Classifying AGN and PISN within Fink

Etienne Russeil - *LPC Université Clermont Auvergne, France*
Emille Ishida - *LPC Université Clermont Auvergne, France*
Roman Le Montagner - *IJCLab Université Paris-Saclay, Orsay, France*
Julien Peloton - *IJCLab Université Paris-Saclay, Orsay, France*
Anais Möller - *Centre for Astrophysics and Supercomputing, Swinburne University of Technology, , Australia*
Emmanuel Gangler - *LPC Université Clermont Auvergne, France*
Stephane Blondin - *LAM, Université Aix Marseille, France*

AGN classifier

ZTF database

The active galactic nuclei (AGN)

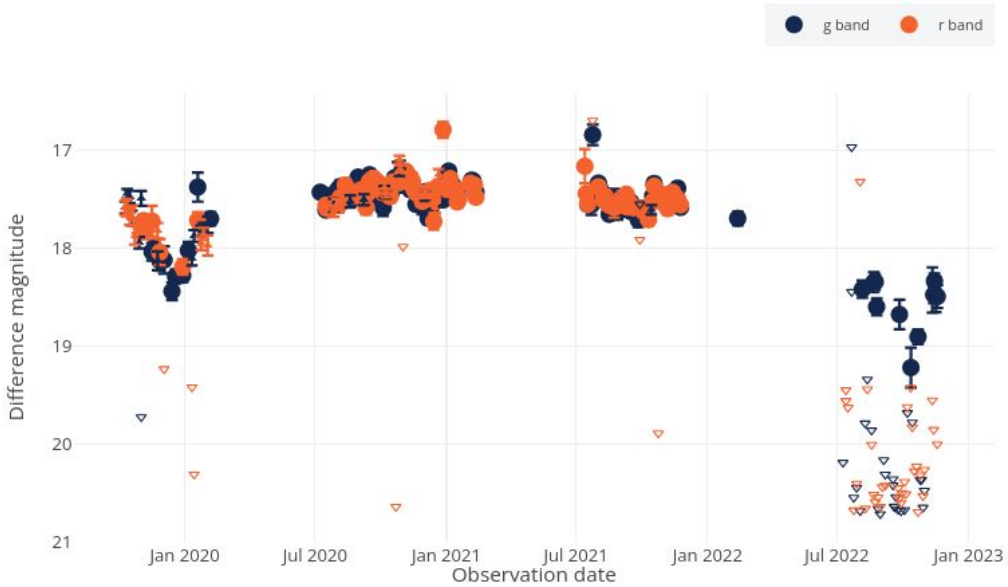
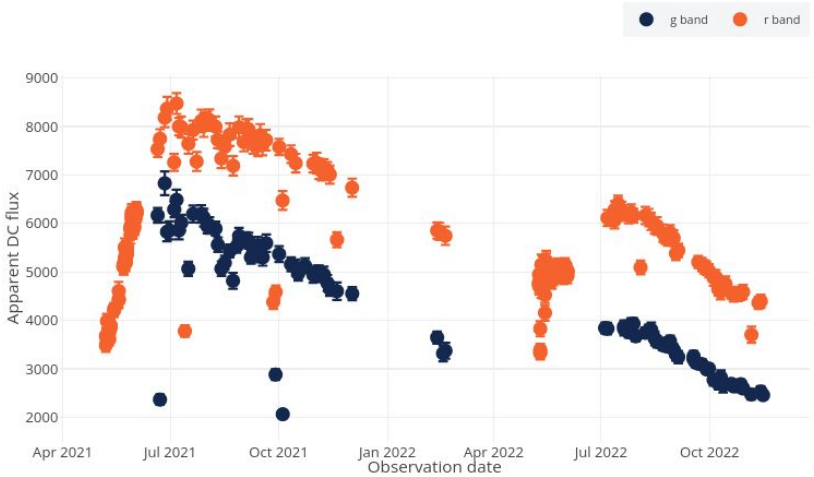
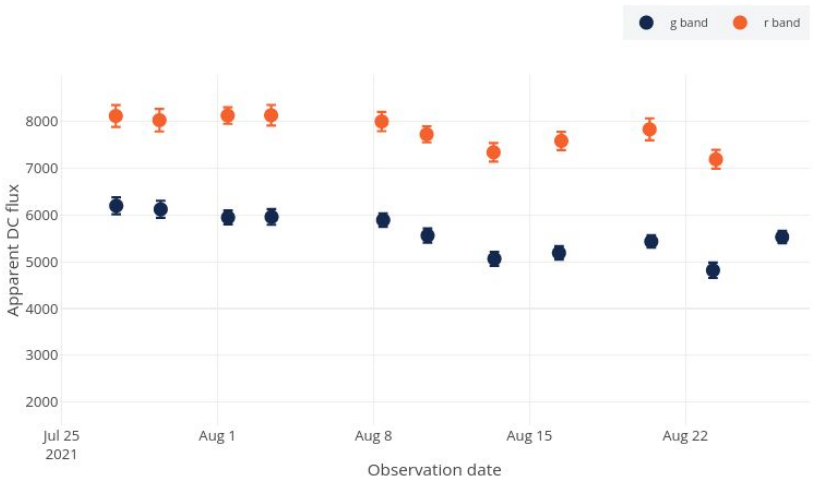


***Image**: Centaurus A : ESO/WFI (Optical); MPIfR/ESO/APEX/A.Weiss et al. (Submillimetre); NASA/CXC/CfA/R.Kraft et al. (X-ray)*

Databases : ZTF data

- ZTF data from 2019 to 2021 with label (TNS or Simbad)
- Available passbands : g and r
- Dealing with **alerts** : 30 days light curves
- We sample 1 000 000 non-AGNs VS 100 000 AGNs





Etienne Russeil et al. Finding active galactic nuclei through Fink : https://arxiv.org/abs/2211.10987

Feature extraction ZTF : statistical features

Per passband:

- Standard deviation of the flux
- Mean signal to noise ratio
- Number of points
- Maximum flux

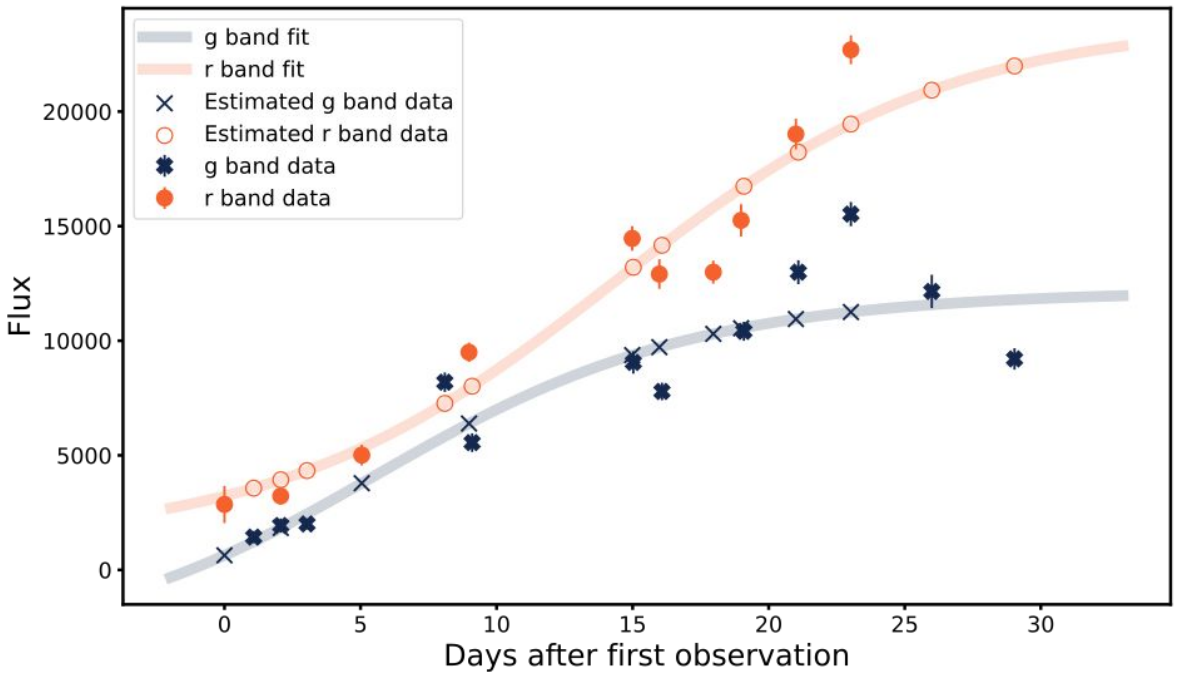
Additionally:

- Right ascension and declination
- Maximum color (g-r)
- Color (g-r) standard deviation

12 features

Etienne Russeil et al. Finding active galactic nuclei through Fink : https://arxiv.org/abs/2211.10987

Color estimation



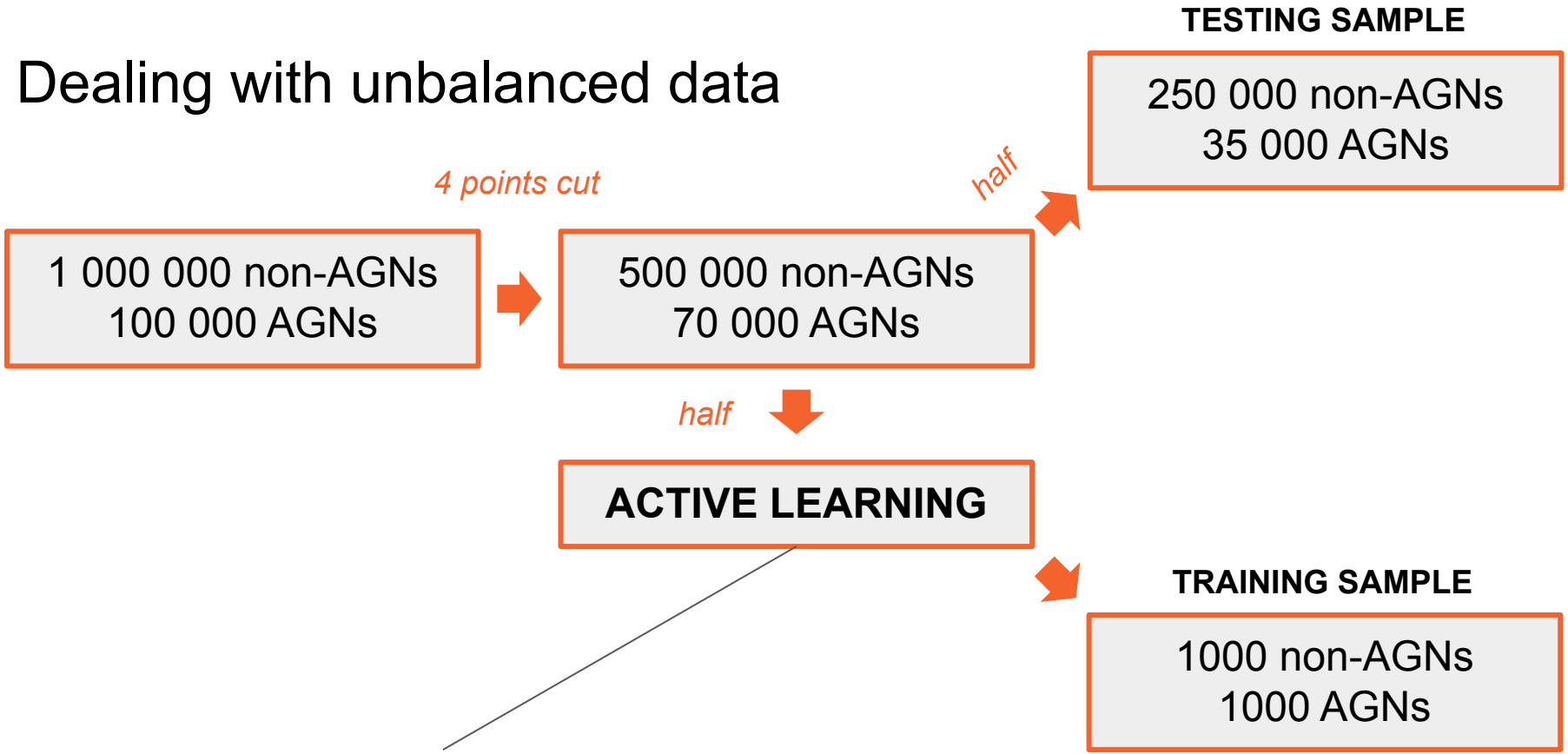
Symbolic Regression :

$$f(t) = \frac{1}{1 + exp^{-(At - exp^{Bt} + C)}} + D$$

(We require 4 points per passband minimum)

Etienne Russeil et al. Finding active galactic nuclei through Fink : https://arxiv.org/abs/2211.10987

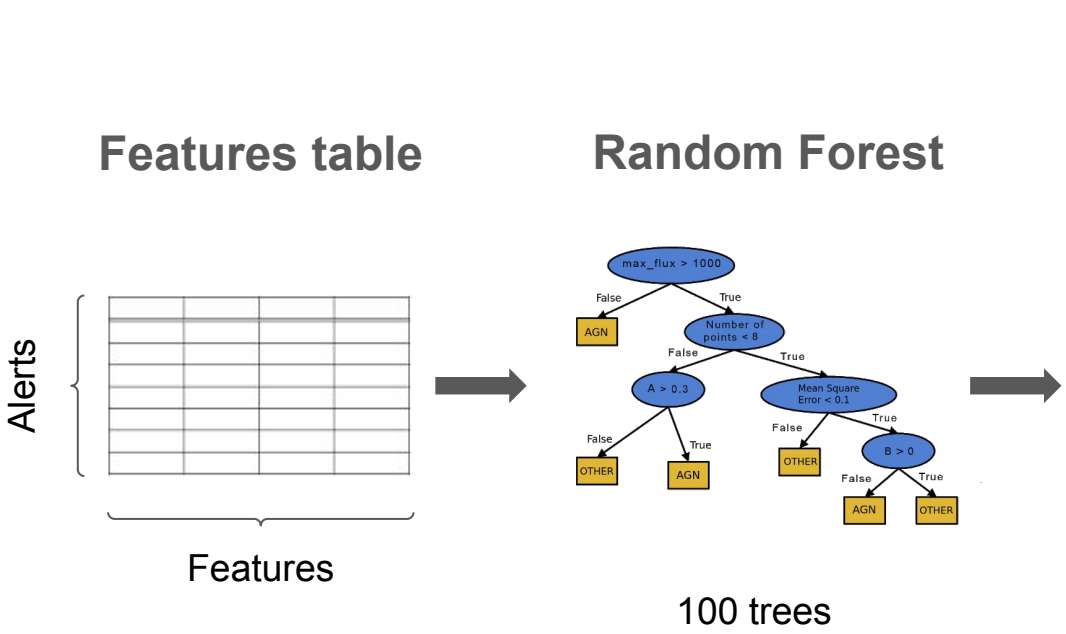
Dealing with unbalanced data

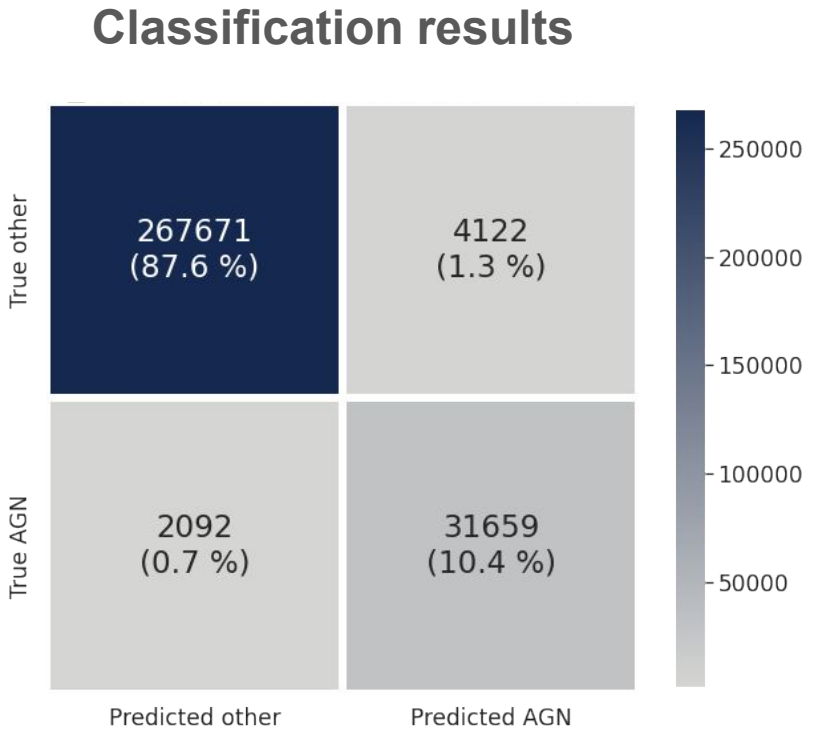


See Marco Leoni : *Early supernovae Ia classification using active learning*

Etienne Russeil et al. Finding active galactic nuclei through Fink : https://arxiv.org/abs/2211.10987

Random forest results





Unbalanced testing dataset with :

88.3 % non-AGNs
11.7 % AGNs

Etienne Russeil et al. Finding active galactic nuclei through Fink : https://arxiv.org/abs/2211.10987

AGN classifier

ELAsTiCC data

Extended LSST Astronomical Time-series Classification Challenge (ELASTiCC) challenge

- Available passbands : u, g, r, i, z, Y
- Training sample → whole light curves
- Additional metadata : host galaxy photometric redshift and distance to host galaxy
- We sample 100 000 non-AGNs VS 100 000 AGNs

See Emille Ishida : *The ELAsTiCC data challenge: preparing the Fink broker for LSST*

Etienne Russeil et al. Finding active galactic nuclei through Fink : https://arxiv.org/abs/2211.10987

Feature extraction ZTF : statistical features

Per passband:

- Standard deviation of the flux
- Mean signal to noise ratio
- Maximum flux
- Number of points

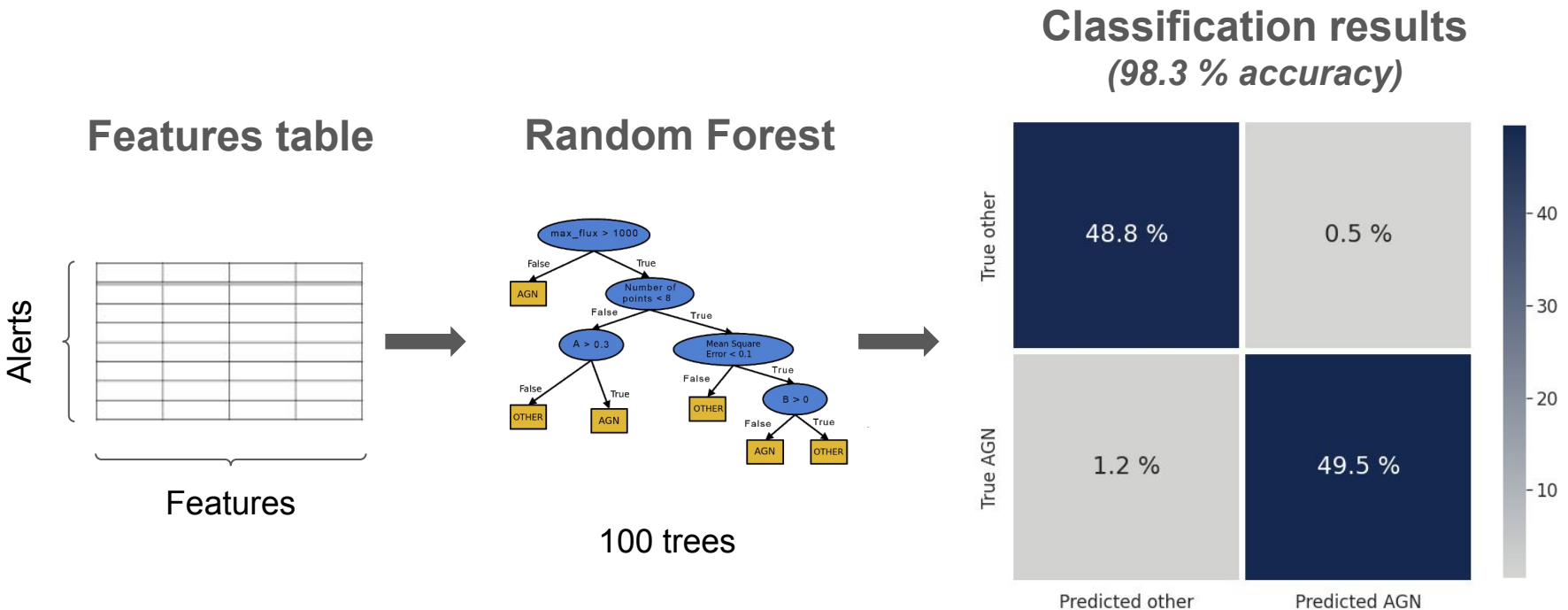
Additionally:

- Right ascension and declination
- Maximum color
- Color standard deviation
- Host galaxy metadata

32 features

Etienne Russeil et al. Finding active galactic nuclei through Fink : https://arxiv.org/abs/2211.10987

Random forest results



Etienne Russeil et al. Finding active galactic nuclei through Fink : https://arxiv.org/abs/2211.10987

Conclusion

AGN classifier : ZTF

- Internally deployed within Fink
- Promising results on direct Fink alerts
- To be deployed ..

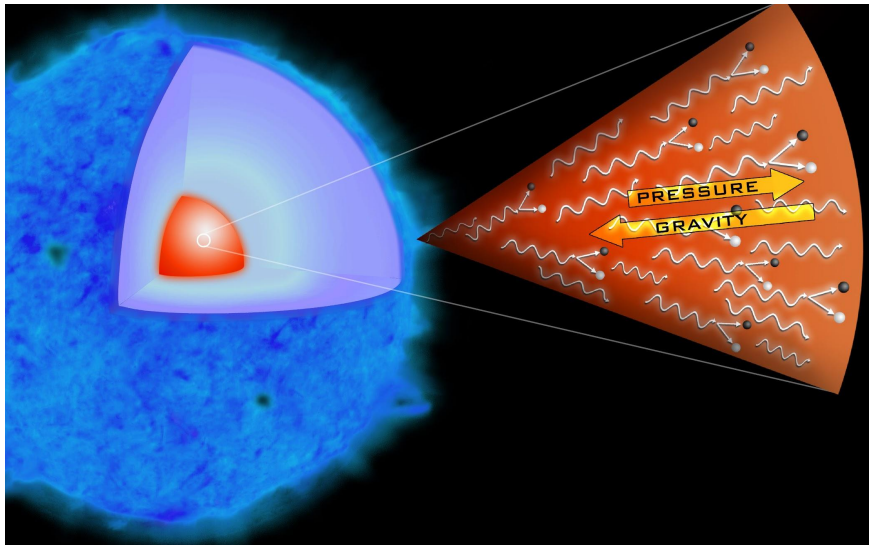
AGN classifier : ELASTiCC

- Efficient on training data
- Seems to apply to testing alerts

PISN classifier

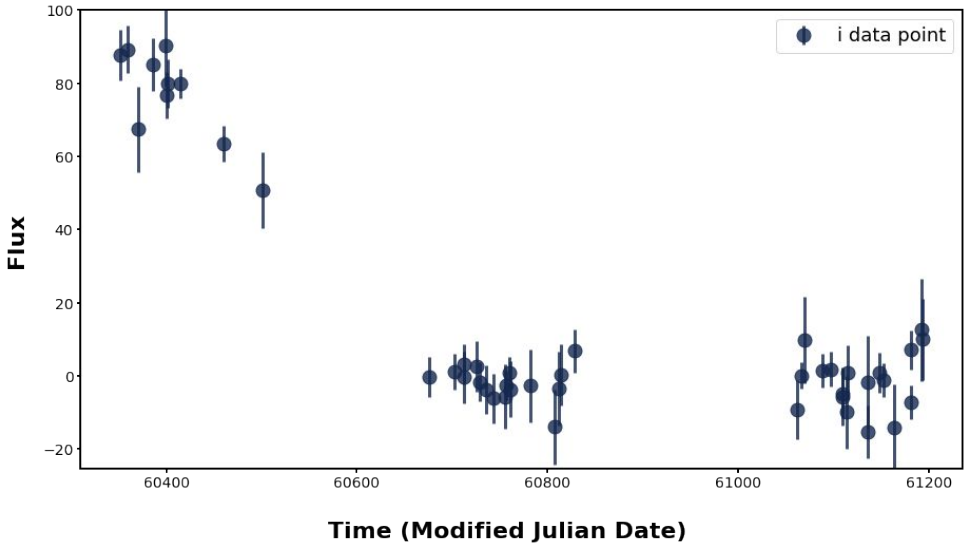
ELAsTiCC data

Pair instability supernova (PISN)



https://en.wikipedia.org/wiki/Pair-instability_supernova

PISN example in ELASTiCC



Feature extraction ZTF : statistical features

Per passband:

- Standard deviation of the flux
- Mean signal to noise ratio
- Maximum flux
- Number of points

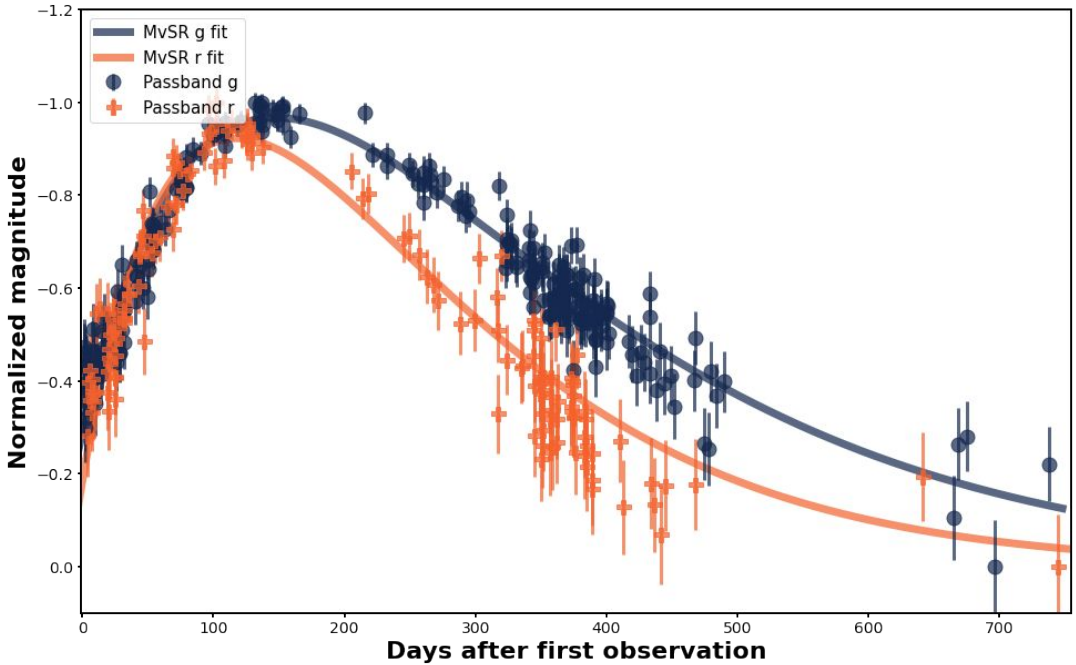
Additionally:

- Right ascension and declination
- Maximum color (r-i)
- Color (r-i) standard deviation
- Host galaxy metadata
- Fitted parameters

40 features

Parameters fit + color estimation

Reminder : SNAD 160

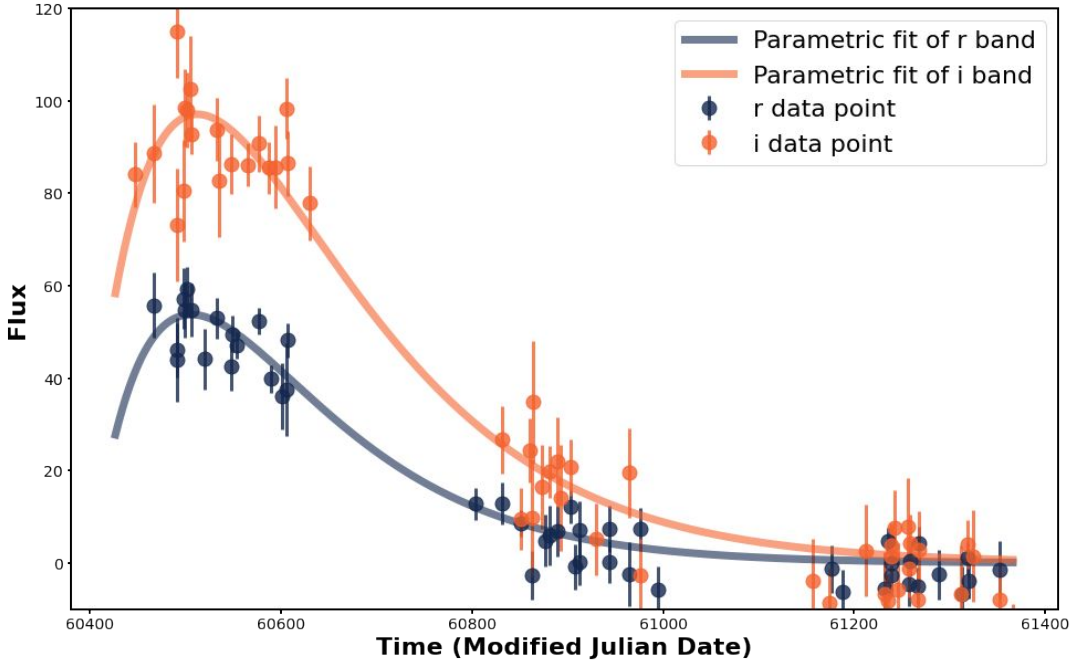


Multiview Symbolic Regression :

$$f(t) = P_2(t - P_3) \times e^{P_1(t-P_3)}$$

Maria Pruzhinskaya et al. Could SNAD160 be a Pair-instability Supernova? https://iopscience.iop.org/article/10.3847/2515-5172/ac76cf

Parameters fit + color estimation



Multiview Symbolic Regression :

$$f(t) = P_2(t - P_3) \times e^{P_1(t-P_3)}$$

We require only 3 points !

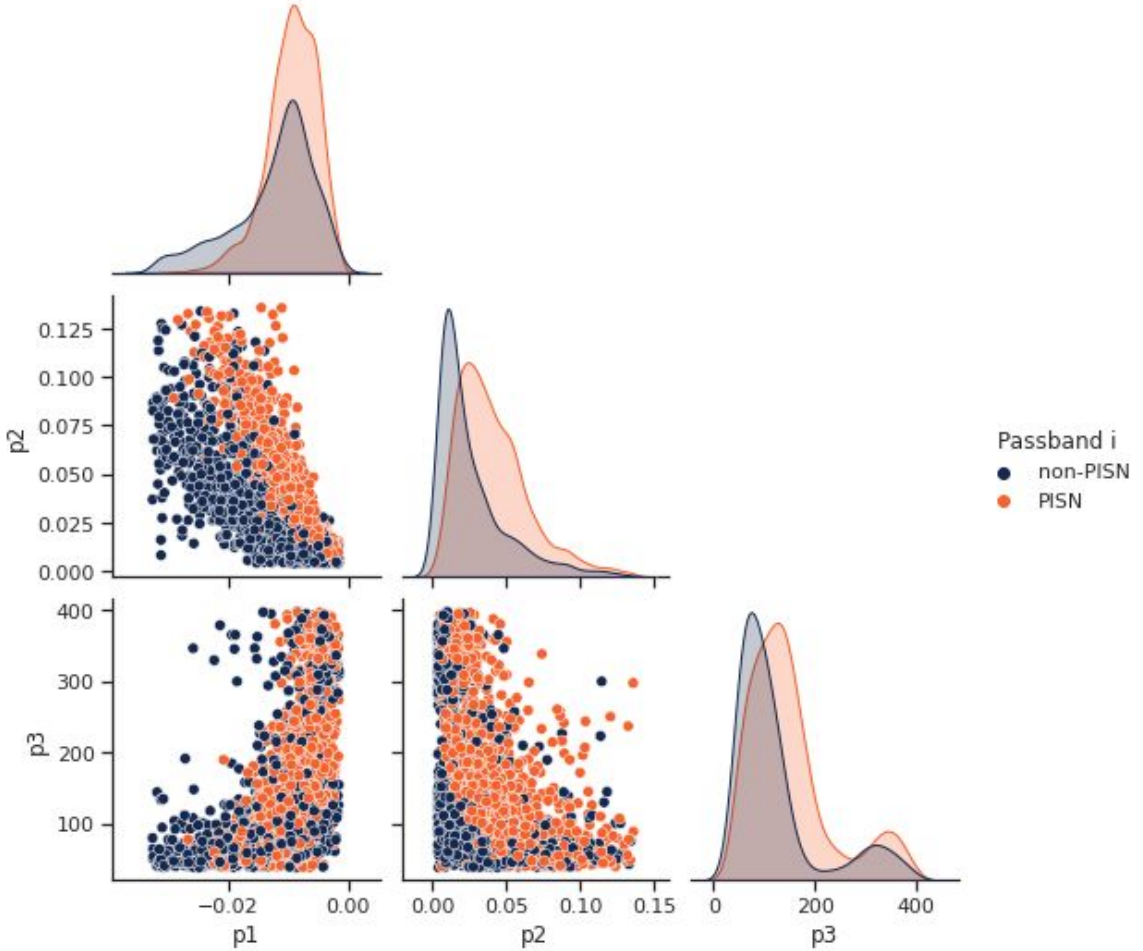
Fitted parameters distribution

100 000 non-PISN

VS

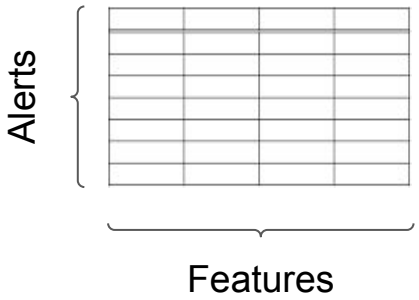
100 000 PISN

Features parameter space for passband “i”

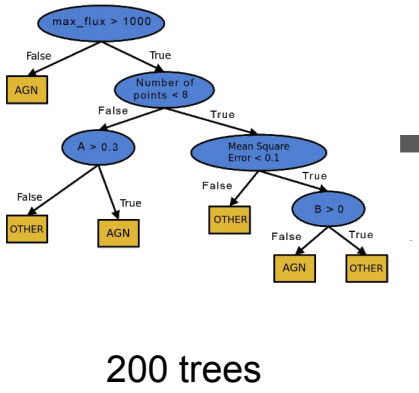


Random forest results

Features table



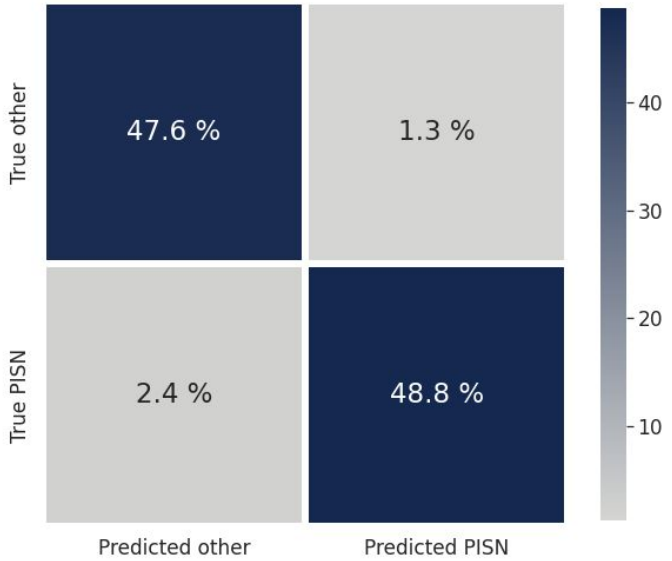
Random Forest



200 trees

Classification results

(96.3 % accuracy)



Conclusion

PISN classifier : ELASTiCC

- Efficient on training data
- Currently bad on testing alerts
- Need to cut training data

Thank you for your attention

Color computation priority

[g-r] ***1st***

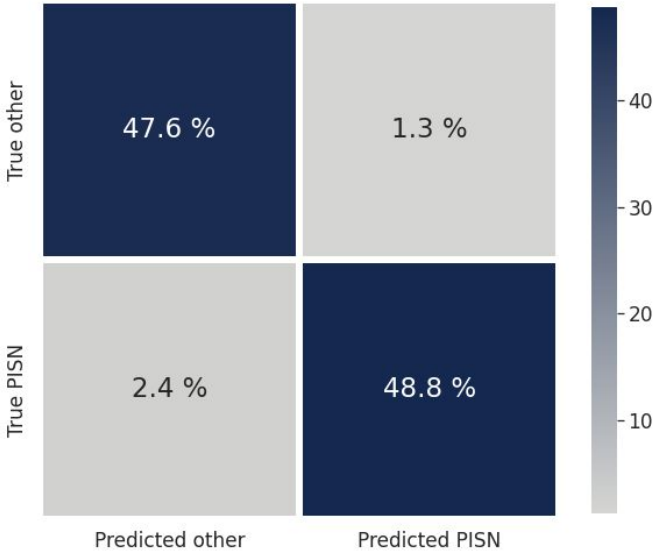
[r-i] ***2nd***

[z-Y] ***3rd***

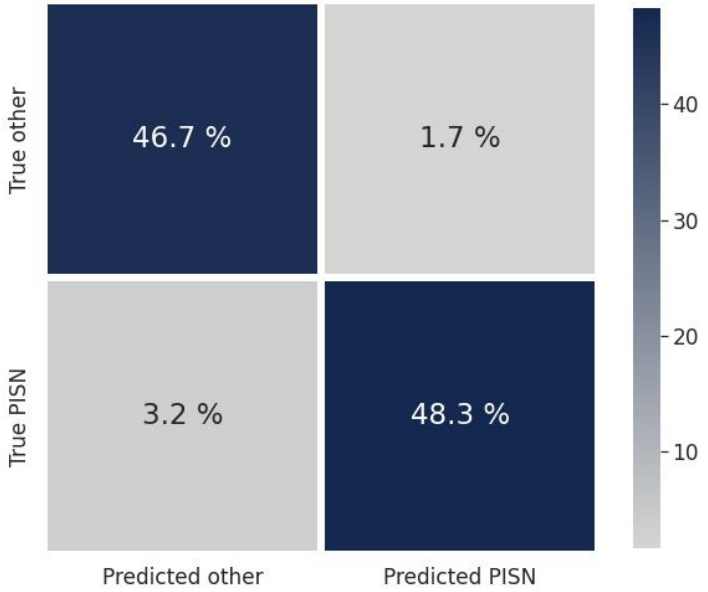
[u-g] ***4th***

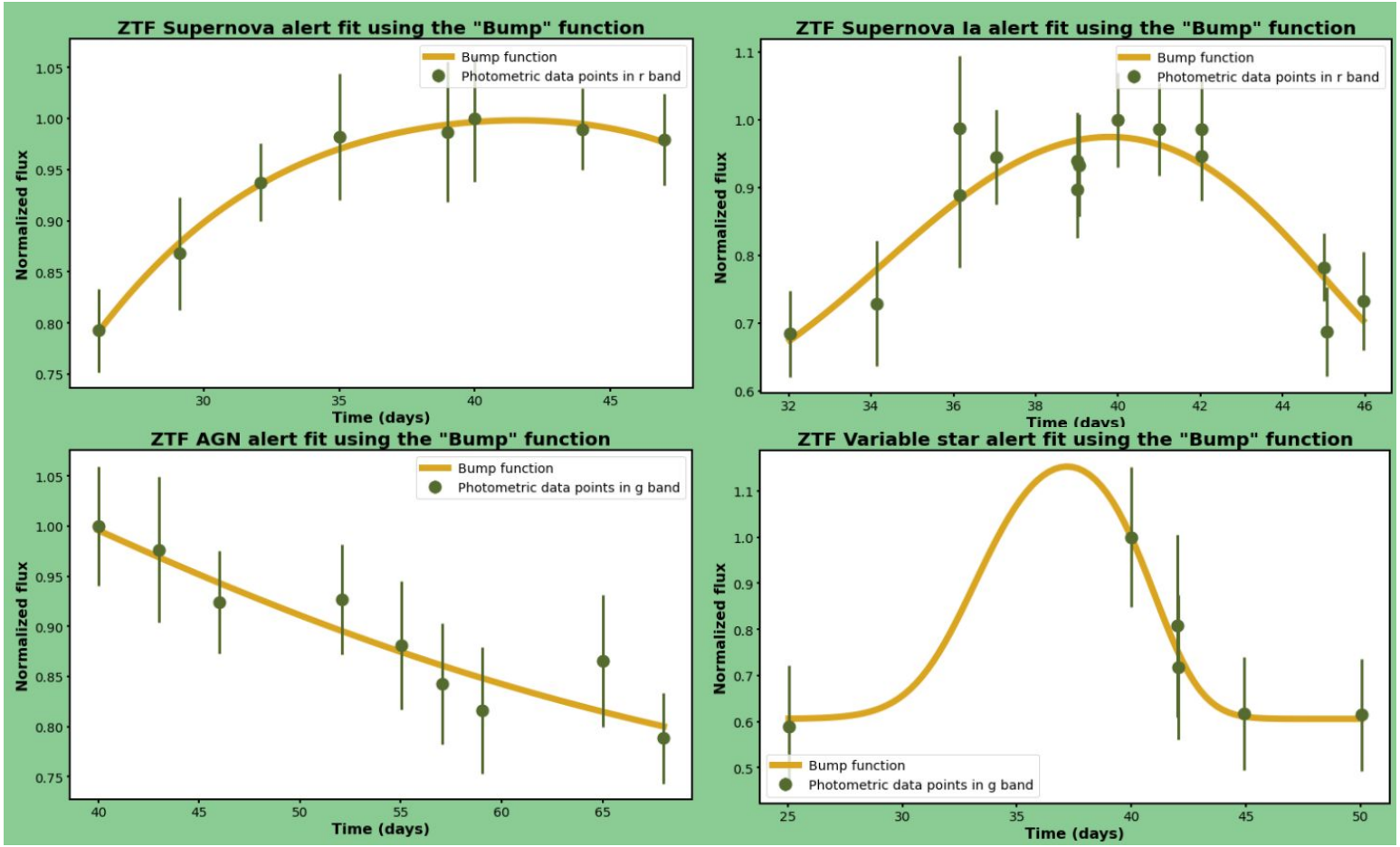
[i-z] ***5th***

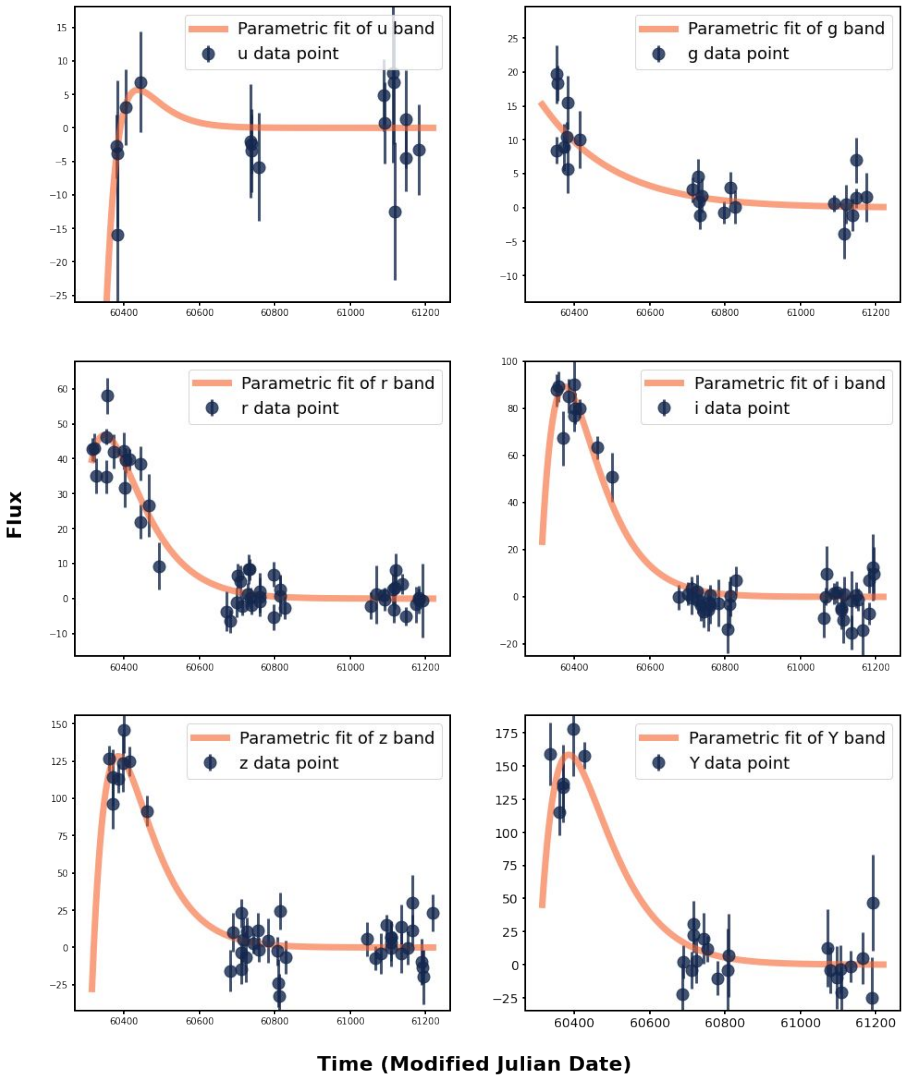
With fitted parameters

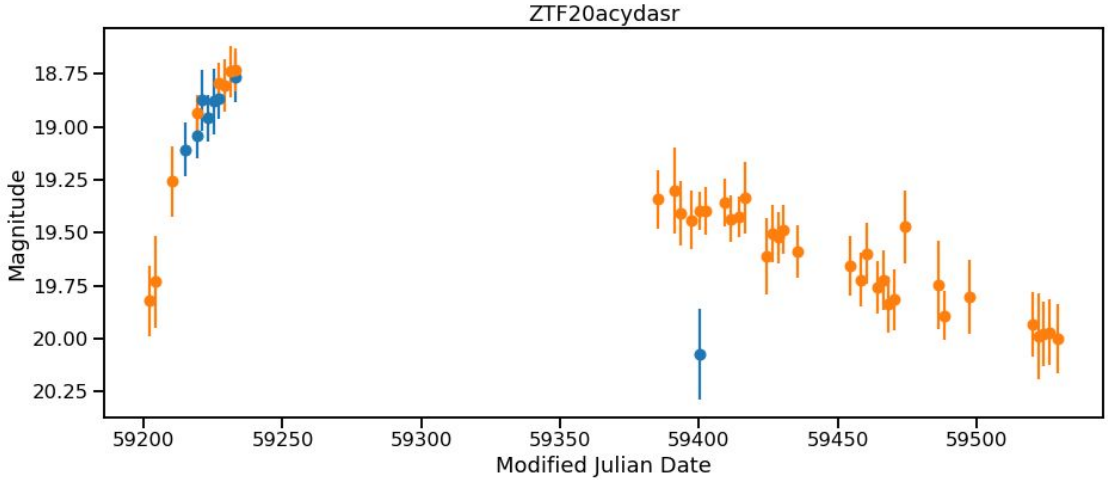


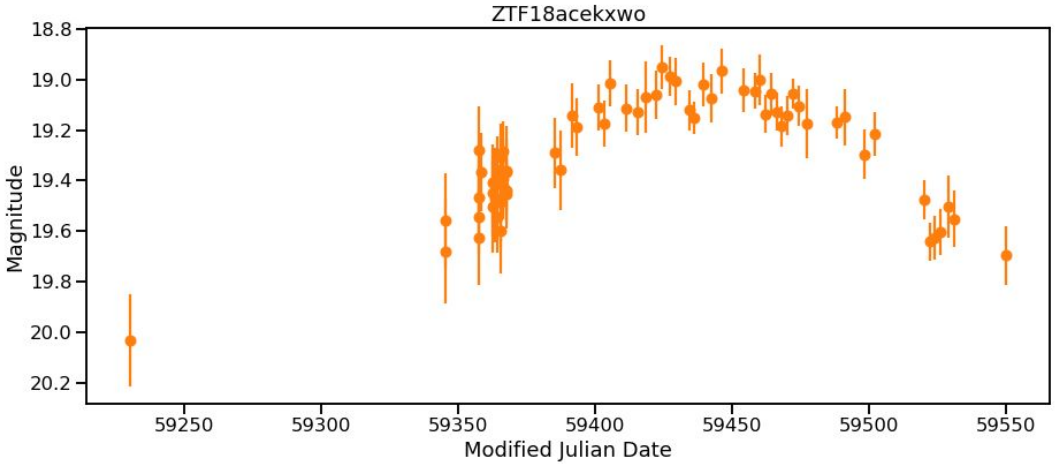
Without fitted parameters




ZTF Supernova alert fit using the "Bump" function
Bump function
Photometric data points in r band
Normalized flux
Time (days)
ZTF Supernova Ia alert fit using the "Bump" function
Bump function
Photometric data points in r band
Normalized flux
Time (days)
ZTF AGN alert fit using the "Bump" function
Bump function
Photometric data points in g band
Normalized flux
Time (days)
ZTF Variable star alert fit using the "Bump" function
Bump function
Photometric data points in g band
Normalized flux
Time (days)

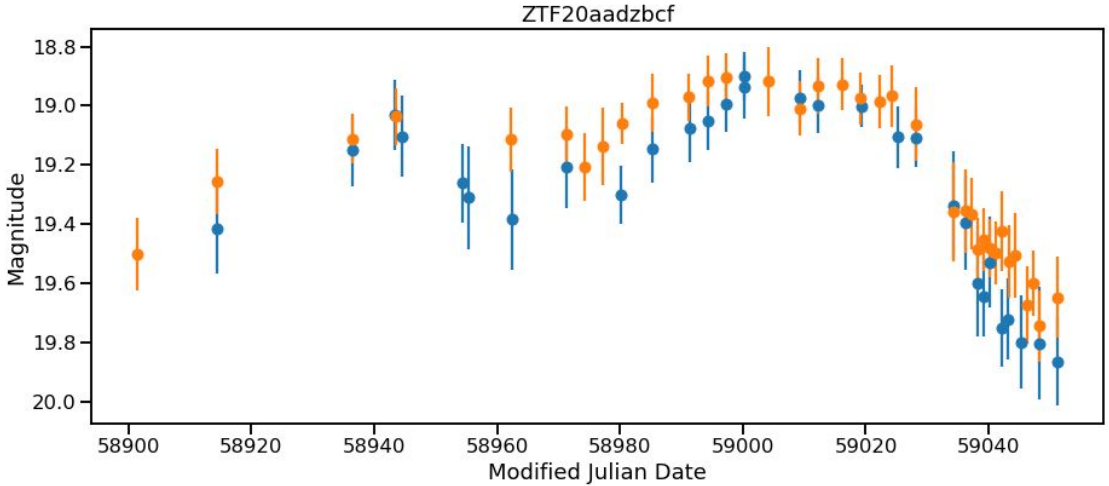

Parametric fit of u band
u data point
Parametric fit of g band
g data point
Parametric fit of r band
r data point
Parametric fit of i band
i data point
Parametric fit of z band
z data point
Parametric fit of Y band
Y data point
Flux
Time (Modified Julian Date)

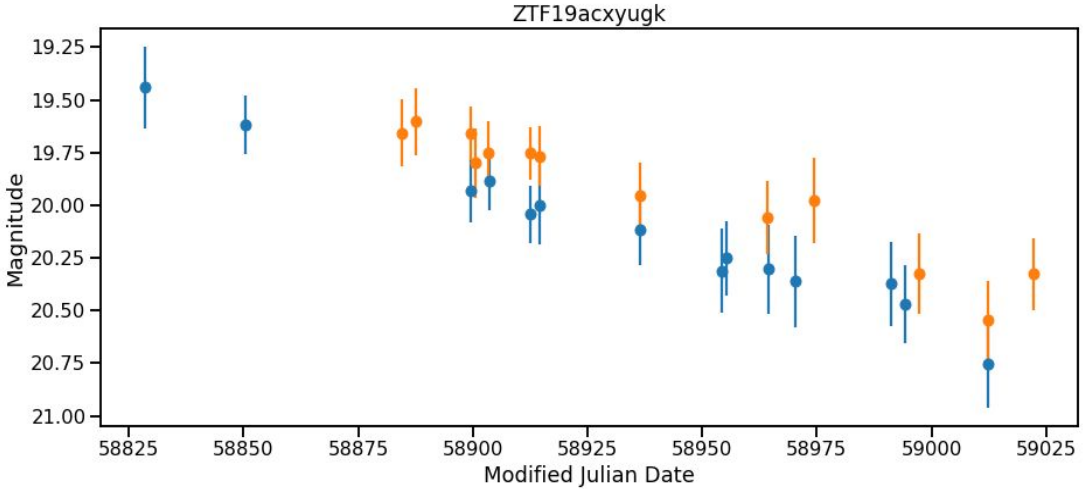

ZTF20acydasr
Magnitude
18.75
19.00
19.25
19.50
19.75
20.00
20.25
59200
59250
59300
59350
59400
59450
59500
Modified Julian Date


ZTF18acekxwo
Magnitude
18.8
19.0
19.2
19.4
19.6
19.8
20.0
20.2
59250
59300
59350
59400
59450
59500
59550
Modified Julian Date

g band

r band


ZTF20aadzbcf
Magnitude
18.8
19.0
19.2
19.4
19.6
19.8
20.0
58900
58920
58940
58960
58980
59000
59020
59040
Modified Julian Date


ZTF19acxyugk
Magnitude
19.25
19.50
19.75
20.00
20.25
20.50
20.75
21.00
58825
58850
58875
58900
58925
58950
58975
59000
59025
Modified Julian Date

g band

r band