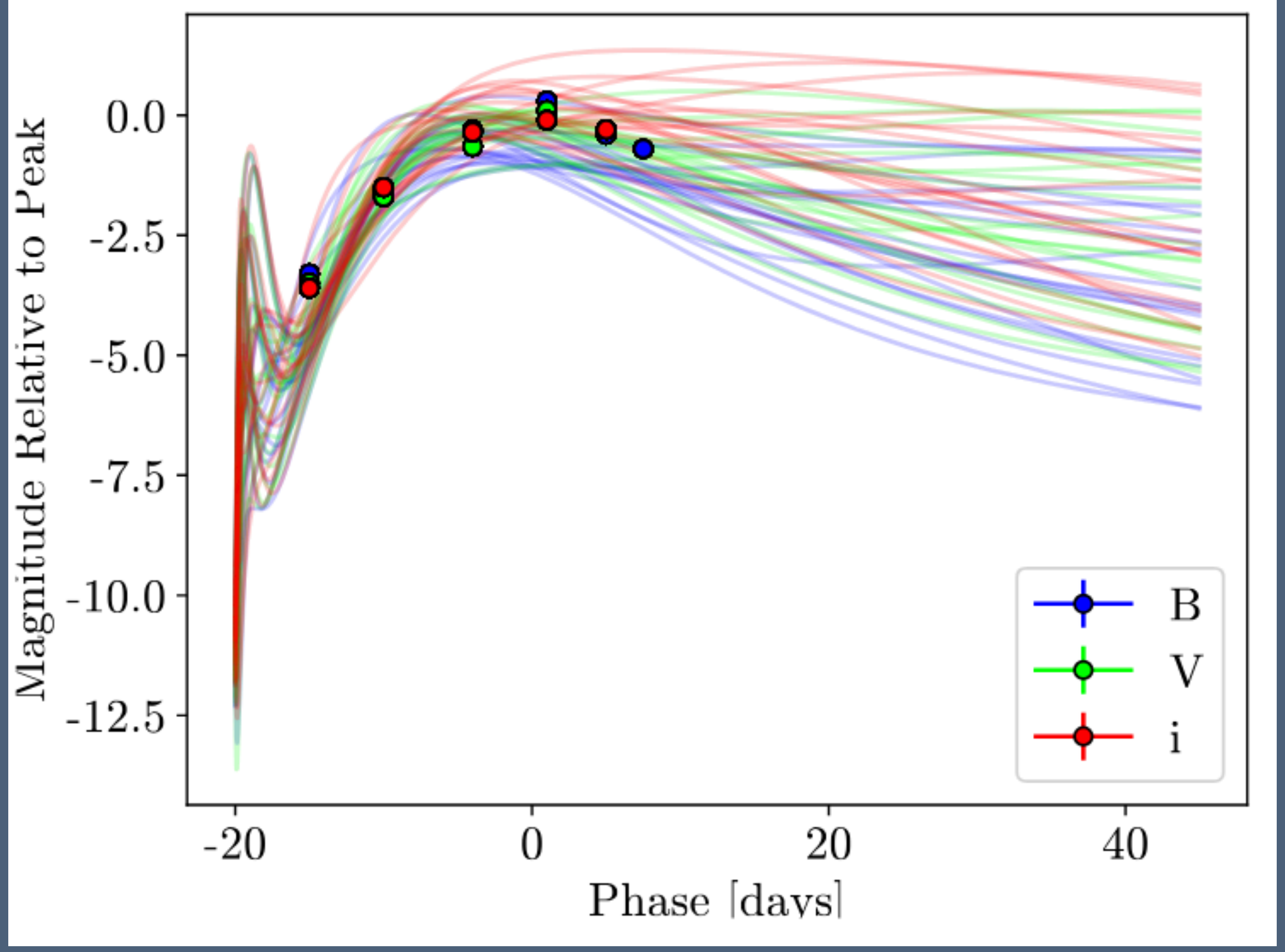
How can I make a bunch of fake objects that look like these real objects? But sampled at different times? Filters? Extinctions? Redshifts?

The Tools









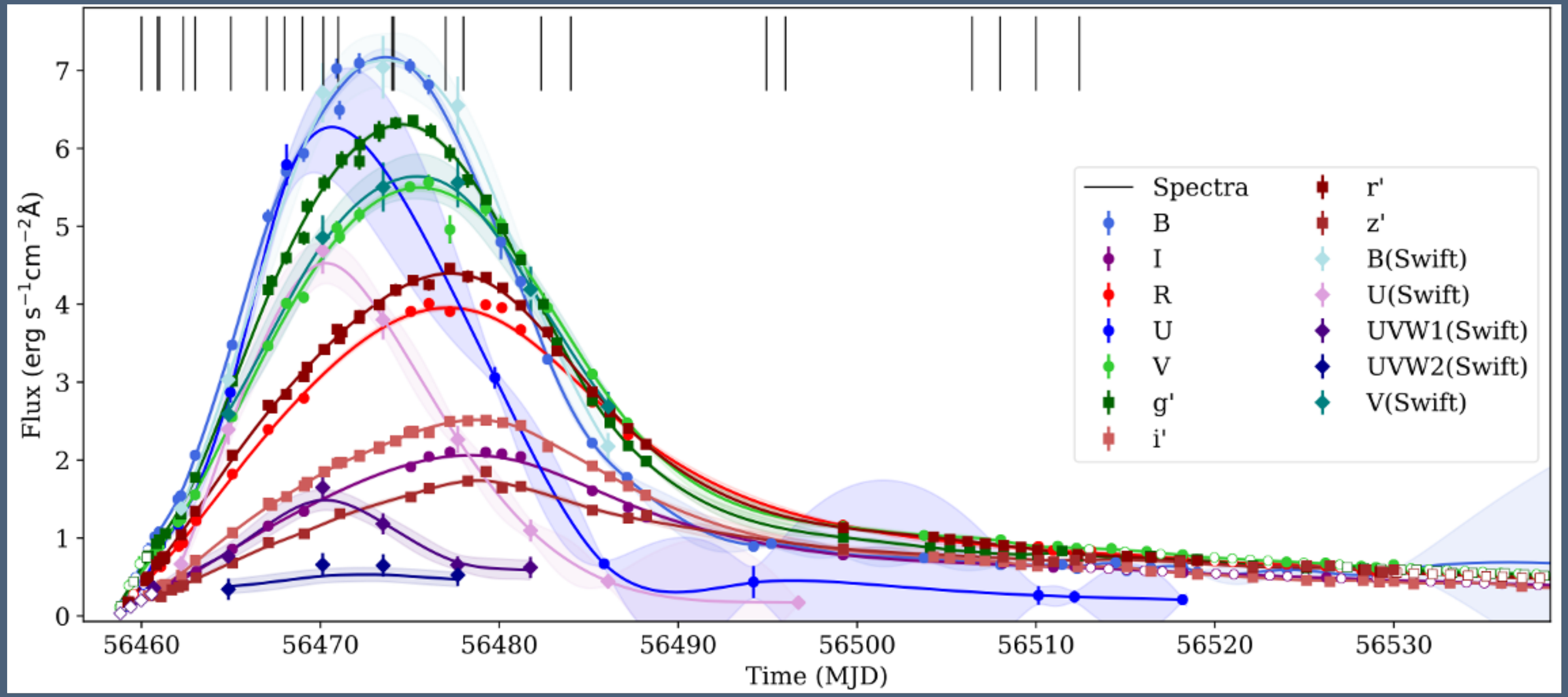
For a single object, or a collection of objects

2-D Gaussian Process Regression for SED Surfaces

…. but why?

GPR has several advantages when applied to observations of SNe:

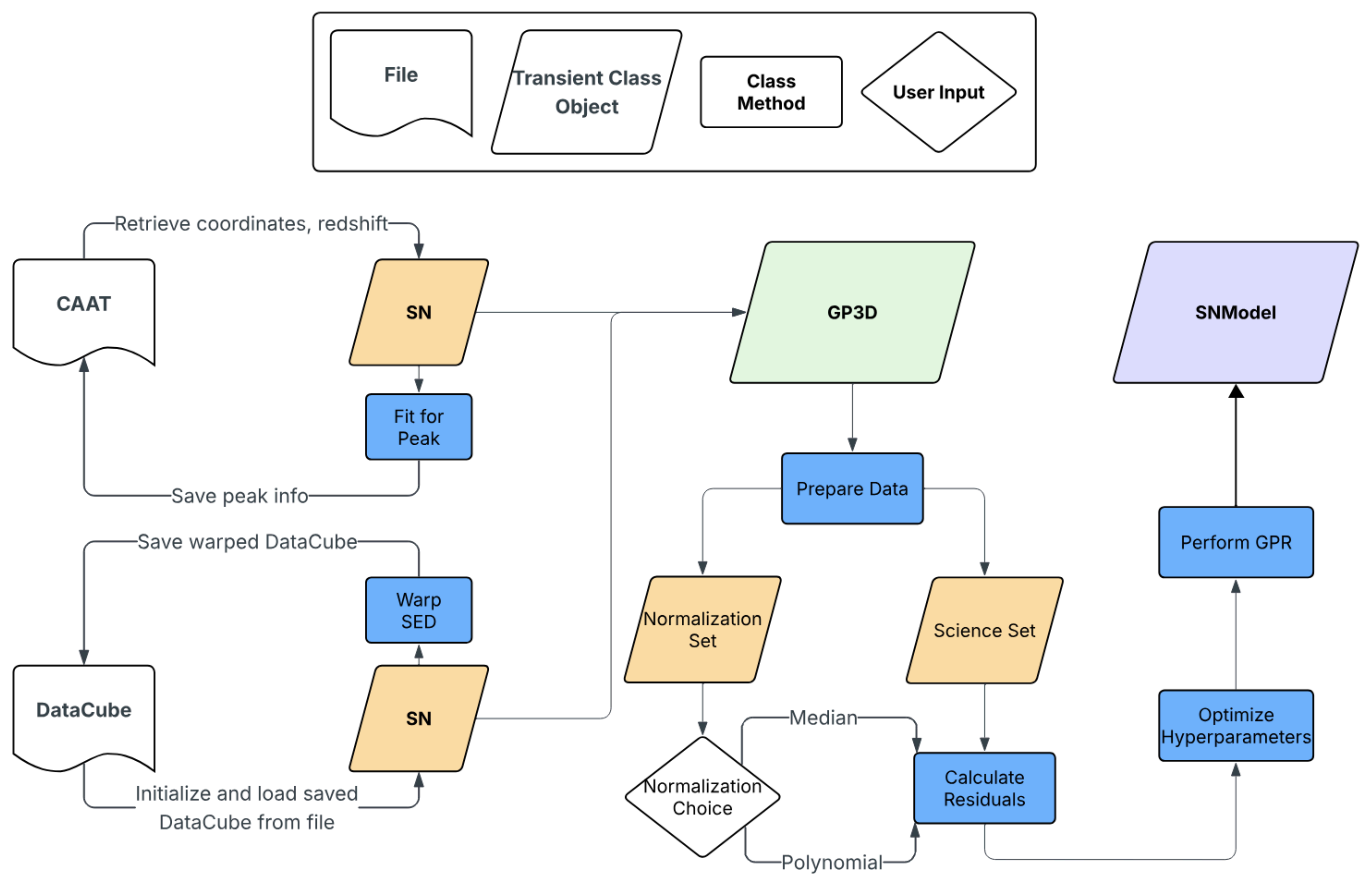
- Multi-dimensional -> allows for fitting across phase and wavelength simultaneously
- Ideal for modeling sparse data
- No physics assumed, e.g. no prior knowledge about the progenitor/explosion properties
- In log space, kernel length scales match typical SN evolution across phase and wavelength
- Expensive to run but can perfectly handle small sample sizes
- Enables uniform modeling across SN subtypes—never been done before



Vincenzi et al. (2019)

The Tools

The Process



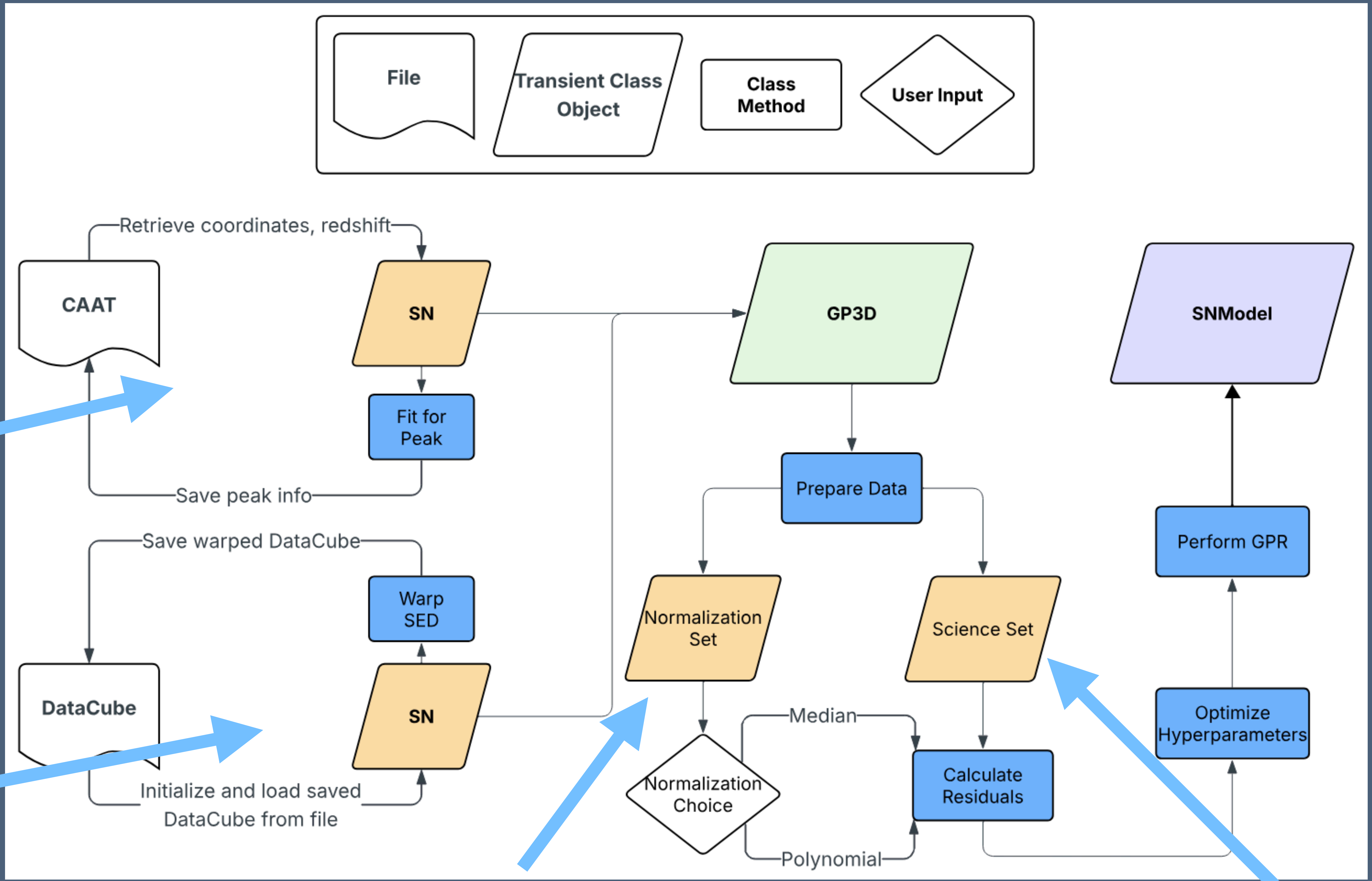
Pellegrino et al. 2026

The Tools

The Process

Get Some Metadata

Turn Input Photometry into SED



Then:

- Remove Mean Model Behavior
- Condition GP's on Sample
- GP Interpolate Model-Normalization Residuals
- Construct Model From Normalization Set + Interpolated Residuals

Choose objects to Normalize Against and Object(s) to model

These should be related, e.g.:

- Normalize ALL SESNe, Model All Type IIb SNe -> Type IIb Template
- Normalize ALL FBOTs, Model SN2018gep - > Single Object Model

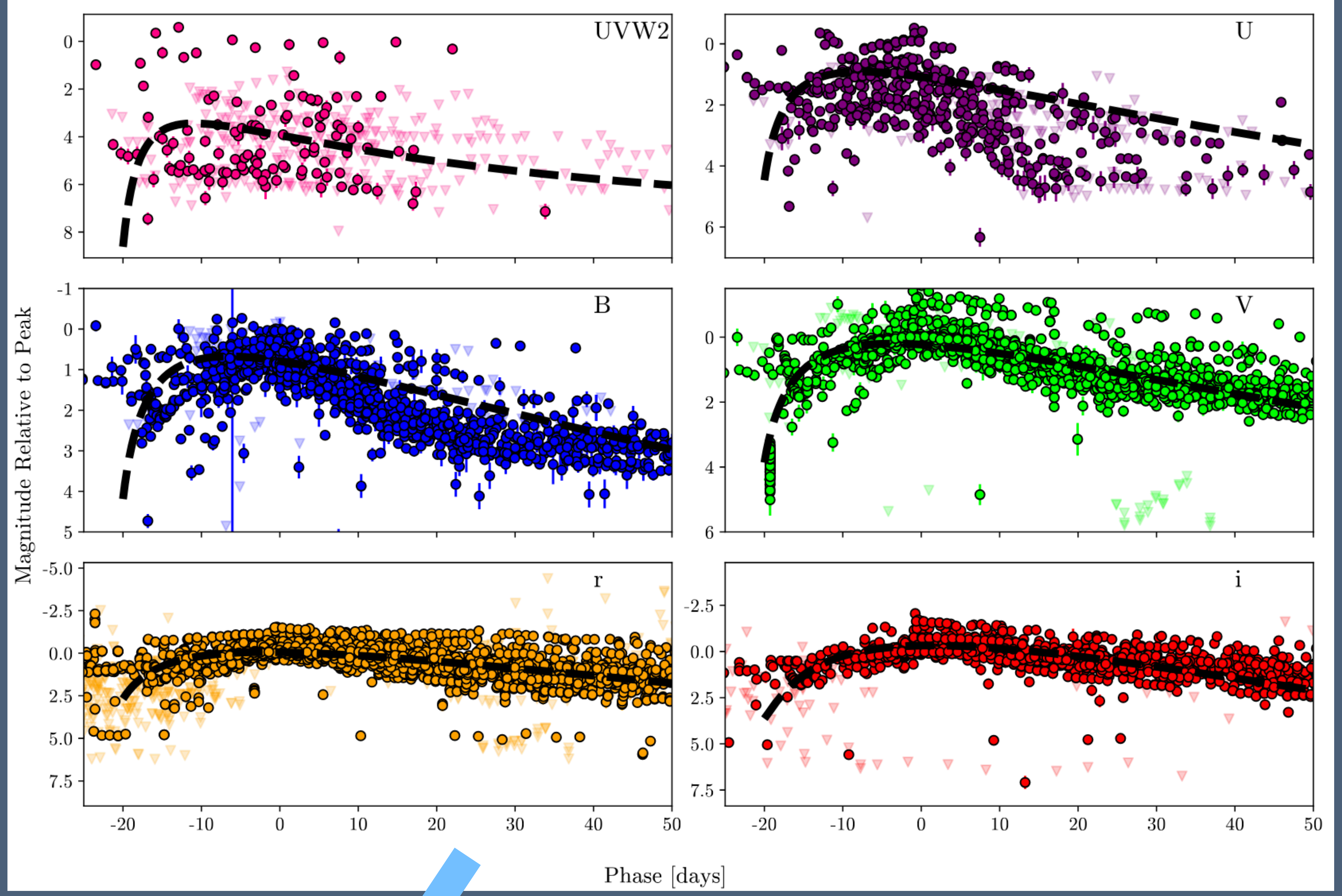
The Process

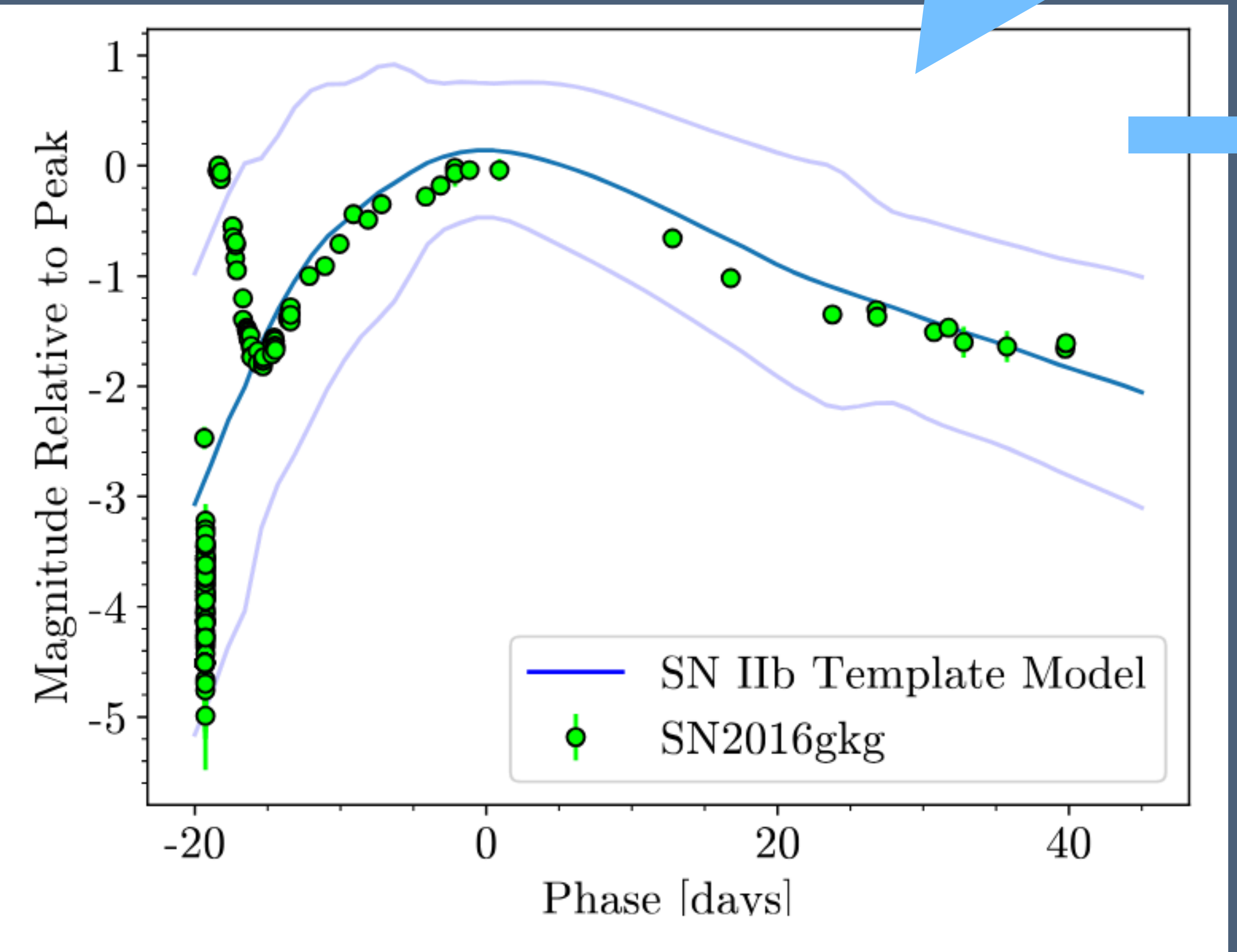
Normalization Set

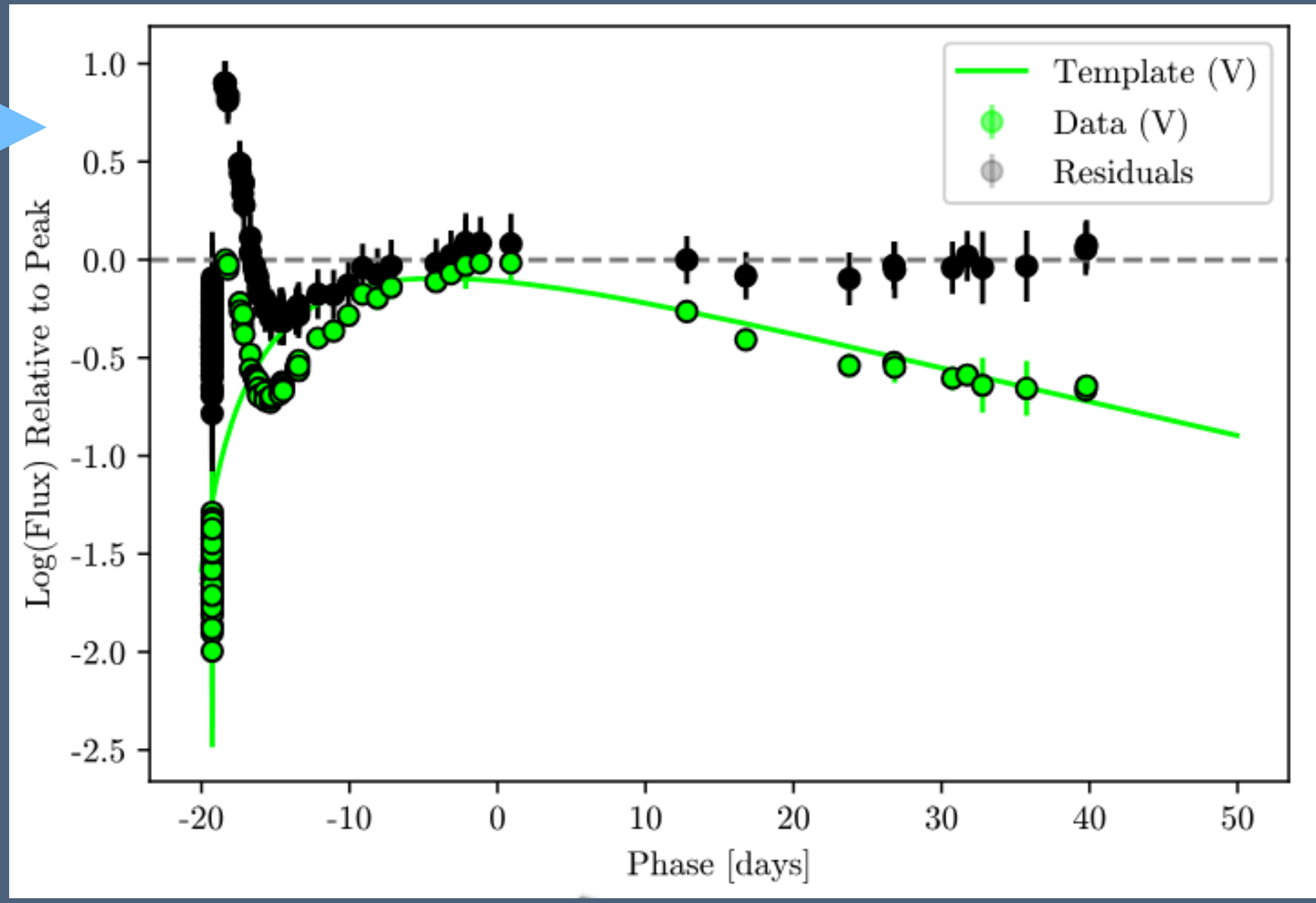
GP's Work Best with a mean of Zero

Describe your model object(s) as the residuals about some sample mean

See Khakpash et al. 2024 for single-filter photometry vs SED models here





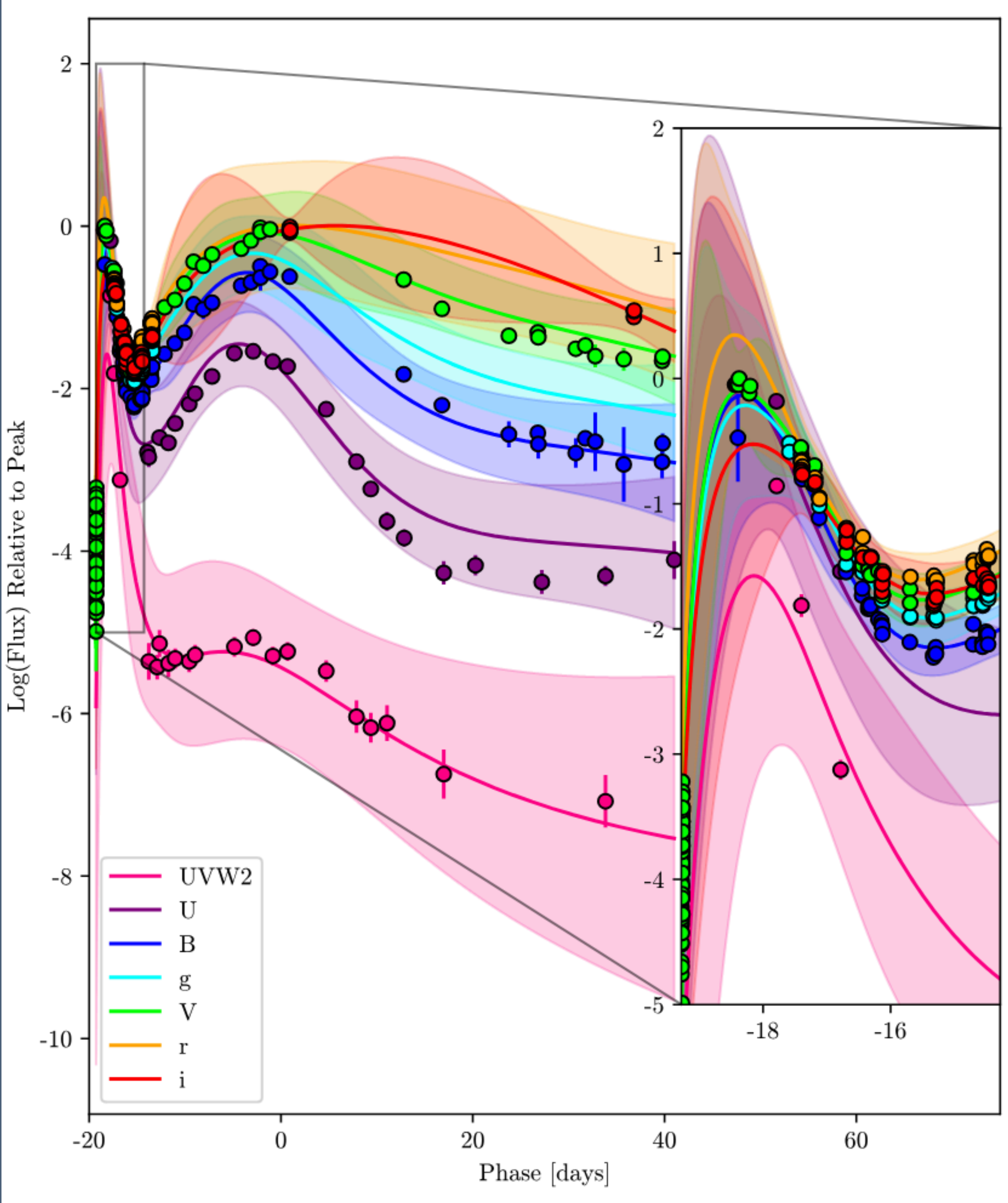


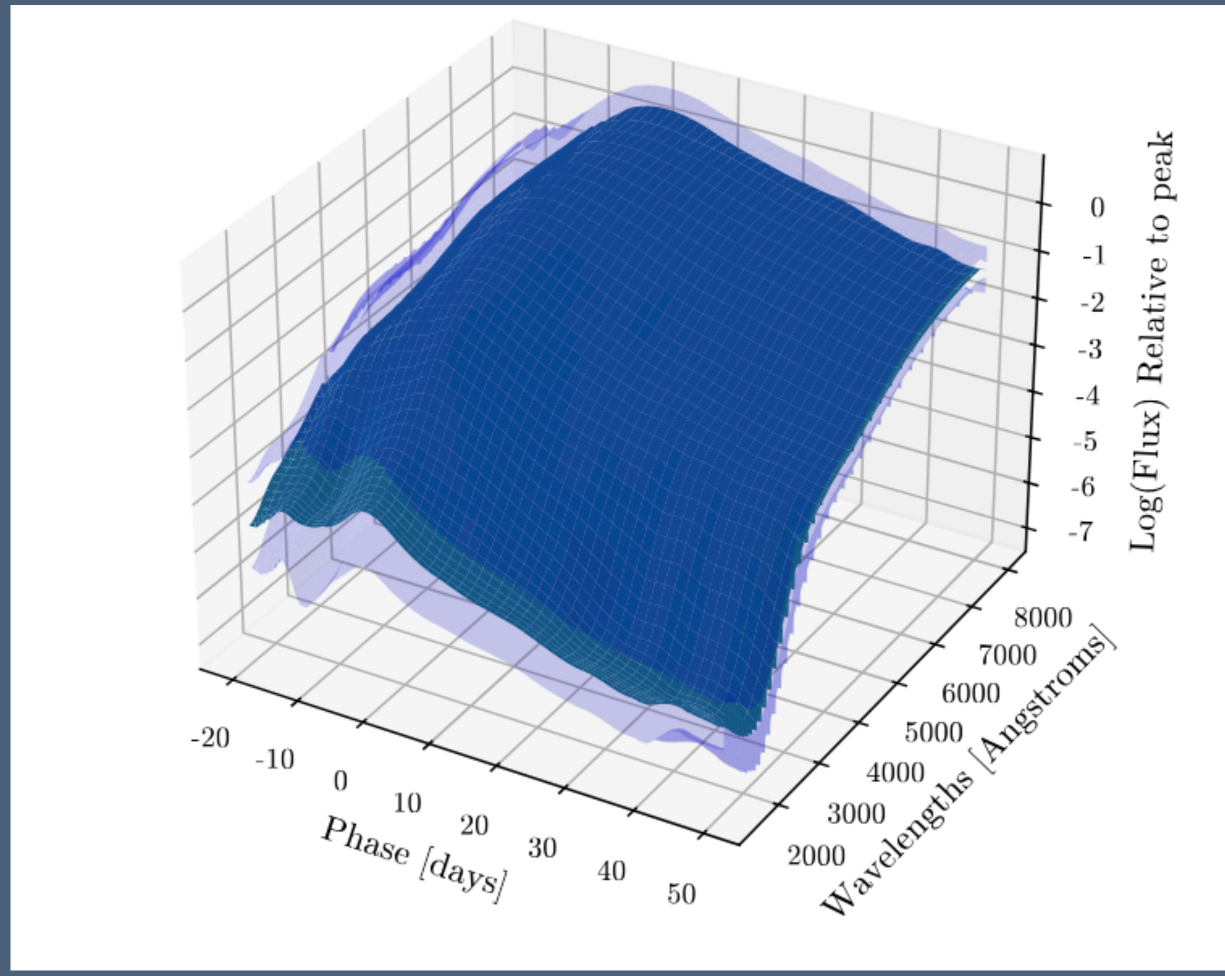
Pellegrino et al. 2026

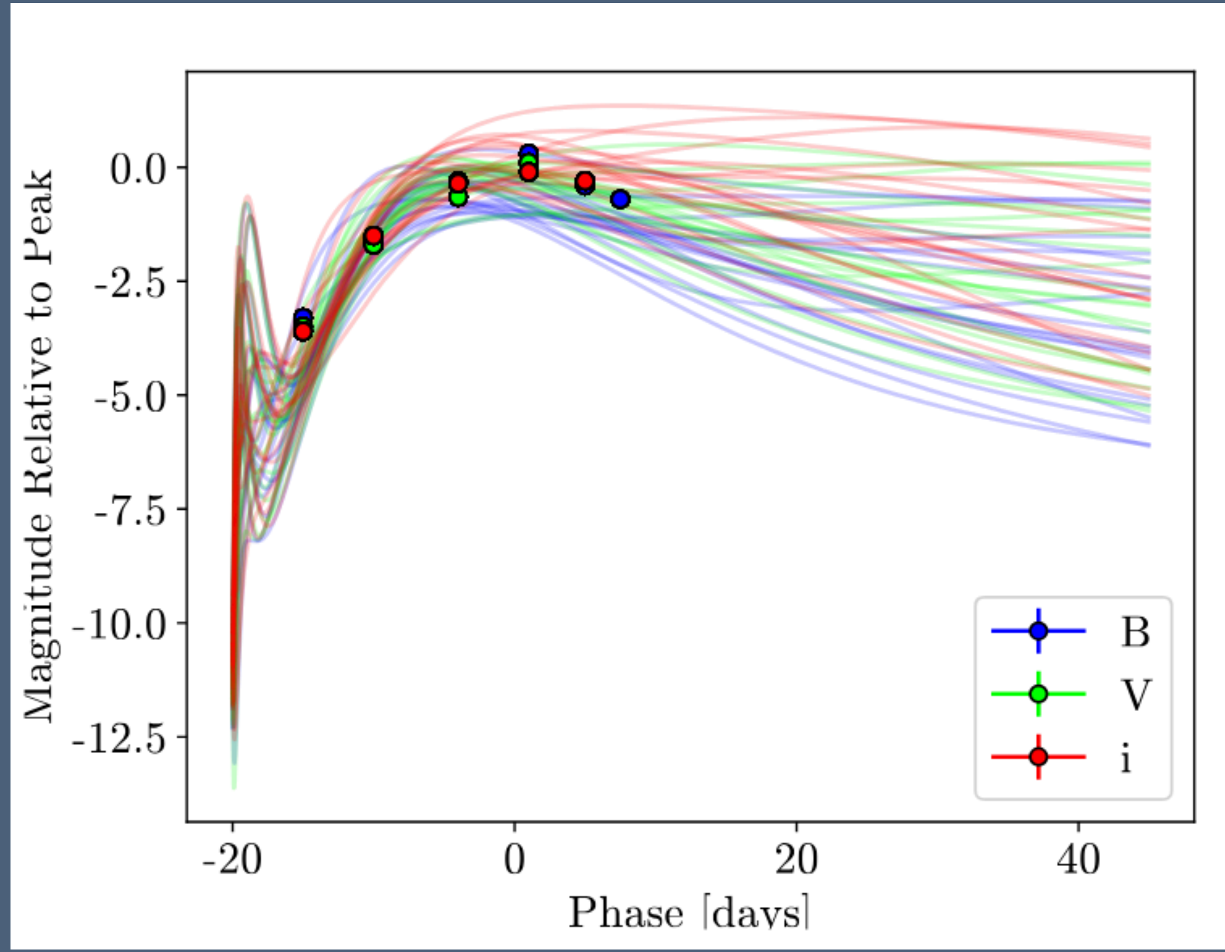
Tools
Outputs

Outputs include:

- Best-fit model and error's for model object(s)
- Broadband SED from GP2D Interpolation that can be used to model objects at different filters, redshifts, extinctions, etc.
- GP Kernels that can be sampled to create varied observed lightkurves from model fits



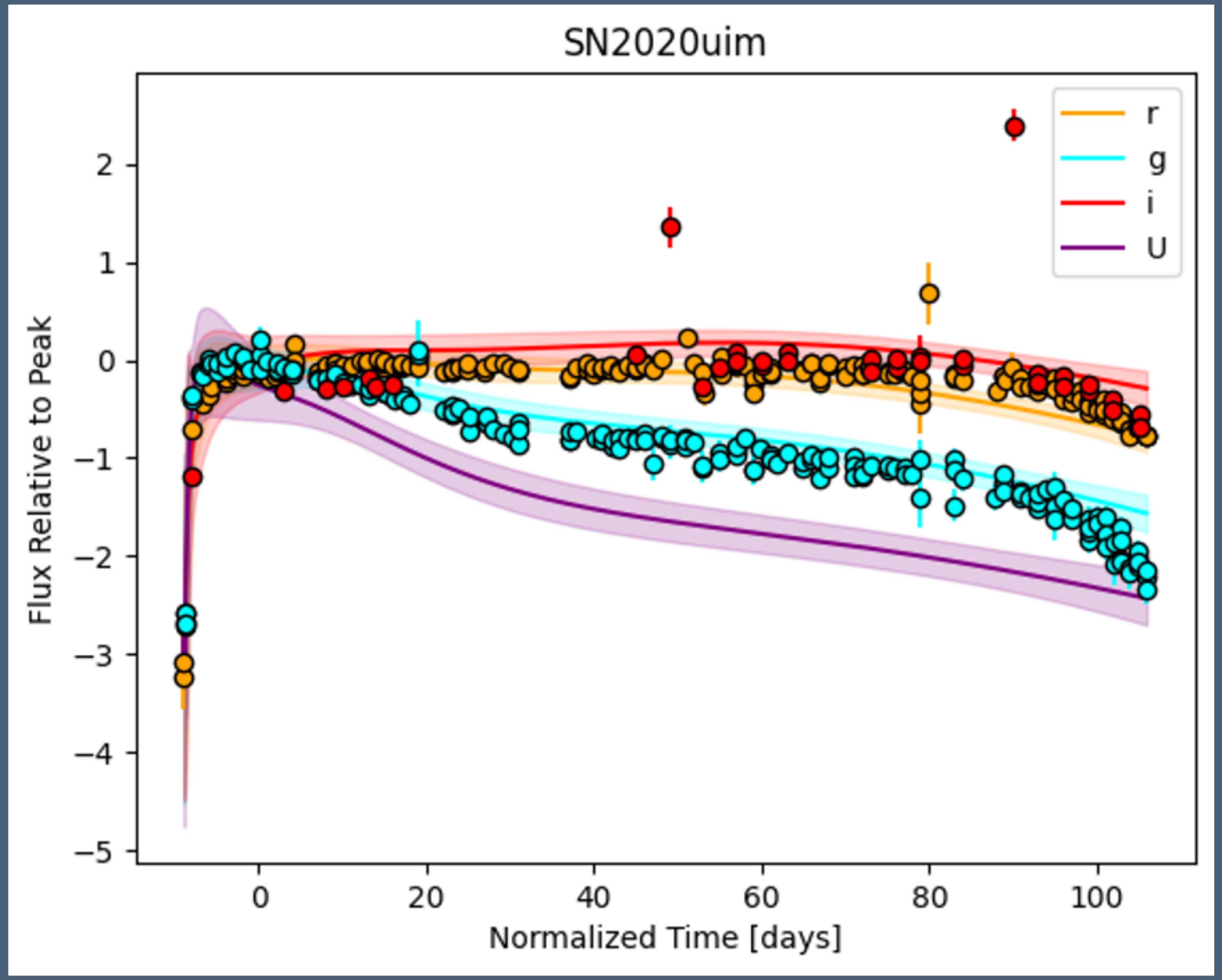


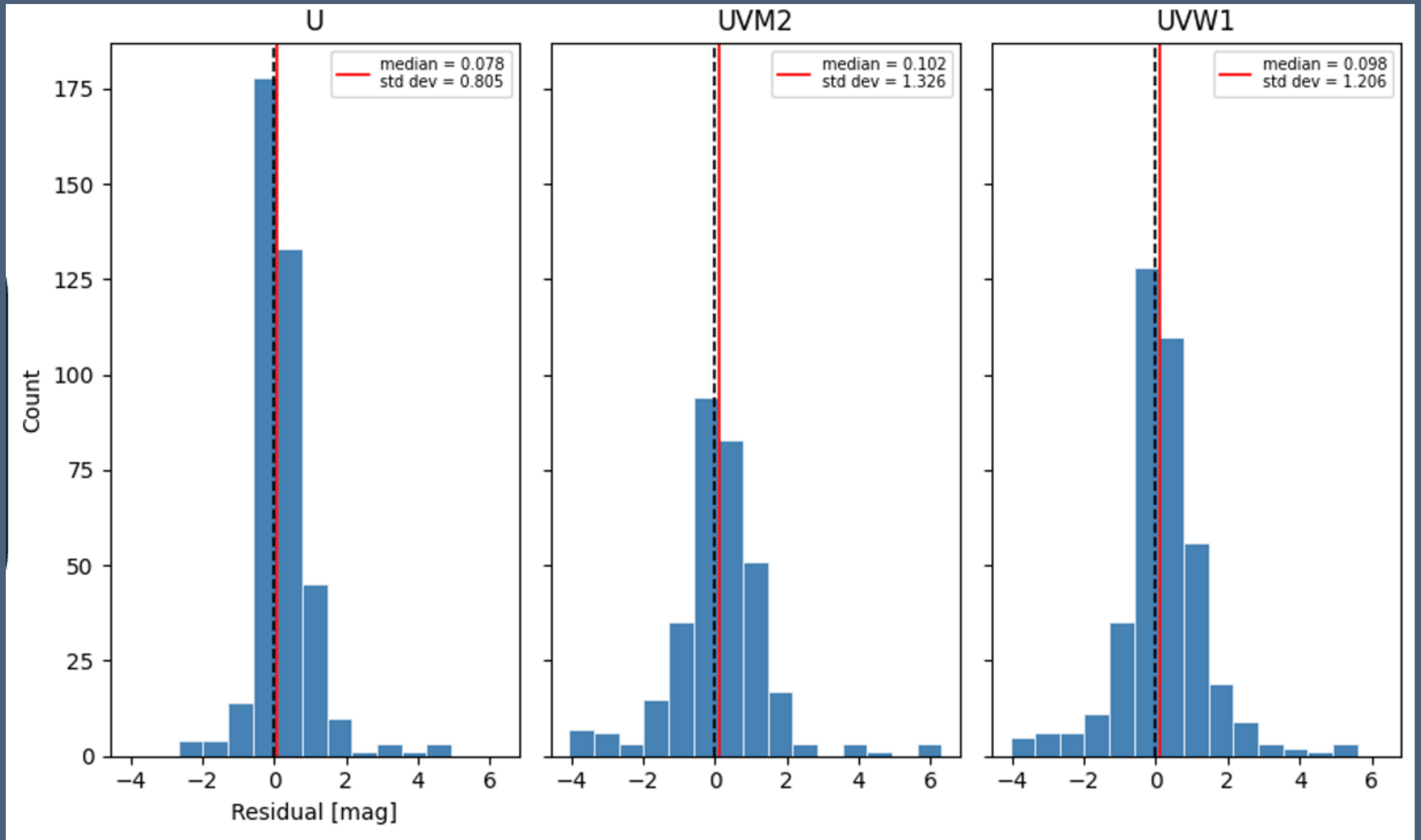


Pellegrino et al. 2026

How well can we Recover Sample Behaviour?

A Test on the ZTF Bright Transient Survey





- Identified ZTF BTS objects in GOPREAUX sample with UV + gri photometry
- Used fits to predict UV photometry and compared with the true values from the observations
- Fit these SNe in gri-bands

GOPREAUX can be used to extrapolate light curves for missing filters, enabling better modeling, bolometric luminosity measurements, and physical parameter inference.

gopreaux

Data and Tools are available now!

- All of the data and tools are publicly available on Craig's GitHub now!
  - github.com/crpellegrino/gopreaux/
- Details in Pellegrino et al. 2026 *subm.* (arXiv:2604.03372)
- Tools are effectively open beta - still some rough edges and documentation to work on, but functional, reach out to Craig , Maryam, or I if you are interested or have any questions/comments/concerns!
- Integration into LightcurveLynx and in progress! Talk to us if you're interested in incorporating this into your workflow.