Kilonova detection in Fink

by

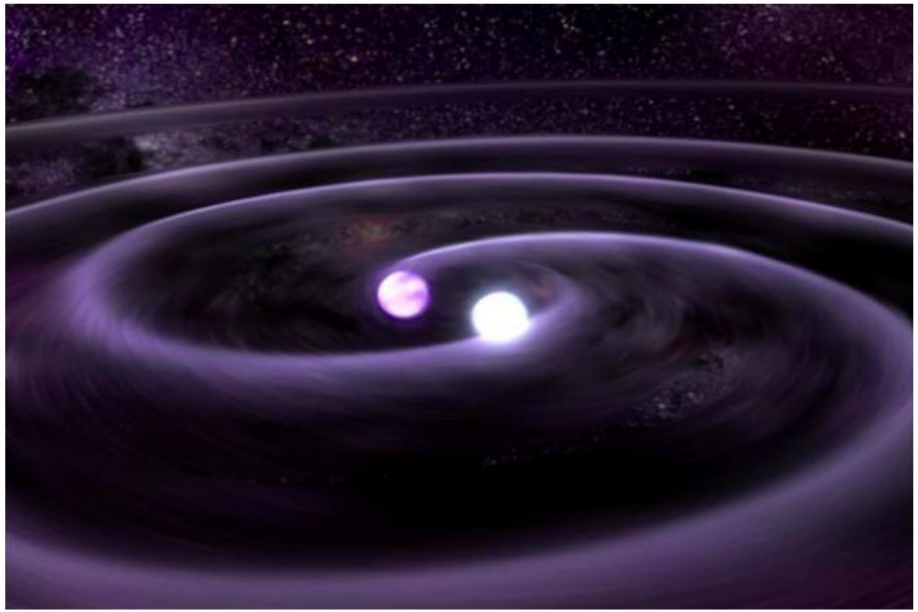
Biswajit Biswas

Fink Collaboration meeting (19th May, 2022)

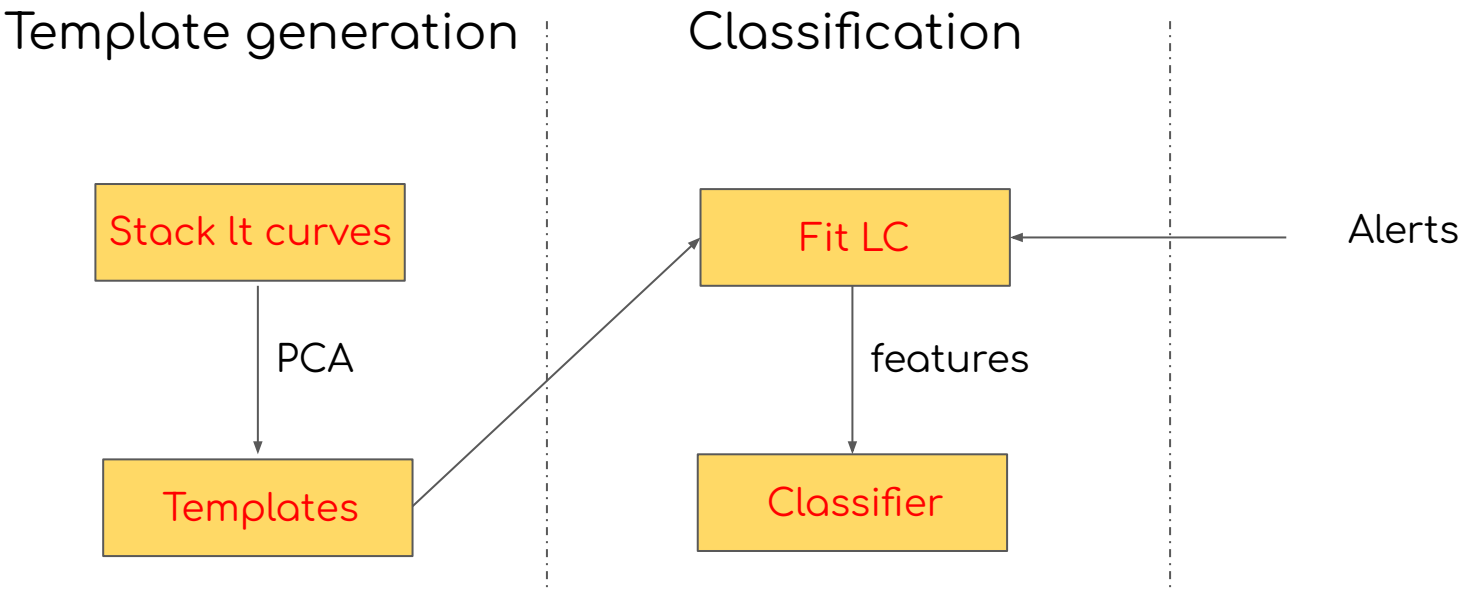
Special thanks to: Emille Ishida (LPC), Anais Moller(LPC), Julien Peloton(IJCLab)

Kilonova

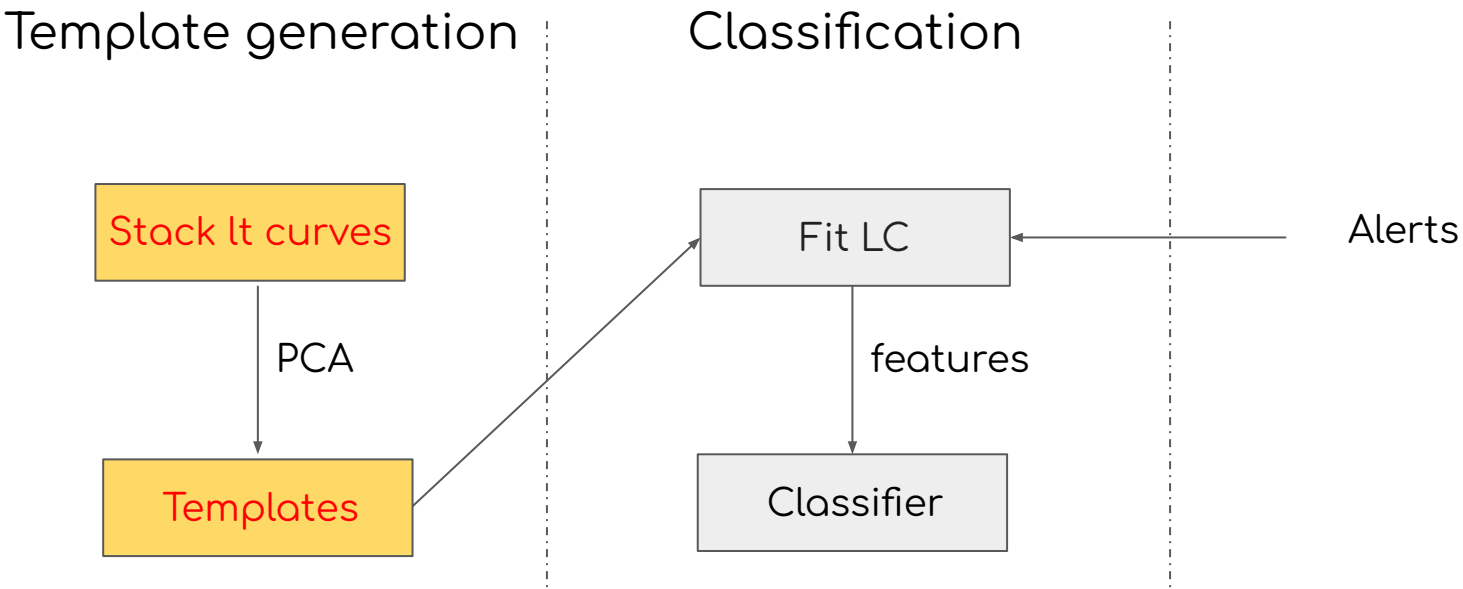
- Binary neutron star mergers
- Possible GW detection (GW170817)



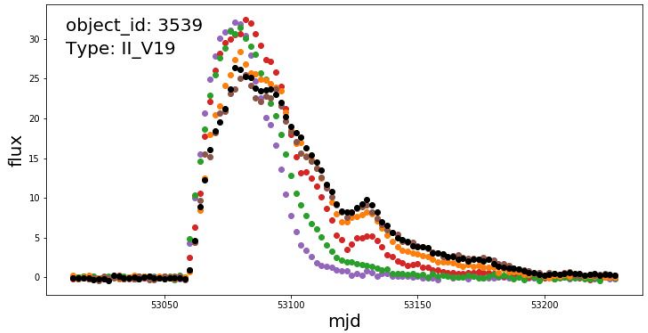
KN module for Fink

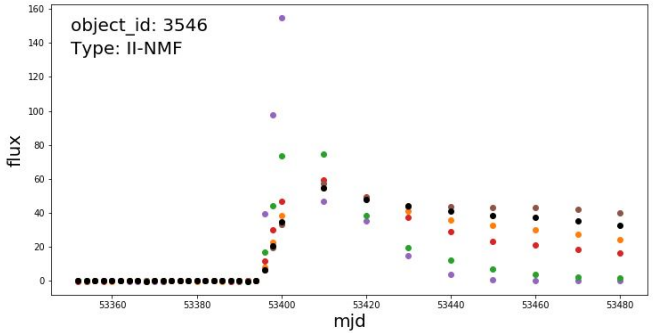


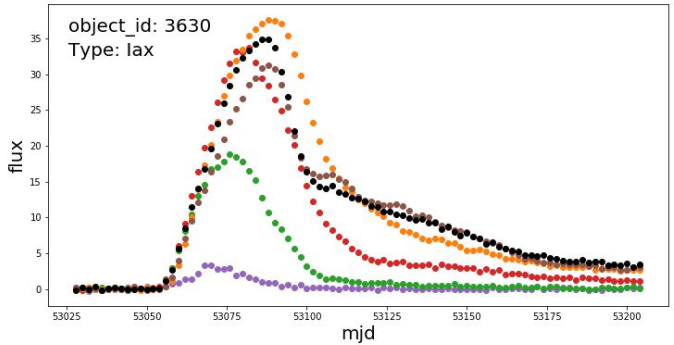
KN module for Fink



Generating PCs: Dataset (RESSPECT 'perfect' simulations)





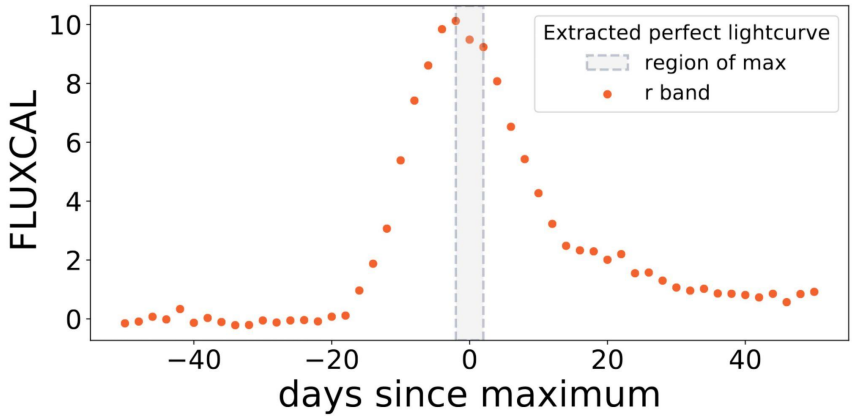


'Perfect' simulations:

- Different types of transient events
- Readings every 2 days
- No missing data points

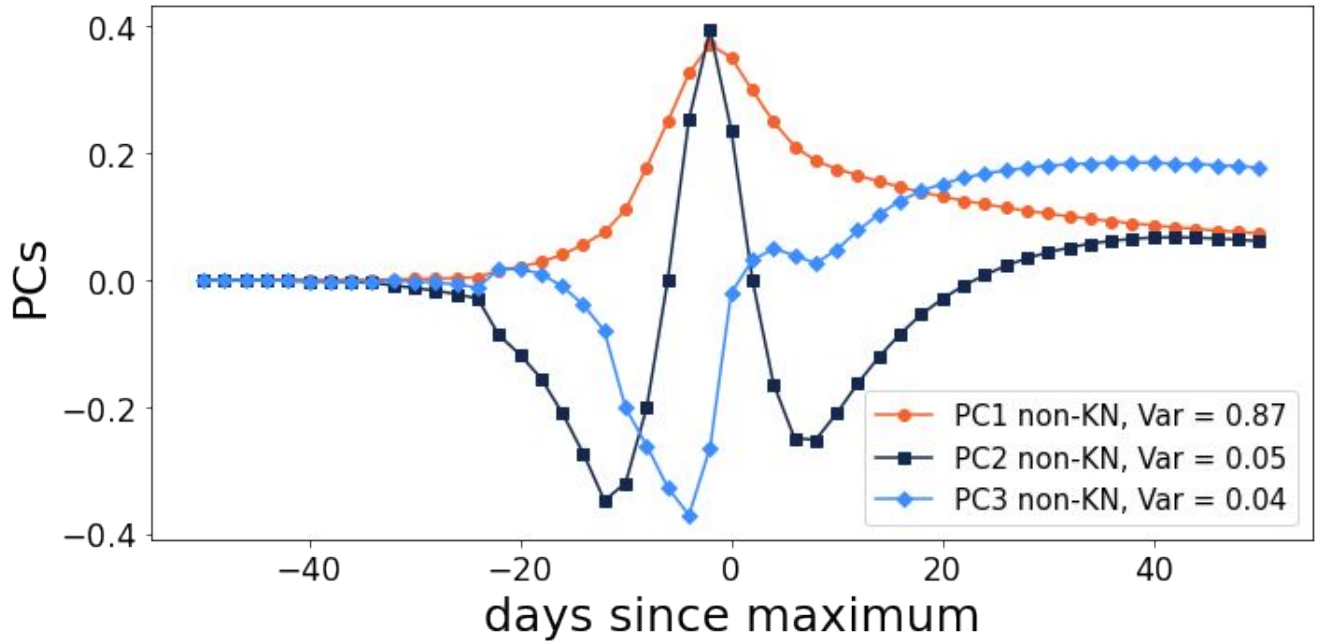
Generating PCs: Anchoring the data

- Events: 1000 KN events + 1000 non-KN events
- Filter only ‘g’ and ‘r’ bands of LSST
- Extract 100 days of data in such a way that Amplitude is placed at day 50.



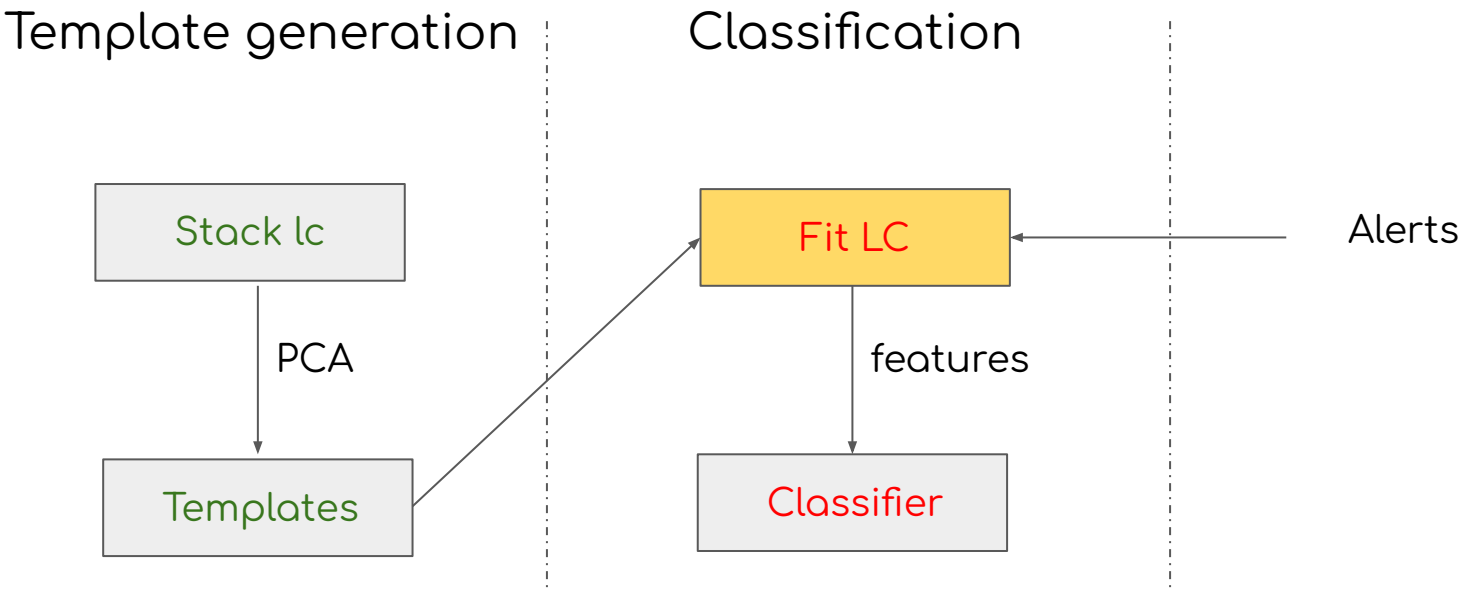
Dataset: RESSPECT sims

Generated templates



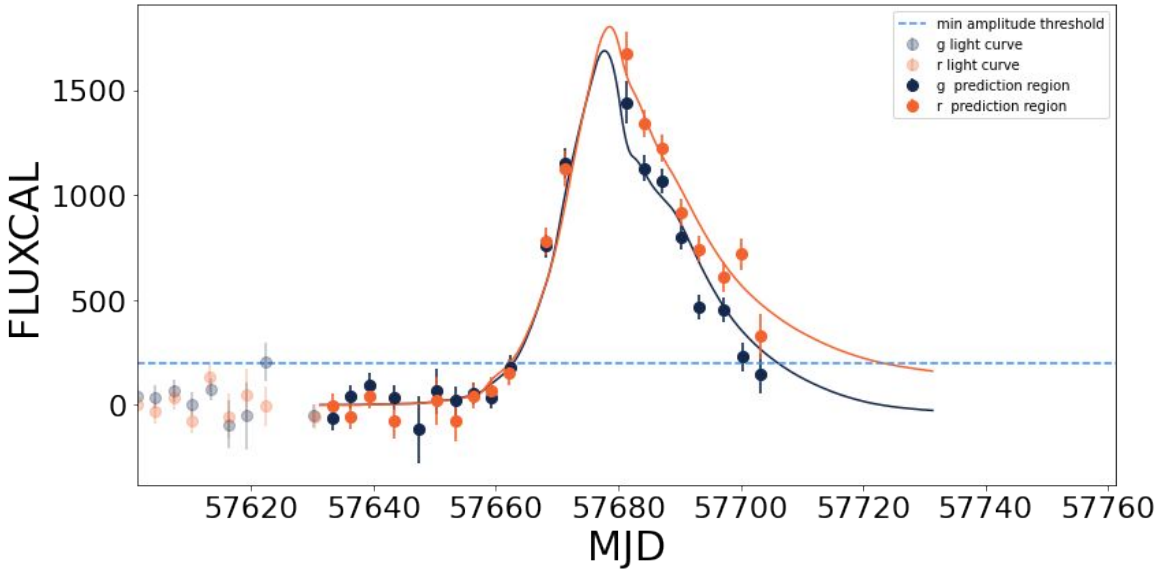
Next we will represent light curves as a linear combination of these 3 curves

Flowchart



Now we will represent light curves as a linear combination of the templates

Fitting the curve: ZTF Simulations



For each band:

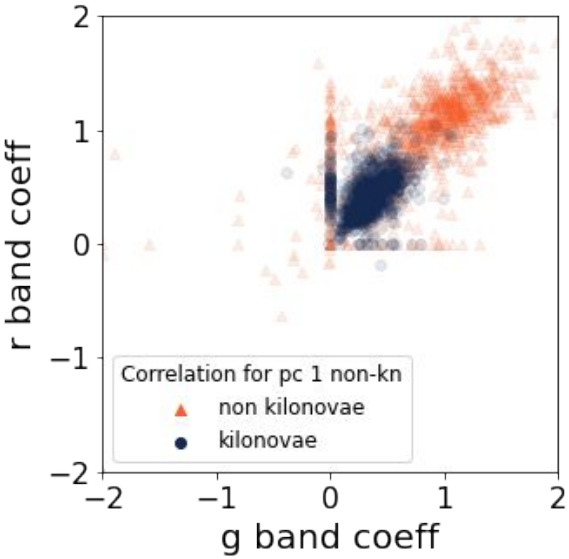
Make fit only if max flux is above a threshold (>200)

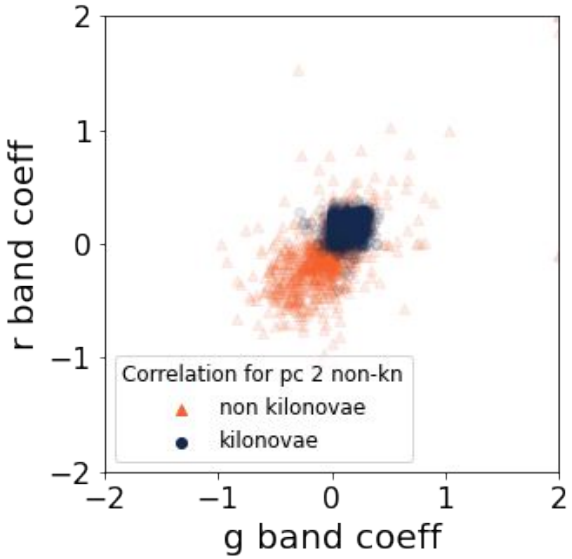
Anchor the data with day 50 of the prediction region as highest flux.

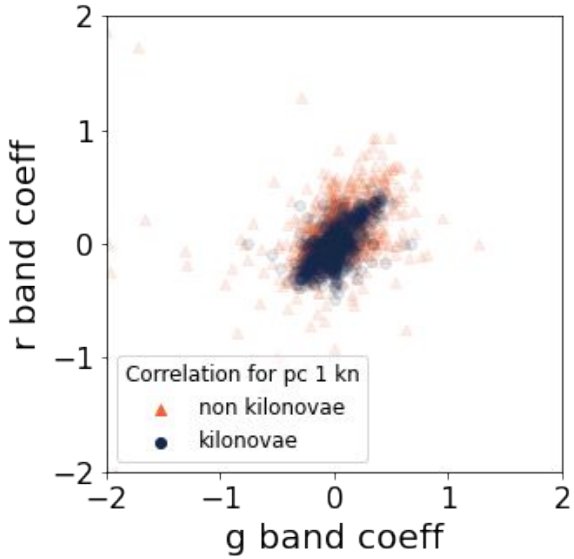
Optimize values of PC coefficients to obtain the fit.

$$loss = \sum_{i}^{N} \frac{(l_{p,i} - l_i)^2}{\sigma_i^2} + \left[\sum_{k=1}^{3} c_k^2 - c_1^2 H(c_1)\right] \frac{f_{\max}^2}{\sigma_{f_{\max}}^2}$$

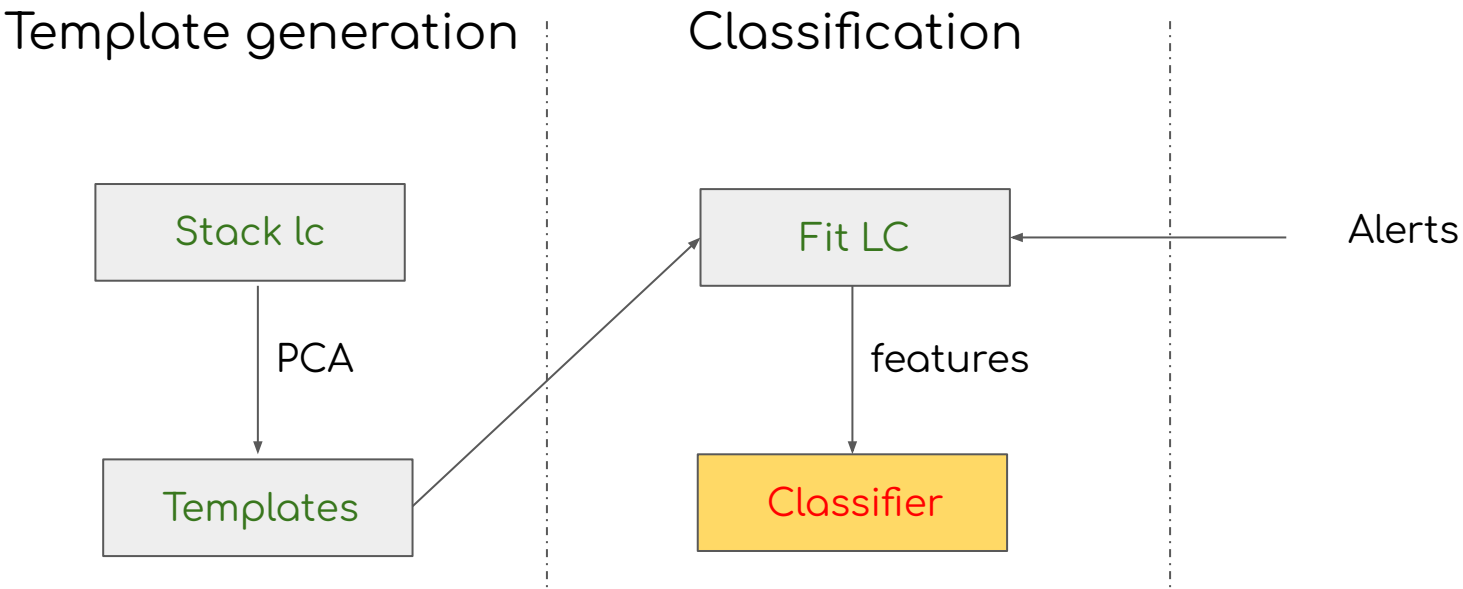
Correlation plots







Flowchart



Now we will represent light curves as a linear combination of the templates

Classification (Random Forest)

Features: fit coefficients, max flux, fit residual

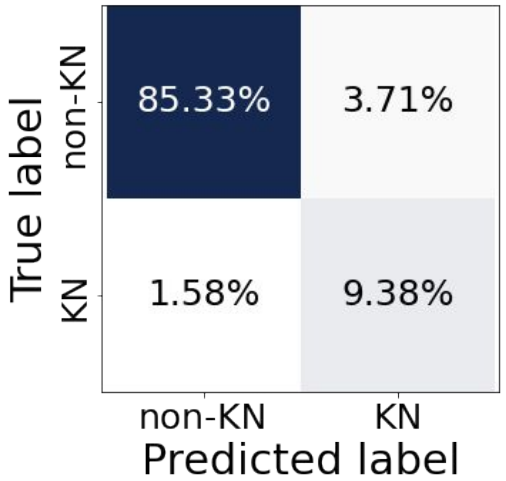
Train dataset:

total number of events: 21205
total number of KN: 3210
total number of non KN: 17995

Test statistics:

total number of events: 20198
total number of KN: 2213
total number of non KN: 17985

results

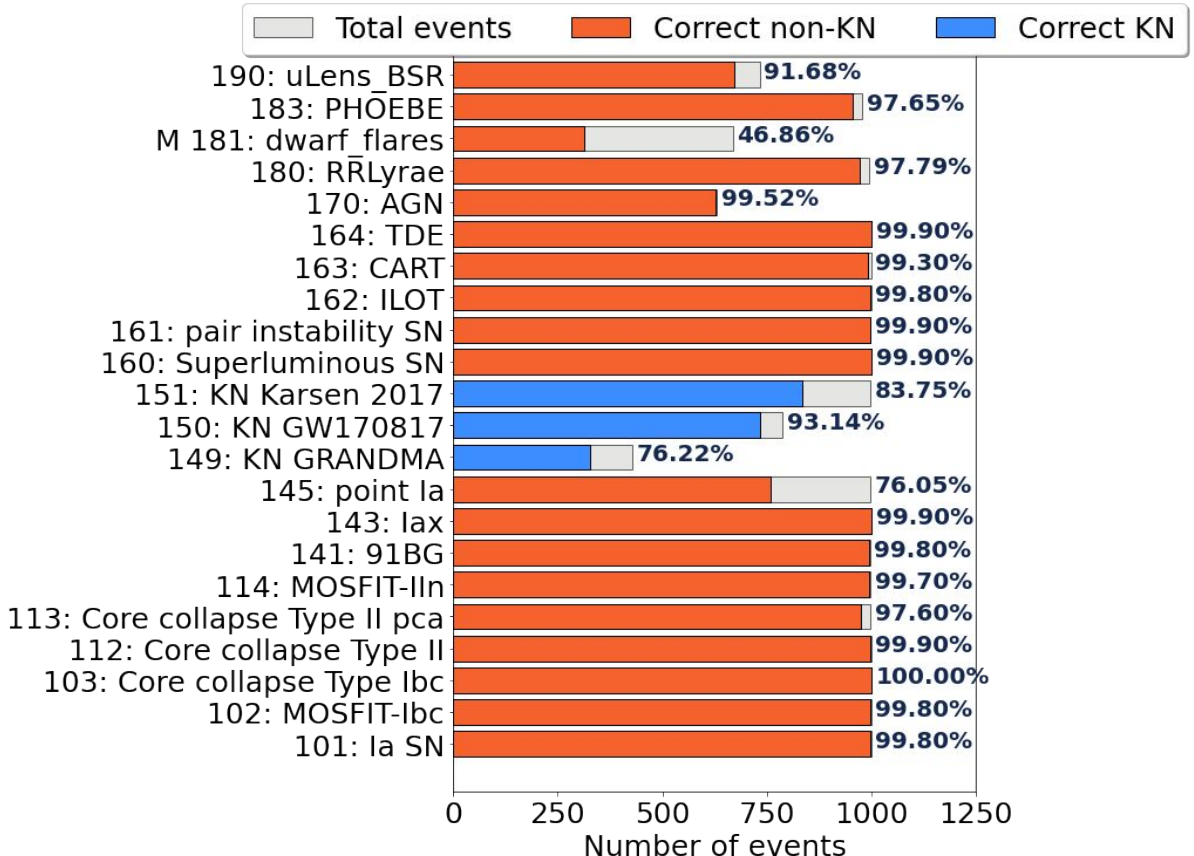


True Positive : 1894
False Positive : 750
True negative : 17235
False negative : 319

Classification (Results)

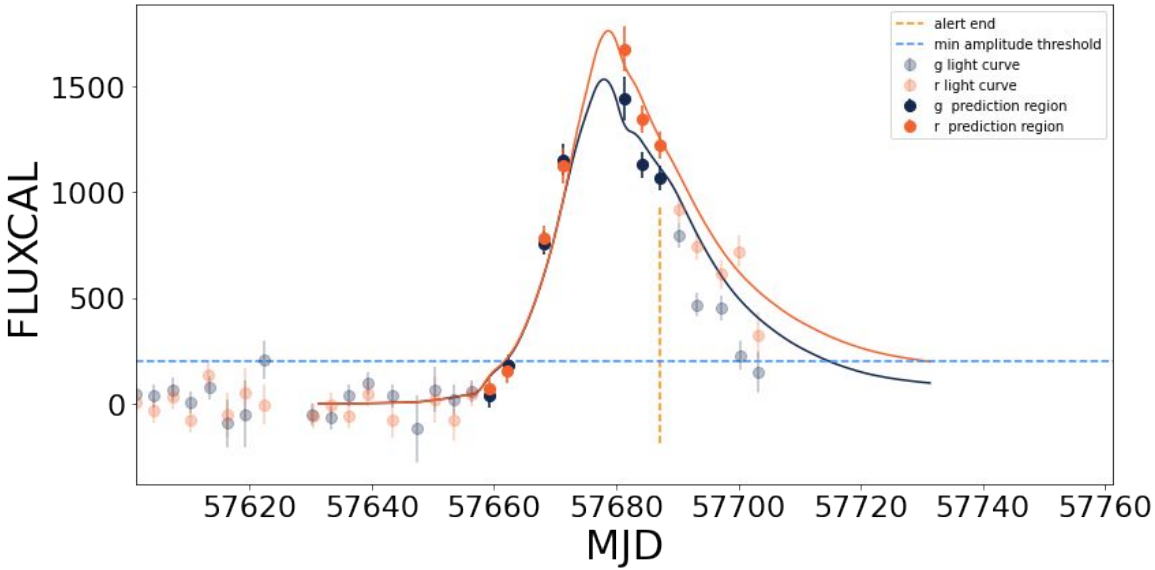
Test Dataset

total number of events: 20198
total number of KN: 2213
total number of non KN: 17985



What about alerts?

Mimic alerts



Pick a point with flux>200

Use only 30 days of data before this date

Classification (Random Forest)

Features: fit coefficients, max flux, fit residual

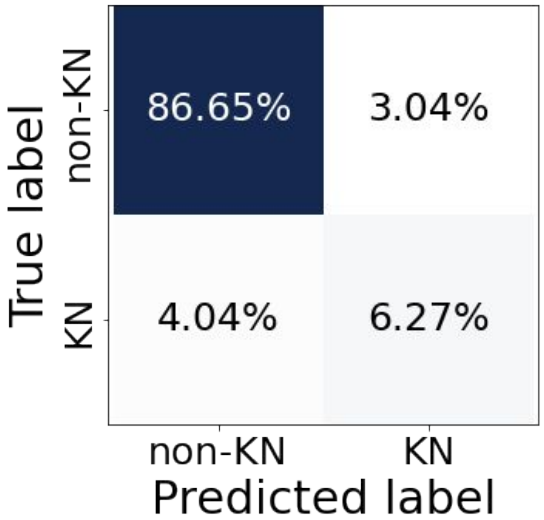
Train dataset:

total number of events: 20463
total number of KN: 2986
total number of non KN: 17477

Test statistics:

total number of events: 19450
total number of KN: 2005
total number of non KN: 17445

results

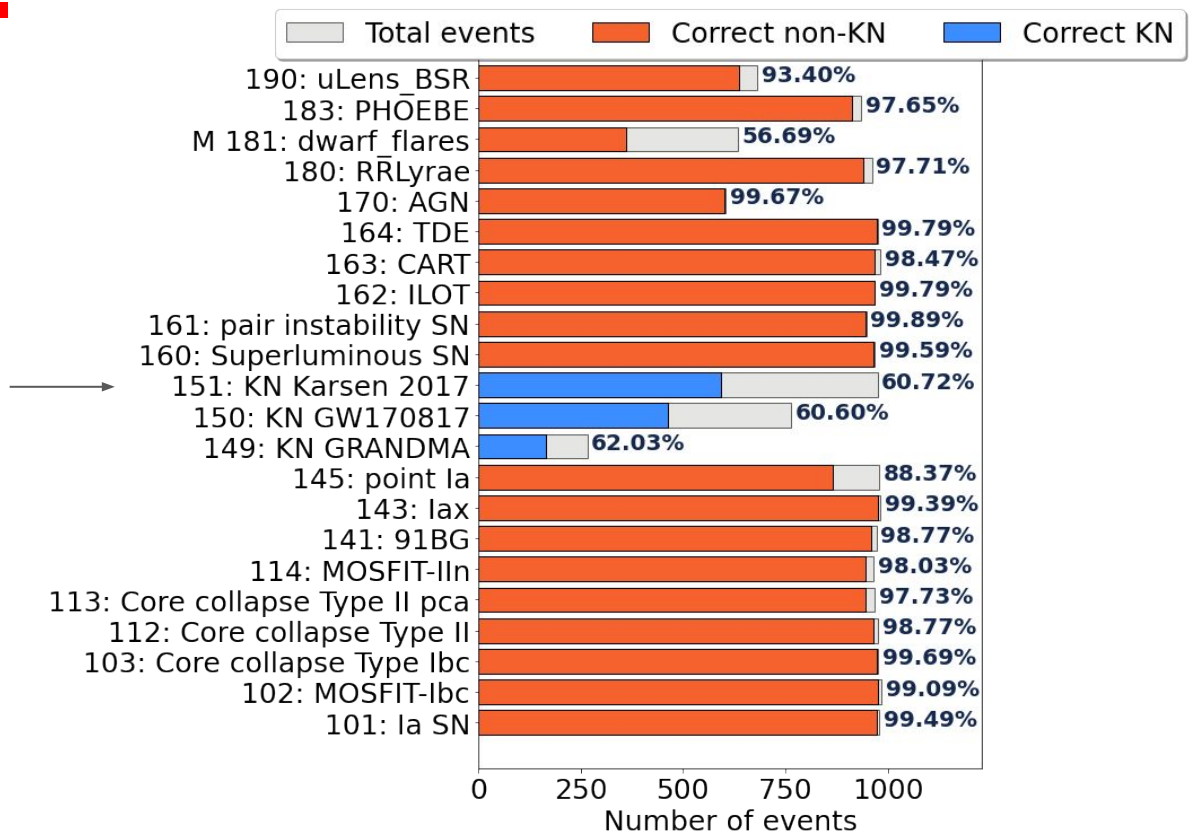


True Positive : 1220
False Positive : 592
True negative : 16853
False negative : 785

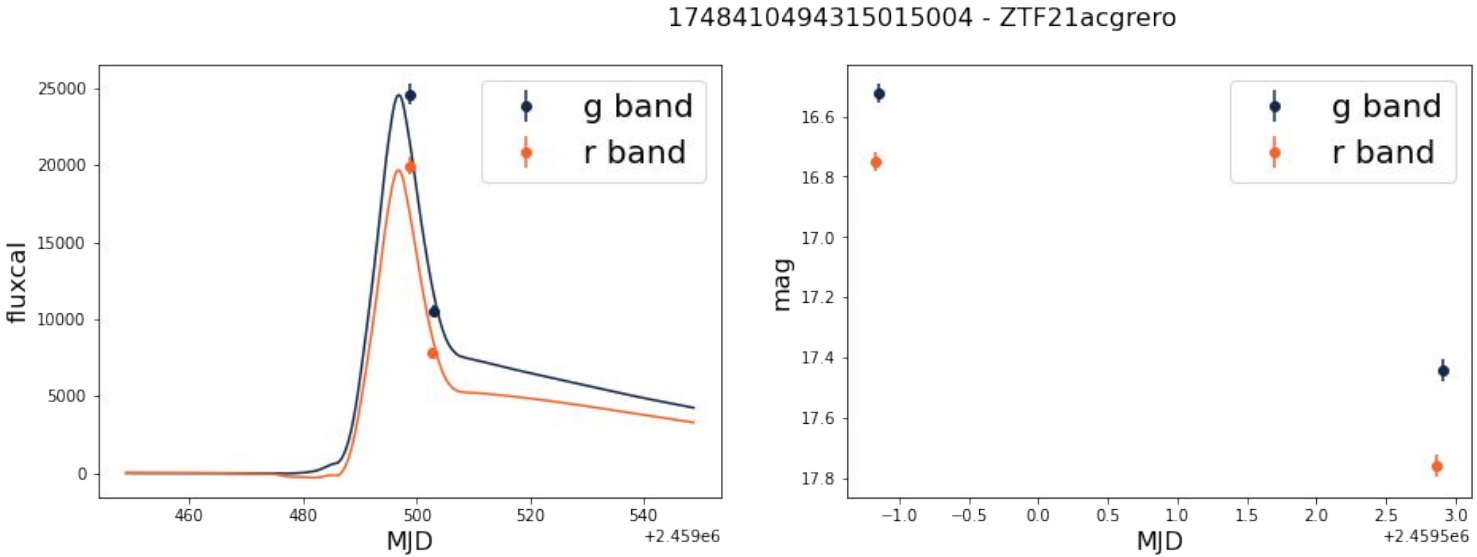
Classification results

Test dataset:

total number of events: 19450
total number of KN: 2005
total number of non KN: 17445



Example of classified KN:



Score = 0.56

Summary and lessons learnt

- Requirement:
  At least 2 points in a band with max_flux > 200
- Impurities mostly come from other transients
- Initial results on a smaller set of alerts are very encouraging!
  Time for (another) test run?
- **Common code base for training testing and deployment.**
- Code available on git: https://github.com/b-biswas/kndetect

Thank you!