Rubin Observatory

Image Differencing in Science Pipelines

Eric Bellm (UW/Rubin Obs.)
Alert Production Science Lead

Rubin Observatory Algorithms Workshop | Virtual | March 17-19, 2020

NSF AURA U.S. DEPARTMENT OF ENERGY Office of Science SLAC CHARLES AND LISA SIMONYI FUND FOR ARTS AND SCIENCES LSST CORPORATION

Many people have contributed to image differencing algorithm and pipeline development in DM.

Rubin Observatory

- Yusra AlSayyad
- Andrew Becker
- Steven Bickerton
- Andrew Connolly
- Jim Bosch
- Krzysztof Findeisen
- Josh Hoblitt
- Tim Jenness
- Mario Juric
- Gabor Kovacs
- Simon Krughoff
- Lauren MacArthur
- Chris Morrison
- Fred Moolekamp
- Shu Liu
- Nate Lust
- Robert Lupton
- Russell Owen
- John Parejko
- Paul Price
- Meredith Rawls
- David Reiss
- Pim Schellart
- Jonathan Sick
- Colin Slater
- Ian Sullivan
- John Swinbank
- Michael Wood-Vasey
- *plus many more on the broader DM team*

Image Subtraction Algorithms

presentation and notation follow this Lupton talk

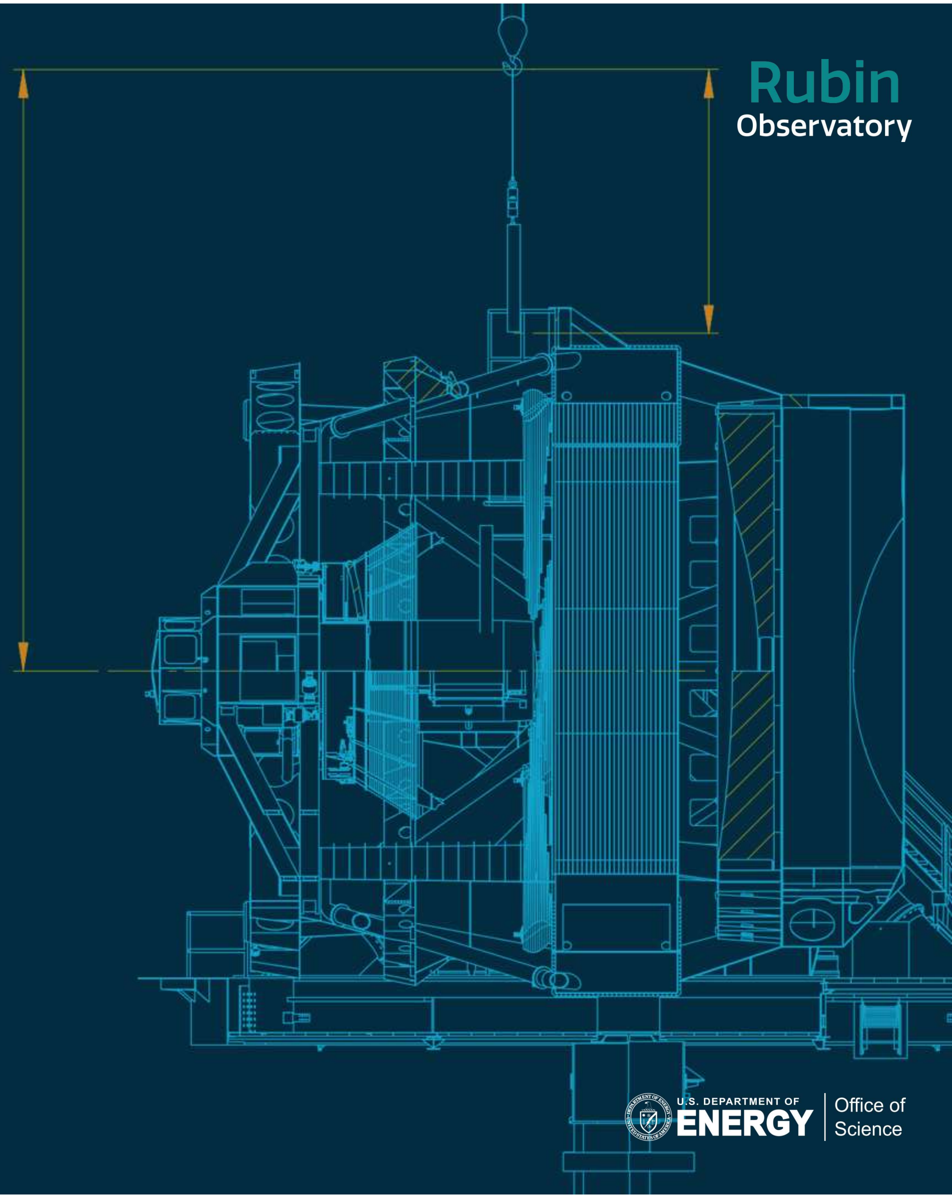
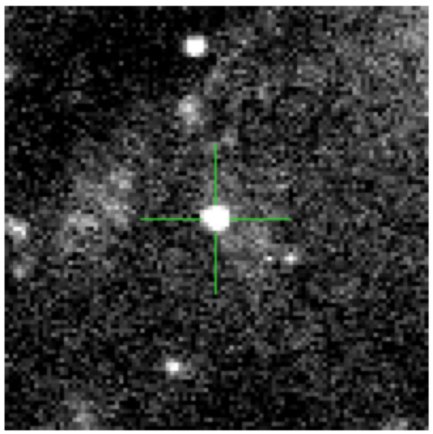
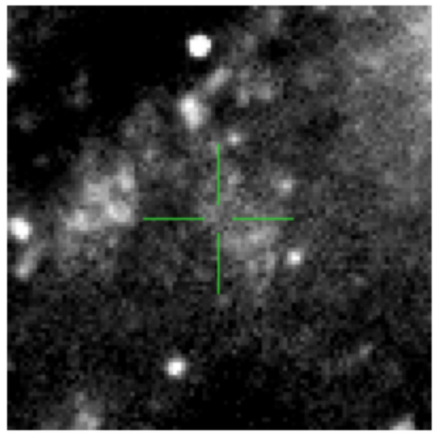


Image subtraction attempts to find sources that are changing.



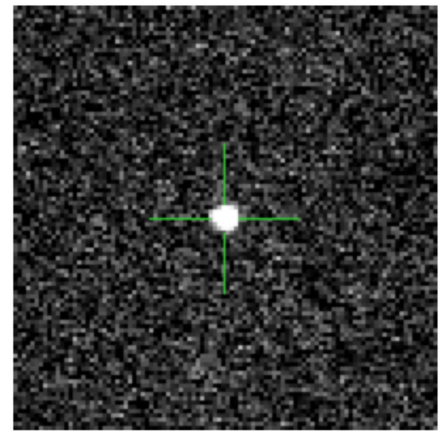
Science image

-



Template image

=



Difference image

Less sensitive to crowding, galaxy light, source association, etc. than measuring fluxes of sources on each image and comparing catalogs.
Less complex than forward modeling.

For ground-based surveys we have to account for PSFs that change from image to image.

Rubin Observatory

One option is to PSF-match the images by computing an appropriate convolution kernel in Fourier space, degrading to the resolution of the worse-seeing image (e.g., Ciardullo+90, Phillips & Davis 95).

$$I'_2(k) = I_2(k)\frac{\phi_1(k)}{\phi_2(k)}$$

$$A_1 - A_2 = \frac{\sum_i \left(I_{1,i} - I'_{2,i}\right)\phi_i}{\sum_i \phi_i^2}$$

But this requires very good knowledge of the PSF, which may be difficult in crowded fields, and care to avoid numerical instability in the division.

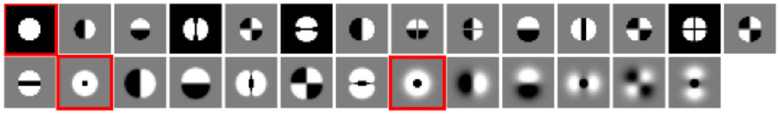
Alard & Lupton 1998 model the convolution kernel as a series of basis functions.

Rubin Observatory

Defining the convolution kernel as:

$$I_2' = \kappa \otimes I_2 \qquad \text{where} \qquad \kappa = \sum_r a_r B^r$$

and Br are a set of basis functions



LDM-227

in the background-dominated regime we can use least-squares fitting to determine the kernel basis function coefficients.

$$\left| \frac{I_1 - \sum_r a_r \left(B^r \otimes I_2\right)}{\sigma} \right|^2$$

This works nicely if the template is noise-free and the science image has better seeing than the template.

It also allows for spatial variations in the PSF as well as background matching.

Zackay, Ofek, and Gal-Yam (2016) recognized that classical A&L is not optimal if the template is noisy.

Rubin Observatory

They proposed an alternative algorithm (now widely called "ZOGY") in Fourier space:

$$\widehat{D} = \frac{F_r \widehat{P_r} \widehat{N} - F_n \widehat{P_n} \widehat{R}}{\sqrt{\sigma_n^2 F_r^2 |\widehat{P_r}|^2 + \sigma_r^2 F_n^2 |\widehat{P_n}|^2}}$$

- fully symmetric between the science and template images
- maximizes the SNR of detected point sources when both images are noisy
- requires knowledge of the PSFs (or their ratio)

The ZOGY noise-whitening approach can also be applied within the A&L framework (Reiss & Lupton 16).

See https://dmtn-021.lsst.io/

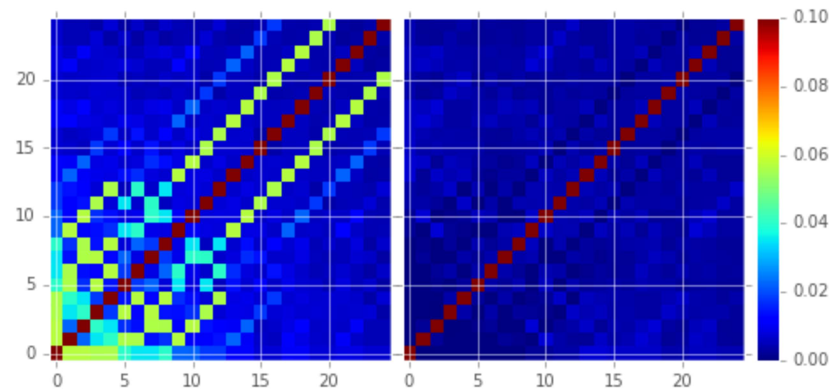
Rather than use the ratio known PSFs, retain the use of basis functions, but whiten the noise in k-space:

$$D(k) = \left[I_1(k) - \kappa(k) I_2(k)\right] \sqrt{\frac{\overline{\sigma}_1^{2} + \overline{\sigma}_2^{2}}{\overline{\sigma}_1^{2} + \kappa^2(k)\overline{\sigma}_2^{2}}}$$

which is identical up to normalizations to the ZOGY expression when

$$\kappa(k) = \phi_1(k)/\phi_2(k)$$

- avoids need to know the PSFs directly
- removes noise covariance as expected



All of these algorithms are implemented in the LSST Science Pipelines.

Rubin Observatory

Classical Alard & Lupton

- kernels: spatially varying sum-of-Gaussians, or regularized delta-function (pixel) basis

ZOGY

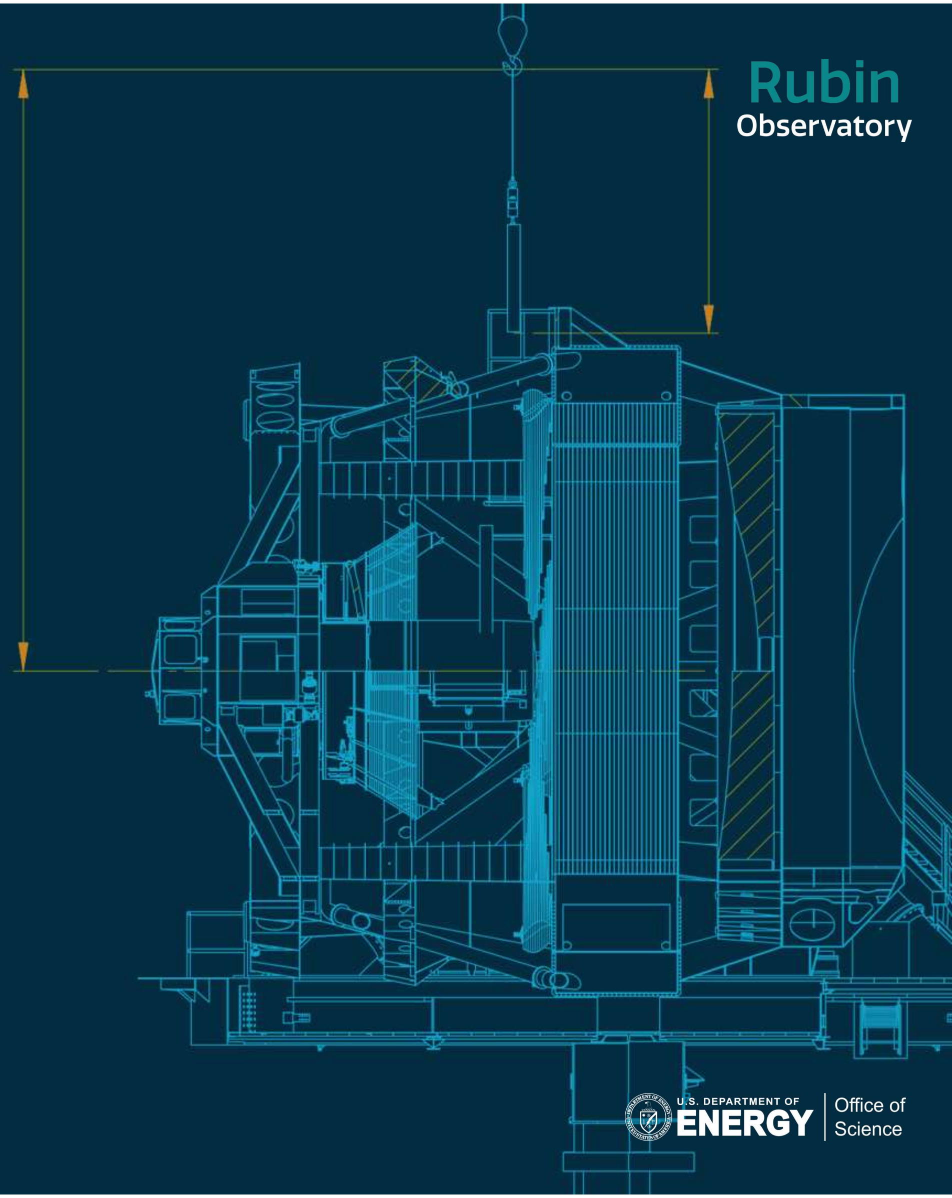
- including piecewise-constant segmentation to handle spatial PSF variation

Decorrelated Alard & Lupton

Until recently implementations of both ZOGY and Decorrelated A&L had bugs in their DFT implementation—fixed in DM-21868

- Results shown today are with the corrected implementation

Other factors



Good templates are required for effective image differencing.

Rubin Observatory

DM can produce a wide variety of coadds (cf. Jim Bosch's talk yesterday).

We are investigating the performance impacts of various choices.

Input image selection

- Number and seeing range (SNR & resolution; impact on artifact rejection (DMTN-080))
- Temporal range (minimizing proper motion effects vs. more inputs for DCR models)

Handling PSF discontinuities

- PSF-matching? careful bookkeeping (e.g., cell-based coadds?)

Algorithm

- direct, PSF-matched, or decorrelated

We are exploring options for template generation in the first year of LSST operations.

Rubin Observatory

No all-sky DRP-produced templates are available prior to DR1.

Several options are outlined in DMTN-107; these options are under review.

- likely to use templates built from any commissioning data
- incremental template generation in Year 1 is being studied


LSST
LARGE SYNOPTIC SURVEY TELESCOPE

Large Synoptic Survey Telescope (LSST)
Data Management

Options for Alert Production in LSST Operations Year 1

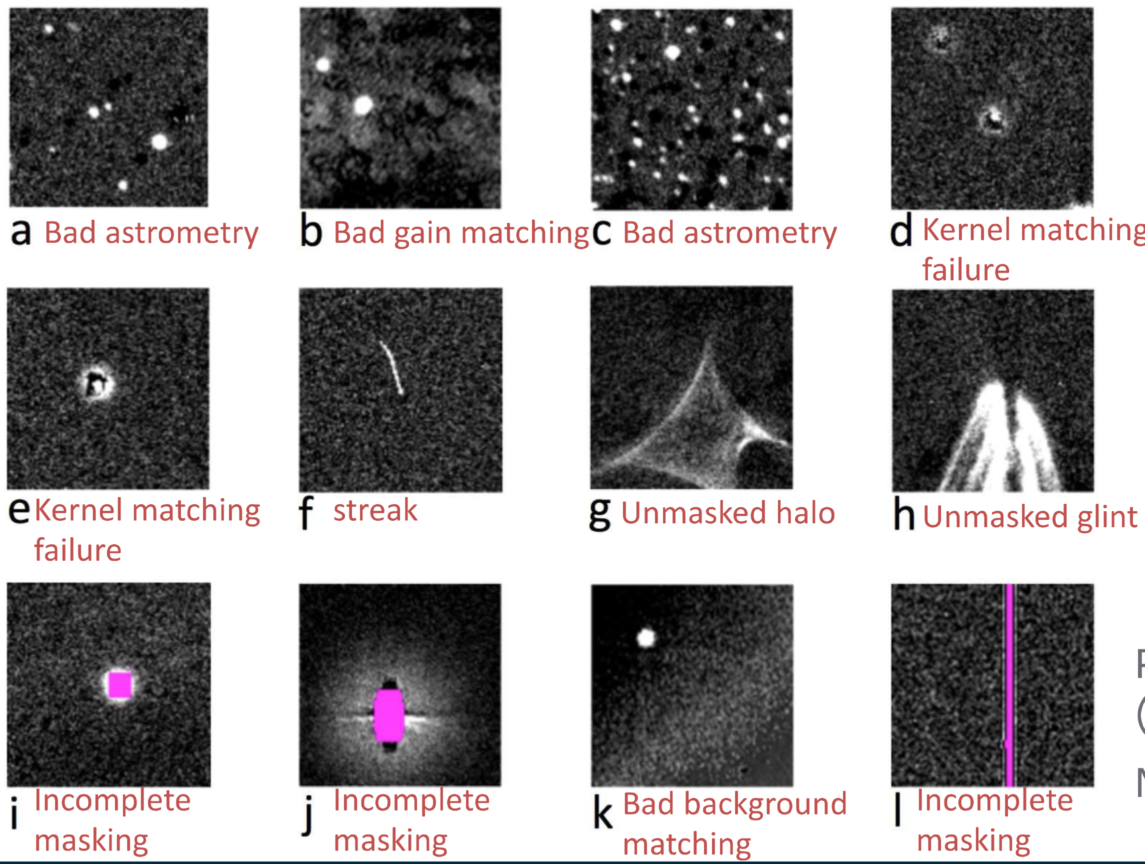
M. L. Graham, E. C. Bellm, C. T. Slater, L. P. Guy, and the DM System Science Team

DMTN-107

Latest Revision: 2019-06-03

Image differencing is also sensitive to ISR, artifact rejection, and astrometric calibration.

Rubin Observatory

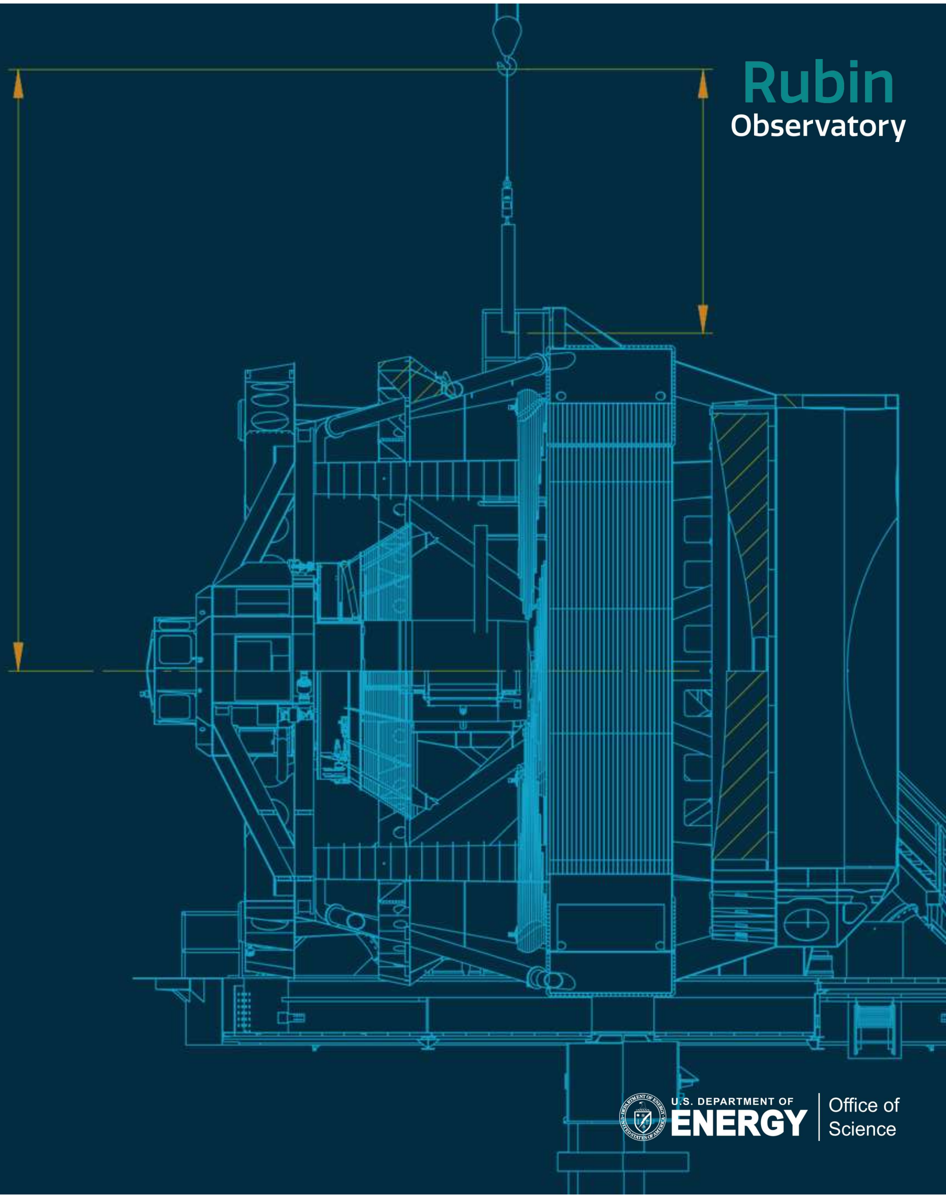


PTFIDE
(A&L w/ pixellated basis)

Masci+17

Rubin Observatory

Current Implementation and Testing



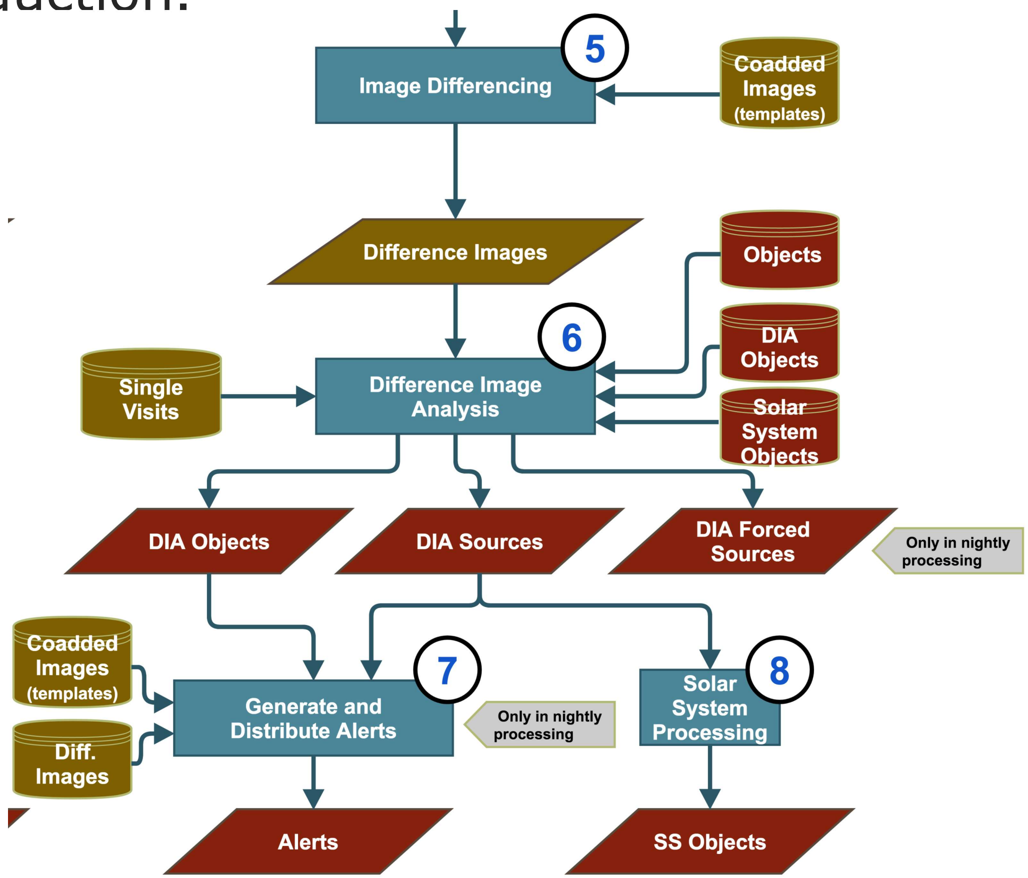
U.S. DEPARTMENT OF ENERGY | Office of Science

Data Management runs Image Differencing in both Alert Production and Data Release Production.

This presentation will focus mainly on implementation and testing for AP

- 60 second latency requirement between camera readout and alert transmission to brokers

DRP is expected to use the same algorithmic codebase, but integration into pipelines will differ (particularly for source association)



We are testing on a variety of precursor datasets.

Rubin Observatory

DECam HITS

- 3 multi-season DECam fields, g-band, 4-5 visits/night for 1 week

HSC SSP Ultra-Deep (COSMOS)

- 1 field, grizy imaging, 30-300 visits/band
- HSC has an ADC—no DCR

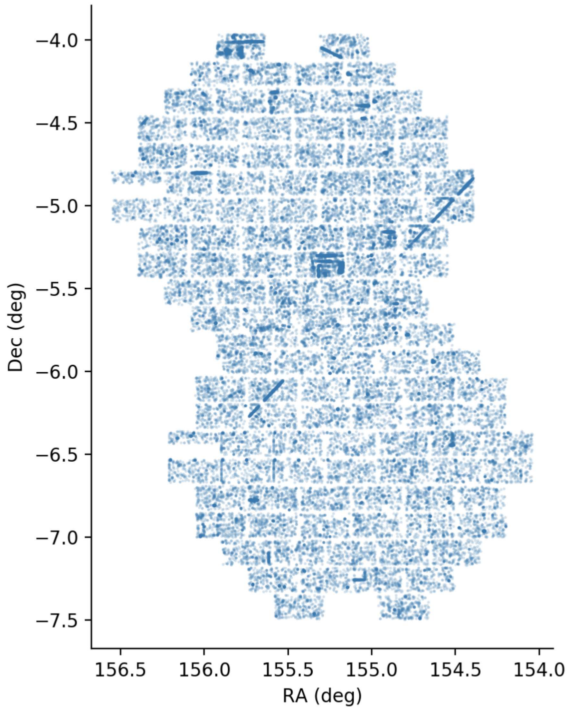
HSC, DECam Bulge data

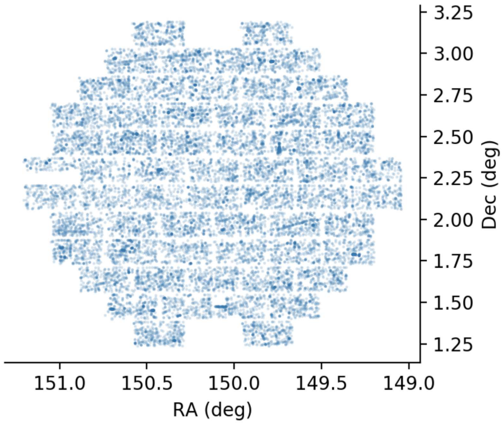
- crowded field analysis

DES SN fields

- weekly griz observation sequences

all datasets have small dithers relative to LSST



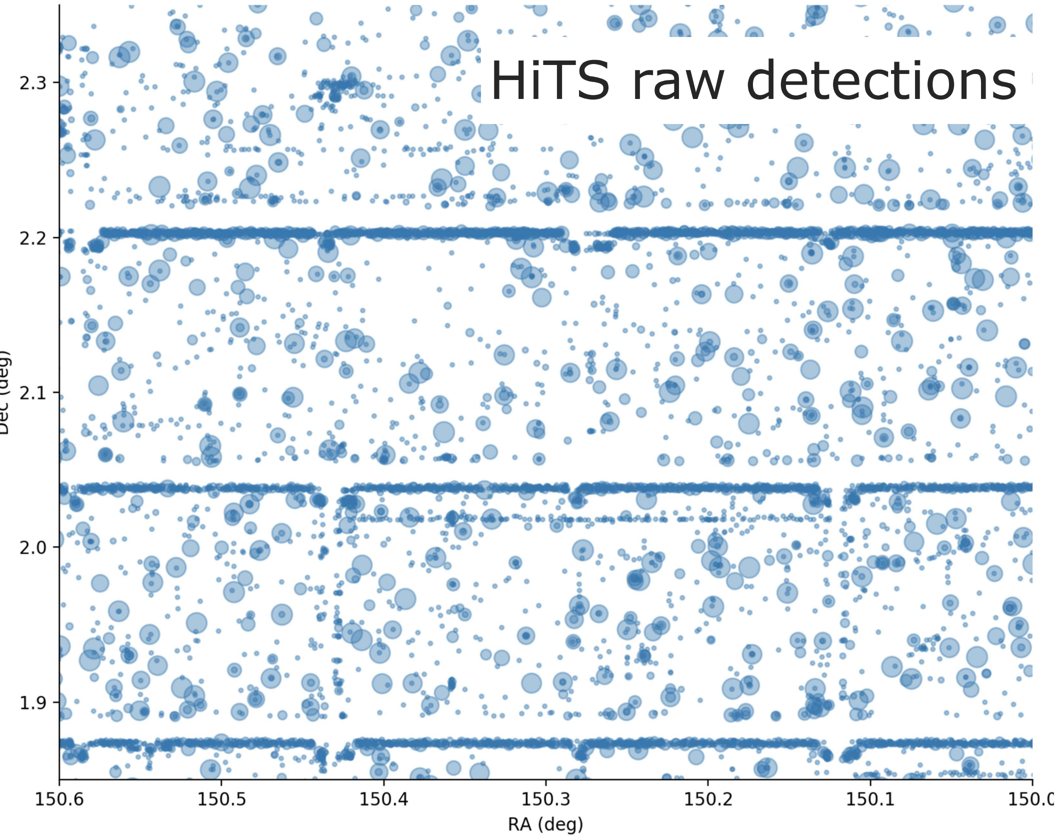


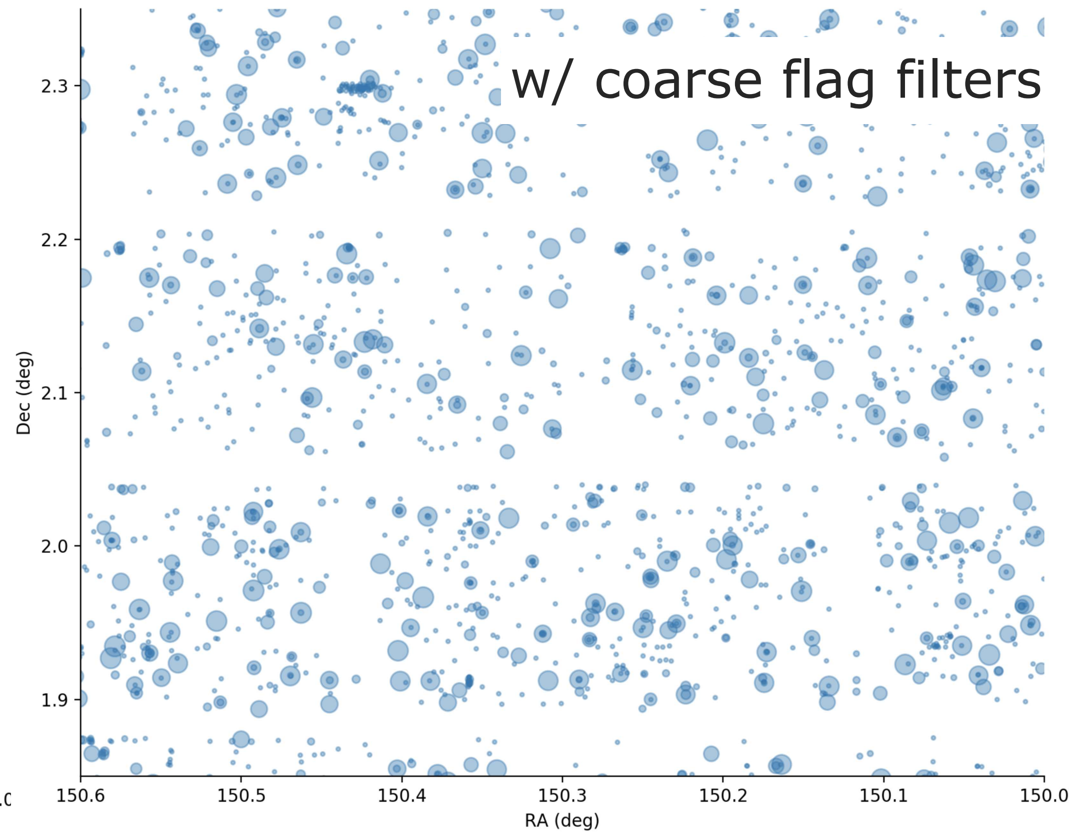
We monitor high-level performance metrics through daily continuous integration.

Rubin
Observatory

Monthly reprocessing enables us to identify larger-scale effects.

Rubin Observatory





Alert Production has high-level requirements on our knowledge of completeness and purity.

Rubin
Observatory

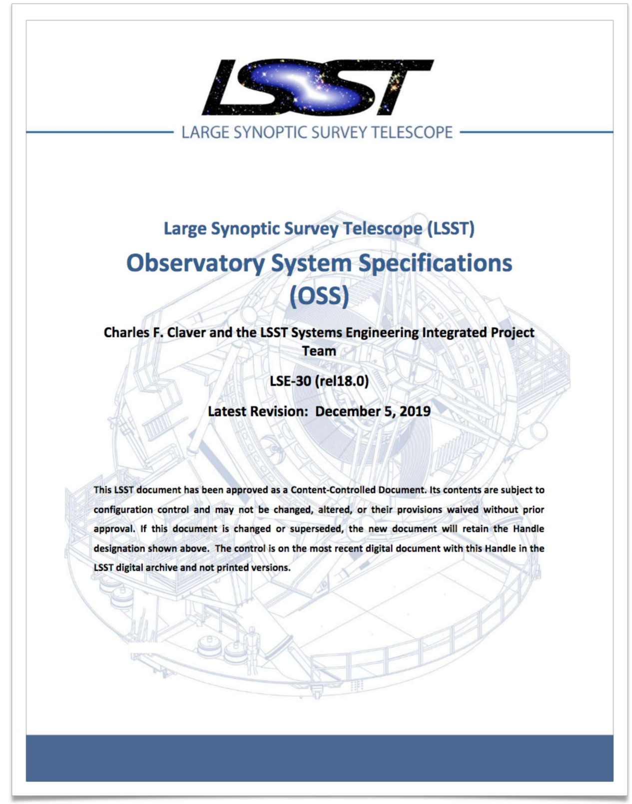
OSS-REQ-0353:

Description	Value	Unit	Name
SNR threshold at which the above are evaluated	6	unitless	transSampleSNR
Minimum average purity for transient science	95	percent	transPurityMin
Minimum average completeness for transient science	90	percent	transCompleteness Min

OSS-REQ-0354:

Description	Value	Unit	Name
Minimum average completeness for Solar System object discovery	99	percent	mopsCompleteness Min
Minimum average purity for Solar System object discovery	50	percent	mopsPurityMin

Verifying these will benefit from an ML algorithm, which can also help us improve the pipeline.


LSST
LARGE SYNOPTIC SURVEY TELESCOPE

Large Synoptic Survey Telescope (LSST)

Observatory System Specifications (OSS)

Charles F. Claver and the LSST Systems Engineering Integrated Project Team

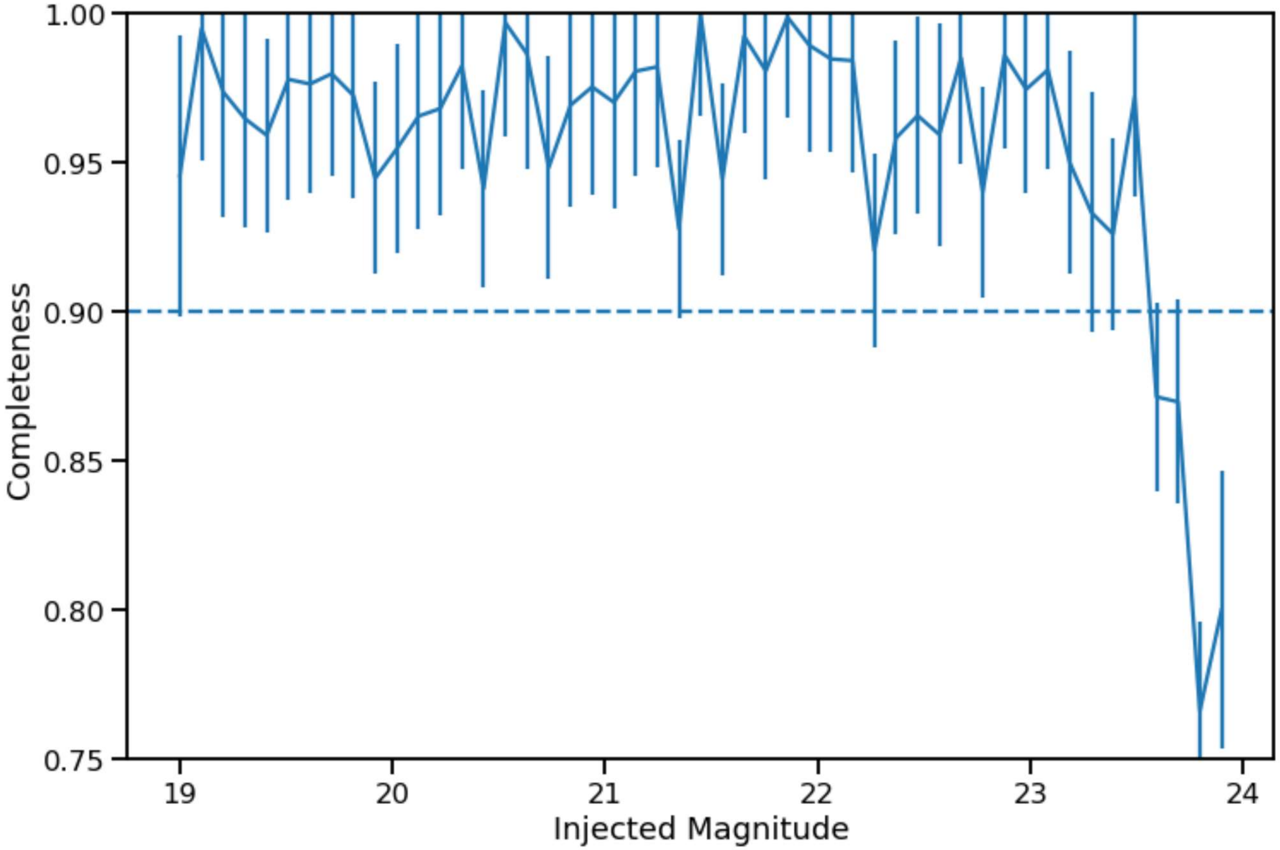
LSE-30 (rel18.0)

Latest Revision: December 5, 2019

This LSST document has been approved as a Content-Controlled Document. Its contents are subject to configuration control and may not be changed, altered, or their provisions waived without prior approval. If this document is changed or superseded, the new document will retain the Handle designation shown above. The control is on the most recent digital document with this Handle in the LSST digital archive and not printed versions.

http://ls.st/LSE-030

We measure completeness by injecting fake point sources.

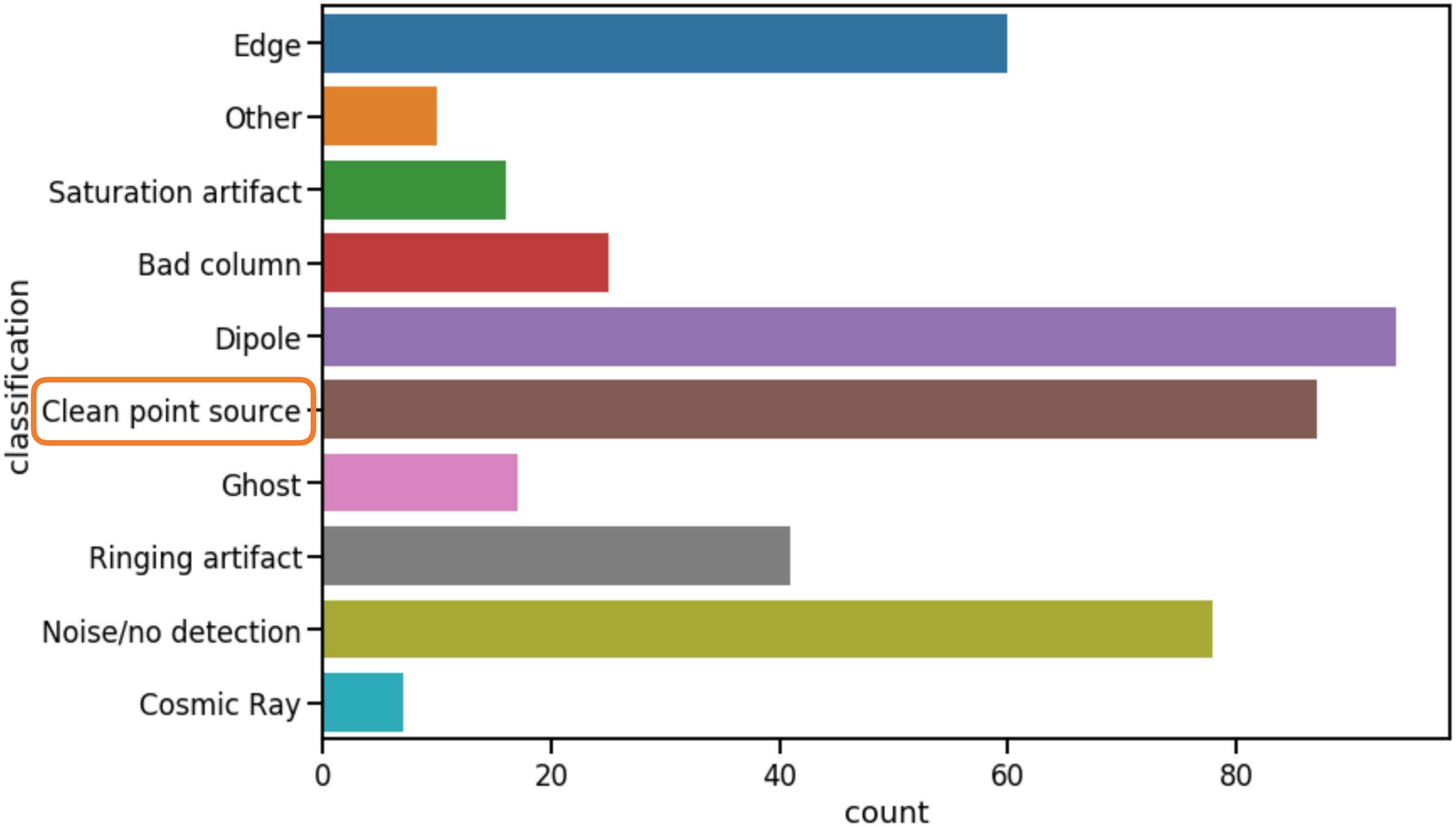


HiTS science image injections

transCompletenessMin

We measured purity by manually classifying several hundred random diaSources from the most recent HiTS processing run.

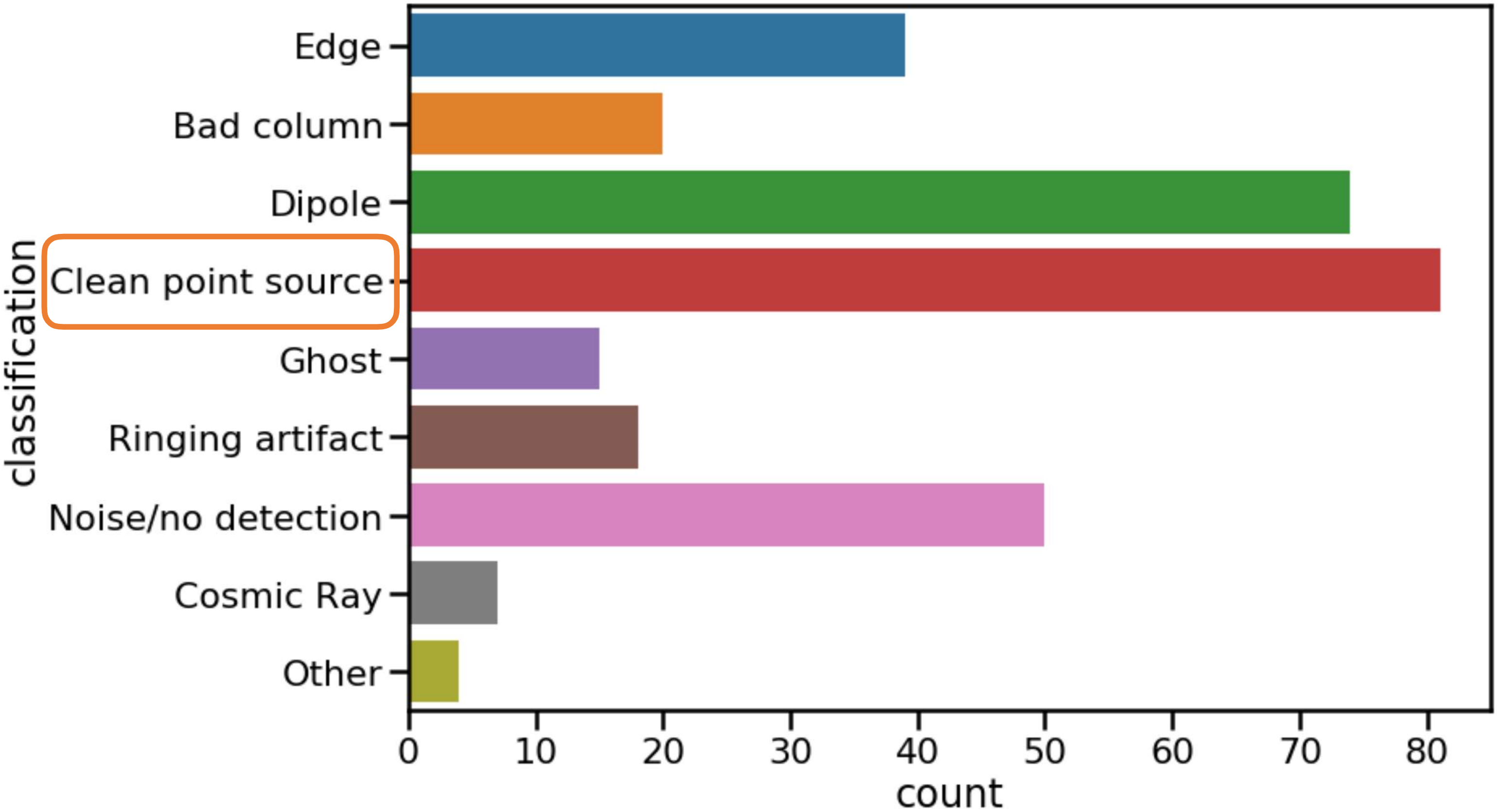
Rubin Observatory



raw detections

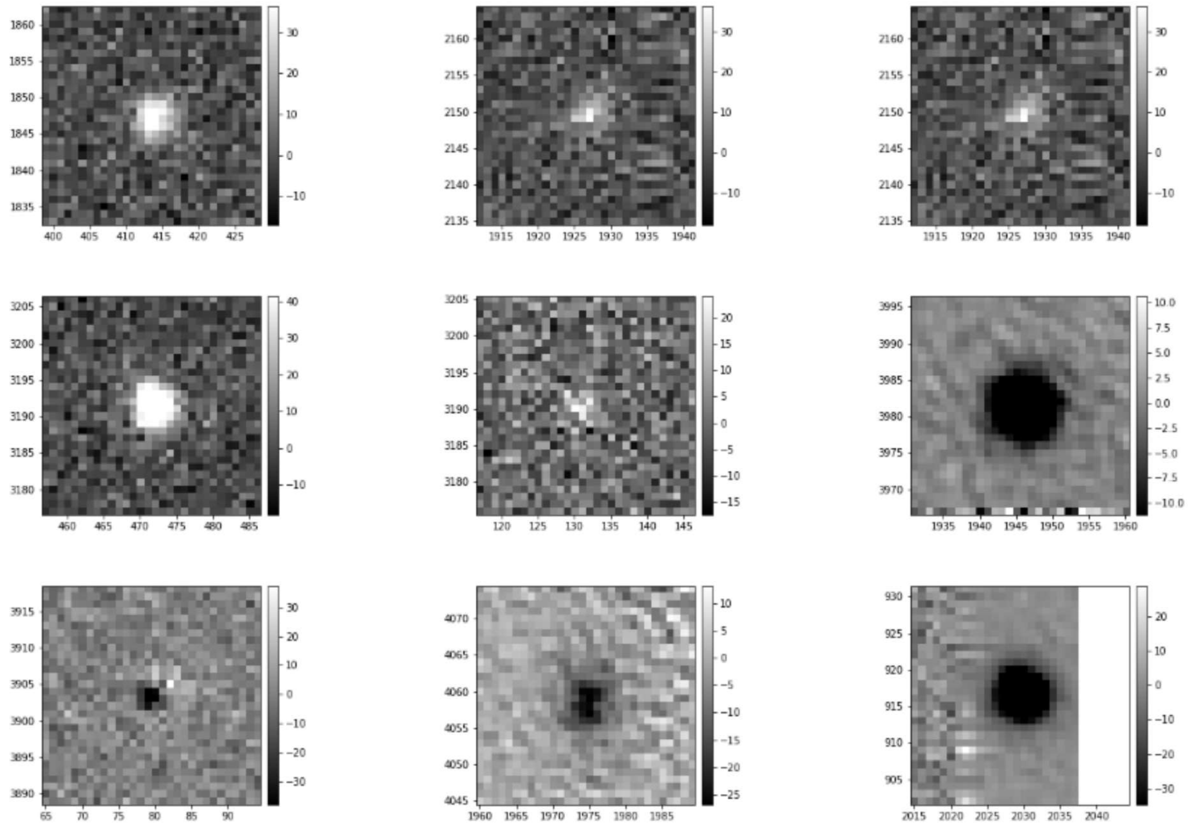
Rubin
Observatory

We measured purity by manually classifying several hundred random diaSources from the most recent HiTS processing run.

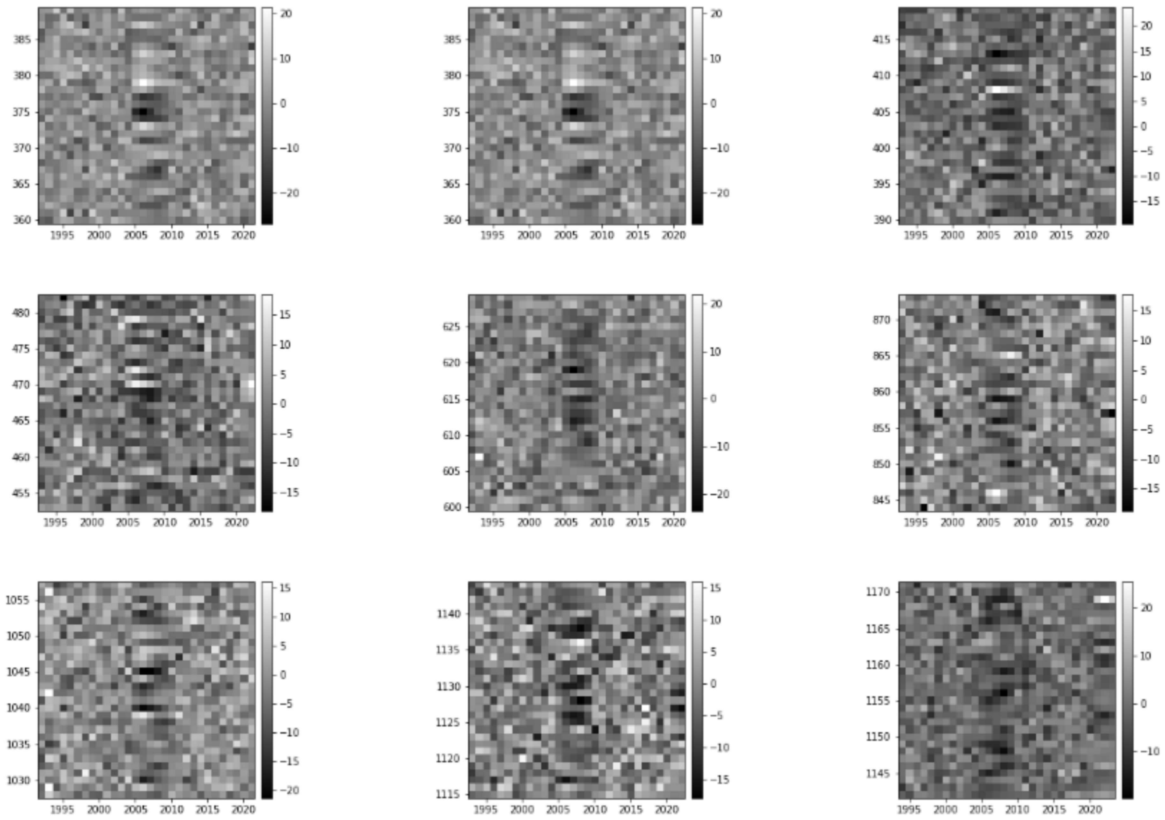


filtered for edge and saturation flags

There are lots of solid detections…

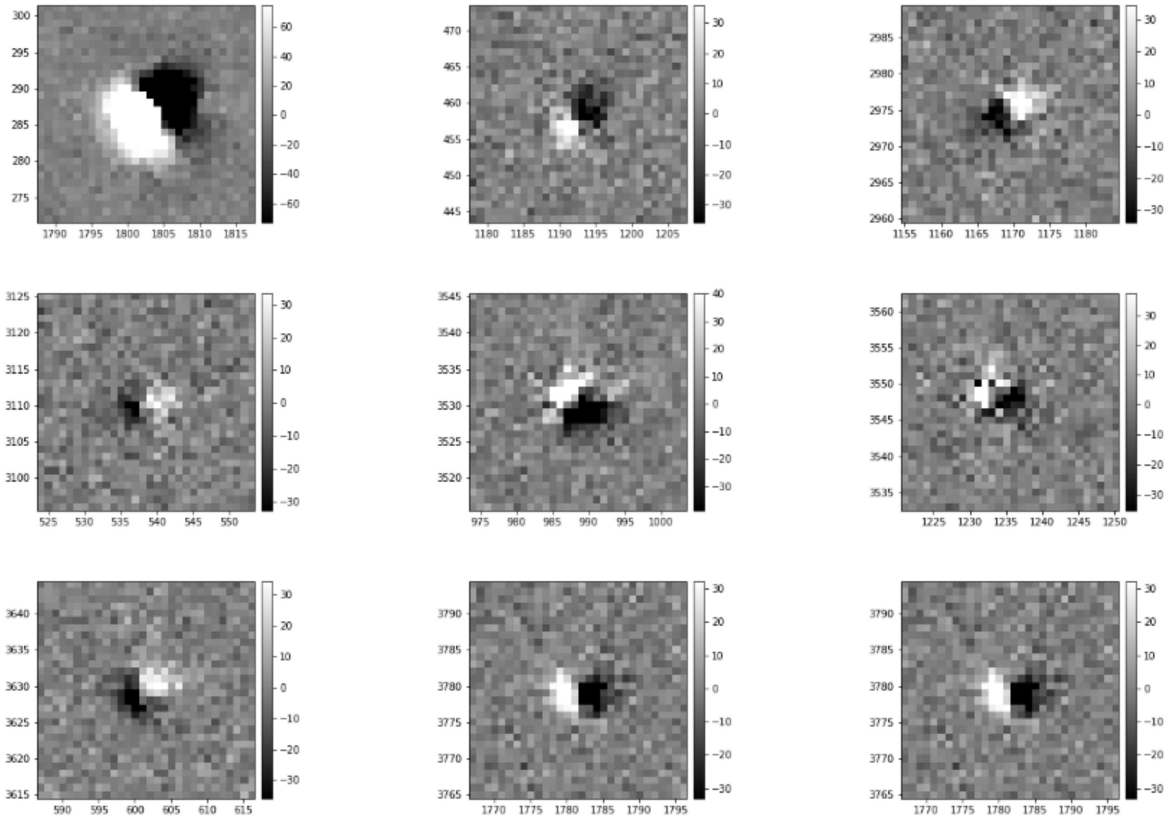


Detections on interpolated bleed columns…



address by masking appropriate flags

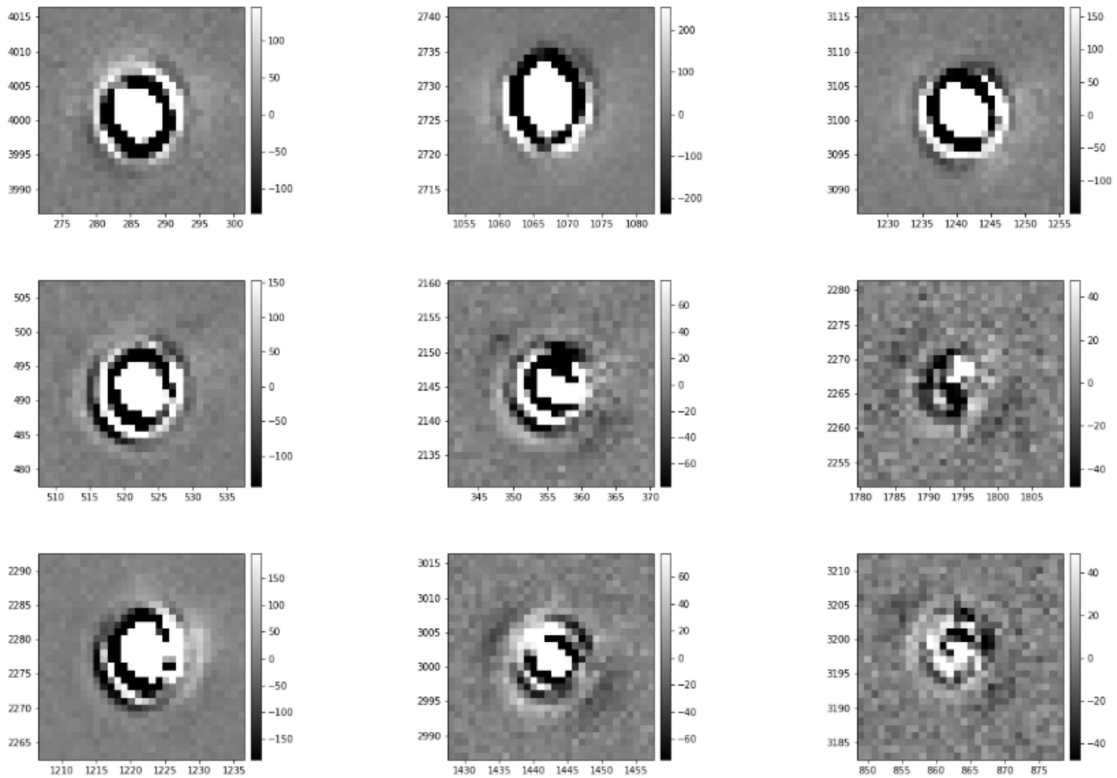
Dipoles...



address by improving astrometric calibration

Kernel deconvolution artifacts…

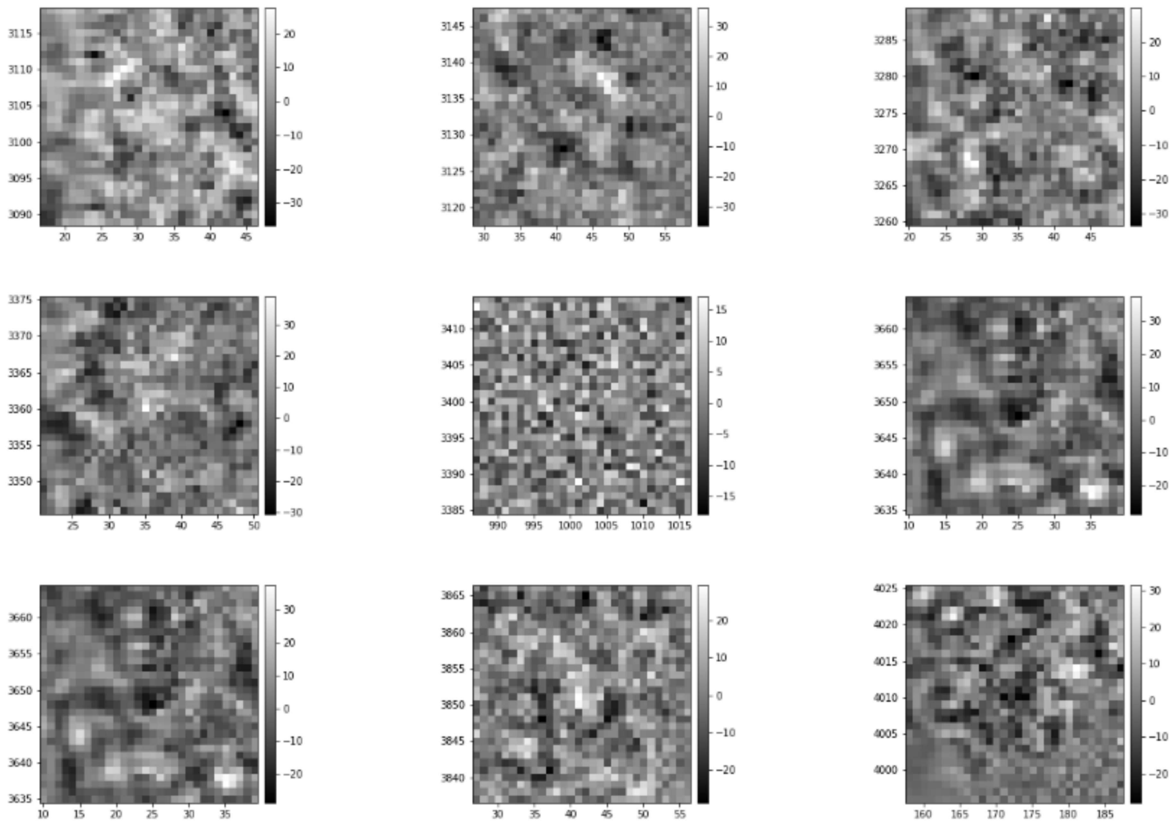
Rubin Observatory



address with algorithmic improvements to differencing

…and noise detections.

Rubin Observatory



debug potential problems with variance planes

How do the raw detection rates compare to other surveys?

Rubin Observatory

PS1 (A&L):

- raw artifact/non-artifact ratio of 10:1-50+:1 (Denneau+13)

DES Y1 (A&L; HOTPANTS):

- raw artifact/non-artifact ratio of 13:1 before ML (Goldstein+15)

ZTF (ZOGY):

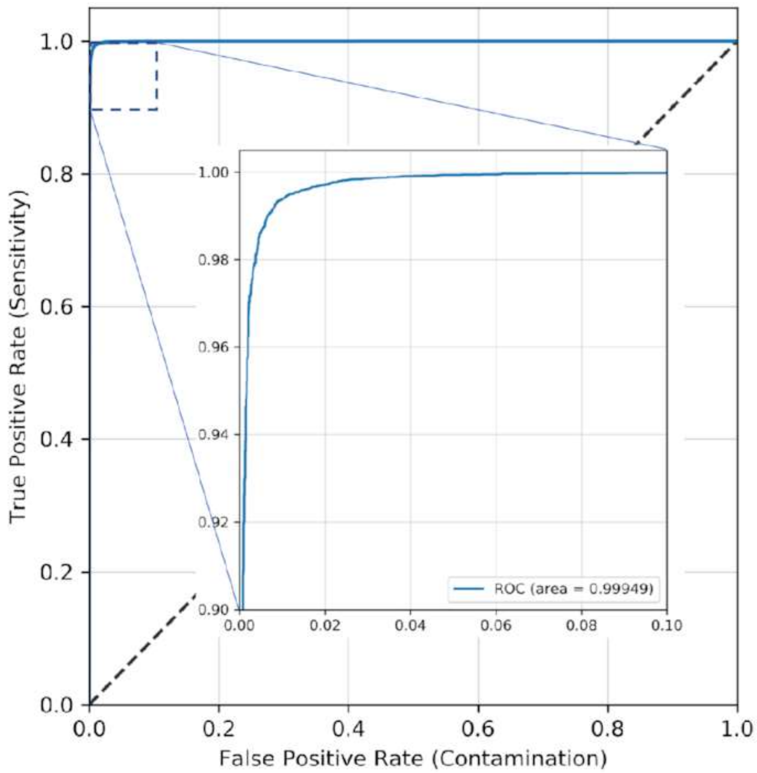
- raw artifact/non-artifact rate between 2.5:1 and 25:1

Current Science Pipelines (decorrelated A&L):

- HiTS: raw artifact/non-artifact rate ~4:1 (no filtering), ~3:1 (coarse flag filtering, minor loss of completeness)

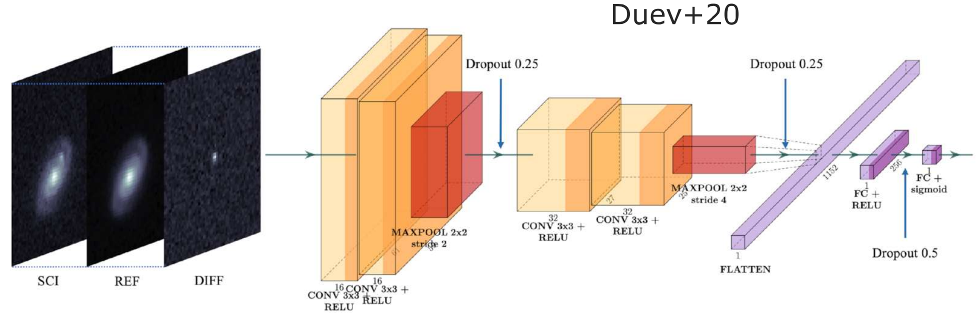
Machine Learning algorithms run on precursor surveys have demonstrated the required performance.

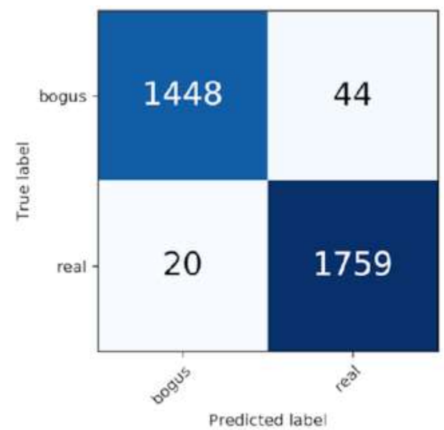
Rubin Observatory



ZTF Braai classifier

Duev+20





Completeness: 98.9%

Purity: 97.6%

We will continue to refine the pipelines on precursor data as we approach commissioning.

Rubin Observatory

Ongoing QA and performance tuning efforts

Process new precursor datasets

- Crowded fields
- larger scale runs
- DESC DC2 simulations

Push on the state of the art for difference image kernels

Production-scale fake insertion

Trailed source fitting & known SSObject attribution

Begin developing Real/Bogus infrastructure

Image Differencing is a key component of our pipelines; algorithmic development and evaluation are ongoing.

Rubin Observatory

- Results are sensitive to effects at all stages of data reduction
- We have flexible pipelines with implementations of the major algorithms.
- We are performing studies with data with precursor survey and commissioning data to choose algorithms and tune parameters for best performance.
- Initial completeness and purity results are promising, although testing on data representative of the full LSST survey is needed.