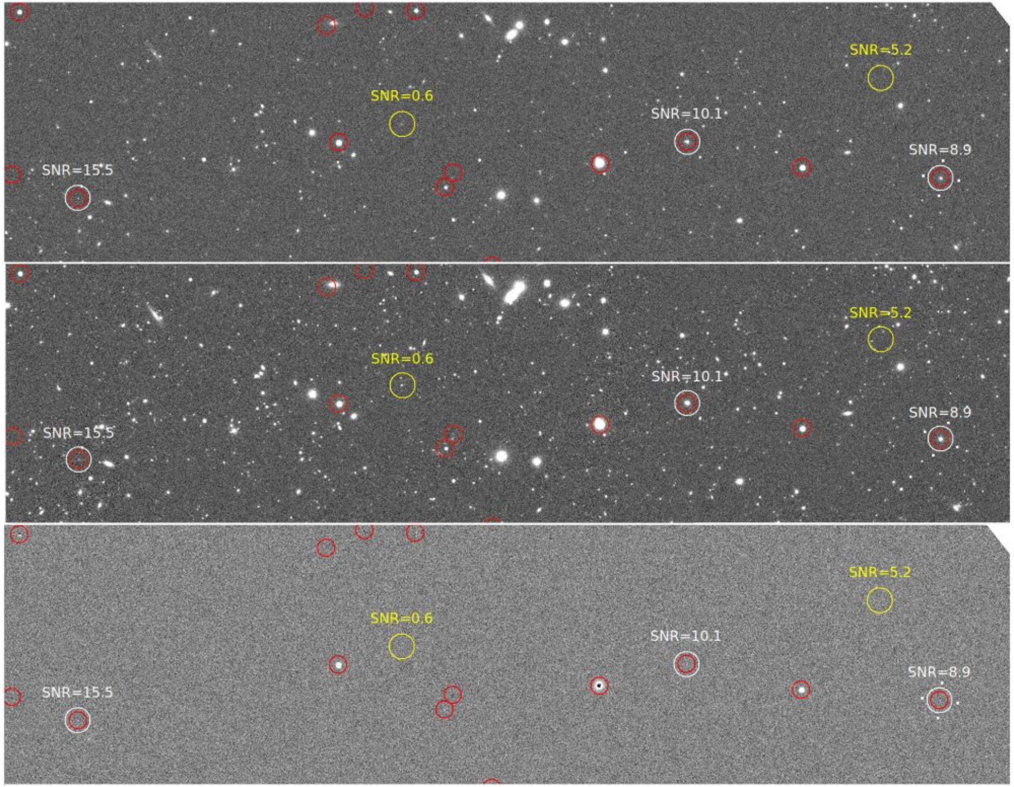
AP Difference Image Analysis

Overall status and future work

Bruno S.

Difference Image Analysis

- DIA is the workhorse behind a large portion of Vera Rubin LSST Time Domain science
- We all know and understand that this is a cornerstone for many science cases
- LSST Data volume will make tiny fractional effects (corner cases) a daily thing
- New and exciting science is going to be near limit cases
  - Very low SNR transients
  - Very noisy environments
  - …



lp_diffim Subtraction tasks

Things to keep in mind about broad perspective:

- The difference image subtraction task is a link in a chain inside a very complex system
- The main objectives of this system are always the detection of transients but many more things depend on it
    - The subtraction results yield products that are used for measurements
    - The subtraction metadata is also valuable piece of information – each implementation might handle this differently
- The core algorithm is to find a transformation kernel that is best suited for finding pixel flux changes, but implementation has to deal with many more things such as masks – this means extra development (such as Zogy case)

lp_diffim A&L Subtraction tasks

Difference Image Analysis consists on pixel-to-pixel comparisons.

- Nominal Alard & Lupton DIA algorithm
  $D = N - k * R$

lp_diffim A&L Subtraction tasks

Difference Image Analysis consists on pixel-to-pixel comparisons.

- Nominal Alard & Lupton DIA algorithm
  $D = N - k * R$

- Auto-convolution mode: perform a re-assignment of the image to transform
  $D = k * N - R$ `if` $\mathrm{PSF}_N < \mathrm{PSF}_R$ `else` $D = N - k * R$

lp_diffim A&L Subtraction tasks

Difference Image Analysis consists on pixel-to-pixel comparisons.

- Nominal Alard & Lupton DIA algorithm
  $D = N - k * R$

- Auto-convolution mode: perform a re-assignment of the image to transform
  $D = k * N - R$ `if` $\mathrm{PSF}_N < \mathrm{PSF}_R$ `else` $D = N - k * R$

- Pre-Convolution mode: convolution of N with a given known kernel v:
  $\mathrm{PSF}_{N'} > \mathrm{PSF}_R$; as result k is no longer a de-convolution kernel

  $v * D = v * N - k * R$

Assessment of DIA performance

A detailed report is available at DMTN-256: Status of Difference Image Analysis

We processed precursor data from HSC Cosmos at USDF using weekly release 07:
“2021 DIA/diffim sprint” dataset:

- g-band visits [11690, 11692, 11694, 11696, 11698, 11700, 11702, 11704, 11706, 11708, 11710, 11712, 29324, 29326, 29336, 29340, 29350]
- r-band visits [1202, 1204, 1206, 1208, 1210, 1212, 1214, 1216, 1218, 1220, 23692, 23694, 23704, 23706, 23716, 23718]
- Detectors in [49, 50, 57, 58, 65, 66]

Assessment of DIA performance

A detailed report is available at [DMTN-256: Status of Difference Image Analysis]

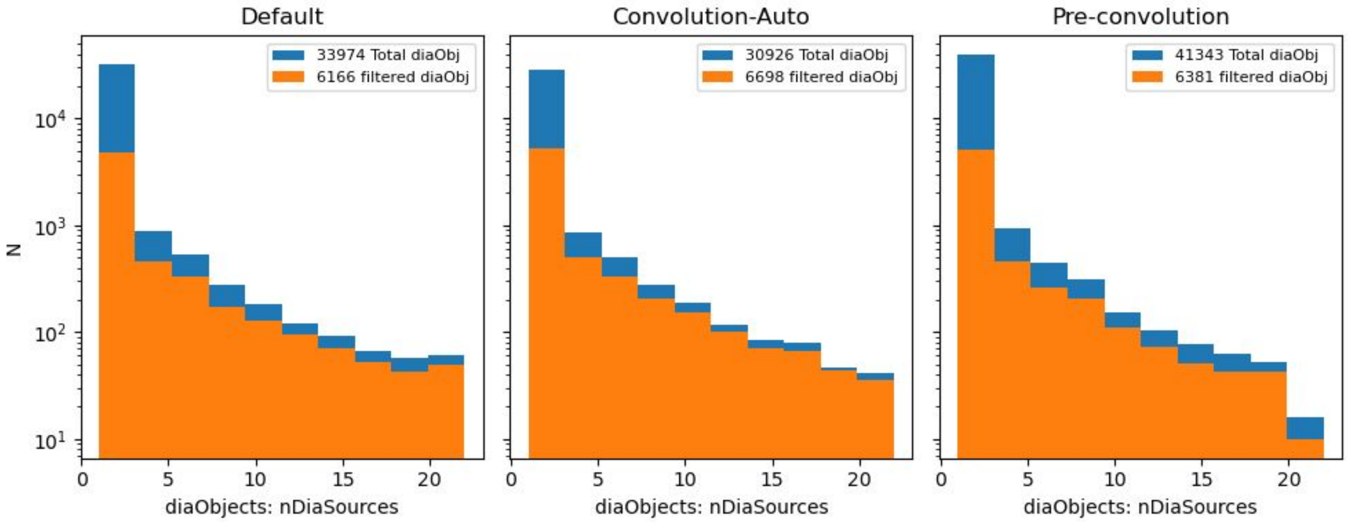
We will use a set of tools to get meaningful information:

1. Synthetic source injection: typically called "fakes", these are the main tool to generate "controlled scenarios" to test and understand algorithm outcomes
2. Global metrics such as number of detections, location of detections, number of flagged sources, etc.
3. Truth catalog values of fakes: useful to evaluate measurements on detections
4. Multiple datasets – same techniques: testing should be done on multiple datasets on varied observing conditions

Runs @ USDF

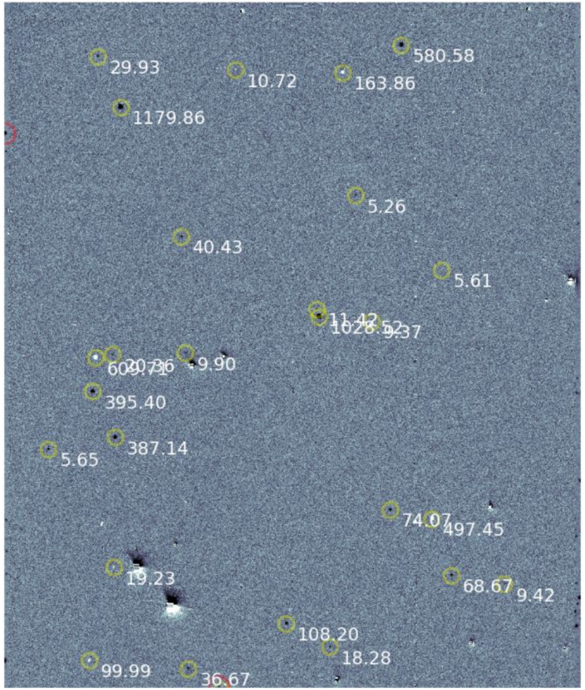
Data is available if you want to look at it.

All the runs are available at the main repo at USDF. Data collection names are listed in the DMTN but new runs are being processed regularly.

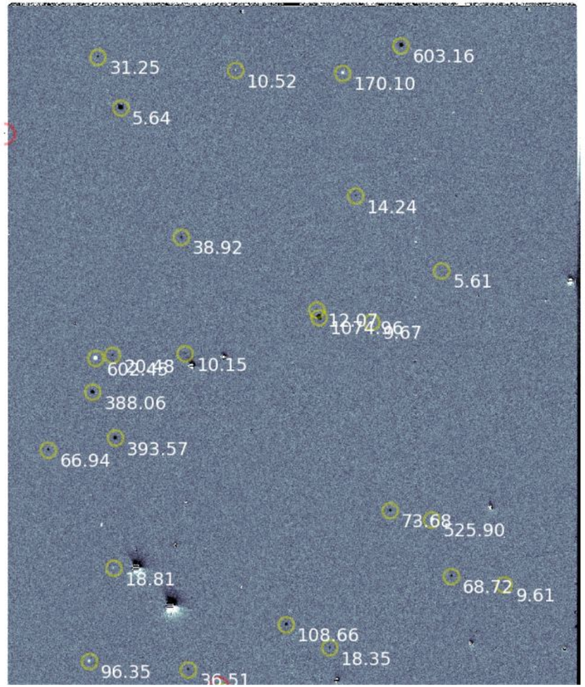


Runs @ USDF

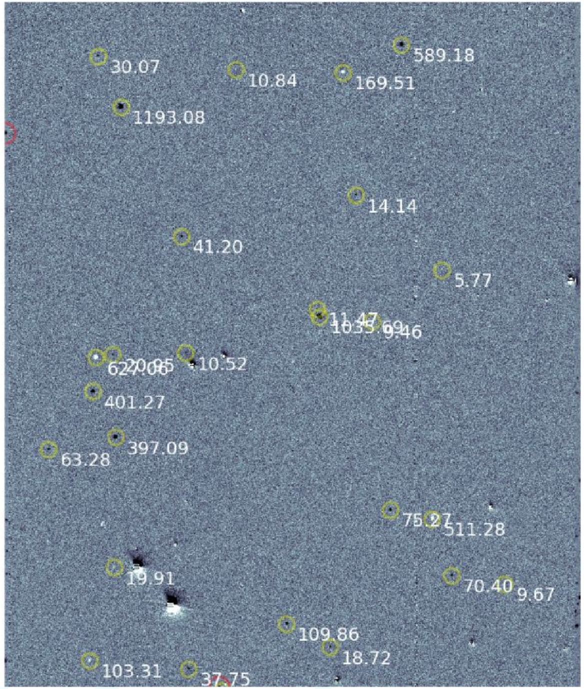
Default



Auto-Convolution



Pre-Convolution



Dia Source detections

We compare sources for the different DIA flavors.
This is a first global metric that AP uses for understanding global changes in results

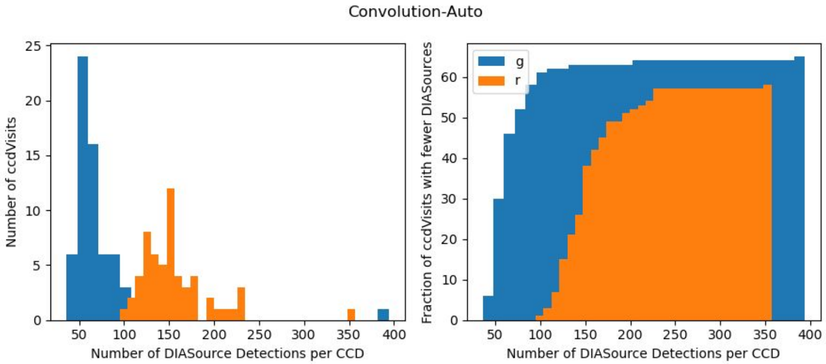
	Good diaSrc	Good diaObject	N. Fake Matches	Contamination	Matches w/flag cuts
Default	13244	6166	4627.0	50172.0	2391.0
Pre-Conv	10909	6381	4594.0	55649.0	1732.0
Auto Mode	13727	6698	4321.0	45728.0	2348.0

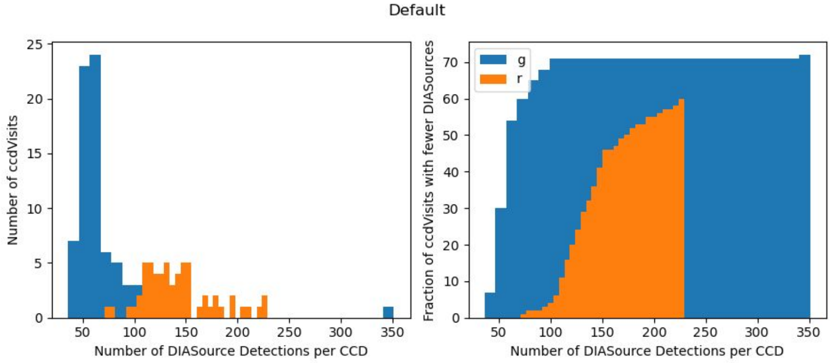
We use diaSources - diaObjects and “good” diaSources and diaObjects i.e. after flag cuts

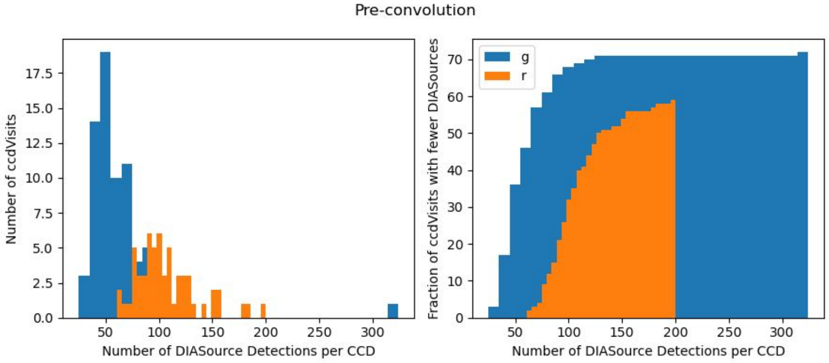
Dia Source detections

Number of diaSources are different for each flavor run. We take this as a global metric for comparisons

We take always ‘default’ as baseline



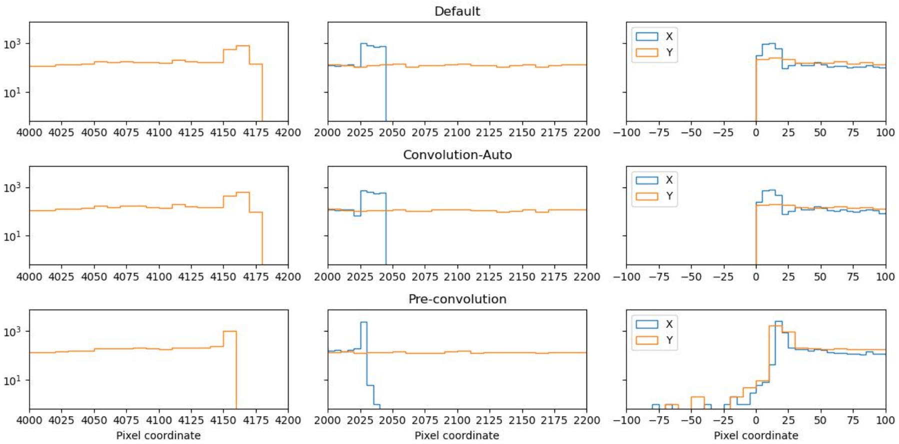


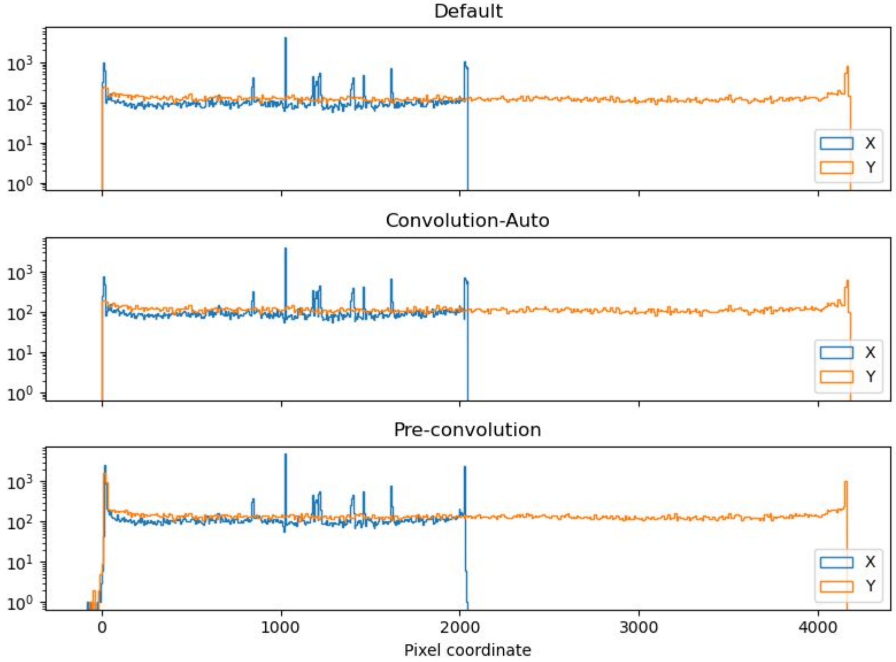


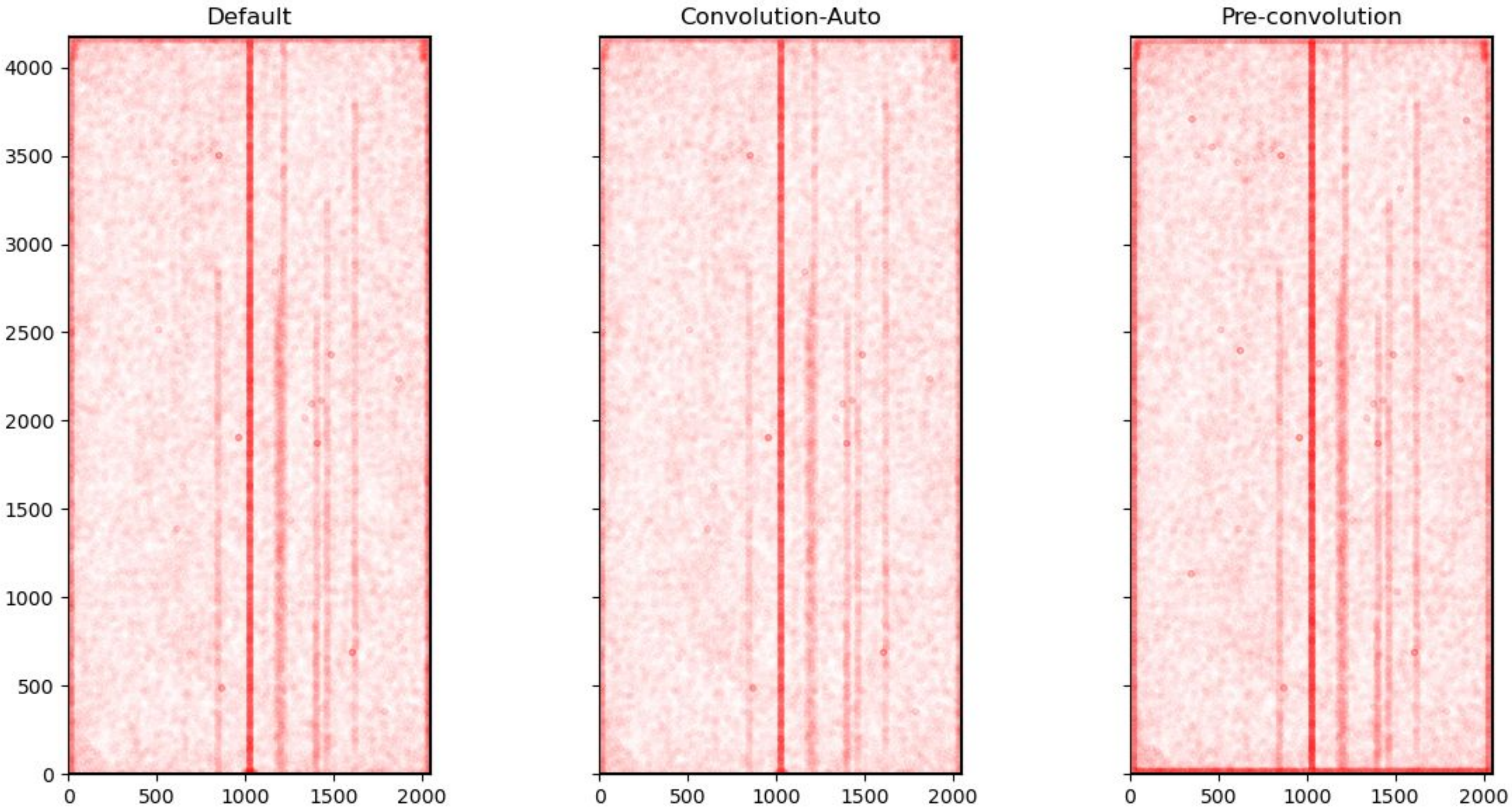
Spatial location of detections

We look at the location of our diaSources in the detector plane.

We can find here edge effects, dead columns, and template overlap effects.

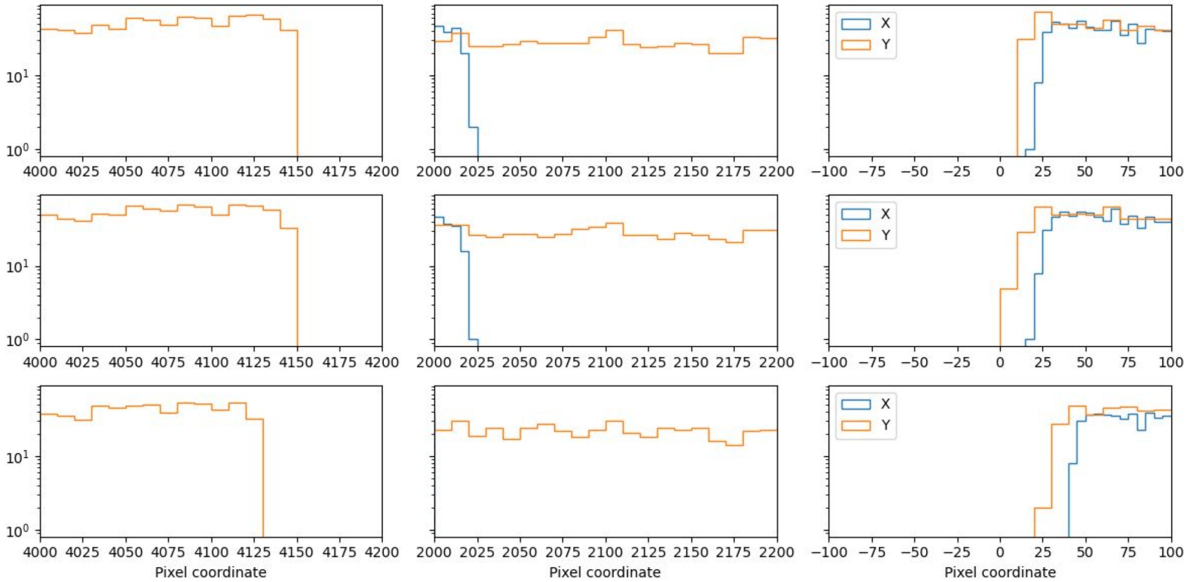





Default
Convolution-Auto
Pre-convolution
0
500
1000
1500
2000
2500
3000
3500
4000
0
500
1000
1500
2000

Spatial location of detections

We can filter these out most of the time.
Caveat: risk of ruling out good transients in some regions if we get too aggressive on flags

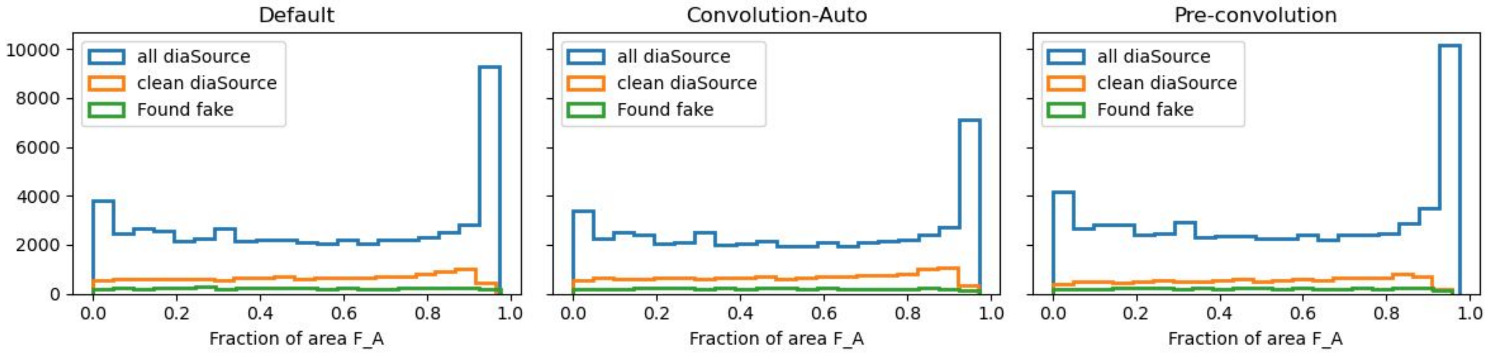


Spatial distribution of diaSources

We estimate the "fraction of area" or γ test on where sources are being detected.

We use transient coordinate (x,y) to get $l{=}max(x',y')$ with $(x',y'){=}(x{-}xc,y{-}yc)$.
The fraction of area value is then: $\gamma{=}l^2/(4096{\times}2048)$

This quantity should in theory be flat distributed

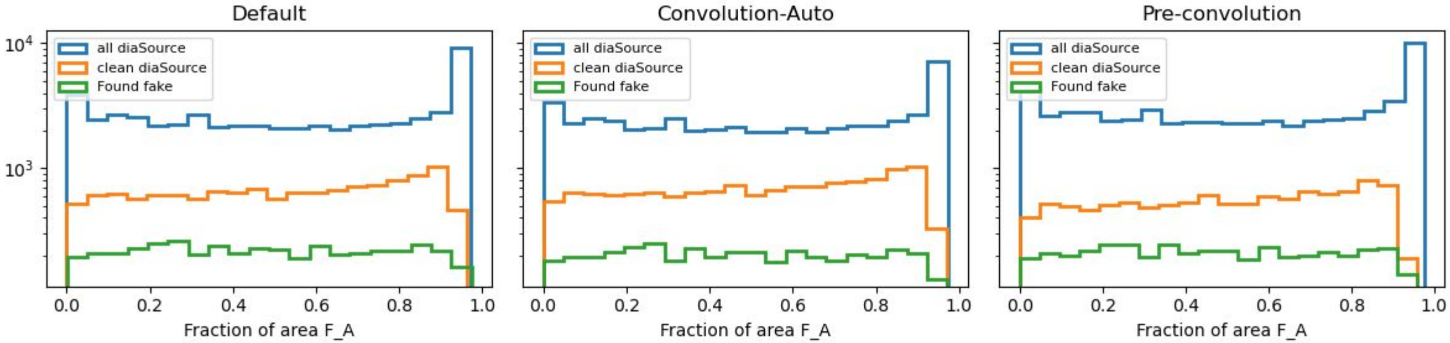


Spatial distribution of diaSources

We estimate the "fraction of area" or γ test on where sources are being detected.

We use transient coordinate (x,y) to get $l=max(x',y')$ with $(x',y')=(x-xc,y-yc)$.
The fraction of area value is then: $\gamma=l^2/(4096\times2048)$

This quantity should in theory be flat distributed

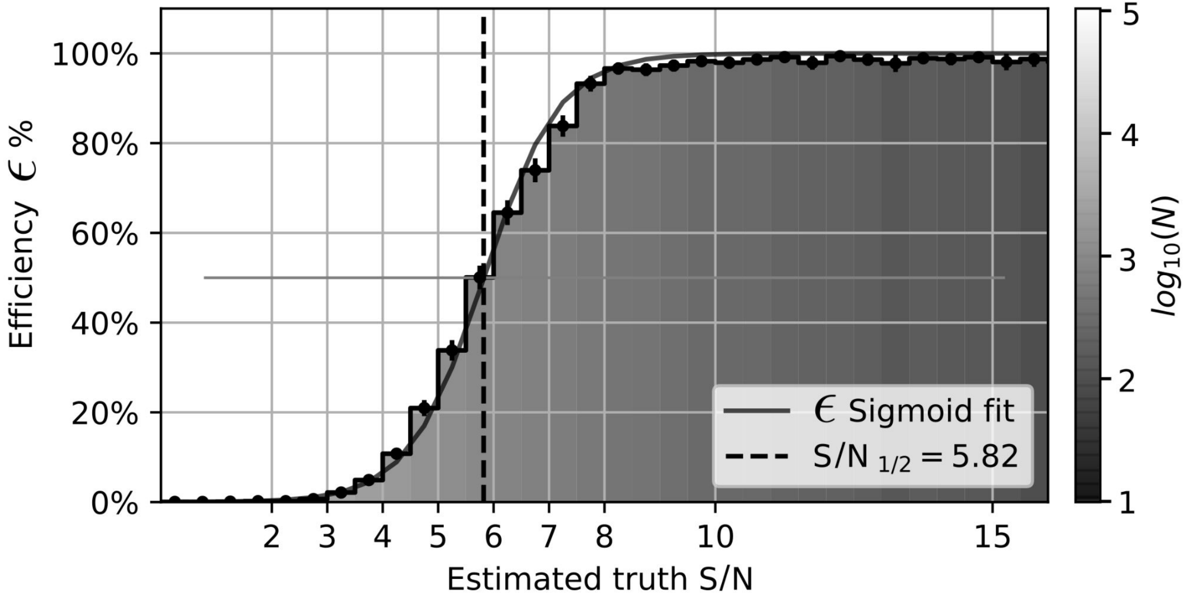


Quantification of efficiency

We would like to reproduce our estimation of Detection efficiency on DC2

We define the value of $S/N_{1/2}$:= SNR at which Recall = 50%.

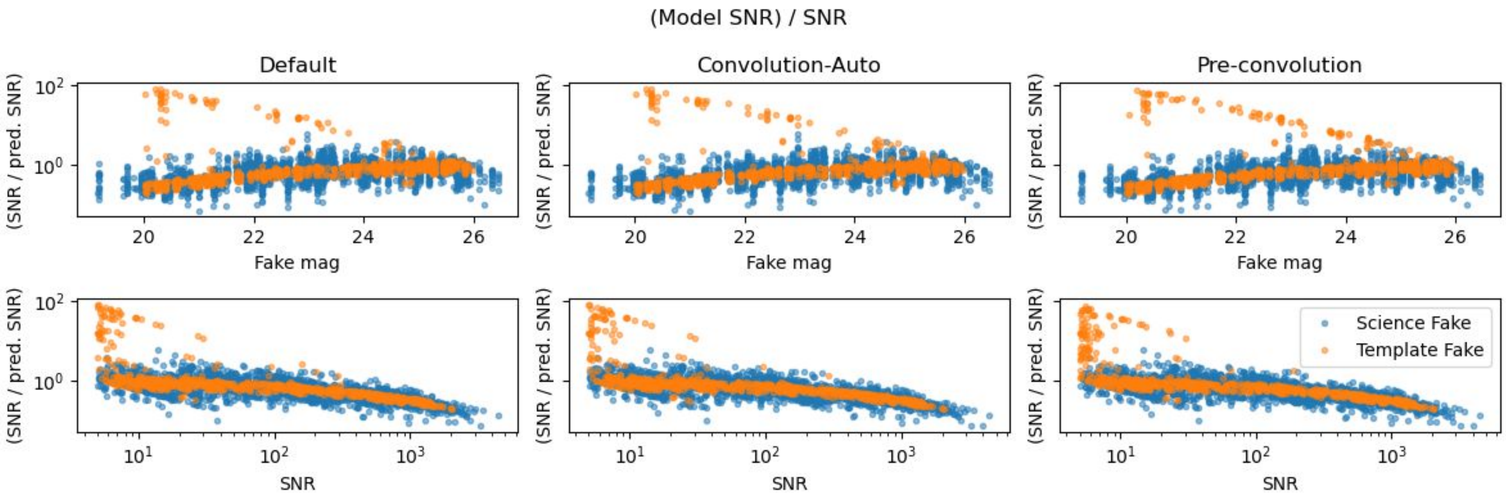
(Lower $SNR_{/2}$ is more sensitive)



Quantification of efficiency

Our first step is to estimate the SNR for our fake sources.

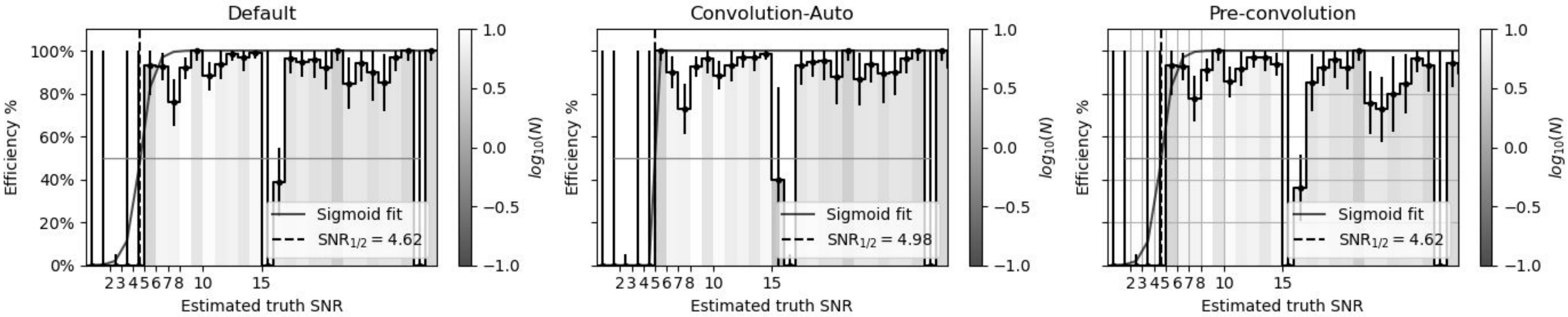
This is using a similar model used for DC2 (using PSF noise equivalent area and sky noise, plus source shot noise)



Quantification of efficiency

Our first step is to estimate the SNR for our fake sources.

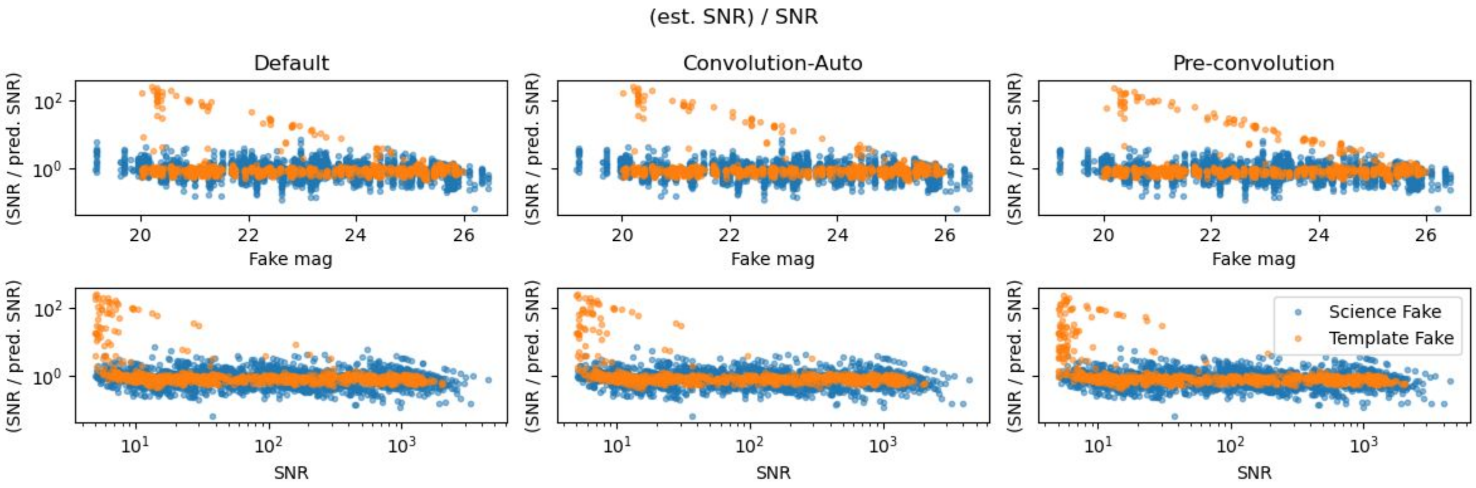
This is using a similar model used for DC2 (using PSF noise equivalent area and sky noise, plus source shot noise)



Quantification of efficiency

Our first step is to estimate the SNR for our fake sources.

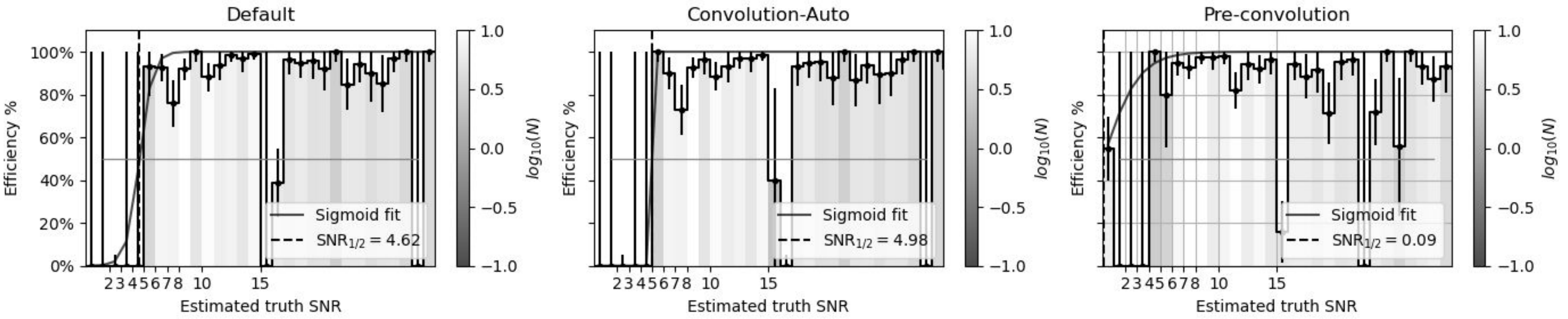
This is using the variance plane instead, using PSF weighted sum (just like PSF phot.)



Quantification of efficiency

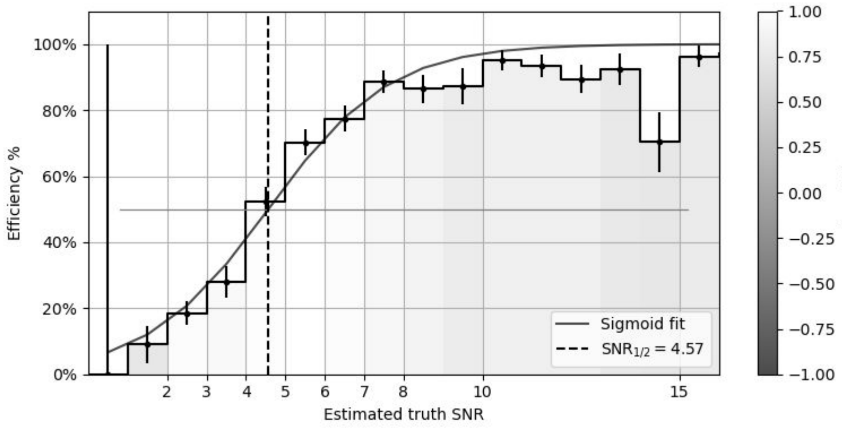
Our first step is to estimate the SNR for our fake sources.

This is using the variance plane instead, using PSF weighted sum (just like PSF phot.)



More recent Run

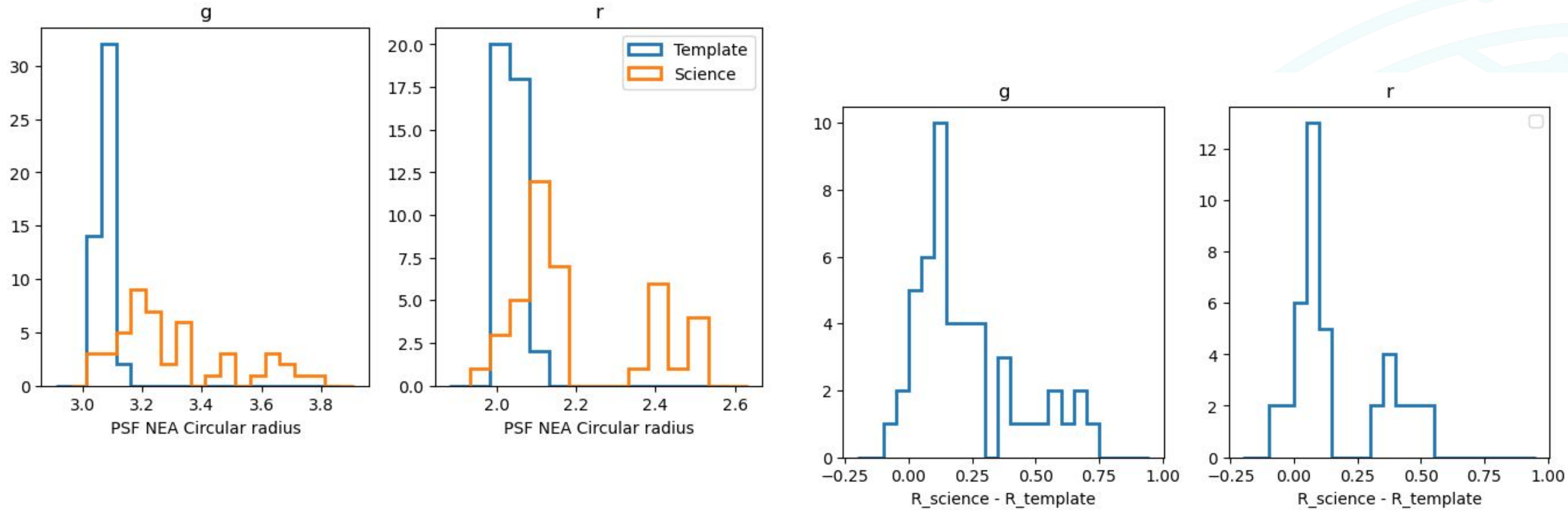
- We increased the fake density
  - From 1500 → 5000 per sq deg
- We increased the magnitude range
  - Maximum magnitude: 26 → 27
- We processed subtractions only in default mode as of now
- We obtained a $SNR_{1/2}$ ~= 4.6 (DC2 had 5.8)
- Runs with auto-conv and pre-conv coming soon



PSF Sizes

As general preference Template PSF should be smaller.

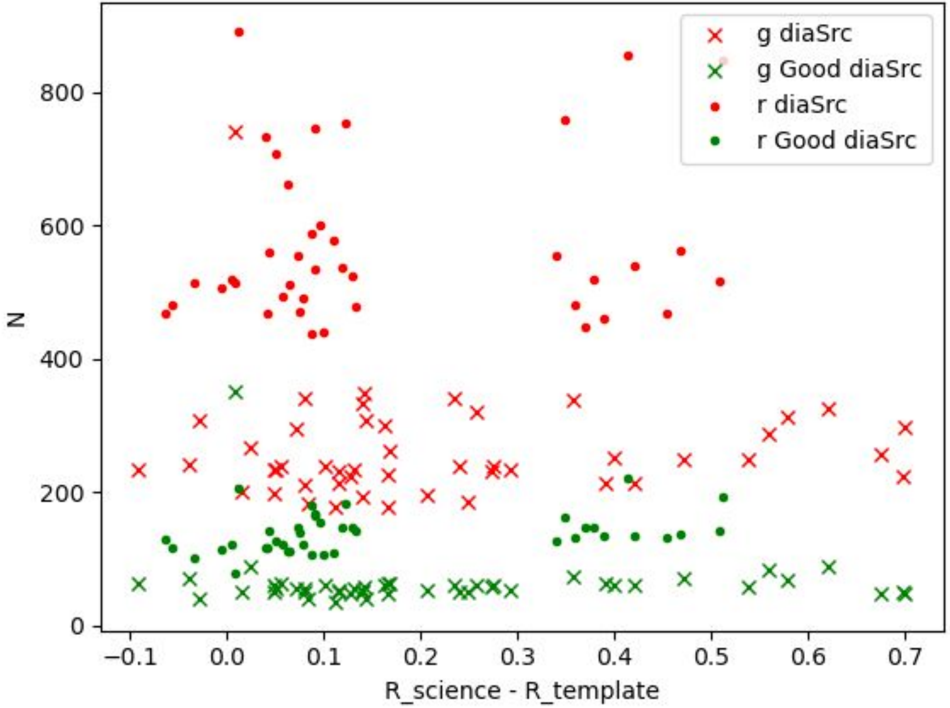
Worse case equal to the science PSF. This is almost always the case.



PSF Sizes

Number of detections (also after flag cuts) vs PSF difference

Not clear relation, it is arguable that this difference in PSF is not cause for changes in number of detections



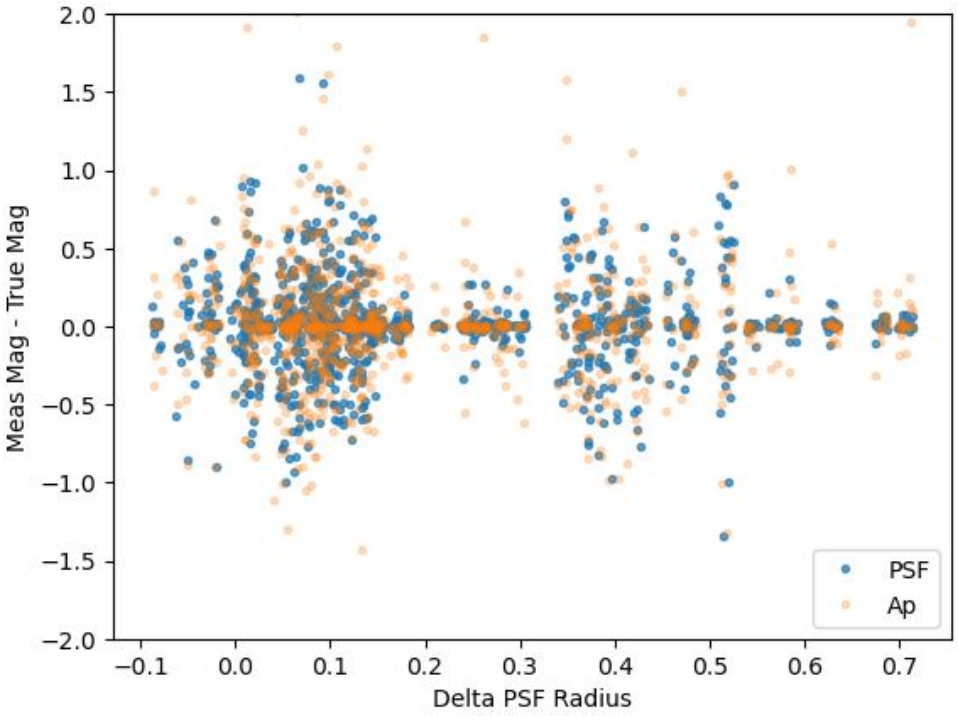
Fakes

What about the photometry?

This might be due to many reasons.

Template quality influences this and it could be a reason

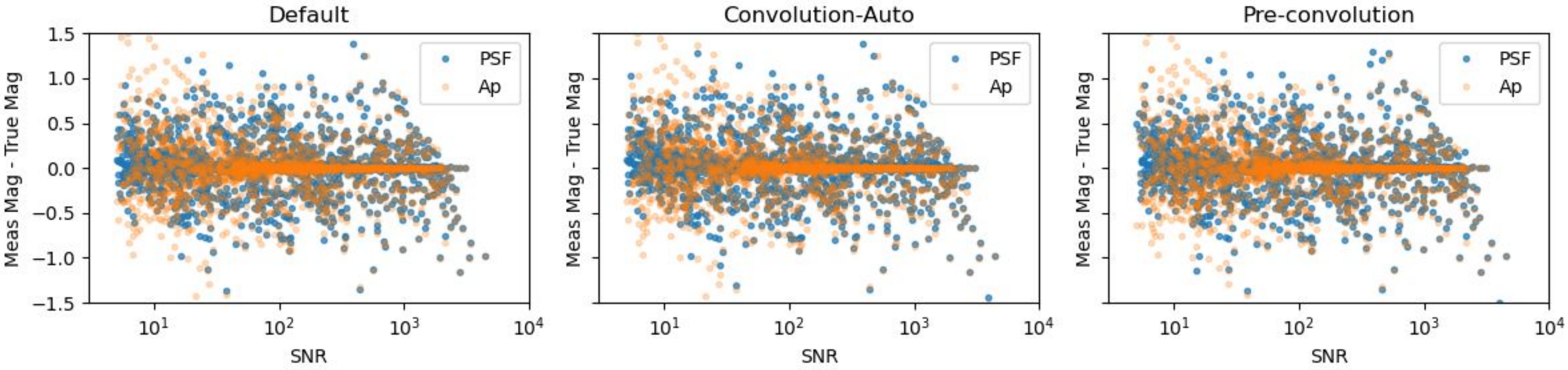
There is better photometry when the PSF radius differences are between 0.2 and 0.35 for some reason, or differ more than 0.5



Photometric performance

We quantify the recovery of fluxes for transients

For this, again, we rely on fakes

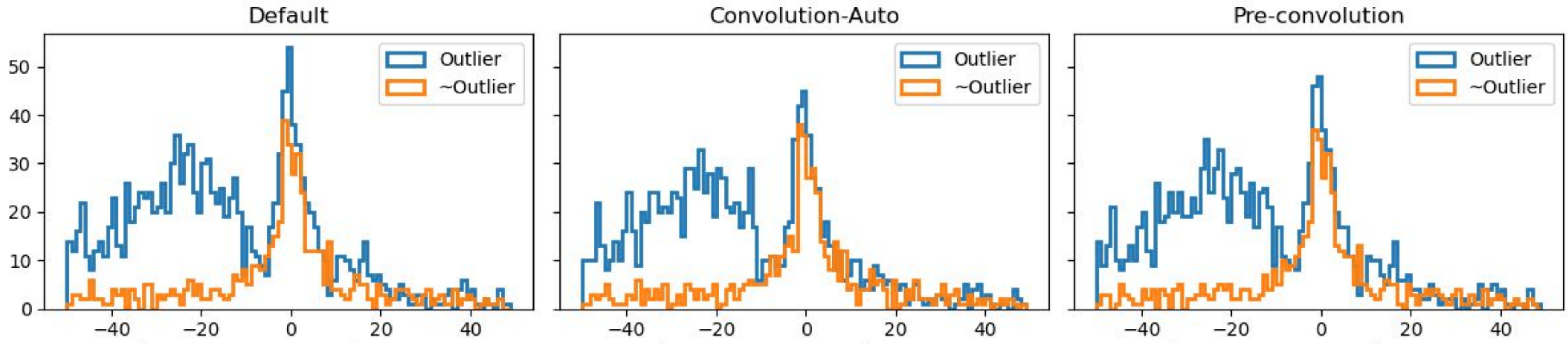


Photometric performance

We quantify the recovery of fluxes for transients

For this, again, we rely on fakes

We define outliers to be sources with reported SNR < 0.5 → we know they should be bad



(PSF Flux meas - True flux)/PSF Flux err

Summary with thoughts and Conclusions

- We processed a small batch of g,r HSC images through ApPipe (wFakes)
    - Tested flavors of DIA on common grounds
- The number of fake injections was too low to get statistical signal for efficiency measurement.
    - Additional runs are needed with higher fake density
- The flag cuts conventionally applied are too aggressive risk of discard good transient sources.
    - Maybe some work needs to be done to optimize this
- Our dataset for this test does not contain some specific scenarios in which the various DIA flavors should out-perform the default Alard & Lupton technique.
  Probably incorporating some cases with bad seeing on the templates is of value for testing purposes.
- Understanding correctly how to model the SNR for point sources of a given magnitude could yield value in the assessment of the limit of detections.
- Flux recovery has weak points that we can investigate and see if they trace back to algorithm features or data features

Future work and other ideas

- Migration from current Fake injection in ap_pipe
    - Refactoring to adopt new package source_injection
    - We want to have a controlled scenario for testing but a relevant scenario
- Simulations like DC2 are a bit too perfect
    - Real data like HSC has its own quirks, which we want to ignore in case we develop new algorithms, so it is important to have a basic idea of metrics in data sets with a comparable procedure
- Developing new ideas in the algorithms and improve their performance has to rely on baseline metrics of single subtraction, but also, mid-size and large size runs that encompass association and potentially light-curve building stage
- We would like to include the desired group of fake based metrics in analysis_ap and analysis_tools so we can quickly iterate on prototype algorithms