Status of AP Image Differencing

Eric Bellm
LSST DM Alert Production Science Lead
University of Washington/LSST

AP requires control (or at least knowledge) of the false positive rate.

3.1.5.1.1.7.7 Difference Source Spuriousness Threshold - Transients

ID: OSS-REQ-0353

Specification: There shall exist a spuriousness threshold *T* for which the completeness and purity of selected difference sources are higher than **transCompletenessMin** and **transPurityMin**, respectively, at the SNR detection threshold **transSampleSNR**. This requirement is to be interpreted as an average over the entire survey.

This specification will be tested using simulations, by insertion and recovery of artificial sources, and comparisons to ground truth where known (i.e., asteroids, known variable stars, known variable quasars, etc).

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

AP false positive rates are increasing with greater realism.

2013-2014: simulation studies of false positive rates, DCR

- Becker, Krughoff, Connolly; DMTN-069, DMTN-070
- *FPR within a few percent of theoretical estimates w/o DCR effects*

2014-2015: tests on DECam instcals

- Slater, Juric, Ivezic, Jones: DMTN-006
- *Detect DIASources at 100x theoretical FP estimates (correlated noise).* ***Note not all are false positives!***

2016: ZOGY & decorrelated A&L on DECam instcals

- Reiss & Lupton: DMTN-021
- *Improved noise handling reduces FPs by 10x*

2017-present: decorrelated A&L on DECam HiTS dataset

- Bellm, Rawls, Kovacs, & the ap_pipe team
- *False positive rate increases by several factors relative to DMTN-021*

Several versions of image differencing are implemented in the stack.

A&L

- “standard” (although no implementation of pixel basis)
- with decorrelation afterburner
- w/ preconvolution

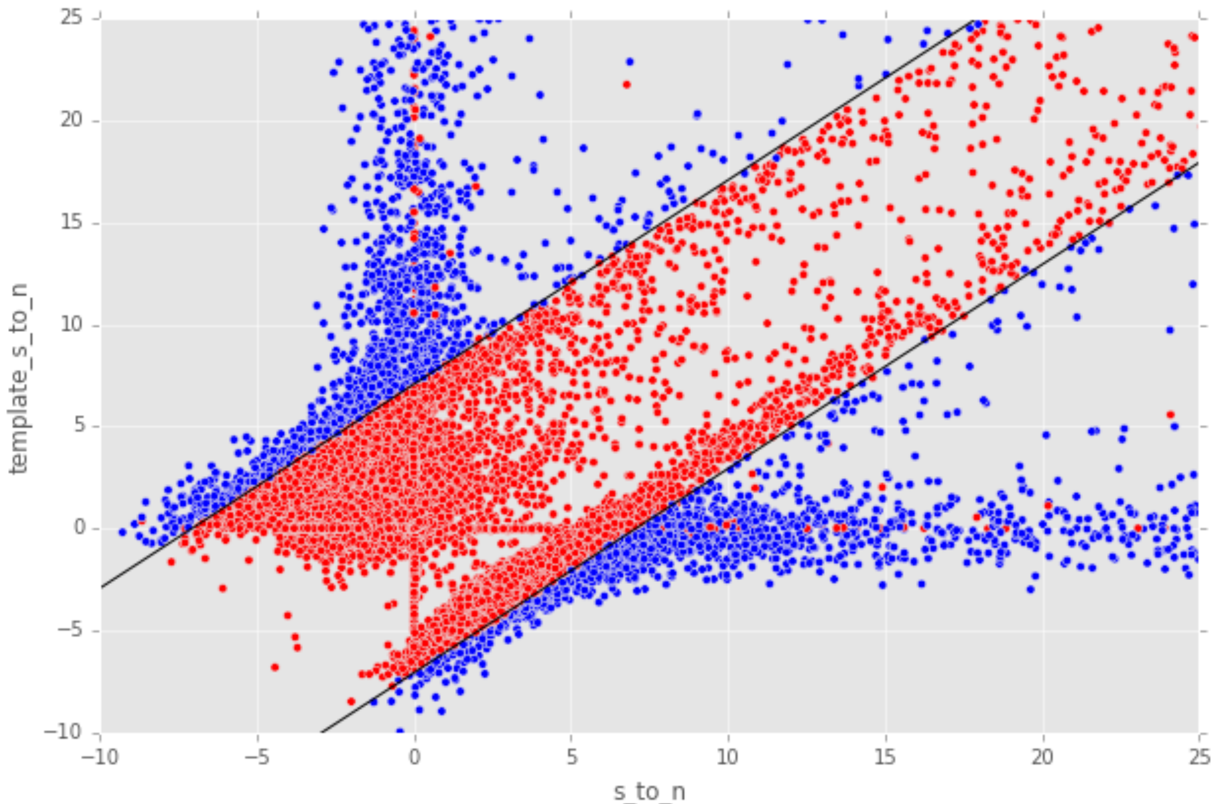
ZOGY

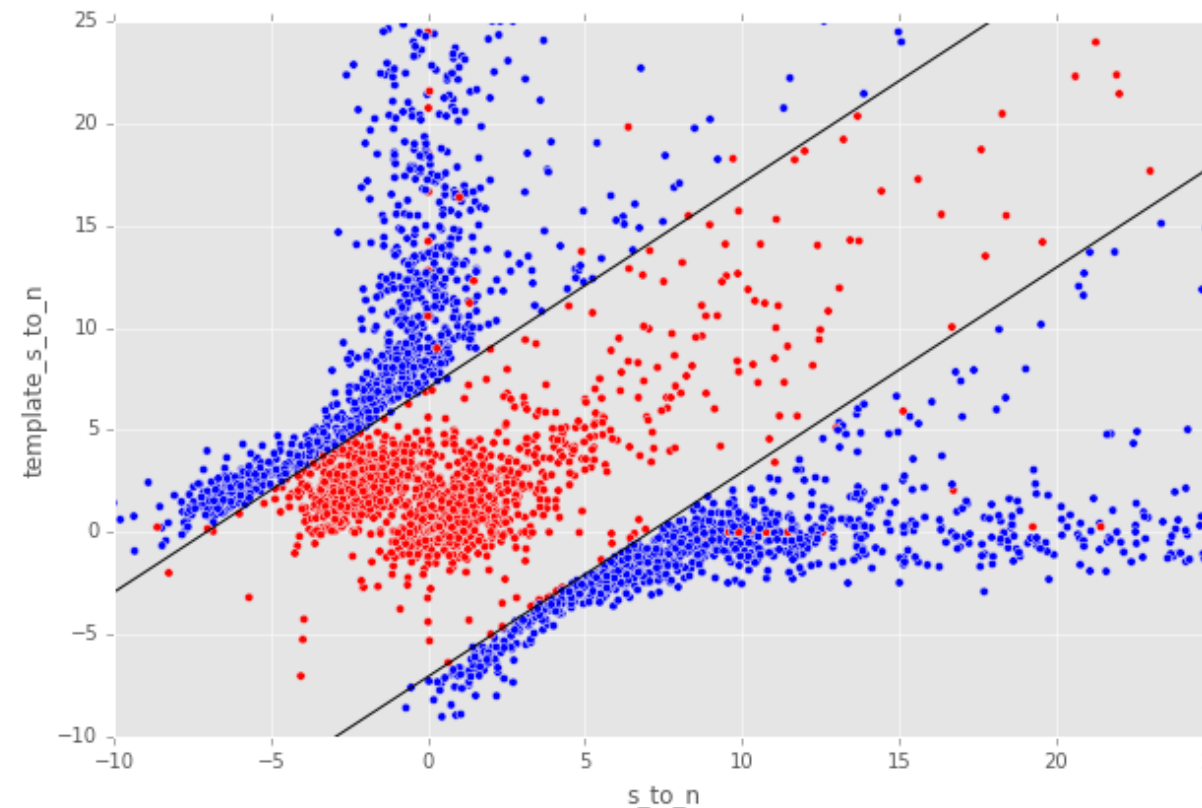
- image domain
- Fourier domain
- “Image map-reduce” (Reiss) for variable PSFs

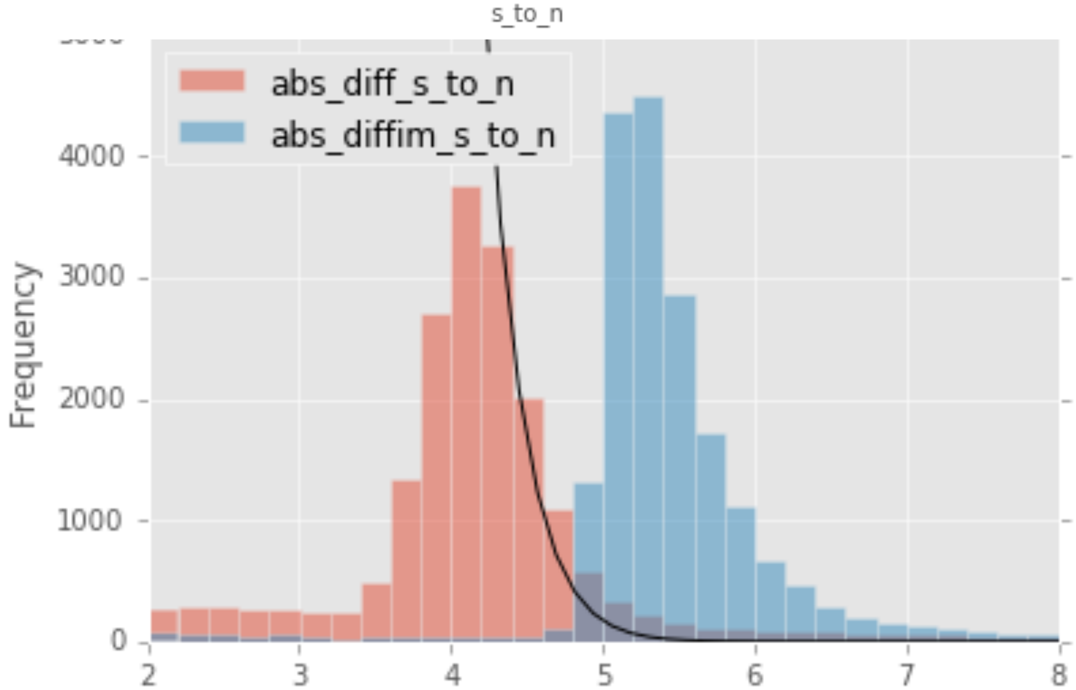
ip_diffim code does not fully conform to modern stack conventions; orphaned 2017-era Reiss pull request to restructure code and fix bugs

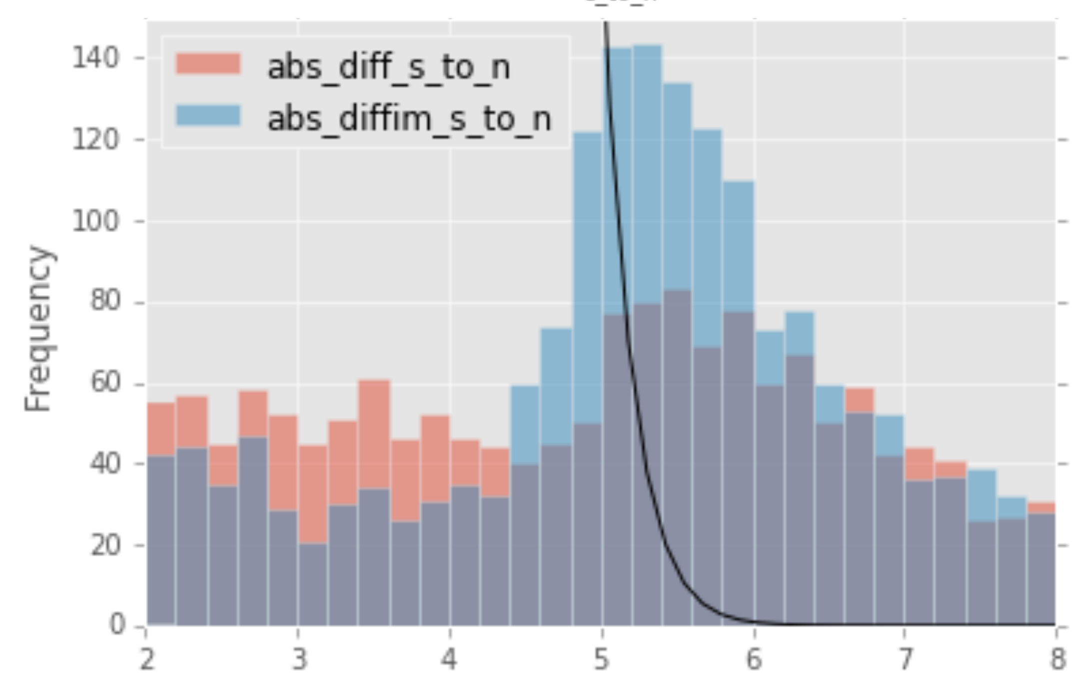
Performance shootout on hold

DMTN-021 provides the most useful false positive reference point, but it's not a perfect comparison.









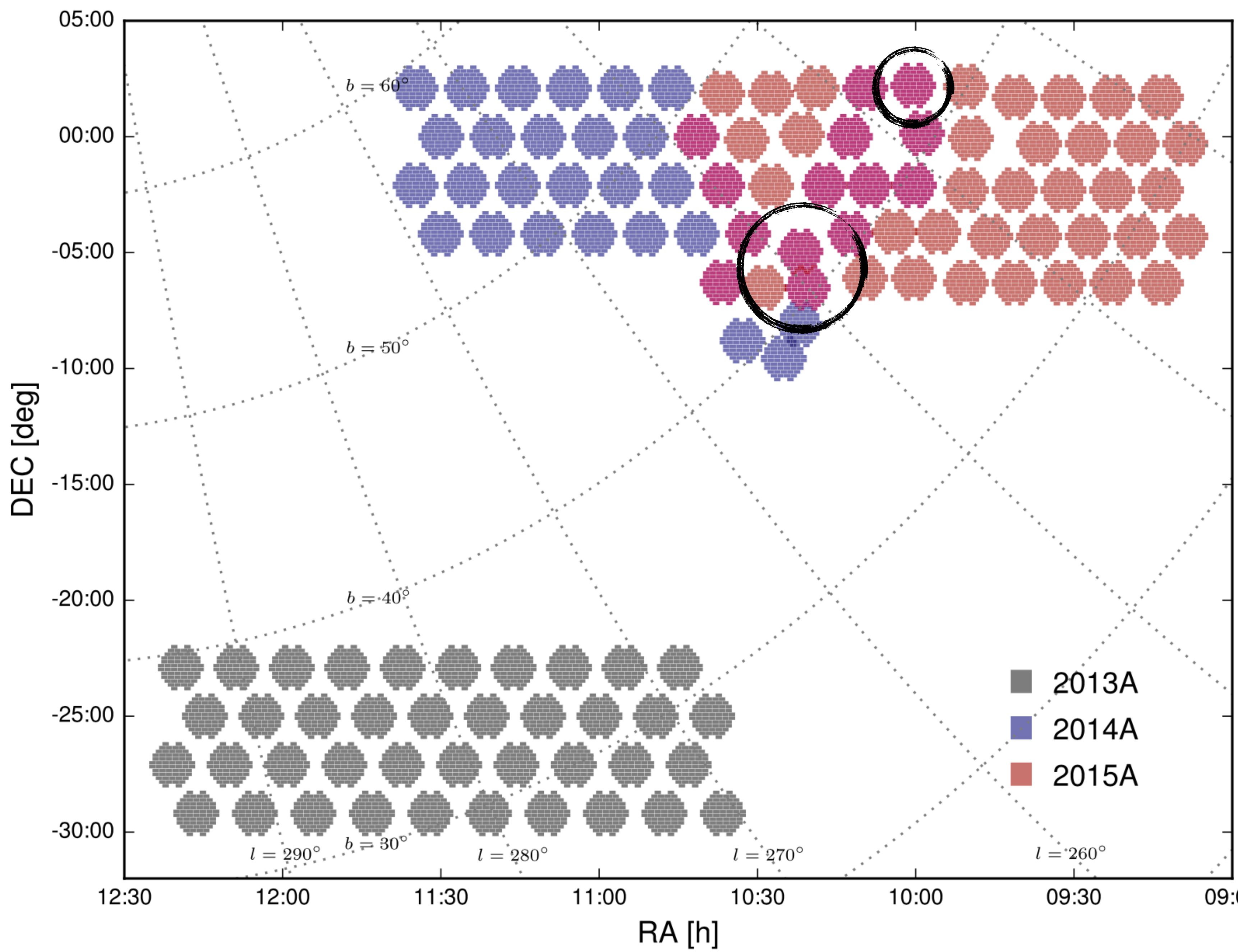
But: used DECam data processed by DECam community pipeline ("instcals"), only did image-image differencing.

ap_pipe development has focused on the DECam HiTS dataset.

Three fields

coadd templates from 2014 observing

~30 g-band epochs



ap_pipe & ap_association provide increasingly realistic LSST data products and processing.

differencing vs. historical coadd templates

DIAObjects with basic timeseries features

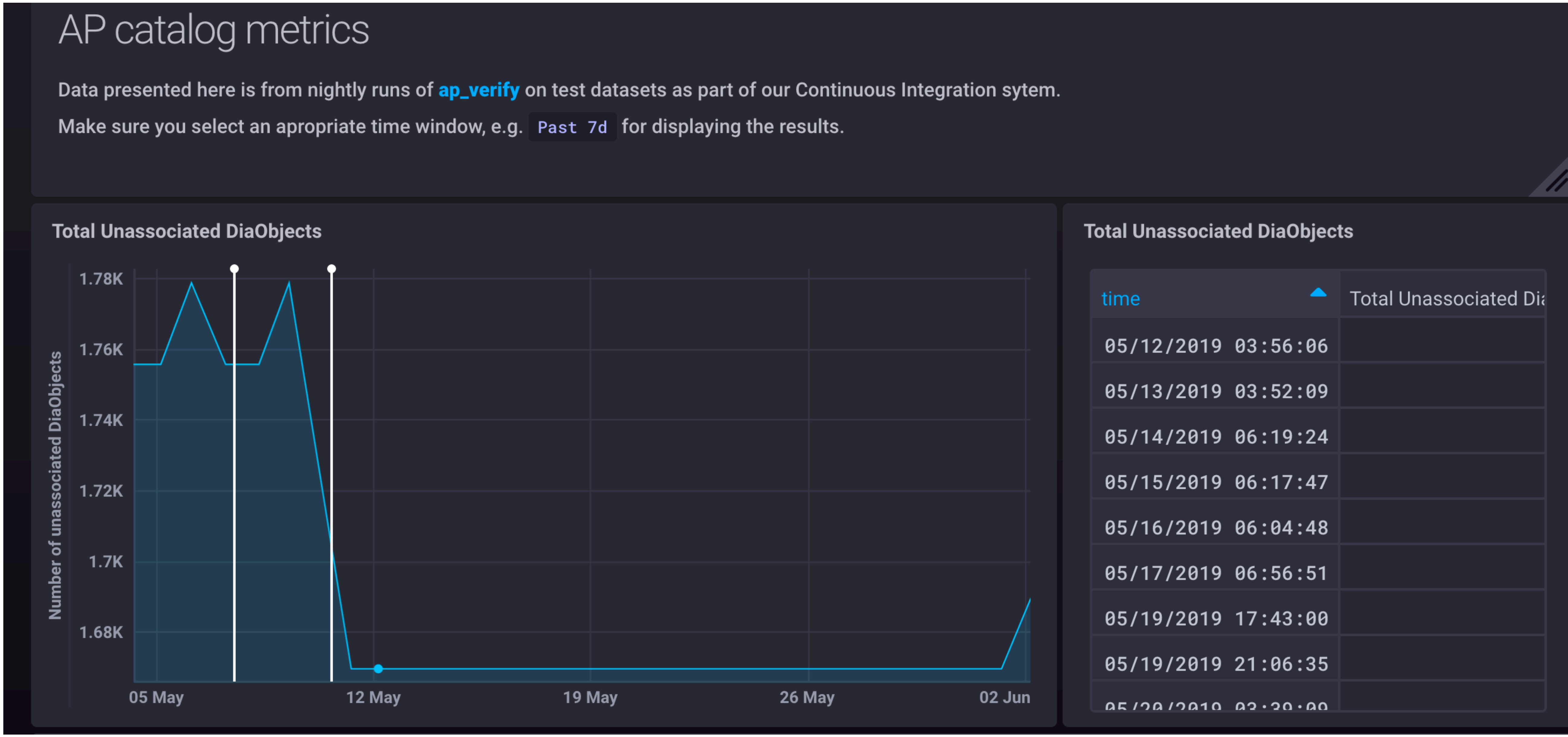
PPDB integration

still to come:

- more DPDD fields
- alert packets
- SSObject attribution
- Precovery forced photometry

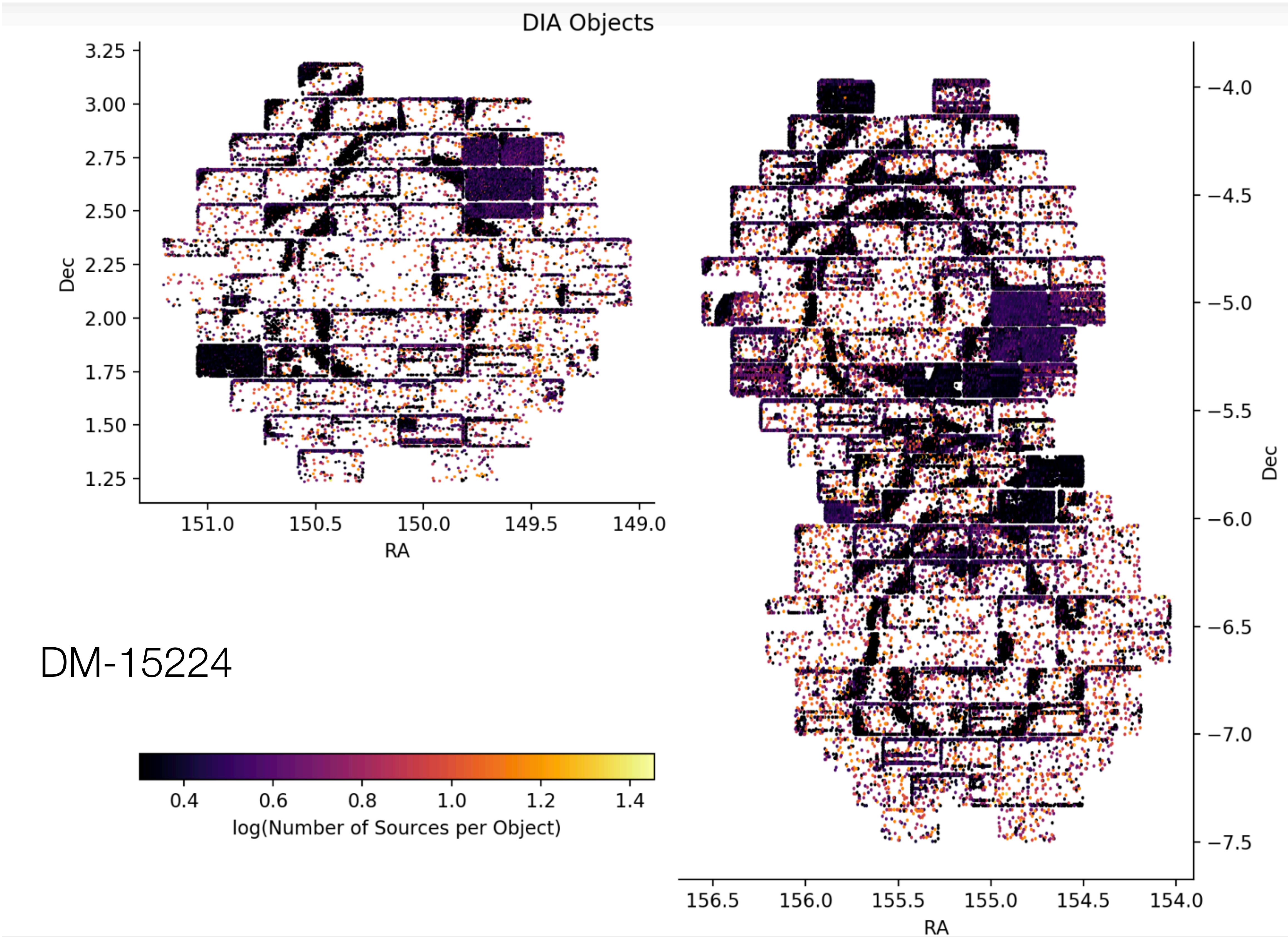
We regularly ap_pipe performance and use it to guide our development and debugging.

Weekly examination of runtime and performance metrics on a CI subset through SQuaSH (via ap_verify)



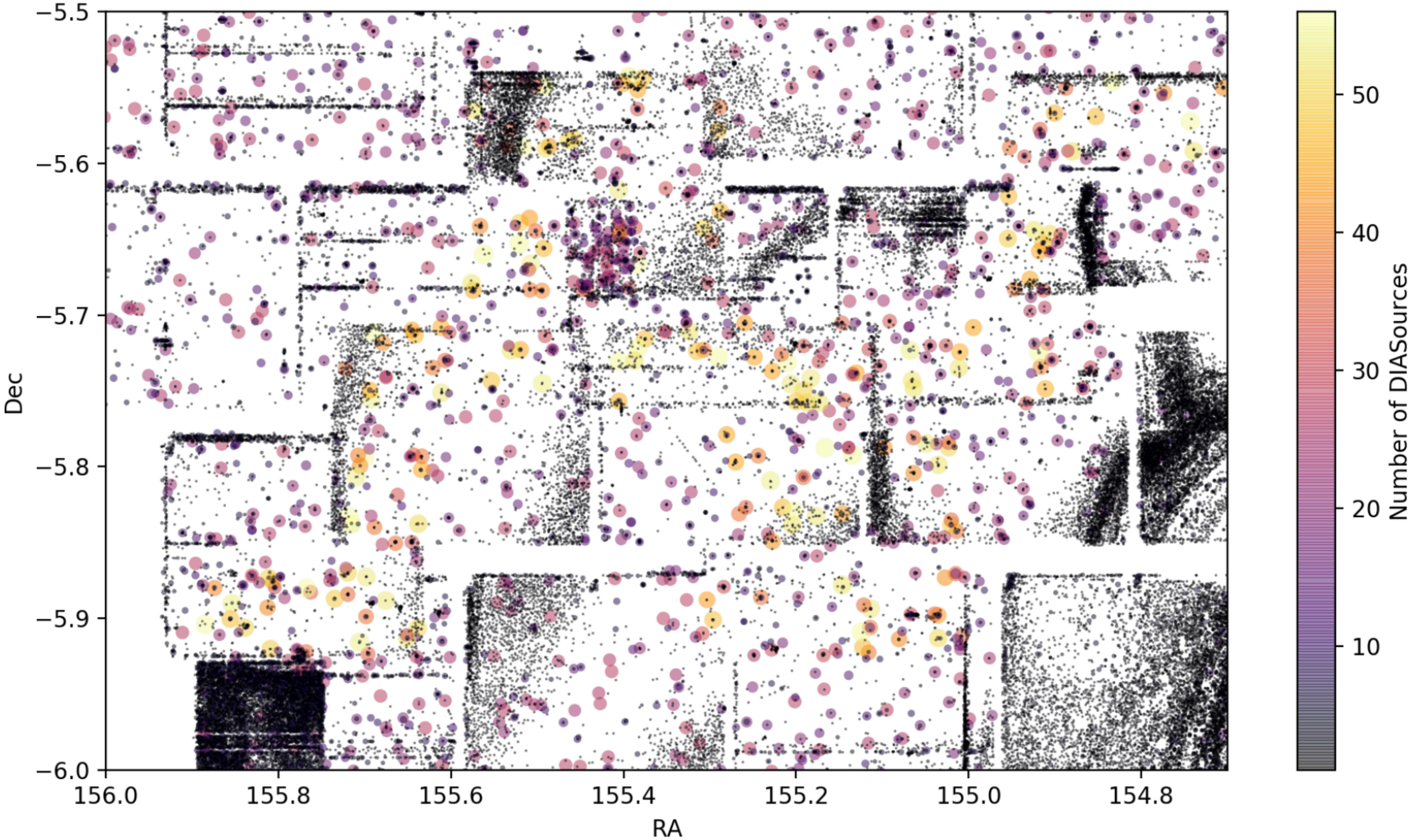
Monthly reruns on the complete HiTS dataset are ~manually examined to identify problems through the Science Platform/notebooks.

While ap_pipe is increasingly feature complete the image differencing performance is not great.

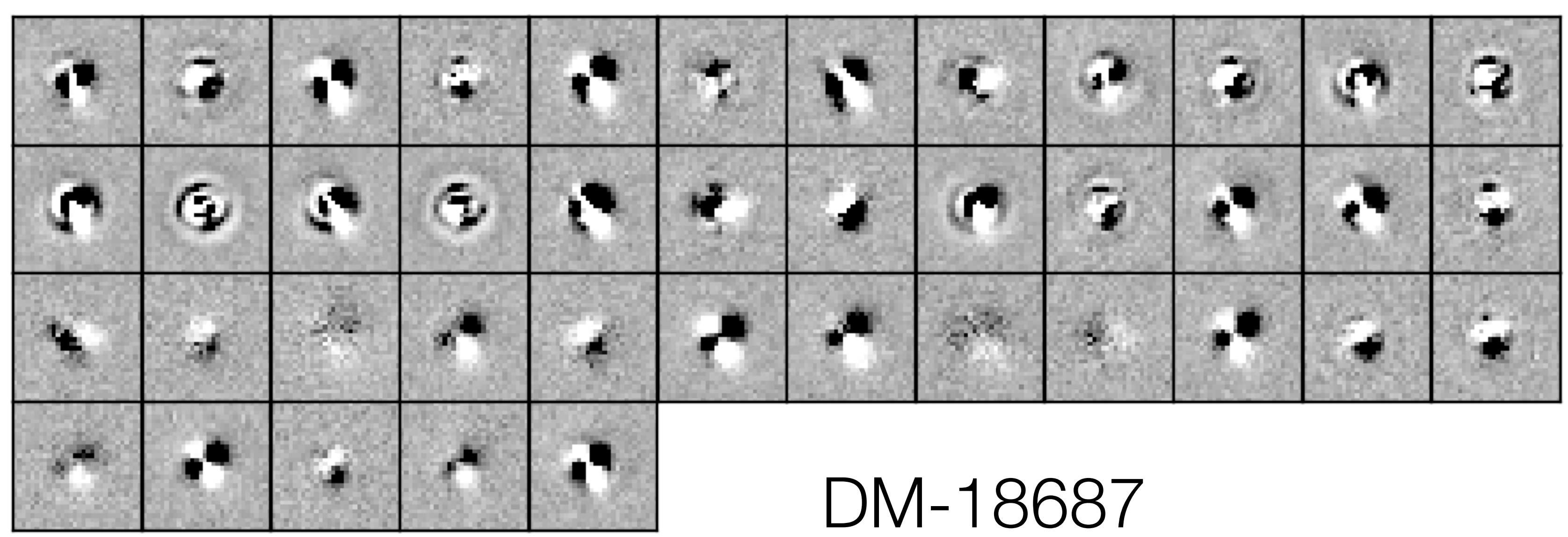


DM-15224

While ap_pipe is increasingly feature complete the image differencing performance is not great.



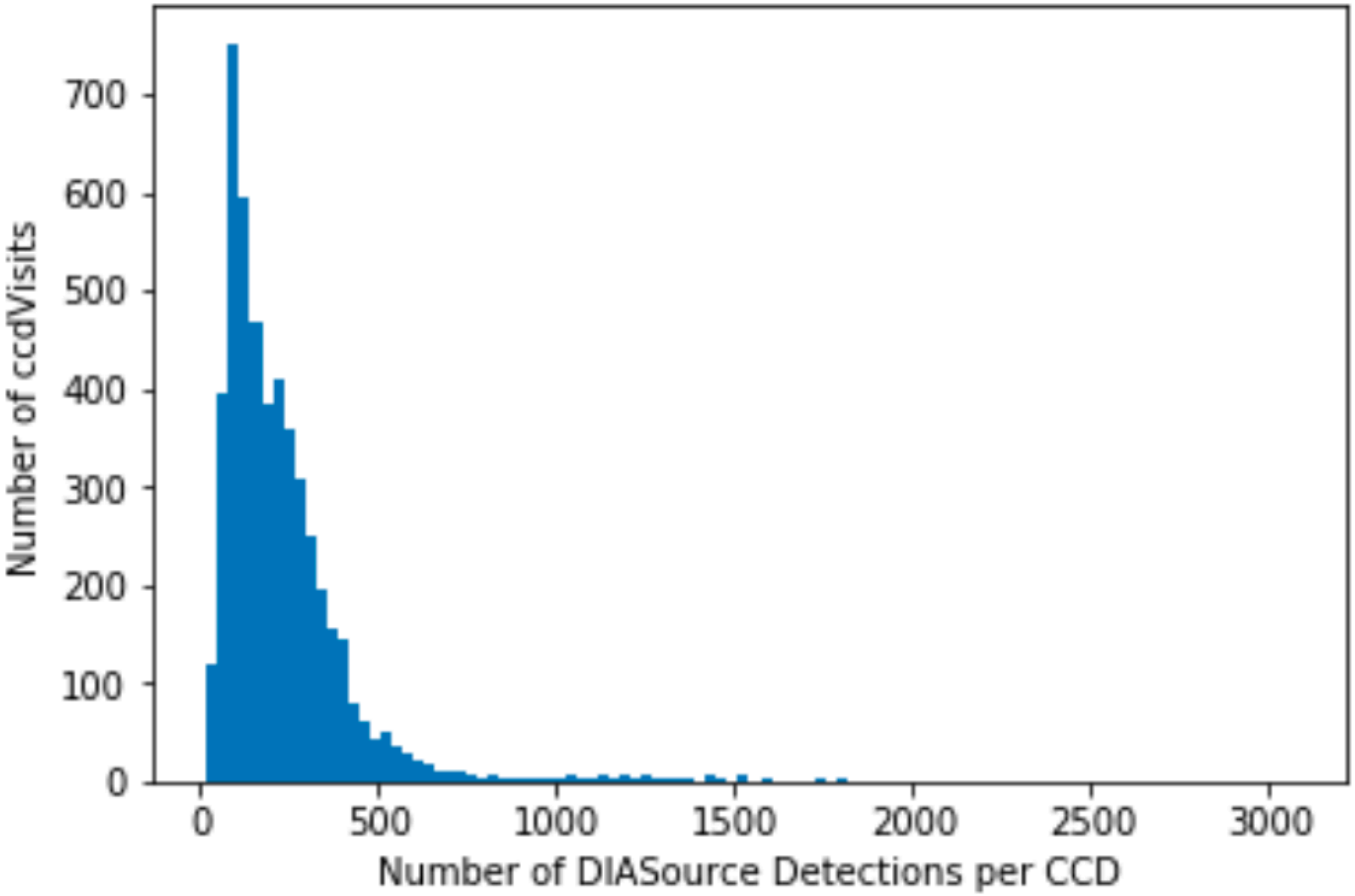
DM-15224



DM-18687

While ap_pipe is increasingly feature complete the image differencing performance is not great.

DM-19292



Number of noise detections per CCD at 5.0 sigma: ~1.5

DMTN-021 detections per CCD: ~50

Lots of problems may be contributing.

~~Bad photometric calibration~~

~~Lack of illumination correction~~

~~Bad background subtraction~~

Edge effects—need to use flags.

Overfitting kernel basis functions?

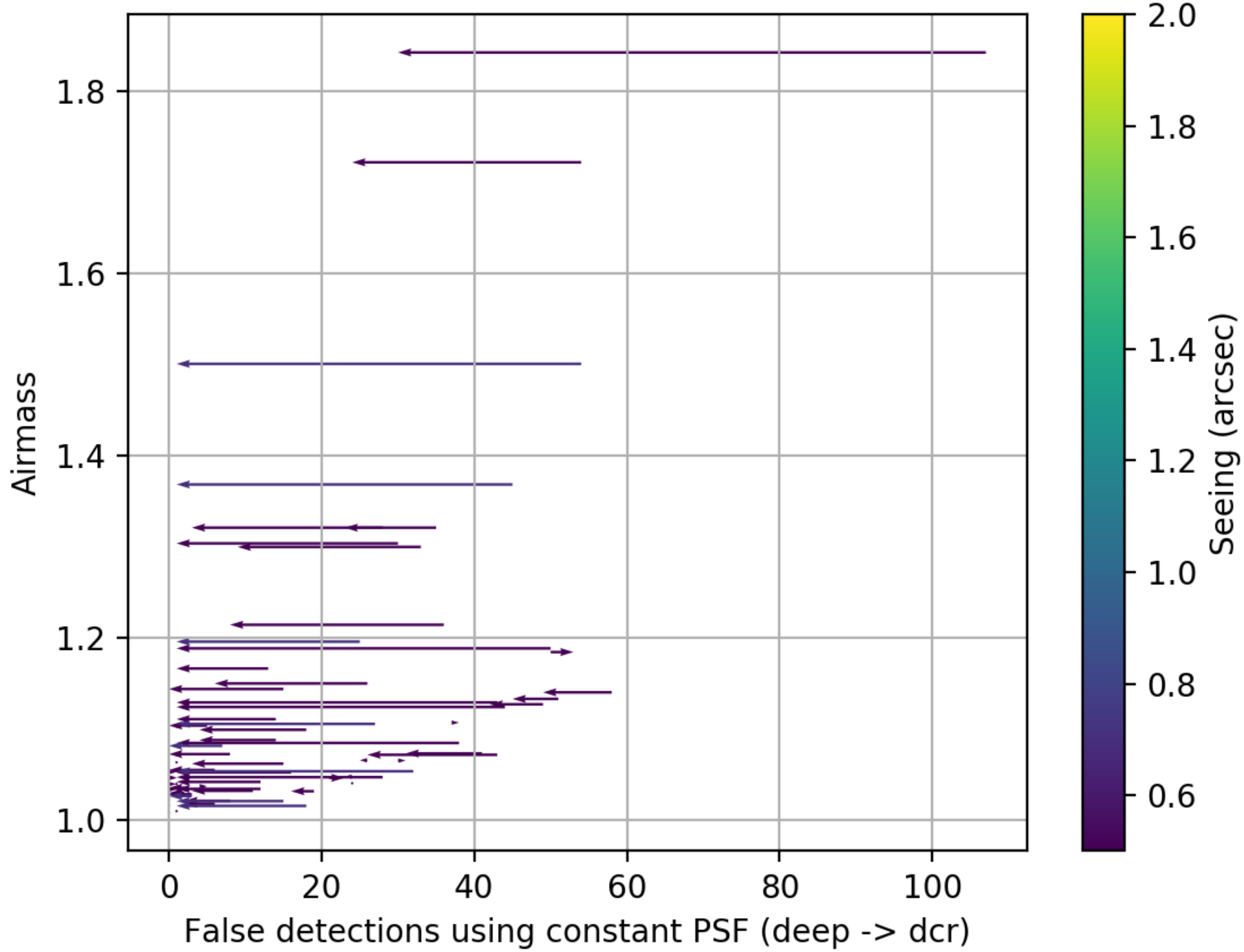
Deconvolution?

Use of DECam CP calibs

DCR modeling is integrated in the stack and being tested on real data.

We see a large reduction in DIASource detections with DCR model coadds.

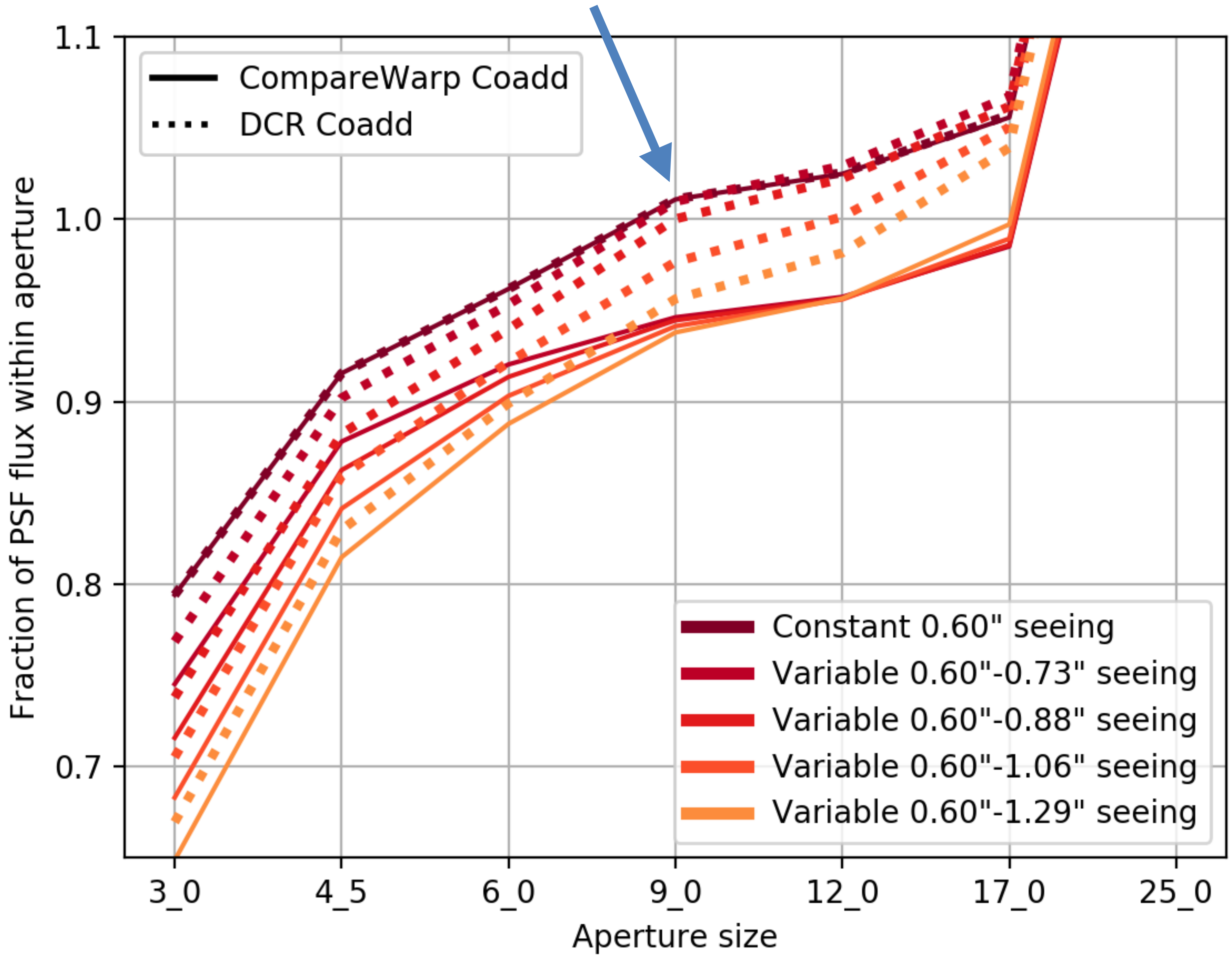
DM-17528



The current DCR algorithm degrades with variable seeing.

Sources in DCR coadds are more spatially compact than CompareWarp, but that degrades as the range of seeing included increases.

Moreover there are unexplained trends in DIASource detections as the input seeing range increases.



It is challenging to get enough u&g band exposures to train the DCR models in the early survey.

We are trying to assess over what range of seeing the current constant seeing algorithm provides useful improvement. In principle if the DCR model were built with a small range of seeing the current pixel-based model will be adequate.

But there just aren't that many exposures to play with: current OpSim runs give something like 4-6 (TBC) g-band exposures in Y1. This seems challenging even with a hypothetical variable seeing algorithm.

Moreover the current simulations give only a handful of airmass and parallactic angle values per field, which hinders model training.

What about Real-Bogus?

Almost certainly we'll need machine learning to hit completeness & purity requirements

Analysis required is the same even if we didn't deploy the model

Random Forest RB is straightforward to implement; likely future sprint. May aid in debugging diffim at scale.

Clear performance improvements from deep learning architectures.

Constructing and maintaining training sets through commissioning and ops is a large & complex task; *fakes alone won't be enough*.

We are doing a False Positive Sprint in June 2019.

Bellm, Rawls, Kovacs, Al Sayyad (but all < 1 FTE)

False positive census

- thorough & systematic accounting of DIASource detection rates by visit, CCD, night, seeing, flag, etc. to

DMTN-021 replication

- repeat DMTN-021 (instcal differences of Lori Allen data)
- iteratively add new elements to improve comparison to current processing
  - HiTS data (instcal differences)
  - HiTS data (CP calibs, stack ISR, calexp differences)
  - HiTS data (coadd differences)

Plus other elements as identified

Future directions

expand to HSC data, larger DECam datasets

stack-built DECam calibs

fake injection

algorithm shootout

Real-Bogus

variable PSF DCR development