

BALROG for DES DR1 (Years 1-3):

**Digging deep to understand biases and
completeness in accurate
Measurement of galaxies in imaging data.**

Brian Yanny

Fermilab, DES experiment

March 14, 2019 Webinar

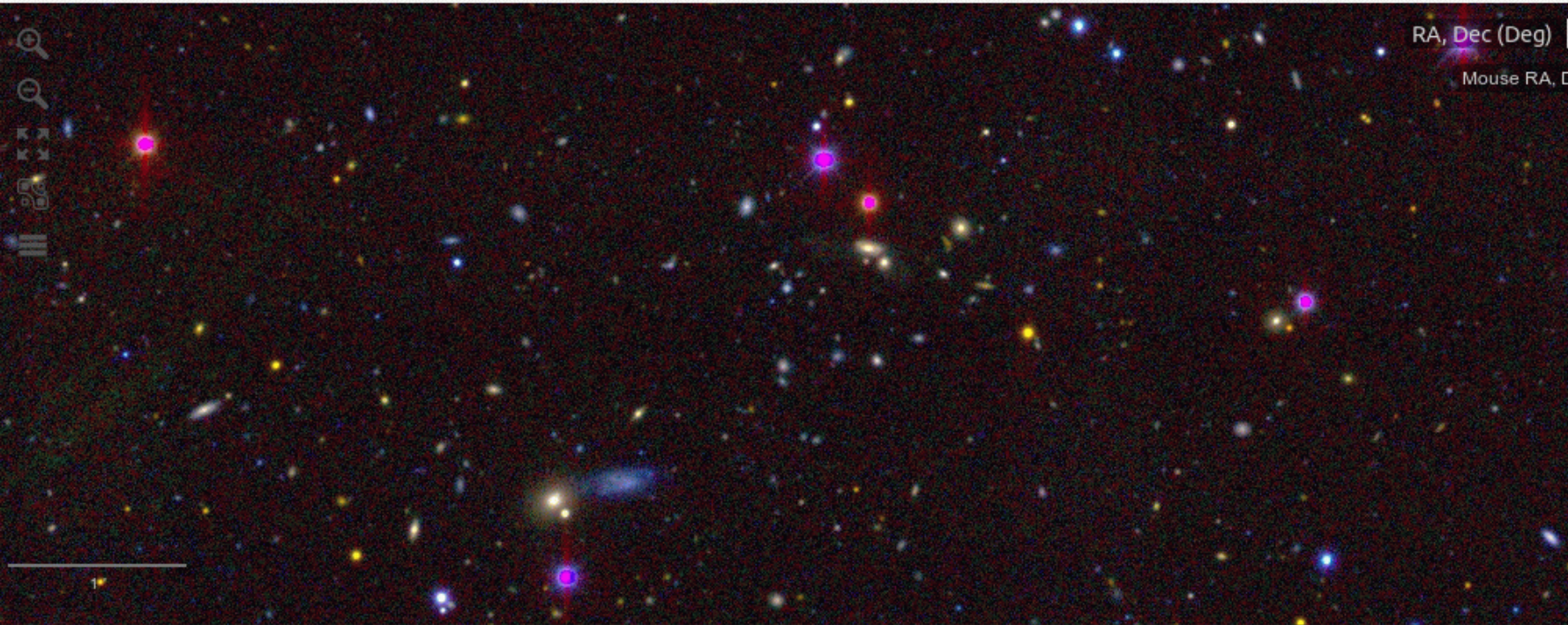
What is DES DR1: Images in grizY optical and near infrared bands of 1/8 of the sky (5000 sq deg) to magnitude of ~24th at ~10 sigma, with a catalog of ~320M Galaxies and ~80M stars.

The screenshot shows the Sky Viewer interface. The browser address bar displays the URL: <https://desportal2.cosmology.illinois.edu/dri/apps/sky/#sky/23>. The application header includes the "Sky Viewer" logo, a search bar with "RA, Dec (Deg)" selected, and a user name "yanny". On the left, a sidebar contains navigation icons and a data panel for the selected region:

- DR1
- DES0127-3914
- RA, Dec (deg): 22.30757, -39.30973
- Mouse RA, Dec (deg): 307.71401, -5.40093

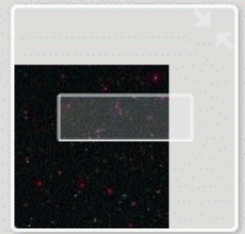
The main view is a circular field of view (FoV) of 180 degrees, showing a dark field with a white outline of the survey region. A purple crosshair is centered on the field. The bottom right corner features the ALADDIN logo.

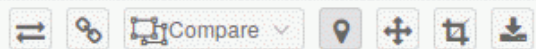
DR1 - All - DES0239+0126



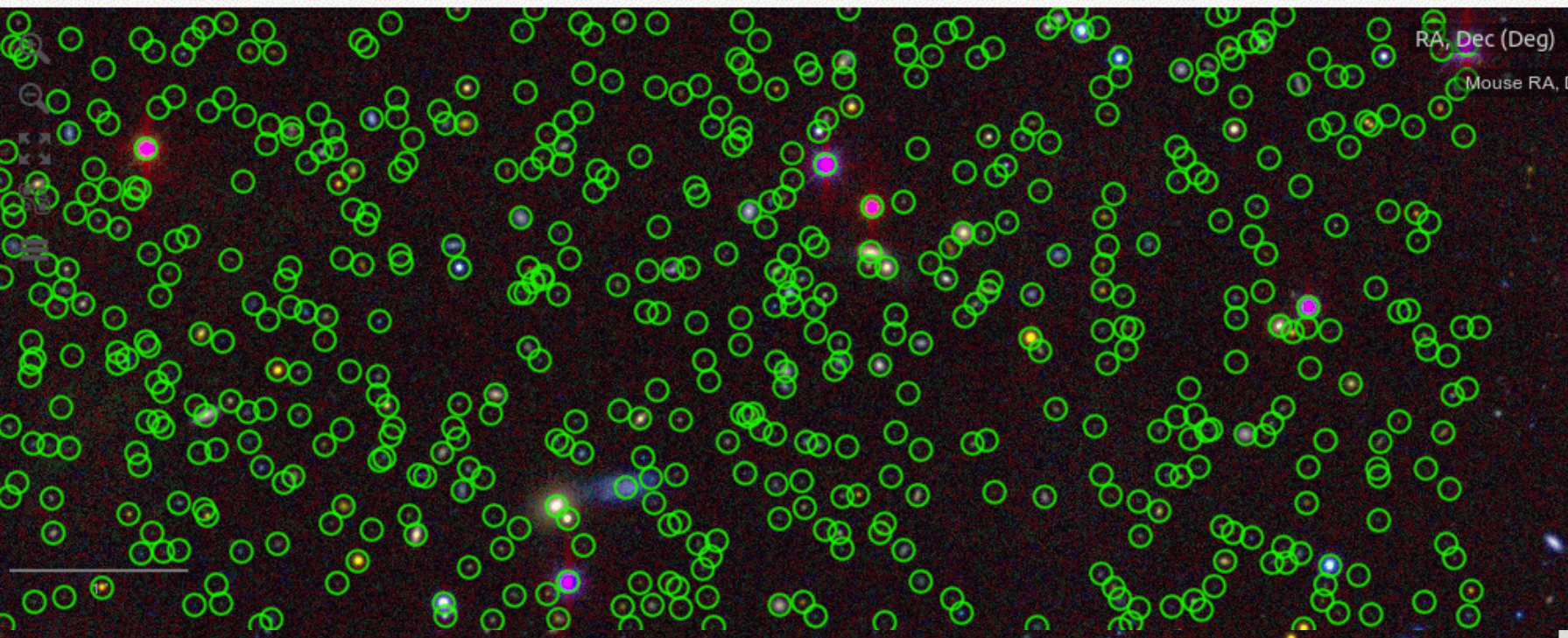
RA, Dec (Deg) 39.66432, 1.74007

Mouse RA, Dec: 39.69619, 1.74819



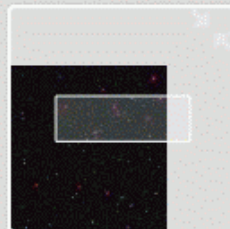


DR1 - All - DES0239+0126



RA, Dec (Deg) ▾ 39.66432 , 1.74007

Mouse RA, Dec: 39.66096, 1.76878



What is DES DR1: Images in grizY optical and near infrared bands of 1/8 of the sky (5000 sq deg) to magnitude of ~24th at ~10 sigma, with a catalog of ~320M Galaxies and ~80M stars.



DR1
DES0127-3914
RA, Dec (deg): 22.30757, -39.30973
Mouse RA, Dec (deg): 307.71401, -5.40093

FoV: 180°

Powered by LineA | Dark Energy Survey | NCSA

ALADDIN

BUT What are the errors on the objects Really?? What is the detection limit and completeness, Really?

How do the errors scale with survey depth, or location within the survey?



DEEP DEPTH (Coadd of 30 exps)



Why the need to run a Balrog at all, can't we roughly estimate the completeness and detection limits and errors ourselves?

We can to some extent: If we have a coadd that's 5x the number of combined the exposures than another, then we can estimate the errors will be $\sqrt{5}$ smaller in some sense, and we'll reach a depth (at the same signal to noise) that is about $2.5 \cdot \log_{10}(\sqrt{5}) = 0.9$ mags fainter. This is for sky limited observations and is approximate. Can also determine this empirically.

But biases or shifts in color and magnitude vs. what's 'true', are very very hard to estimate from first principles. Much better to empirically lay down realistic objects of known magnitude, color, shape and size and (re)measure them, and then count.

Side-bar: Measuring objects in astronomy:

Objects are generally classified into point sources (stars) and extended sources (galaxies).

Point sources, (which all look alike, being to first order scaled versions of the Point Spread Function, or PSF) are easiest to measure: One has a model of the PSF and fits for the height of the peak for each object of interest. Depending on how 'flat' the sky is, one may also have to fit the background sky level with one or more parameters. Errors for non-blended objects approach the Poisson limit: $\text{photons}(\text{object})/\sqrt{\text{photons}(\text{object}+\text{sky})}$. The PSF generally falls off like a Gaussian with radius (i.e. fast) for small radii, giving objects limited size. PSF can vary from place to place however, as can the background.

Extended sources (galaxies) are much more complicated to measure, since if one tries to add up the flux, say in an aperture, the aperture may need to be very large, as the flux falls off only very slowly with radius, while the sky variance grows as the square of the radius of the aperture. To balance the decreasing signal vs. the increasing noise, generally a model form for the galaxy profile is assumed, which optimizes S/N and allows one to estimate the total flux of the object from a fit within a small aperture. Besides Poisson noise, there is also then an error due to inaccuracies in how well the model really represents the galaxy.

The Sloan Digital Sky Survey came up with an empirical way to model galaxy profiles, using knowledge that most galaxies were spirals (exponential disk profile) or ellipticals (DeVauc profiles $r^{1/4}$ law). Their model magnitude is called the 'composite model' (CM) which is a linear superposition of a co-centered spiral and an elliptical model.

DES, in its galaxy fitting, has adopted this CM_MAG magnitude for galaxies (and it works well for stars too (with size=0), having relatively few free parameters per object: (x,y) position, CM_FLUX[4] (or CM_MAG) (the 4 is for the griz bands), CM_FracDeV (a parameter in the range 0:1 expressing the fraction of light in the deVauc component vs. the exponential disk), CM_T (deconvolved area parameter x^2+y^2), and CM_g[2] – a two component reduced eccentricity. Stars have CM_T = 0, 0 shape, 1 fracdev. There are errors (or covariances) on each of these parameters. (x,y,flux[4],size,shape[2],fracdev) = 10 parameters/object (6 for stars).

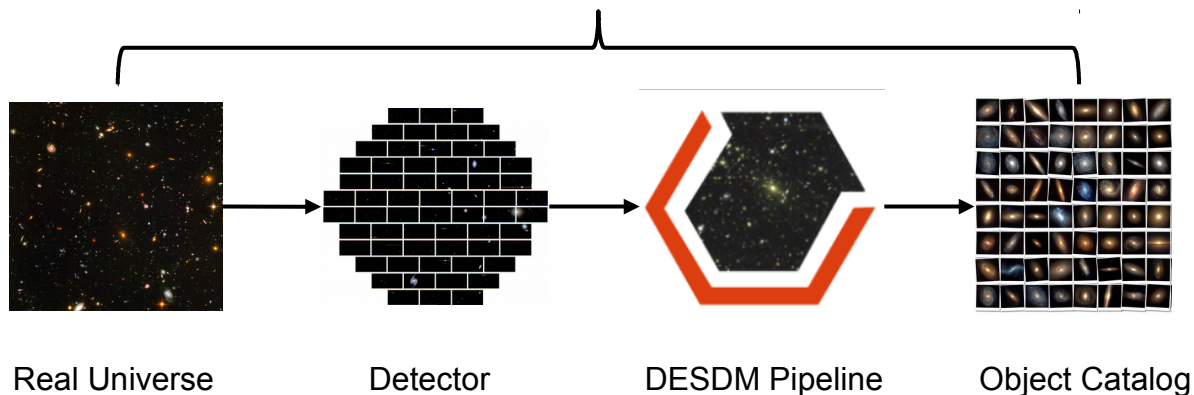
DES Balrog adopts this object model for its simulations. We use a catalog of input galaxies (and stars) taken from special deep DES coadd fields which are 1-2 mags deeper than the DES DR1/Y3 images we are studying and place them down in a hexagonal grid pattern on each exposure that make up a coadd DES tile of interest. Then the usual DES processing software is run on the augmented tile.

Balrog concept:

Characterizes selection effects and measurement bias by injecting a realistic ensemble of fake star and galaxy images in the real survey data

Balrog objects inherit difficult to model systematic effects that can vary across survey footprint

$$P(\text{catalog}|\text{universe})$$



Since some applications (photo-z biases, weak lensing biases) require huge statistics to determine small biases with high significance, millions and possibly hundreds of millions, of galaxies must be laid down and remeasured.

The first Y3 Balrog run chose ~1000 tiles from the 10,000 in DR1/Y3 and added about 11,000 galaxies and 2,500 stars to each tile (about a 25% increase over the number of objects already there). Most of the galaxies were too faint to be detected by Y3 (about 4500 were detected or about 1/3) on each tile, for about 4.5M objects which have matching input and output CM_FLUX and other parameters to compare.

**Balrog Run 1:
Footprint:**

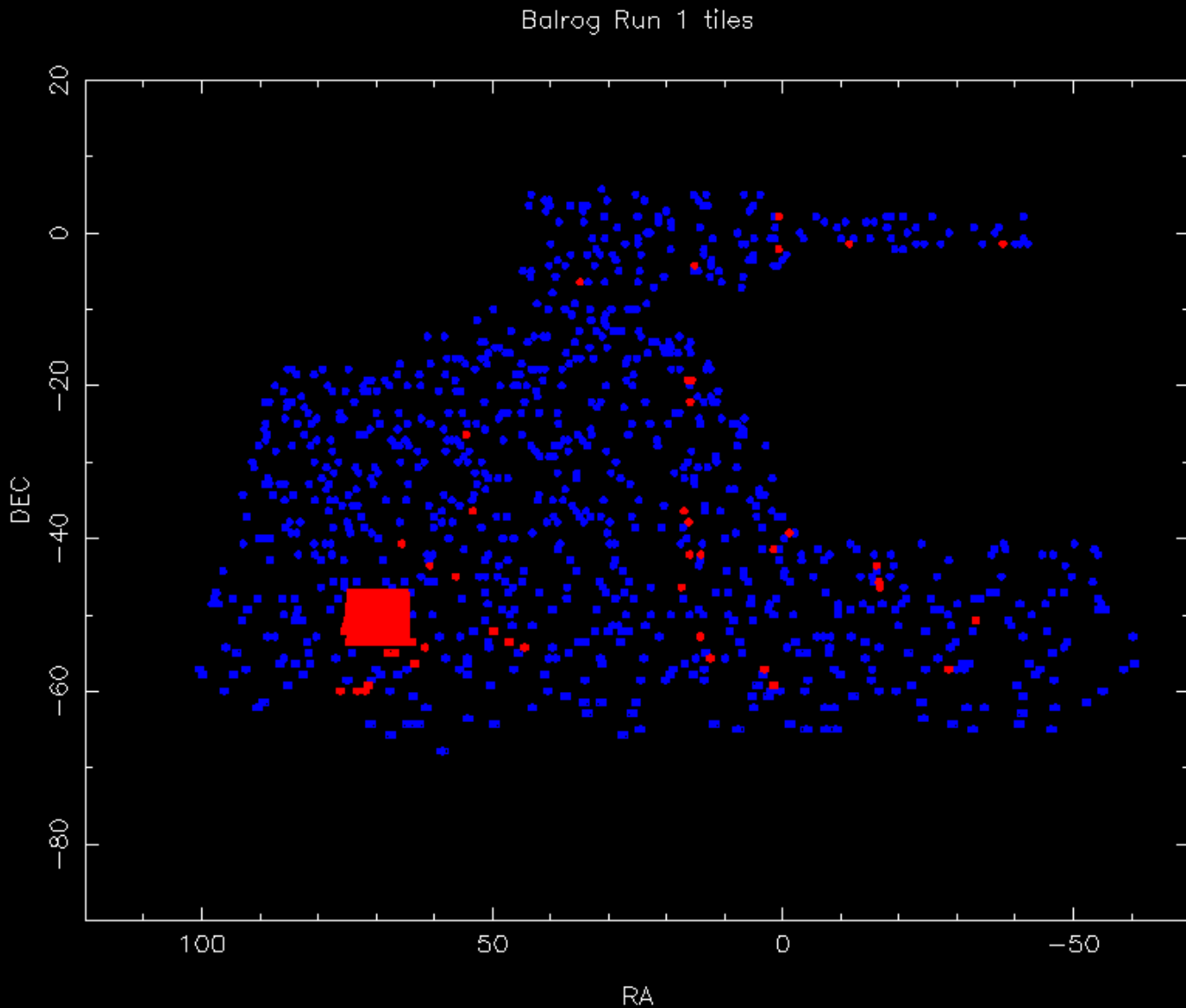
**10% of the
DES DR1**

**Footprint:
1017 tiles**

900 random

99 in block

18 big galaxies



Technical outputs of Balrog:

Object detection completeness vs. magnitude (with errors)

Biases in measured magnitude vs. magnitude, color, size, shape, location within footprint

Bias in measured color vs. magnitude, color, size, shape, location (local galaxy density)

Bias in measured size (magnification bias), in measured shape (shear bias)

Science applications of Balrog outputs (vs. inputs):

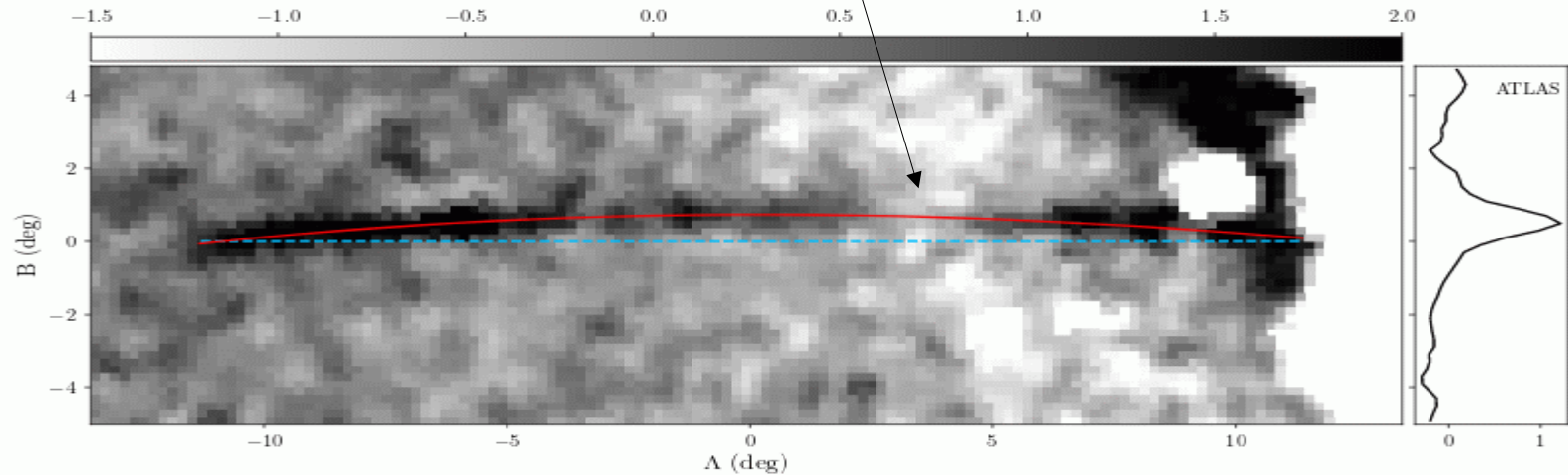
Help refine the photometric redshift relation by allowing corrections for color, magnitude other biases, esp. at the faint end.

Help calibrate the weak lensing science by measuring the bias in shear, magnification vs. magnitude, color, size, shape.

Help calibrate the undercount of galaxies in rich clusters due to incompleteness of detection of objects in crowded fields.

Help calibrate the detection completeness of stars in streams to see if apparent gaps in streams are real.

Sample use of Balrog for Milky Way science:
Shipp+ 2018 (and others) have found possible ‘gaps’ in stellar streams in the halo. But are they real, or an artifact of background subtraction or varying depth of survey?



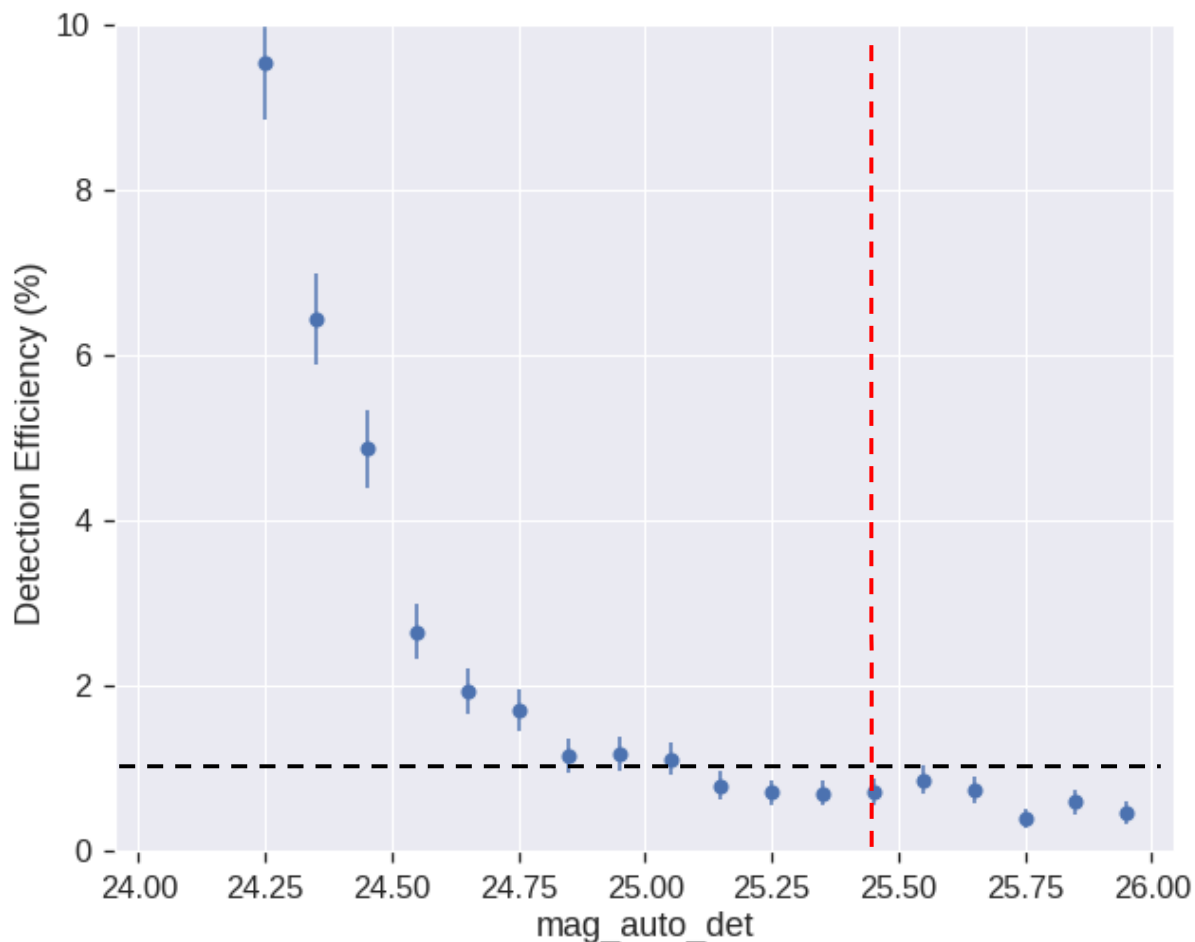
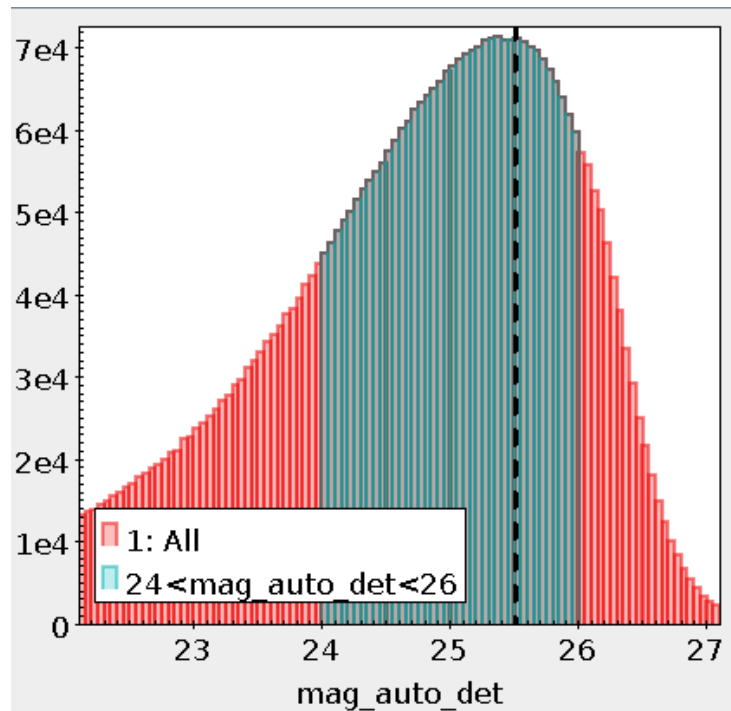
Use Balrog to add faint star catalogs to DR1 Y3 images, and then recover star counts. Note the completeness on faint end as a function of position along the gap.

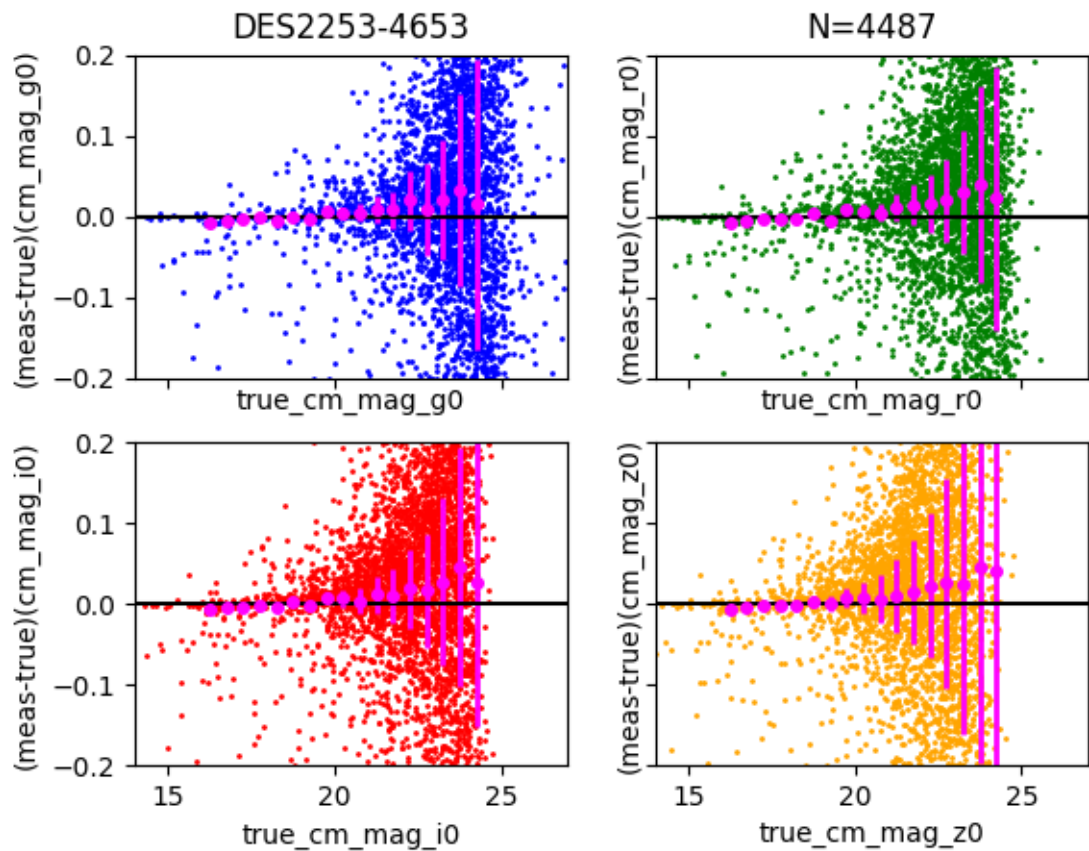




Photo-z's require knowledge of property modulation for very faint objects only detected $\sim 1\%$ of the time

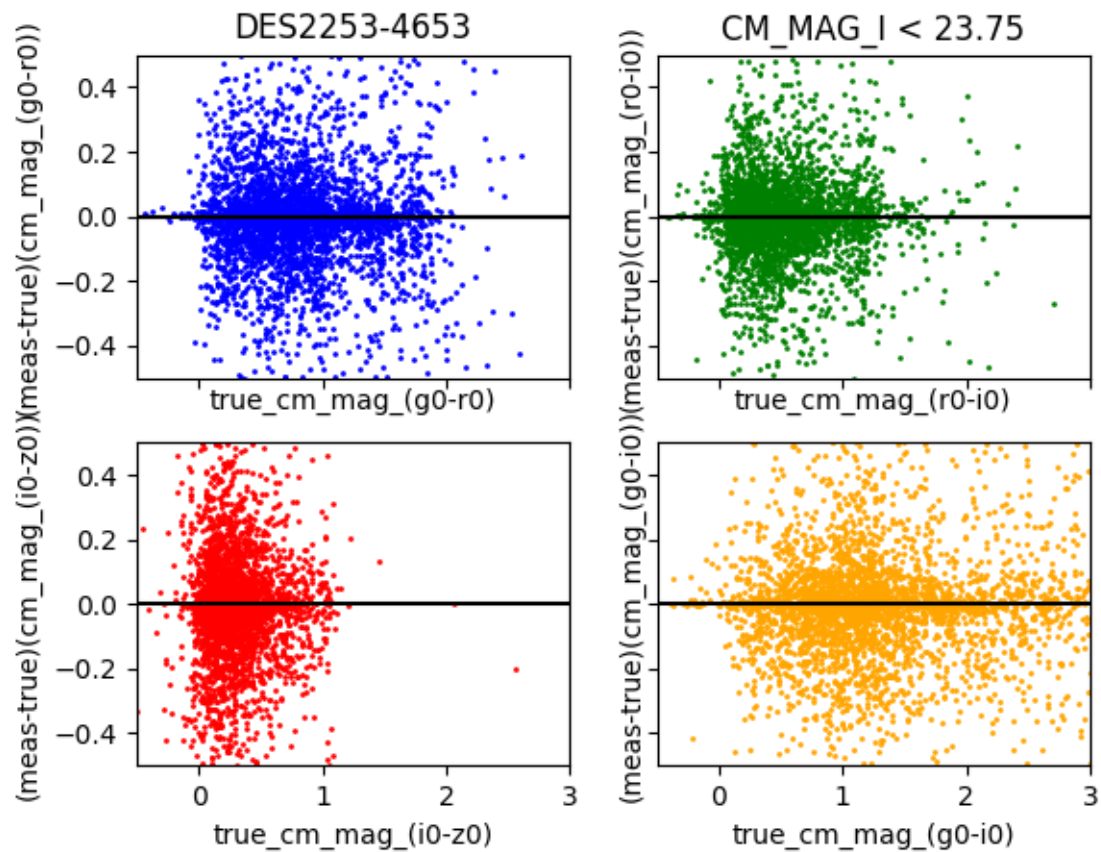
Compromise between recovery rate and completeness ($\sim 1/3$ of injections were detected).

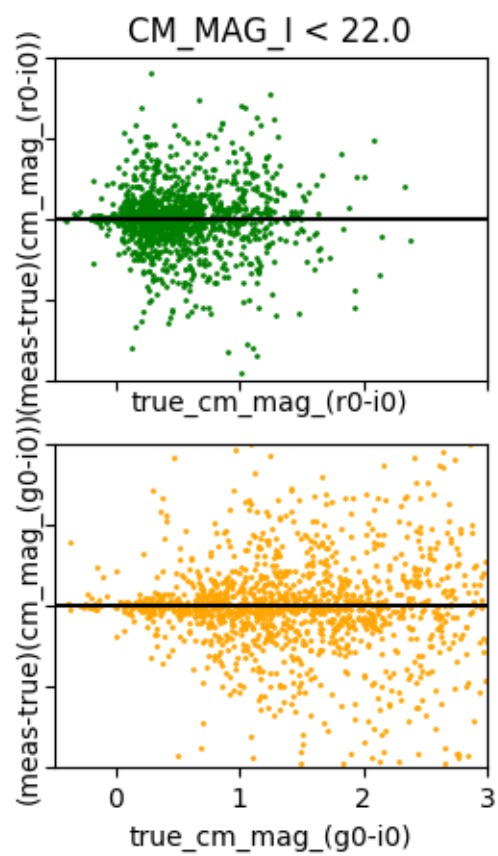
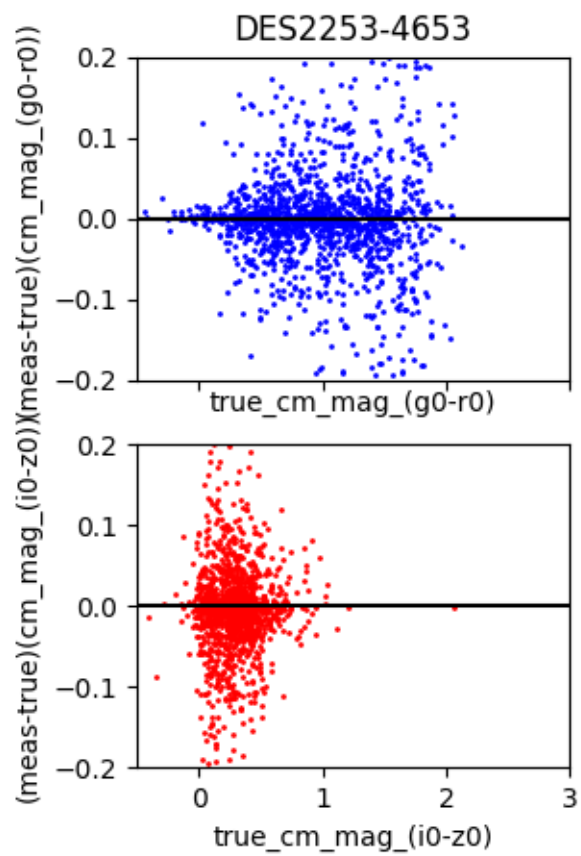




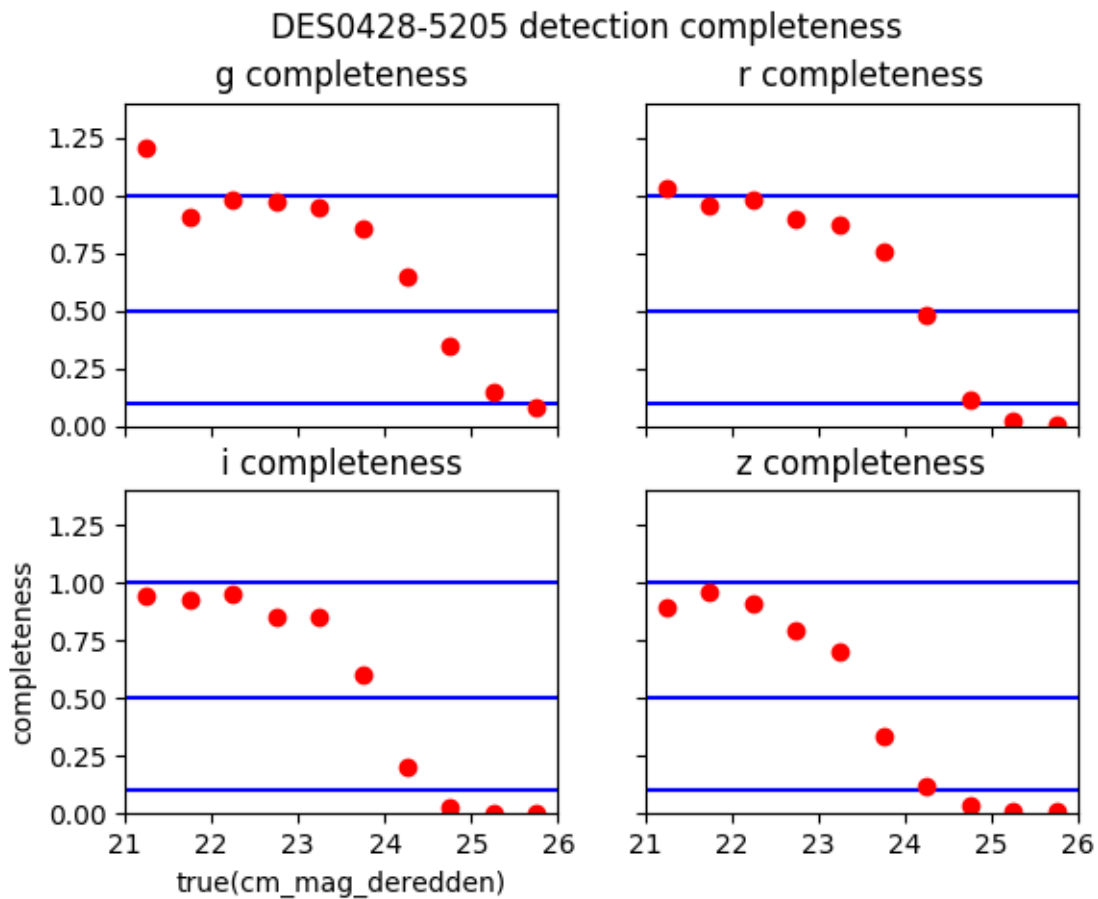
Sample Balrog output vs. input
For magnitudes vs. magnitude

Color input vs. output vs. magnitude





Detection completeness: output vs. input (one tile out of 1000)



Current status of Balrog for DES Y3 March 2019:

Run 1 (1000 tiles is complete) – working now with DES science groups to determine completeness and biases.

Will run some ‘smaller, specialized’ Balrog tests to test specific items requested by science groups (i.e. artificially add a shear, magnification) to a set of ~1M galaxies and then reprocess to determine a bias.

If needed, may run a Balrog which includes all 10,000 Y3 tiles, but may not be necessary, do we need postage stamp injections or are these model params ok?

May consider an updated Balrog run for Y5/Y6 DES data (i.e. the full DES data set).

Will write-up a paper and eventually Balrog tables will become public for wider use of DES data set.