Difference Image Analysis and Machine Learning

Transient detection techniques

Bruno Sánchez - Dan Scolnic
Duke U. - Durham NC

Large Synoptic Survey Telescope

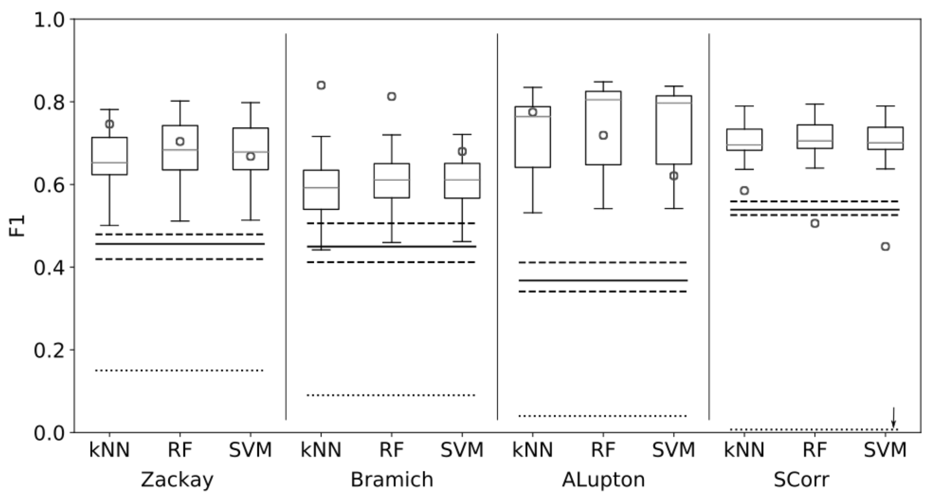
Hello!

Who am I?

- Astronomer (MS degree) 2014 @ UNC - Arg.
- PhD in April 2019 @ UNC - Arg. (advised by Dr. M. Domínguez)
- Started postdoc in June 2019 @ Duke (supervised by Dr. D. Scolnic)

Background in...?

- Difference Image Analysis (DIA) techniques
- Involved in GW counterpart search
- Combined DIA with Machine Learning to clean candidate samples



Difference Image Analysis

Useful as a direct comparison technique between images
This is needed to find *transient variability events*

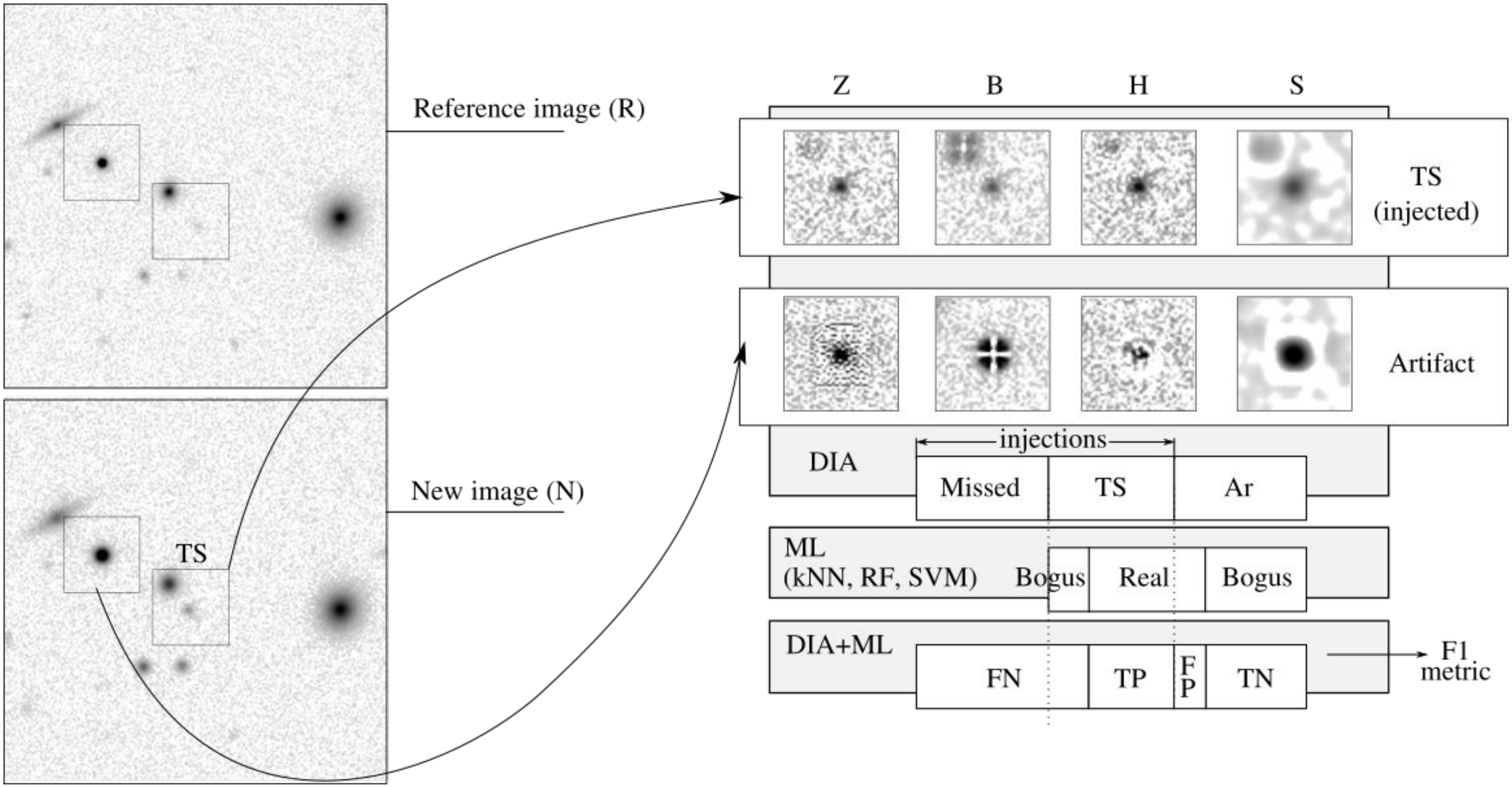
R: reference image (also template, coadd, etc.)
N: new image (science)
Ker: convolution kernel

$$D(x,y) = N(x,y) - R(x,y) * Ker(u,v)$$

There are several subtract techniques:

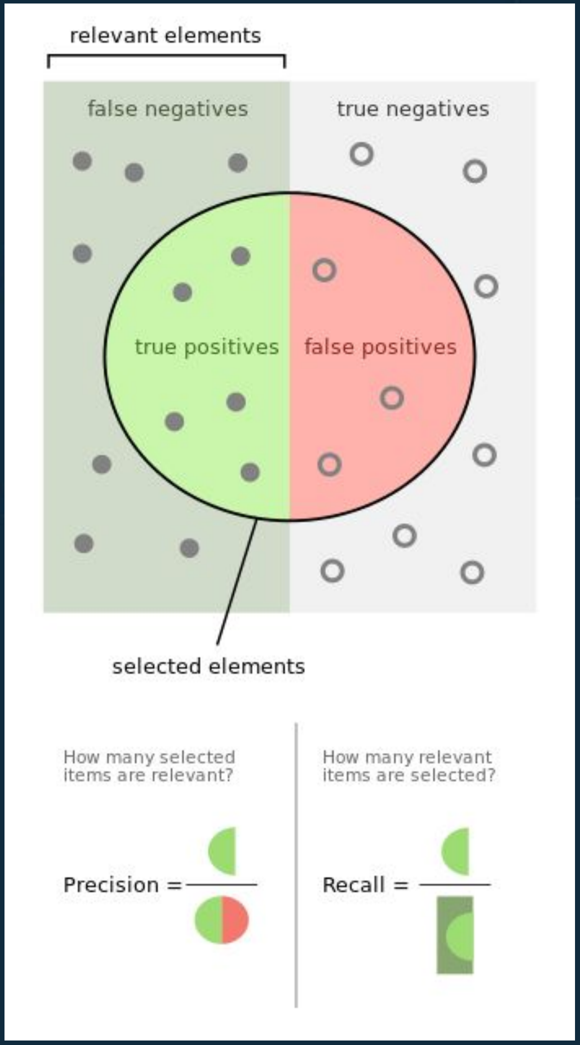
- Alard & Lupton (1998) $\longrightarrow$ $$Ker(u,v) = \sum_i k_i B_i(u,v)$$
- Bramich (2008) $\longrightarrow$ $$Ker(u,v) = \sum_i k_i \delta(u-u_i, v-v_i)$$
- Zackay et al. (2016) $\longrightarrow$ $$\widehat{D} = \frac{z_R\ \widehat{P_R}\ \widehat{N} - z_N\ \widehat{P_N}\ \widehat{R}}{\sqrt{\sigma_N^2 z_R^2 |\widehat{P_R}|^2 + \sigma_R^2 z_N^2 |\widehat{P_N}|^2}}$$
  -also *ZOGY*

Estimate performances...



The metrics used

Using the predictive errors from ML



Can be calculated before using ML or after, and see if it improved...

Before ML:

True Positive: detected true *transient source* (TS)
False Positive: detection produced by bad subtraction *artifact* (Ar)
False Negative: a lost or *missed* transient object

After ML:

True Positive: TS classified correctly
False Positive: *artifact* erroneously classified as a TS
False Negative: a TS discarded by mistake
True Negative: an artifact correctly discarded

Important derived metric:

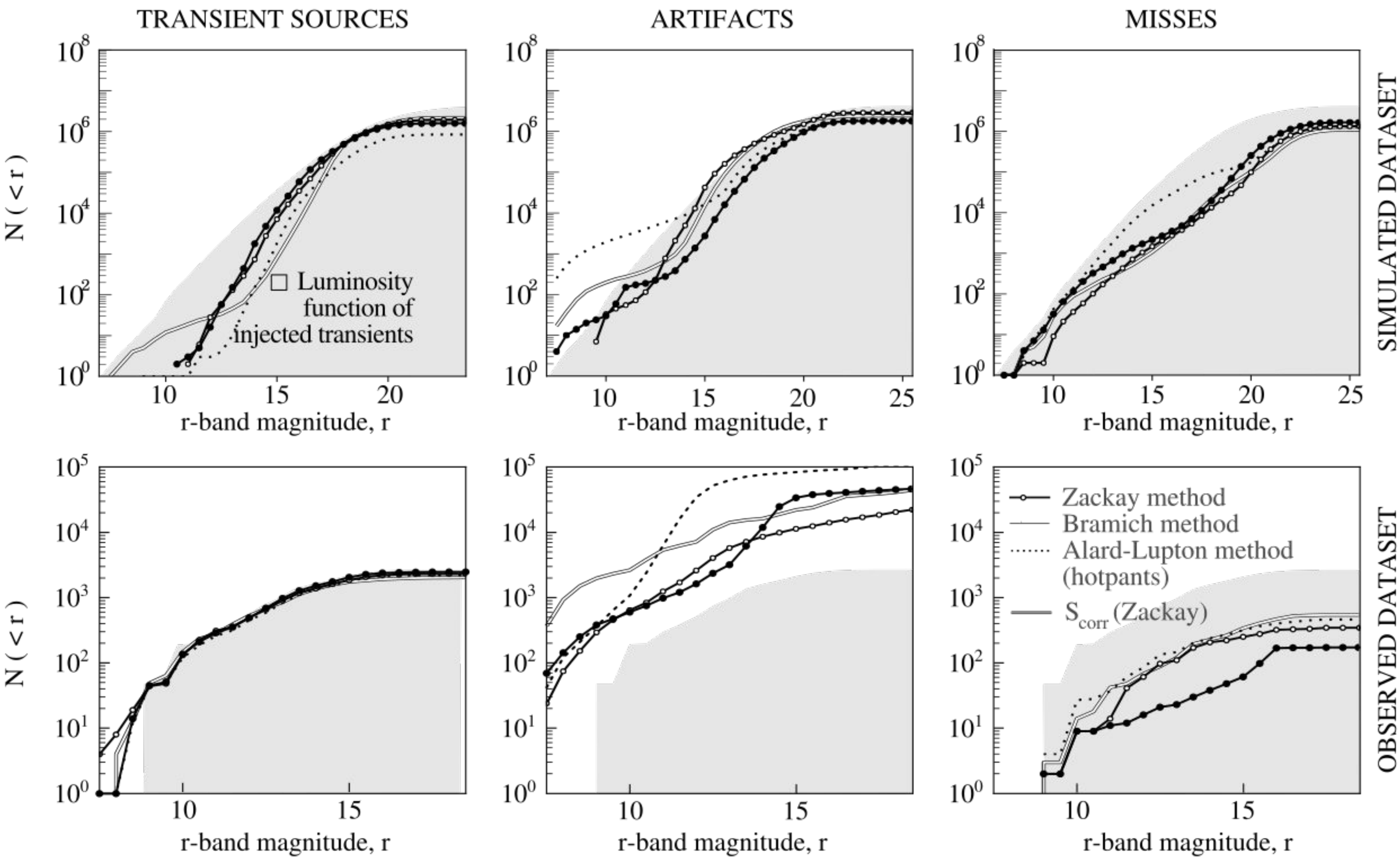
$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Parameters of the simulations for this study

			TOROS instruments		
PARAMETER	UNITS	VALUES	EABA	TOROS	TORITOS
aperture of the telescope	[m]	[0.4, 0.6, 1.54]	1.54	0.6	0.4
reference seeing FWHM	[arcsec]	[0.8, 1, 1.3]	1.3	0.8	0.8
new image seeing FWHM	[arcsec]	[1.3, 1.9, 2.5]	2.5	1.0	1.0
plate scale	[arcsec/pix]	[0.3, 0.7, 1.4]			
exposure times	[sec]	[60, 120, 300]			
stellar density	[stars per sq deg]	[4e3, 8e3, 32e3, 64e3, 128e3, 256e3]			
stellar luminosity distribution exponent	[dexp per mag]	[0.1, 0.5, 0.9]			
background brightness (R)	[mag per arcsec ^{2}]	[20, 21, 22]			
background brightness (N)	[mag per arcsec ^{2}]	[18, 19, 20]			
relative brightness from host	r-band magnitudes	sampled from Unif(-4,1)			
angular distance from host	half light radius	sampled from Unif(0, 5)			

Table 1: Parameter space for simulated images to be explored for transient detection.

Photometric results (before ML)



After using ML: F1 measure

N_{FWHM}	t_{exp}[s]	D [cm]	Zackay kNN	Zackay RF	Zackay SVM	Bramich kNN	Bramich RF	Bramich SVM	A&Lupton kNN	A&Lupton RF	A&Lupton SVM	S_{corr} kNN	S_{corr} RF	S_{corr} SVM
N_{FWHM}=1.3	60	400	629	627	629	587	593	593	634	636	638	731	742	738
		600	637	643	643	599	605	606	648	660	660	734	743	737
		1540	655	658	658	614	623	624	679	681	681	739	746	741
	120	400	715	738	734	640	646	645	818	843	832	732	737	733
		600	734	745	742	653	668	665	835	844	832	737	758	750
		1540	738	756	752	667	683	681	817	848	838	742	754	748
	300	400	754	796	791	683	705	703	760	805	791	774	787	782
		600	764	802	798	703	719	718	770	804	794	782	792	787
		1540	781	802	798	716	720	721	777	818	805	789	795	790
N_{FWHM}=1.9	60	400	564	577	577	511	527	528	583	594	595	688	684	683
		600	577	589	590	526	539	541	606	611	612	690	694	691
		1540	589	594	595	534	557	558	616	626	627	699	704	701
	120	400	663	695	691	592	605	604	814	839	827	686	702	697
		600	671	704	699	599	617	617	807	831	821	692	706	701
		1540	683	712	707	614	637	636	797	845	833	696	715	709
	300	400	713	745	738	622	632	636	764	809	799	720	737	733
		600	706	740	735	629	655	657	771	812	801	728	736	733
		1540	716	755	749	650	678	678	772	823	810	734	745	739
N_{FWHM}=2.5	60	400	501	512	513	442	459	462	532	541	542	636	639	638
		600	509	528	528	450	462	466	555	558	559	645	647	645
		1540	523	534	534	463	472	476	566	576	576	654	663	660
	120	400	619	648	644	540	561	561	793	820	808	656	664	660
		600	628	656	651	539	574	572	784	812	799	667	665	661
		1540	639	663	657	552	580	580	780	828	819	673	673	668
	300	400	649	684	678	555	600	603	756	804	797	680	696	693
		600	650	683	679	574	616	615	743	802	791	688	690	687
		1540	652	690	686	580	611	611	742	793	784	691	694	691

- Groups divided by simulation parameters
- Measure the F1 value *(here we display F1x10^3)*
-
- Groups by:
    - Seeing FWHM (1.3, 1.9, 2.5)
    - Exposure time (1 min, 2 min, 5min)
    - Mirror size (40mm, 60 mm, 1540mm)
- Columns are DIA+ML combination
    - DIA methods: ZOGY D , Bramich, A&L, ZOGY S
    - ML methods: kNN, RF, SVM
- These are final scores, after ML implementation
- It would reflect the final F1 measure of the survey, only calculable *given that we know the population of true transients*

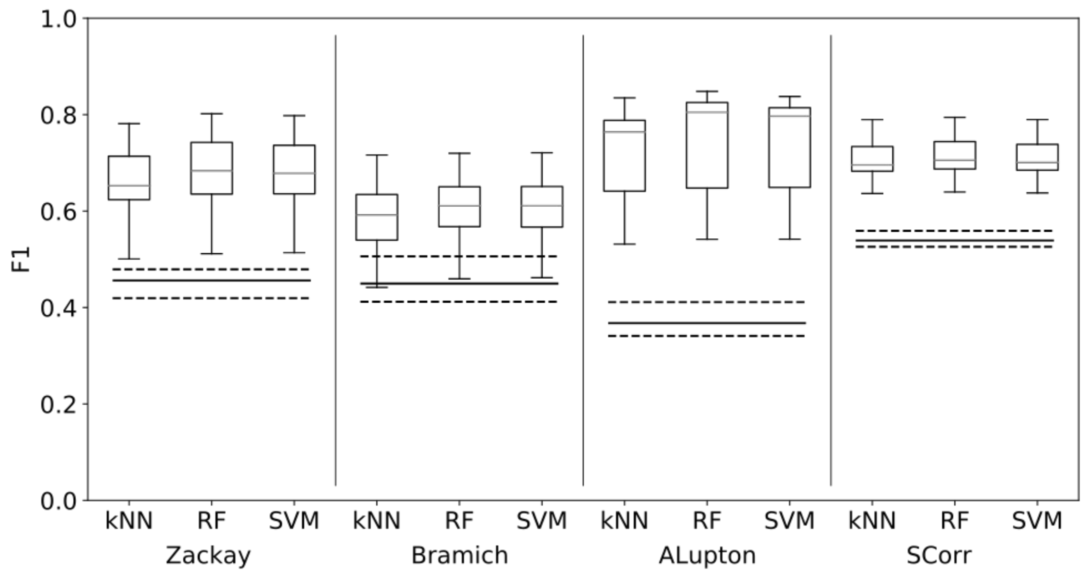
This analysis is applicable to DC2.

We would not have the simulation parameters but we can estimate this grouping by observing conditions and transient true information

How ML changes the picture? F1 measure

	t_{exp}[s]	D [cm]	Zackay kNN	Zackay RF	Zackay SVM	Bramich kNN	Bramich RF	Bramich SVM	A&Lupton kNN	A&Lupton RF	A&Lupton SVM	S_{corr} kNN	S_{corr} RF	S_{corr} SVM
N_{FWHM}=1.3	60	400	629	627	629	587	593	593	634	636	638	731	742	738
		600	637	643	643	599	605	606	648	660	660	734	743	737
		1540	655	658	658	614	623	624	679	681	681	739	746	741
	120	400	715	738	734	640	646	645	818	843	832	732	737	733
		600	734	745	742	653	668	665	835	844	832	737	758	750
		1540	738	756	752	667	683	681	817	848	838	742	754	748
	300	400	754	796	791	683	705	703	760	805	791	774	787	782
		600	764	802	798	703	719	718	770	804	794	782	792	787
		1540	781	802	798	716	720	721	777	818	805	789	795	790
N_{FWHM}=1.9	60	400	564	577	577	511	527	528	583	594	595	688	684	683
		600	577	589	590	526	539	541	606	611	612	690	694	691
		1540	589	594	595	534	557	558	616	626	627	699	704	701
	120	400	663	695	691	592	605	604	814	839	827	686	702	697
		600	671	704	699	599	617	617	807	831	821	692	706	701
		1540	683	712	707	614	637	636	797	845	833	696	715	709
	300	400	713	745	738	622	632	636	764	809	799	720	737	733
		600	706	740	735	629	655	657	771	812	801	728	736	733
		1540	716	755	749	650	678	678	772	823	810	734	745	739
N_{FWHM}=2.5	60	400	501	512	513	442	459	462	532	541	542	636	639	638
		600	509	528	528	450	462	466	555	558	559	645	647	645
		1540	523	534	534	463	472	476	566	576	576	654	663	660
	120	400	619	648	644	540	561	561	793	820	808	656	664	660
		600	628	656	651	539	574	572	784	812	799	667	665	661
		1540	639	663	657	552	580	580	780	828	819	673	673	668
	300	400	649	684	678	555	600	603	756	804	797	680	696	693
		600	650	683	679	574	616	615	743	802	791	688	690	687
		1540	652	690	686	580	611	611	742	793	784	691	694	691

Marginalizing over the parameters (columns)



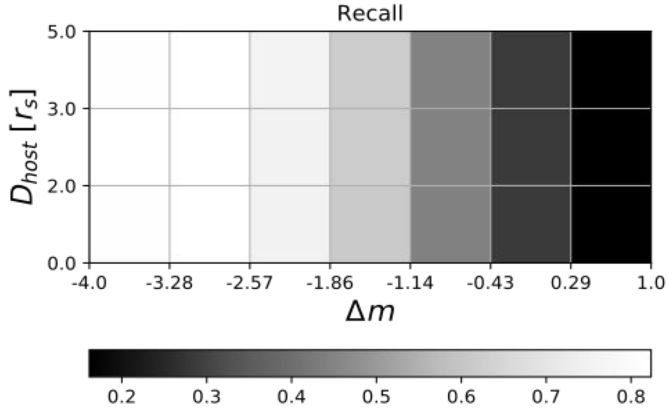
DIA only F1 quartiles as lines -- ***DIA+ML*** F1 quartiles using boxes

Takeaway:

- The ML helps DIA methods to clean samples.
- If DIA results show high contamination this provides a better training set
- The final F1 shows **A&Lupton + RandomForest** as the best contender
  - More complete and clean final sample
  - Real objects easier to distinguish from artifacts

Dependency on transient and image parameters

We can investigate transient and host properties impact on Recall

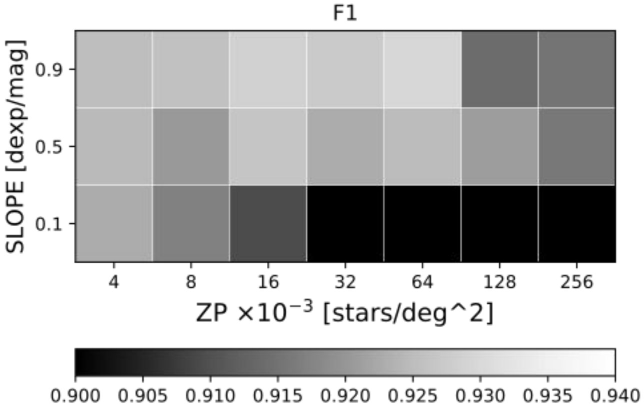


Distance to host center (D) in galaxy scale radius units vs the magnitude difference (Δm)

We find that distance to galaxy host center does not have a big impact on Recall as the magnitude difference does. It's easier to spot if we have larger brighness contrast.

This should be further investigated though, using realistic galaxy brightness profiles

Observed star field parameters also affect the measured F1 metrics



The SLOPE shows the log slope of Star Luminosity Function of the observed field.
ZP is the total stellar density of the observed field.

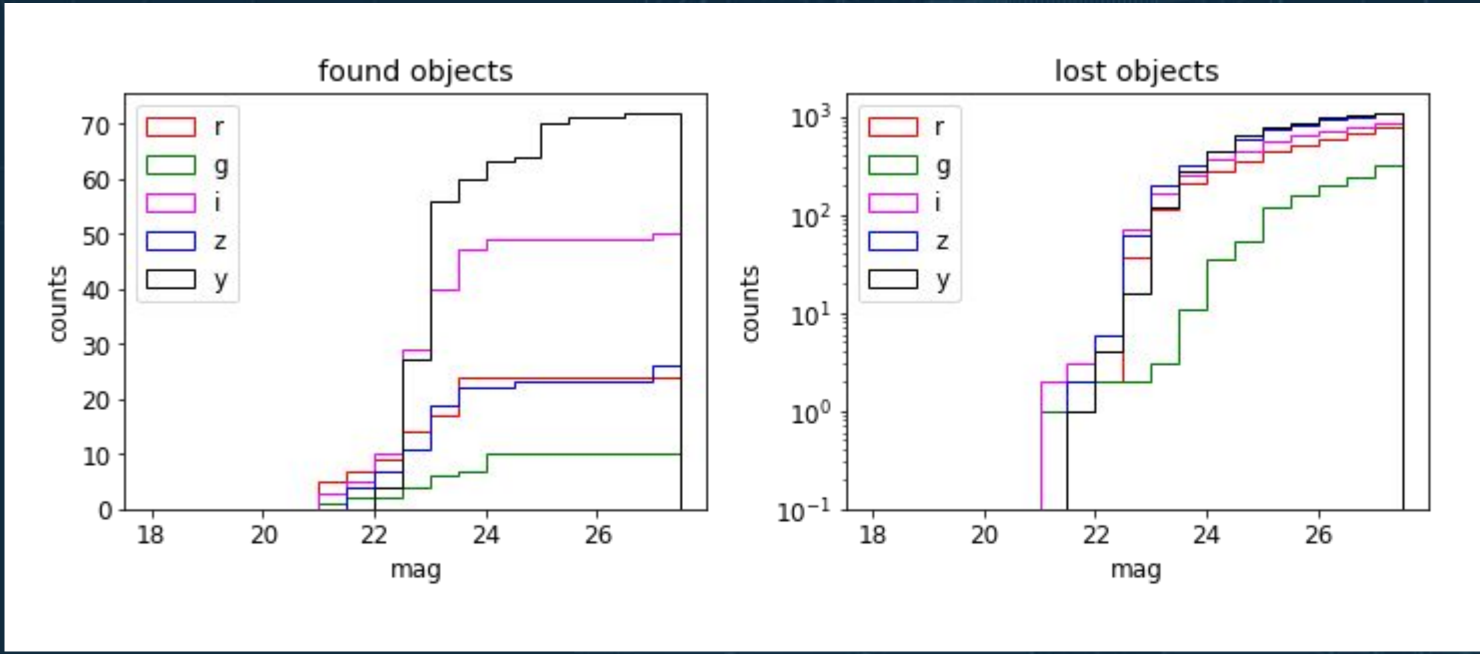
This suggests that if we have more and brighter stars we lose performance.
The more extreme values of SLOPE and ZP correspond to galactic latitudes up to 5° and limiting magnitude of r~19

Would be interesting to see this for LSST

Recent work using DC2

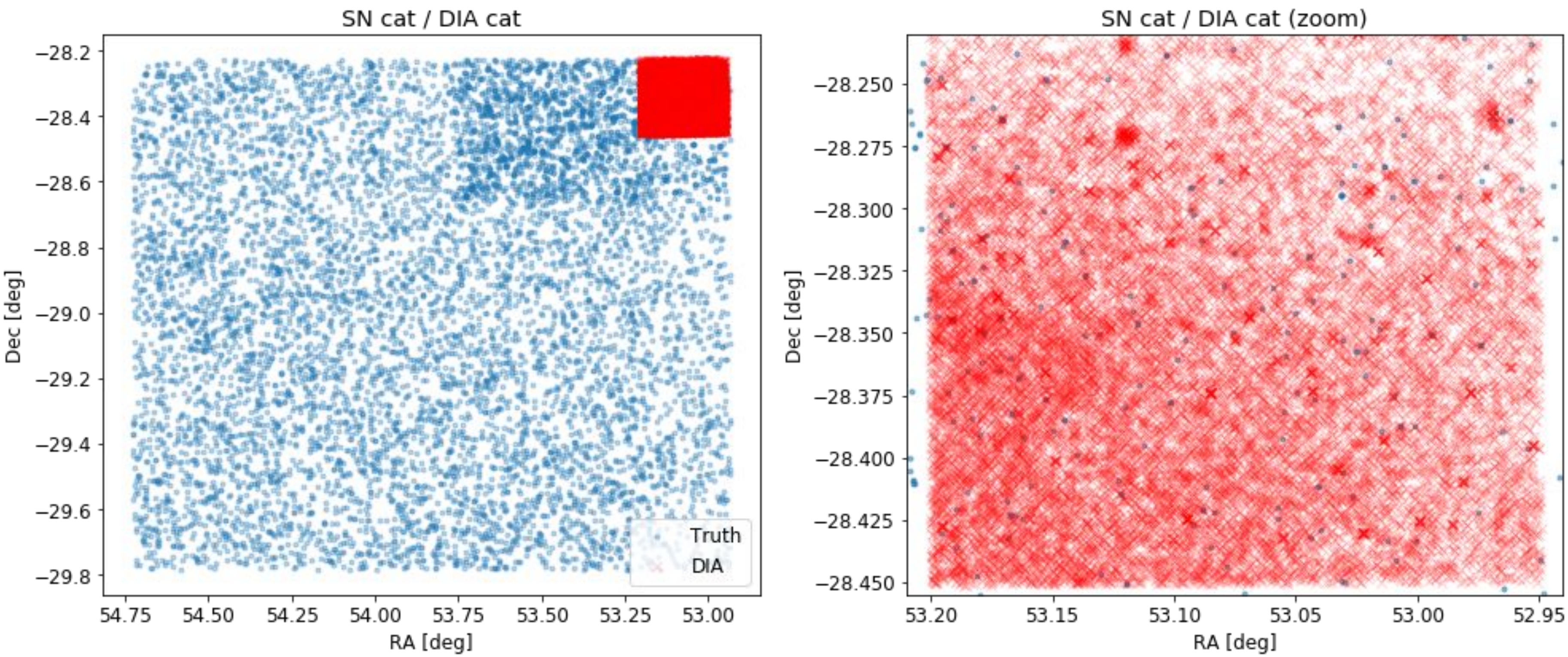
Using the DIA catalogs and truth variable catalogs attempted to measure the metrics. Used a coordinate and time matching strategy

Found really *poor performance*, finding less than a tenth of the true transients present in sample



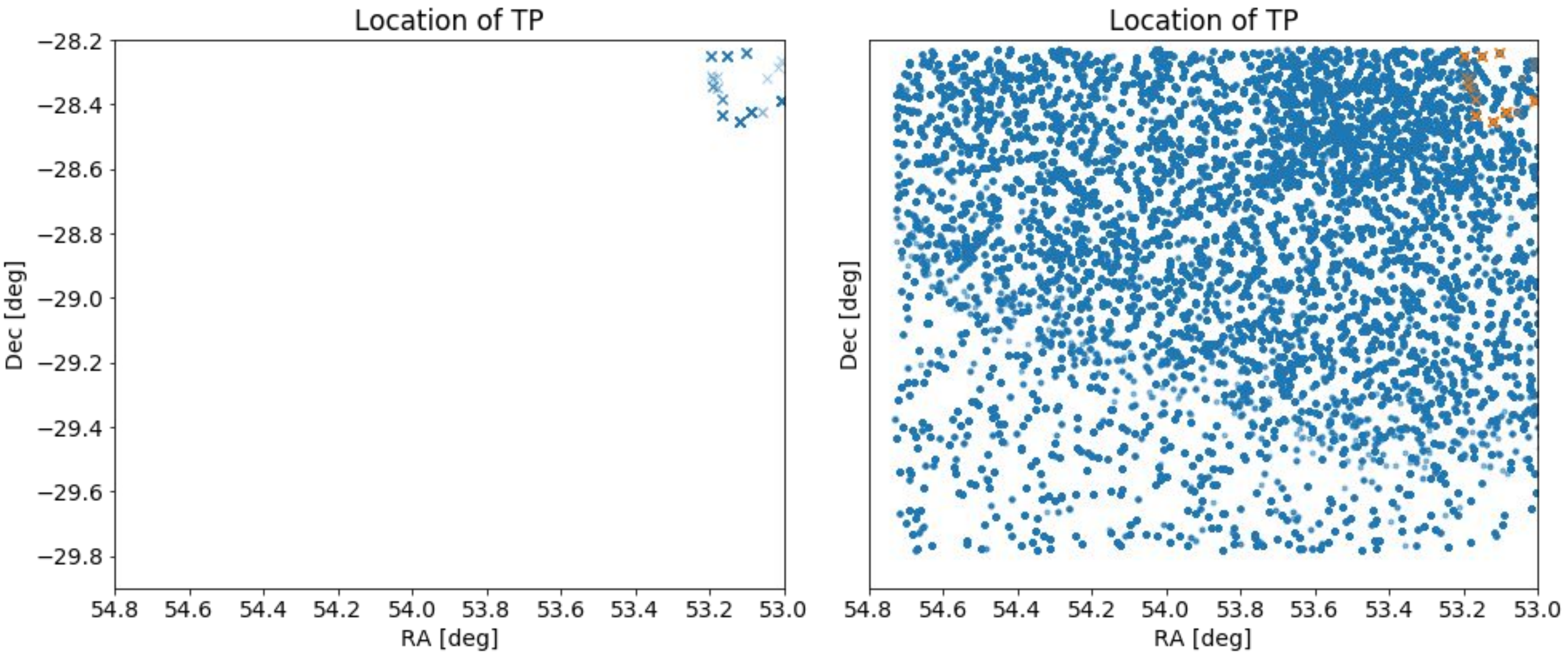
Run 1.2 Truth variable vs DIA Object catalogs

By plotting their positions in the sky we could conclude that coordinates have issues and do not match

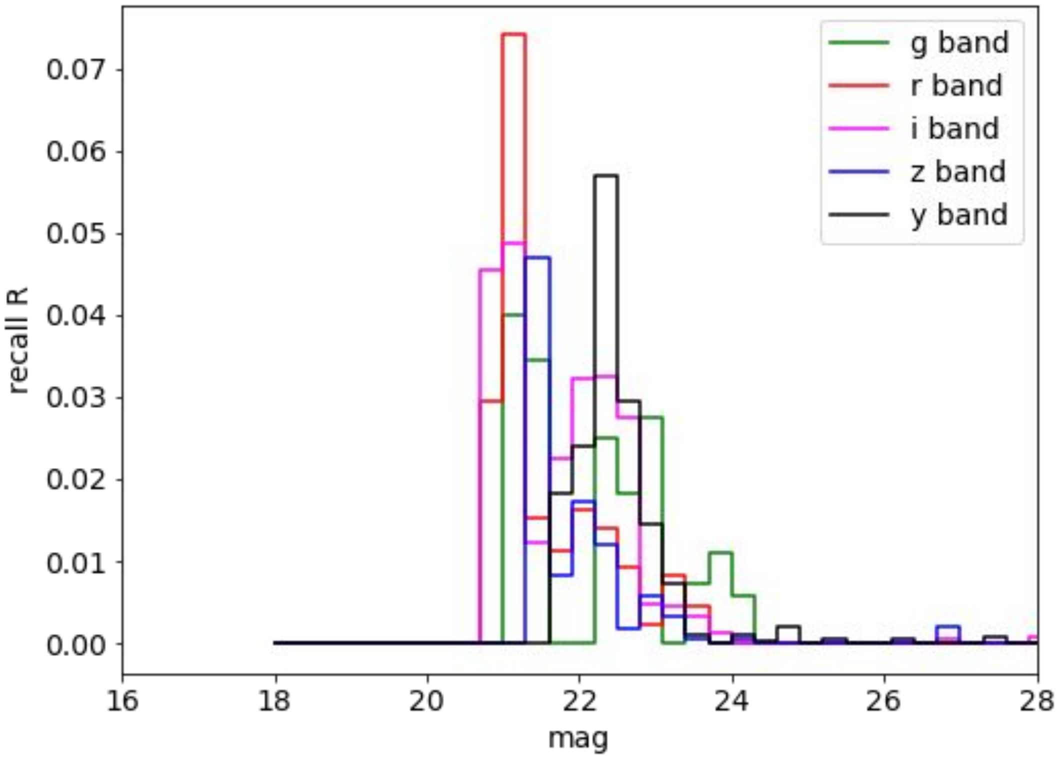


Run 1.2 Truth variable vs DIA Object catalogs

Marking the objects we actually find in crosses



Run 1.2 Truth variable vs DIA Object catalogs



- Again Recall represents a *"recovery rate"*

$$R = \frac{\#detected\ sources}{\#simulated\ sources}$$

- For this sample, in the best case we are **not detecting the 93%** of the transients
- This is extremely low, potentially pointing a **bug** in the catalogs
- This is a good figure to estimate how we are doing transient detection
- Good as a future validation test

Summary and thoughts

- DC2 is an excellent opportunity to test DIA actual implementation results
- Exploring configurations as well as differences among techniques, and results they yield
- Does ML improves the performances? Is it really necessary?

- Influence of photometric properties of transients (colors?)
- Dependencies on the environment (host galaxy brightness profile, distance to its center)
- Dependencies on image quality