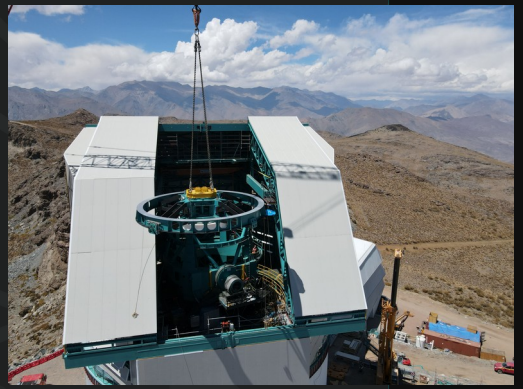
MultiProFit and galaxy photometry update



Dan Taranu

February 28, 2024

Intro & Summary

MultiProFit is a multi-band parametric model fitting code intended to replace meas_modelfit, also using Gaussian mixture models.

MultiProFit showed promising results 2-odd years ago. Since then:

- major performance refactoring (Python classes → gauss2dfit C++)
- new pipetasks split up PSF modelling & source modelling
    - make per-patch catalogs & consolidate to per-tract
- fresh new model rebuilding/fit visualization mode
- Scarlet → scarlet_lite, with flux-conserved models as default

You'll see fresh new DC2 truth match results.

DC2 Truth Match

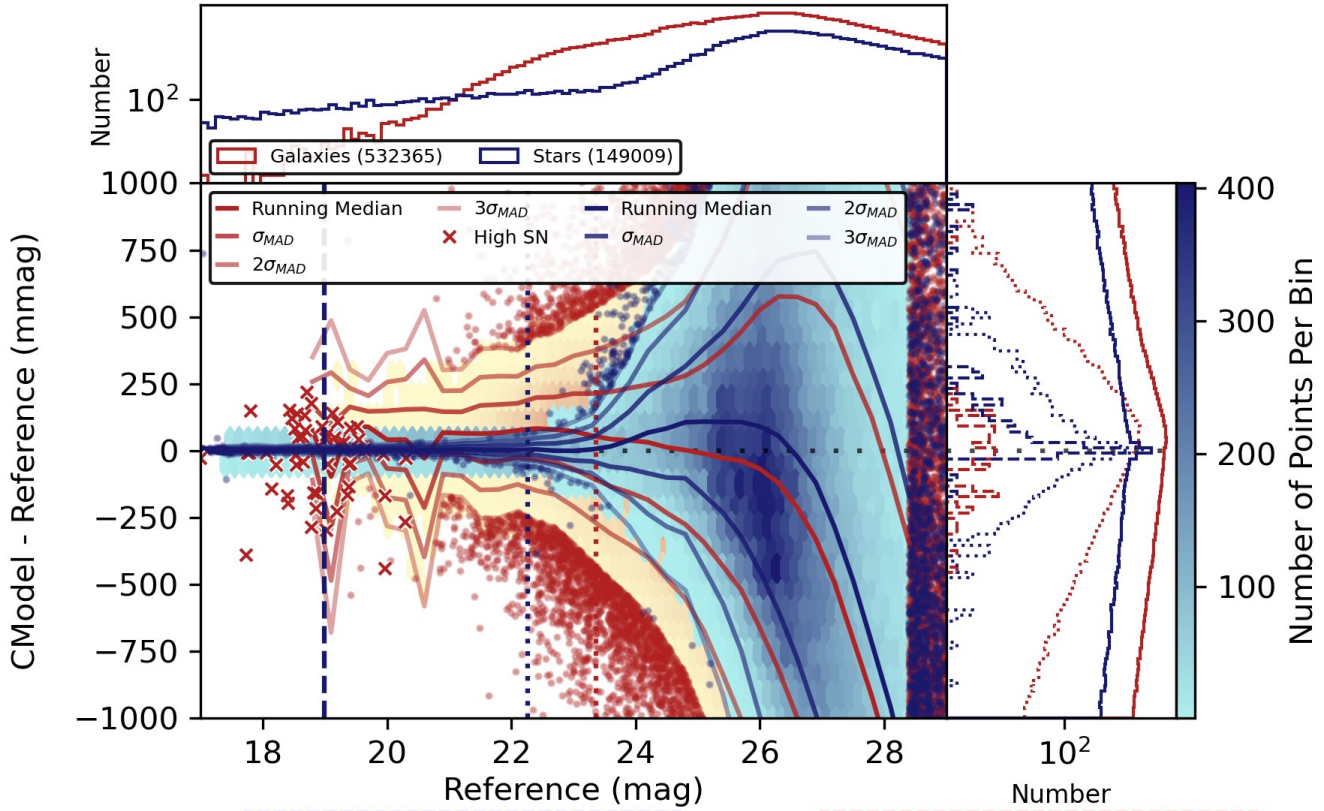
2.1 years ago, I introduced plots & metrics matching to DC2 truth
- the match tables were included in DP0.2, with a tutorial
- use the matcher!! it’s in meas_astrom, with tasks in pipe_tasks

Since then:
- faro/analysis_drp → analysis_tools happened
    - 🔥❓, 2022/01: How do we share plot/metrics code?
- USDF/Sasquatch move – not much visible change
- metrics & plots continue to emerge and look reasonable
    - 2023/10 Scarlet bug clearly visible
- colour diffs pending ticket; completeness/purity TBD

VERA C. RUBIN OBSERVATORY

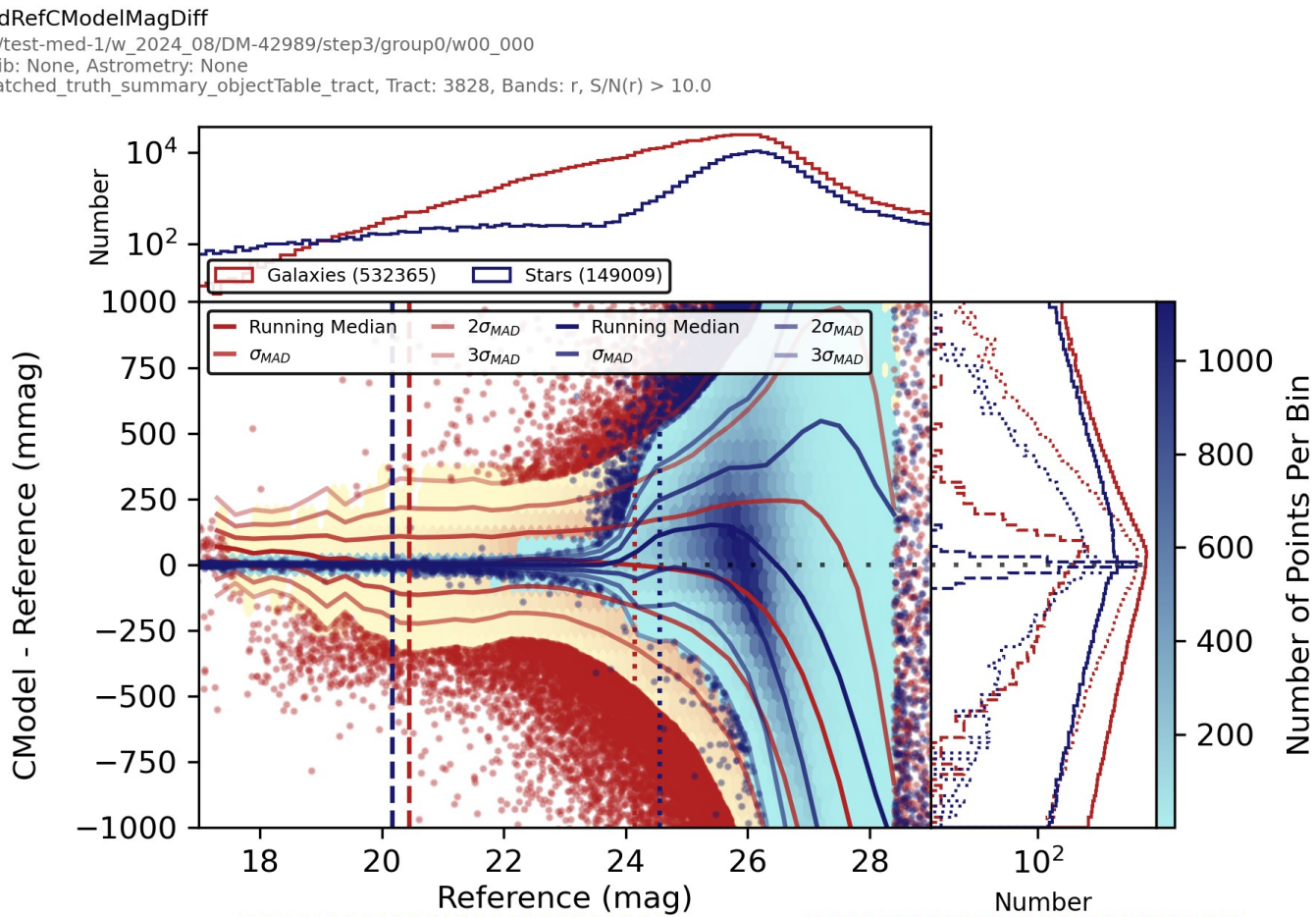
Selecting stars with (ref) Extended (ness) is not really helpful...

matchedRefCModelMagDiff
2.2i/runs/test-med-1/w_2024_08/DM-42989/step3/group0/w00_000
PhotoCalib: None, Astrometry: None
Table: matched_truth_summary_objectTable_tract, Tract: 3828, Bands: u, S/N(u) > 10.0



Same plot but r-band

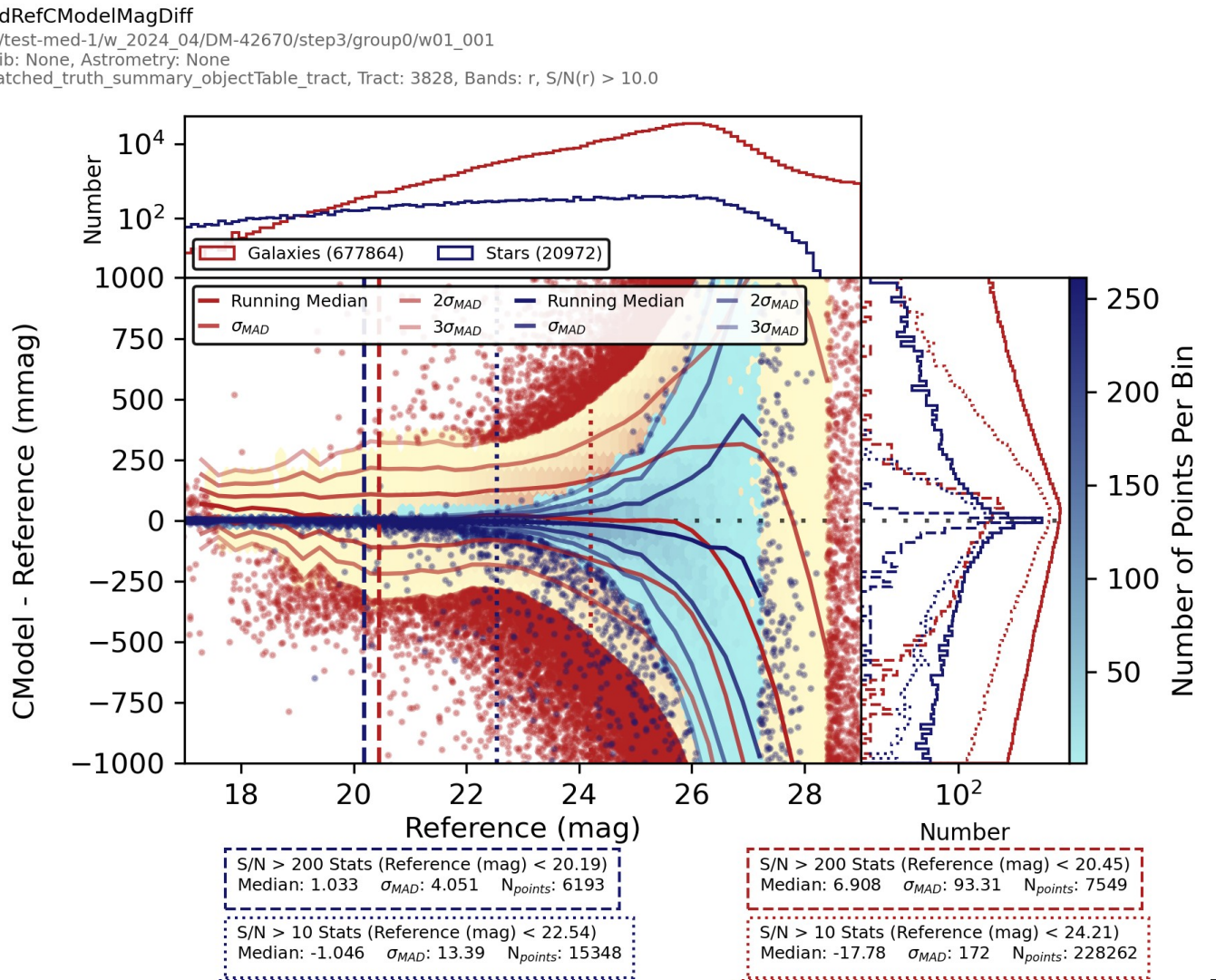
(just to show that I'm not ignoring uzy...)



Now using refcat_is_pointsource

...

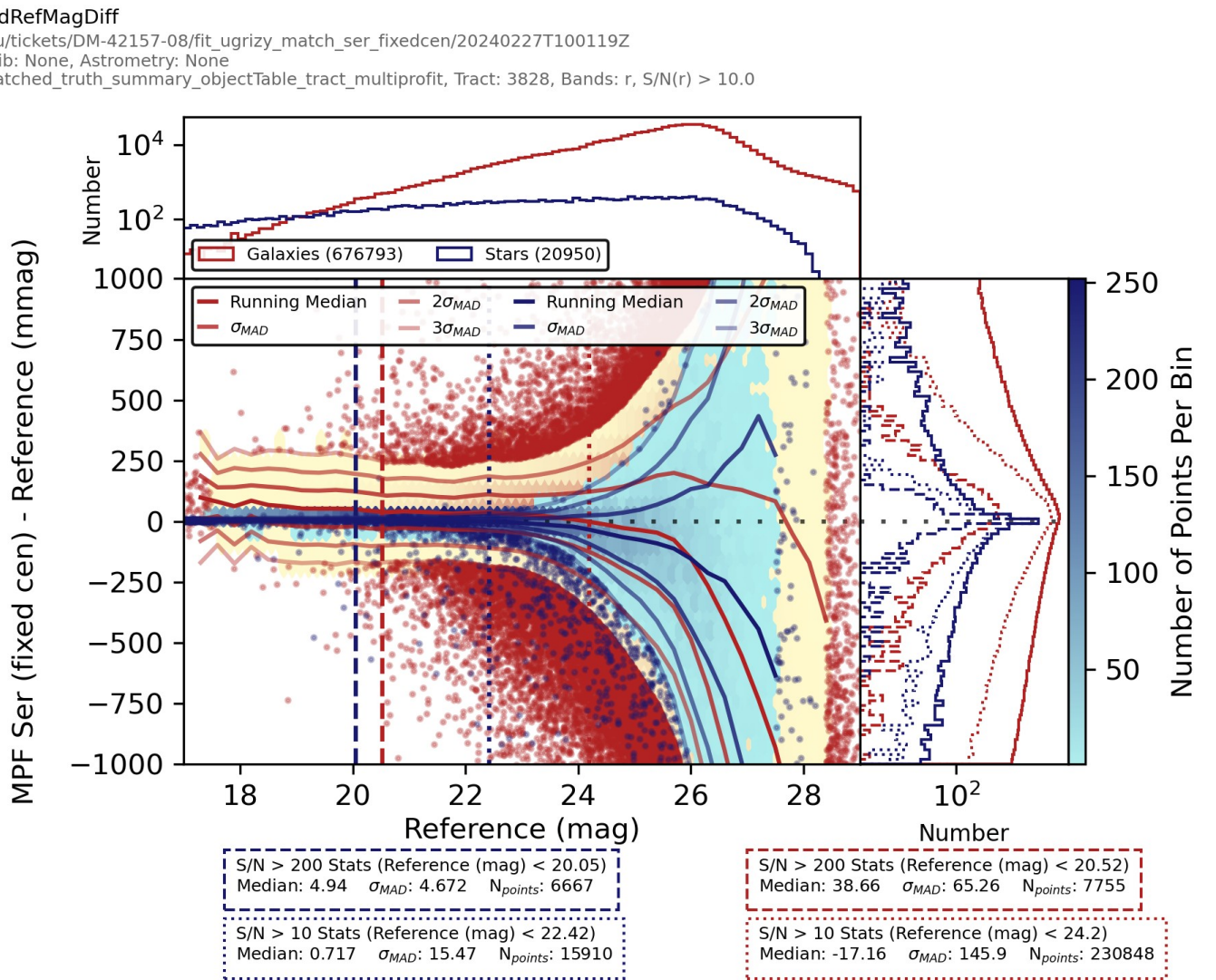
Looks pretty much ok, or at least not pathological



VERA C. RUBIN OBSERVATORY

Fixed centroid Sersic looks good for bright-ish galaxies (20-24)

Bias and flat scatter* at mag<20



VERA C. RUBIN OBSERVATORY

Free centroid Sersic looks very similar

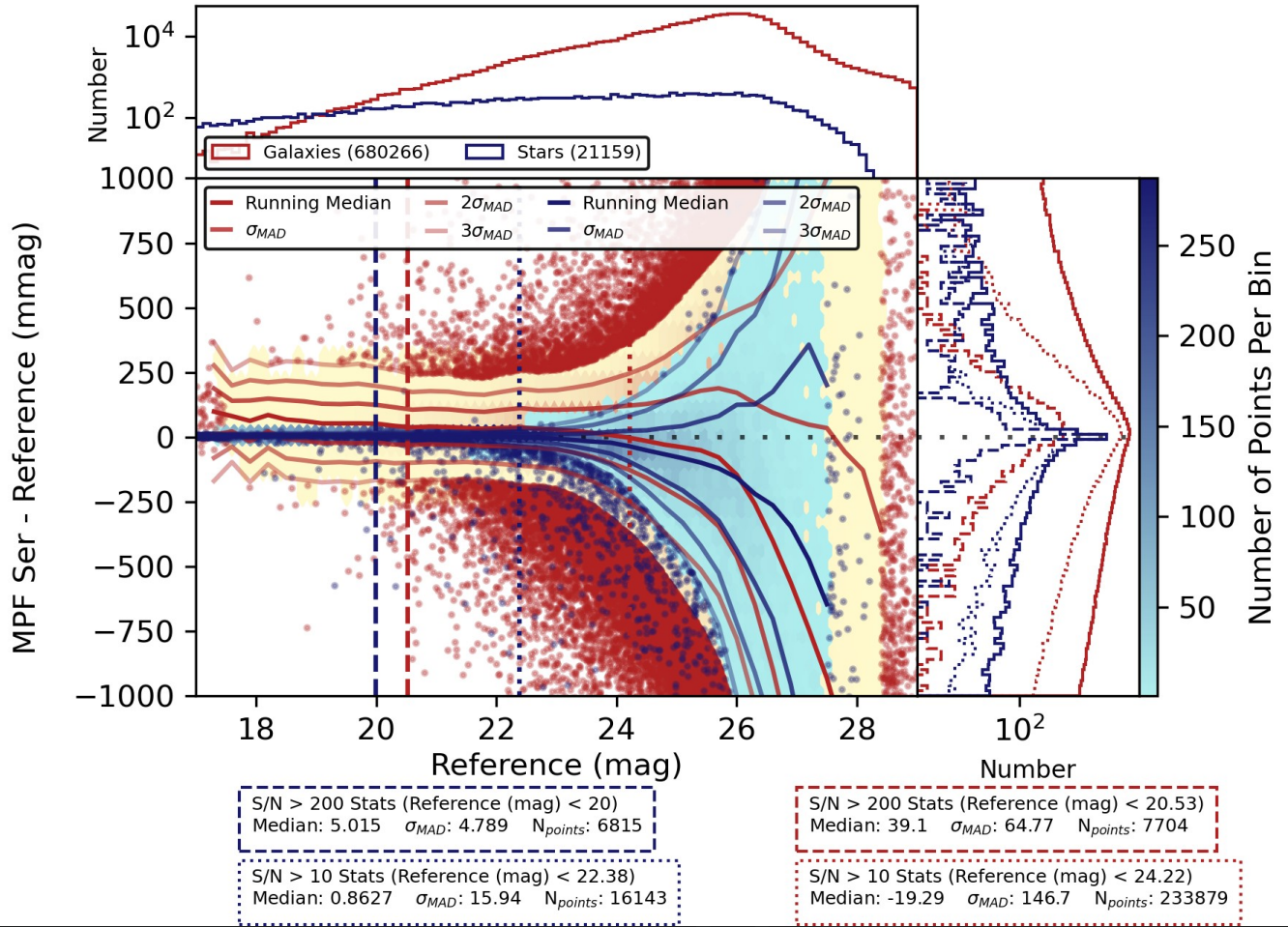
Slightly more stars matched (maybe doing a bit more good than harm?)

matchedRefMagDiff

u/dtaranu/tickets/DM-42157-08/fit_ugrizy_match_ser/20240227T100658Z

PhotoCalib: None, Astrometry: None

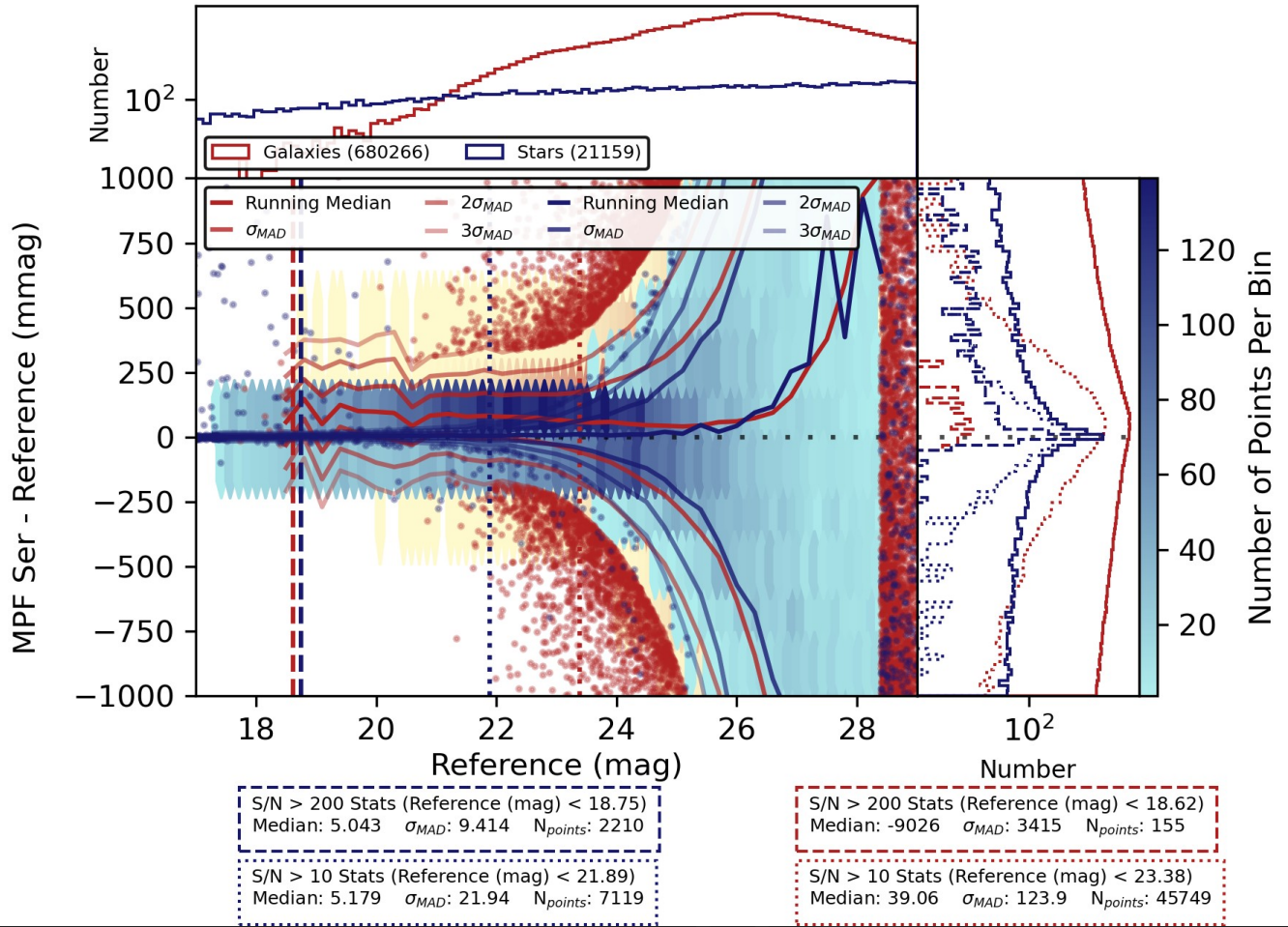
Table: matched_truth_summary_objectTable_tract_multiprofit, Tract: 3828, Bands: r, S/N(r) > 10.0



Free centroid Sersic bias worse in bluer bands like u

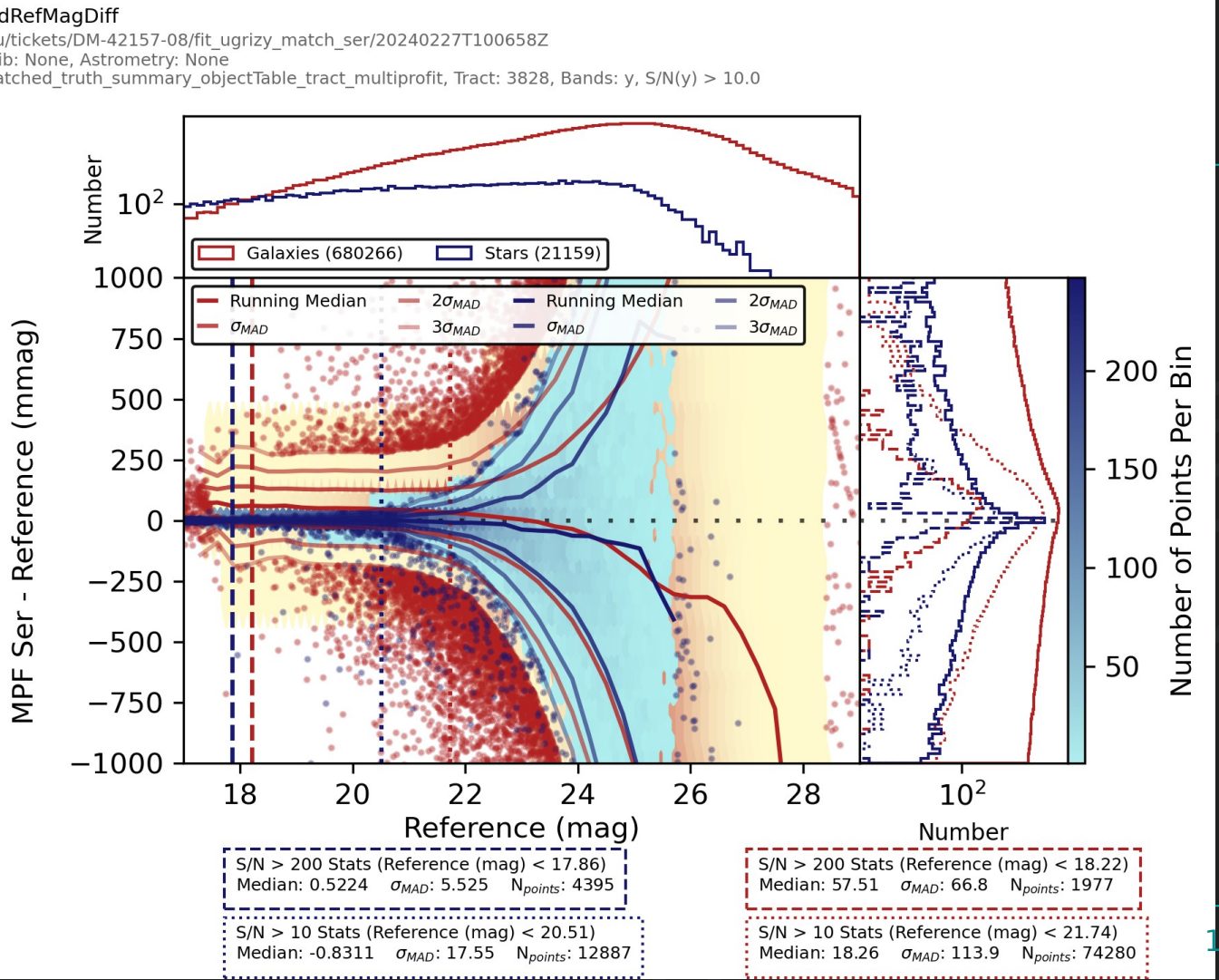
(for galaxies stars are ok, modulo apCorr)

matchedRefMagDiff
u/dtaranu/tickets/DM-42157-08/fit_ugrizy_match_ser/20240227T100658Z
PhotoCalib: None, Astrometry: None
Table: matched_truth_summary_objectTable_tract_multiprofit, Tract: 3828, Bands: u, S/N(u) > 10.0



Free centroid Sersic bias still there in y band...

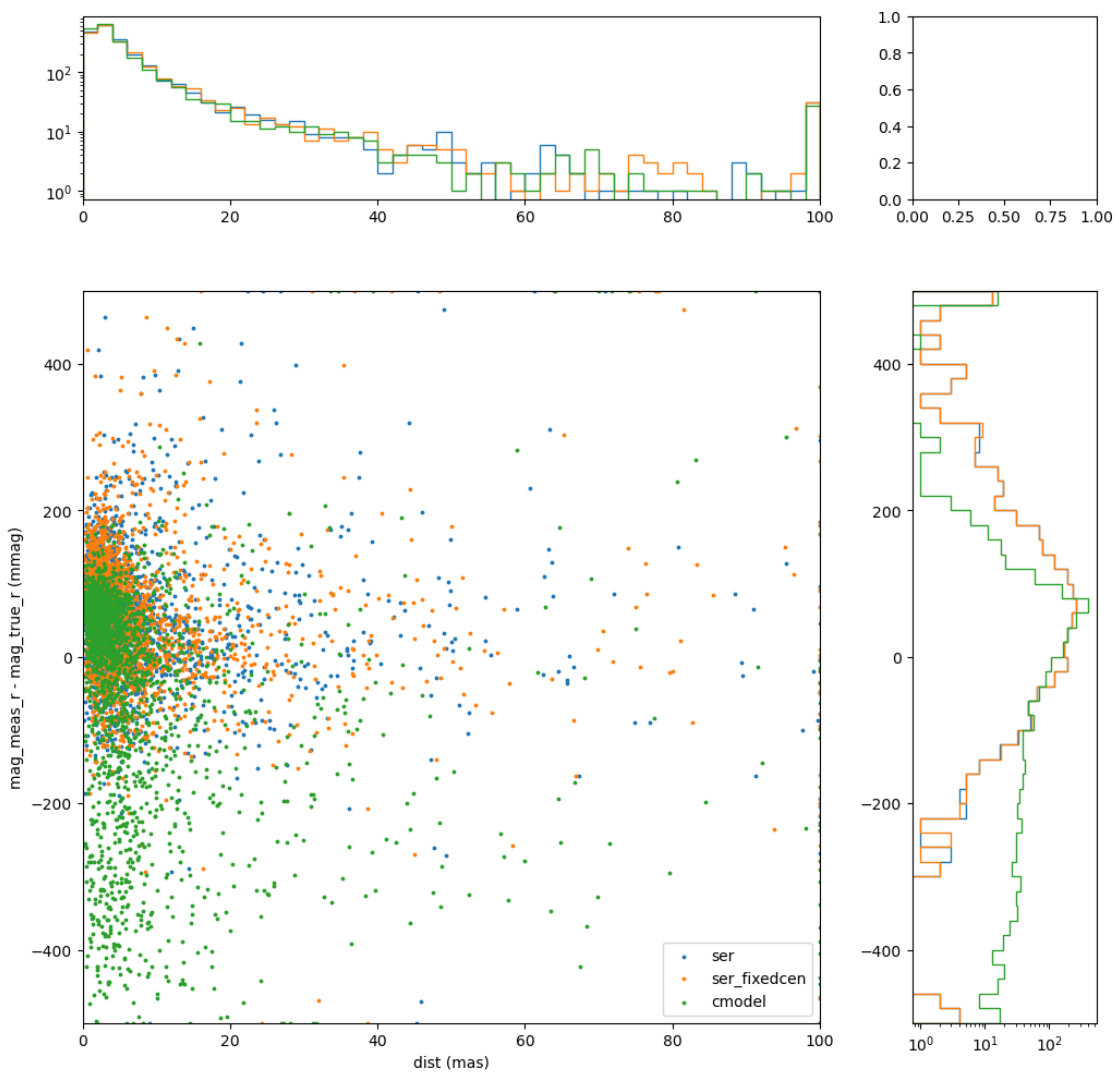
Note that stars are too bright (apCorr would make it worse)



Is cModel bias really “better”? Actually, median is closer to 0 because residuals skew towards “model too bright”. The mode is ~50mmag too faint with both codes.

No connection w/astrometry.

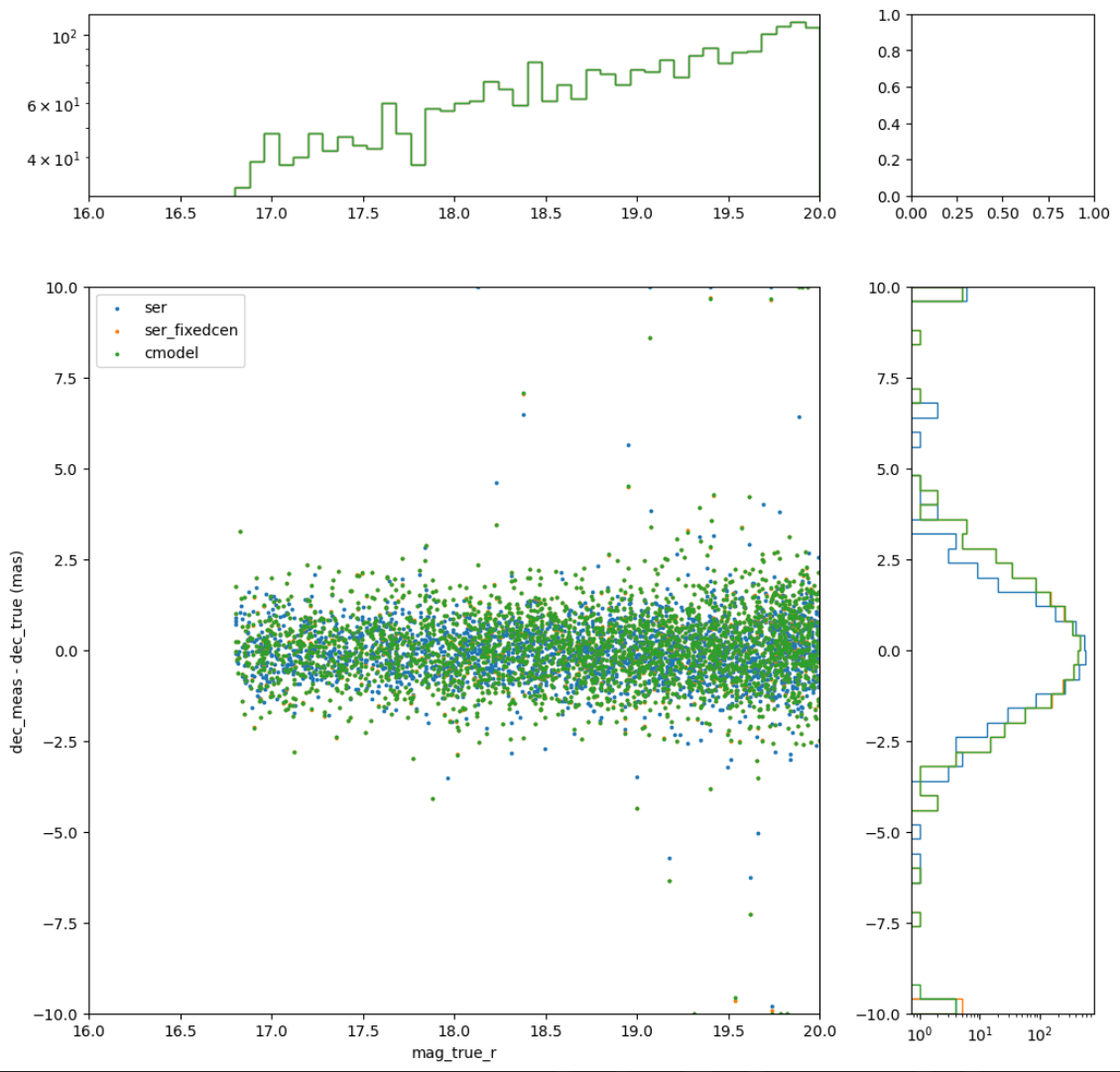
(galaxies, true r < 20)



dec residuals for stars not very interesting.

Free centroid makes no difference.

(stars, true r < 20)

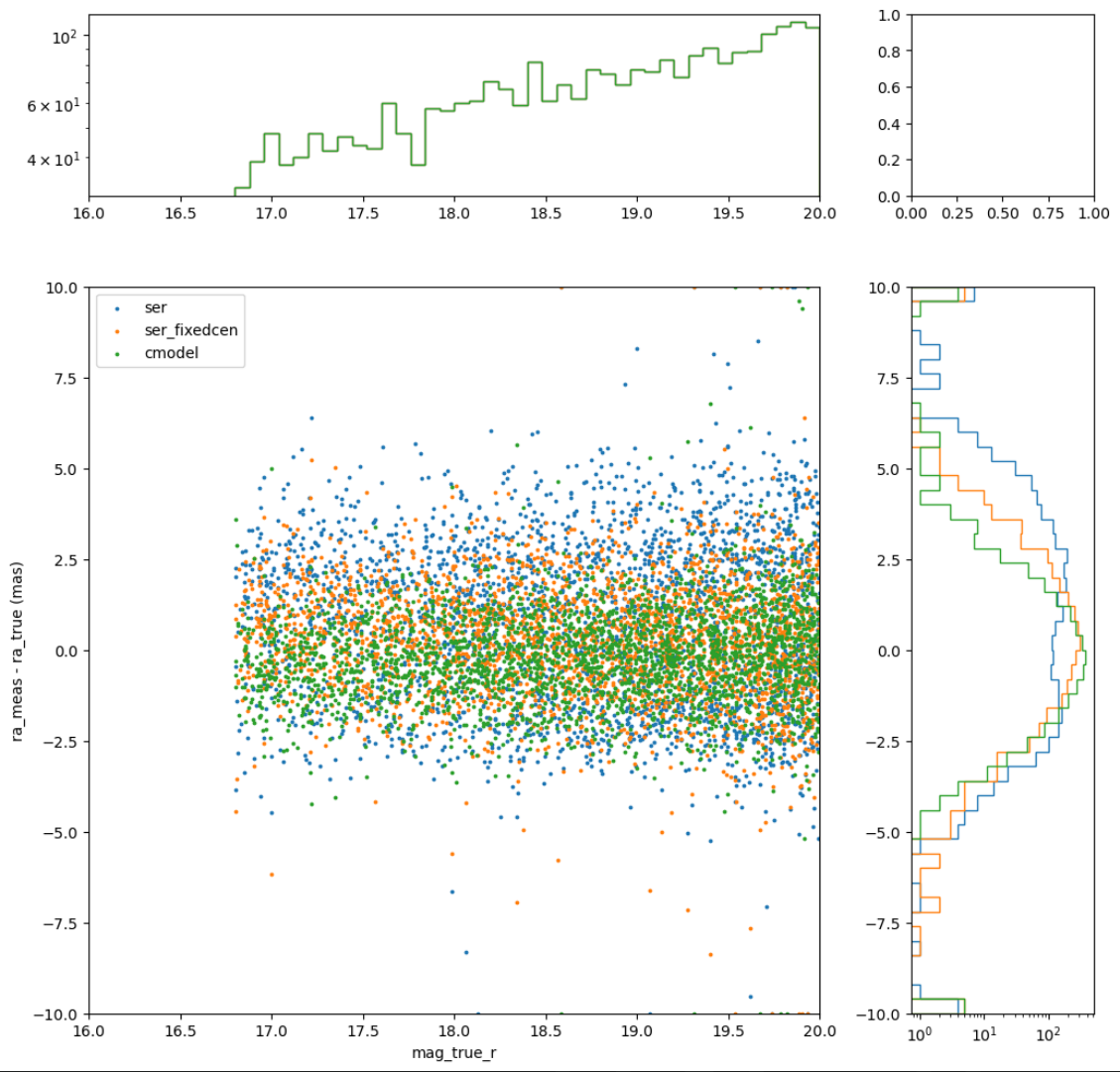


ra residuals for stars slightly baffling.

Why is ser_fixedcen slightly different? (different WCS?)

Why is free cen biased & worse? DCR?

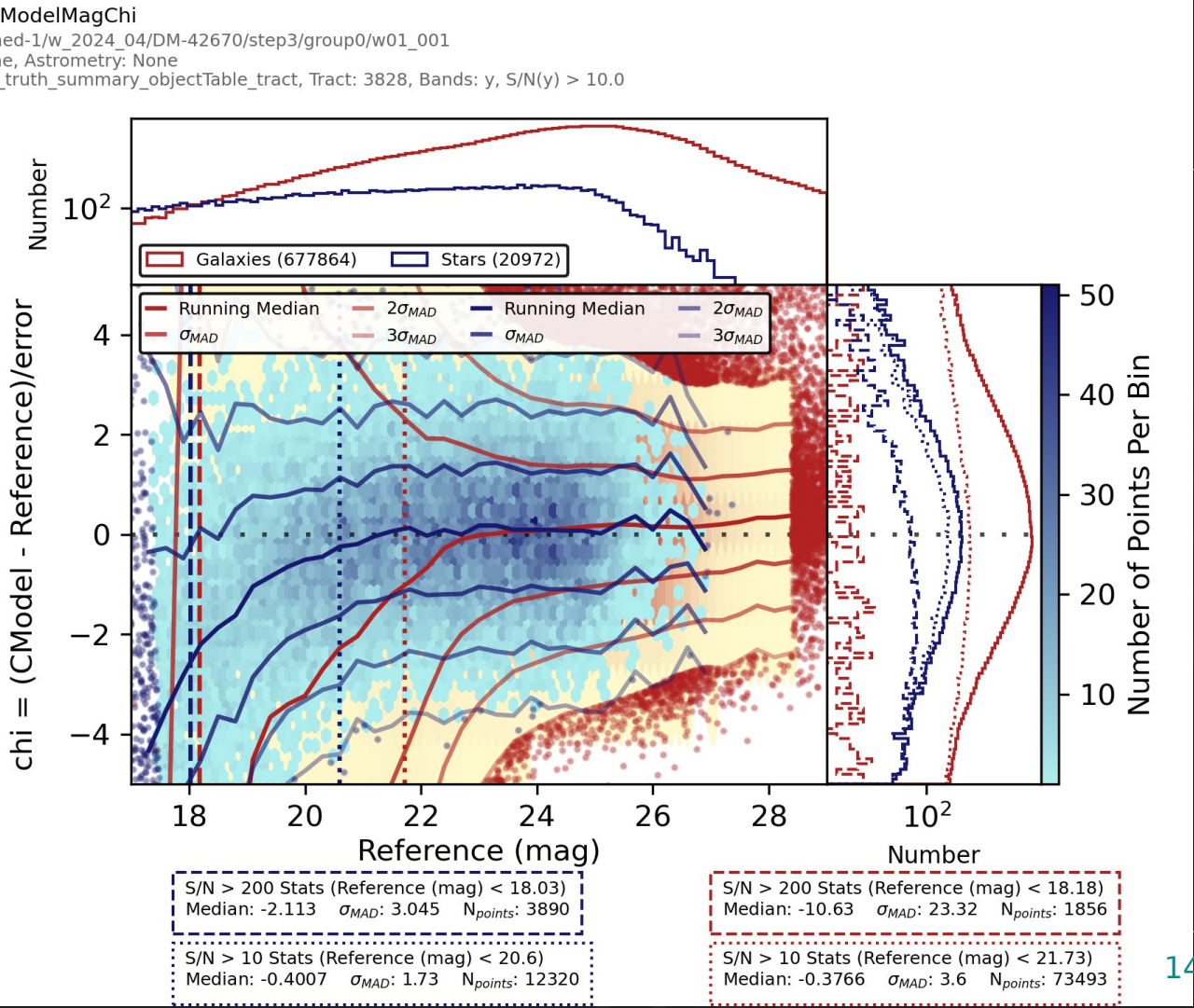
(stars, true r < 20)



Back to CModel for errors:

Only in redder bands ok at faint end.

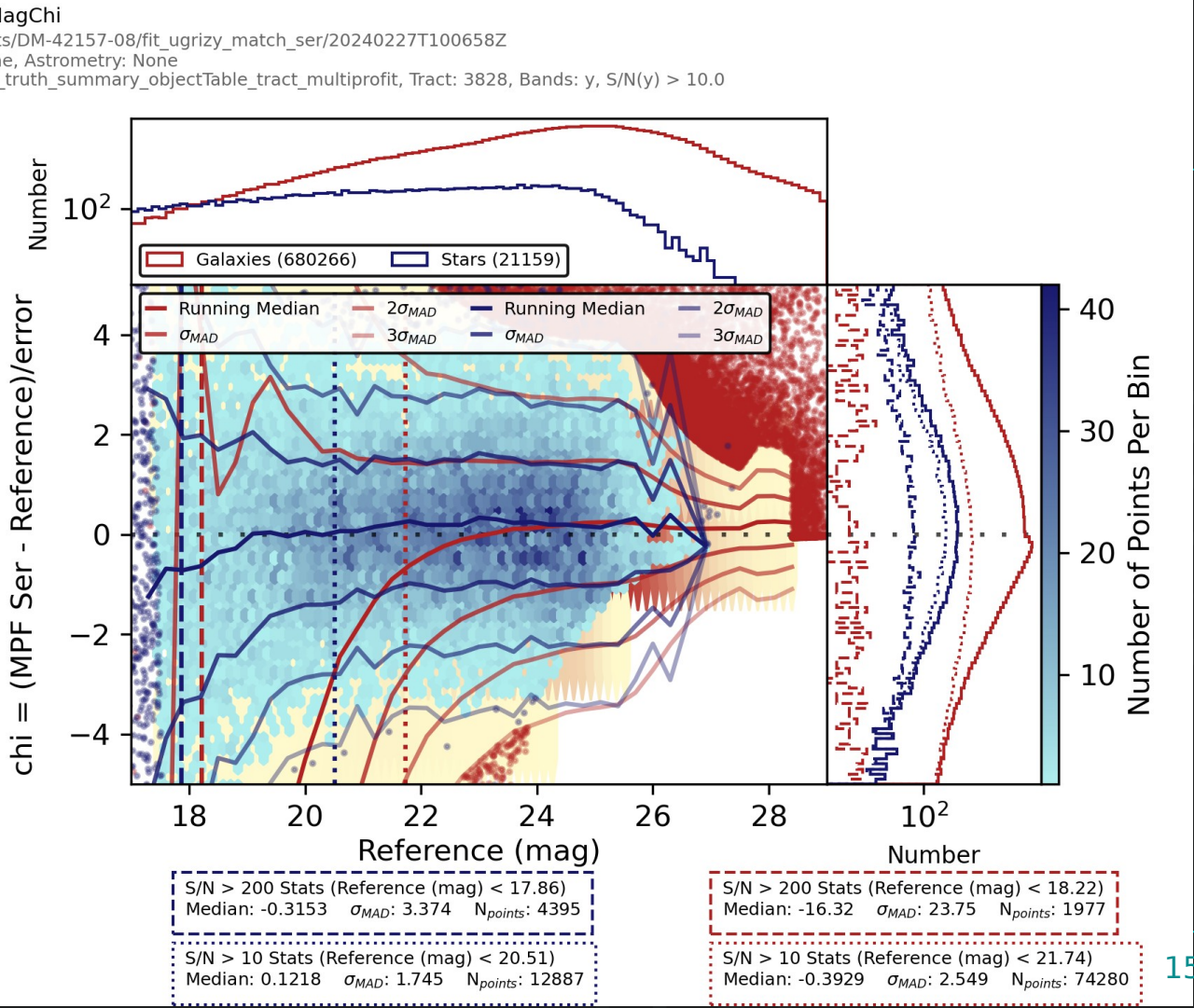
(r band median way worse)



Free-cen Sersic errors:

Better for stars

A little better for galaxies… but the bias still biggest problem



Semi-speculative interlude

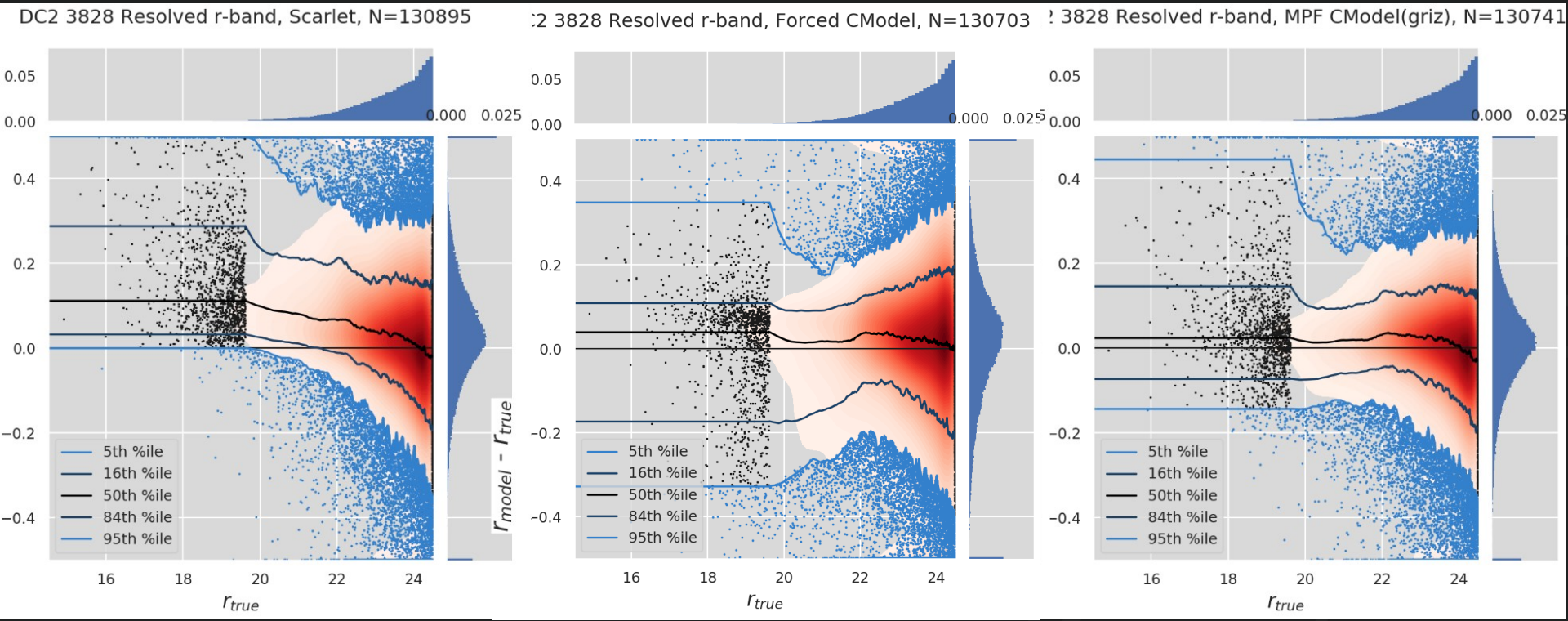
Remember these are fits to deblended models (except for isolated objs)

scarlet(_lite) now has flux-conserving models, which preserve noise…
… sort of. I think it’s impossible to “correctly” assign variances.

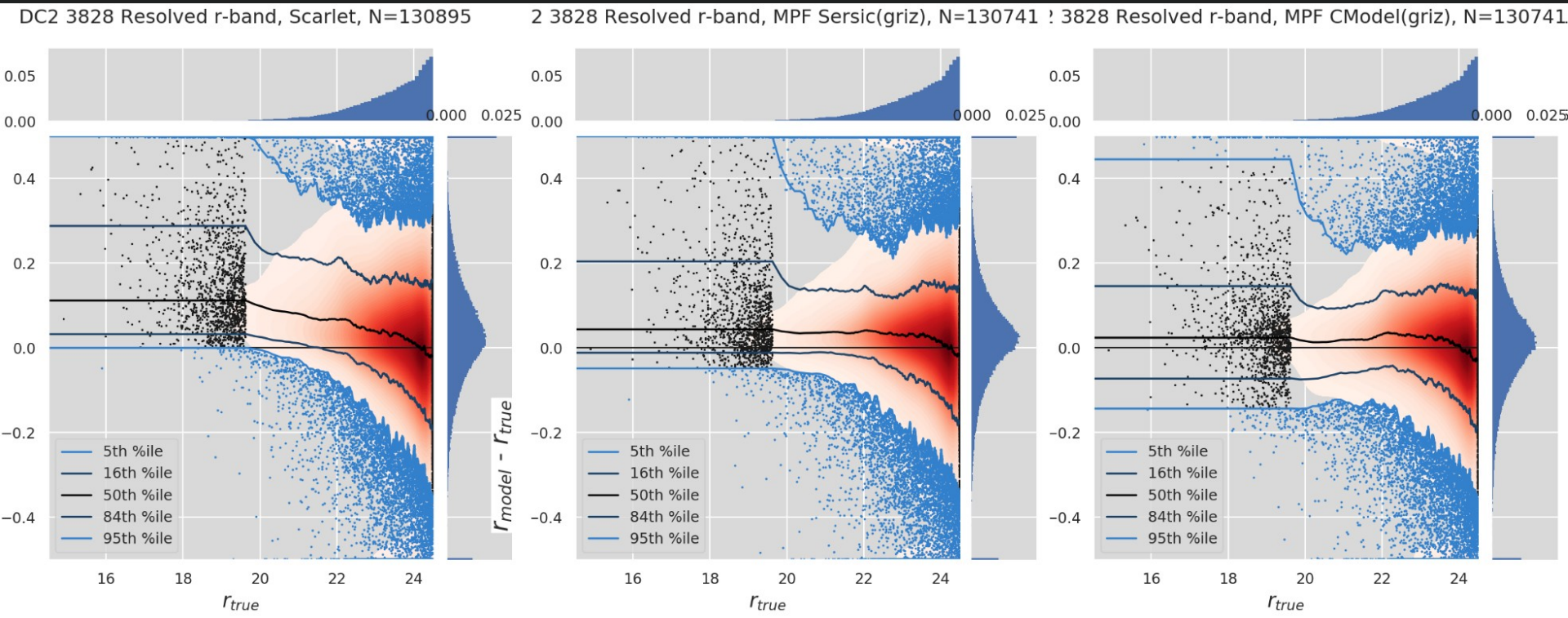
The situation will probably improve after parametric deblending.

Basically a separate task that re-fits blends, taking best-fit params for each child and doing simultaneous fitting. Linear version (fluxes) okay. Nonlinear may be prohibitively slow but maybe just for N<4 or 5 blends? (I tried it before, and it linear deblending had modest benefits)

Older plots pre-flux conserving Scarlet



Older plots pre-flux conserving Scarlet



Sersic bias used to be flatter...

Other models

I had hopes for:
- exponential + deVauc model
- Sersic + point source

ExpDev takes ~2x longer and has worse bias
PS+Ser is not much slower, but also worsens bias
(I made a slight error in implementing the point source but it’s not why)

The central point source is meant to absorb excess flux that would make Sersic n biased high, and also help in star/galaxy separation at low S/N
... but that remains to be demonstrated.

VERA C. RUBIN OBSERVATORY

Free centroid ExpDev

No real benefit to it.

More skew.

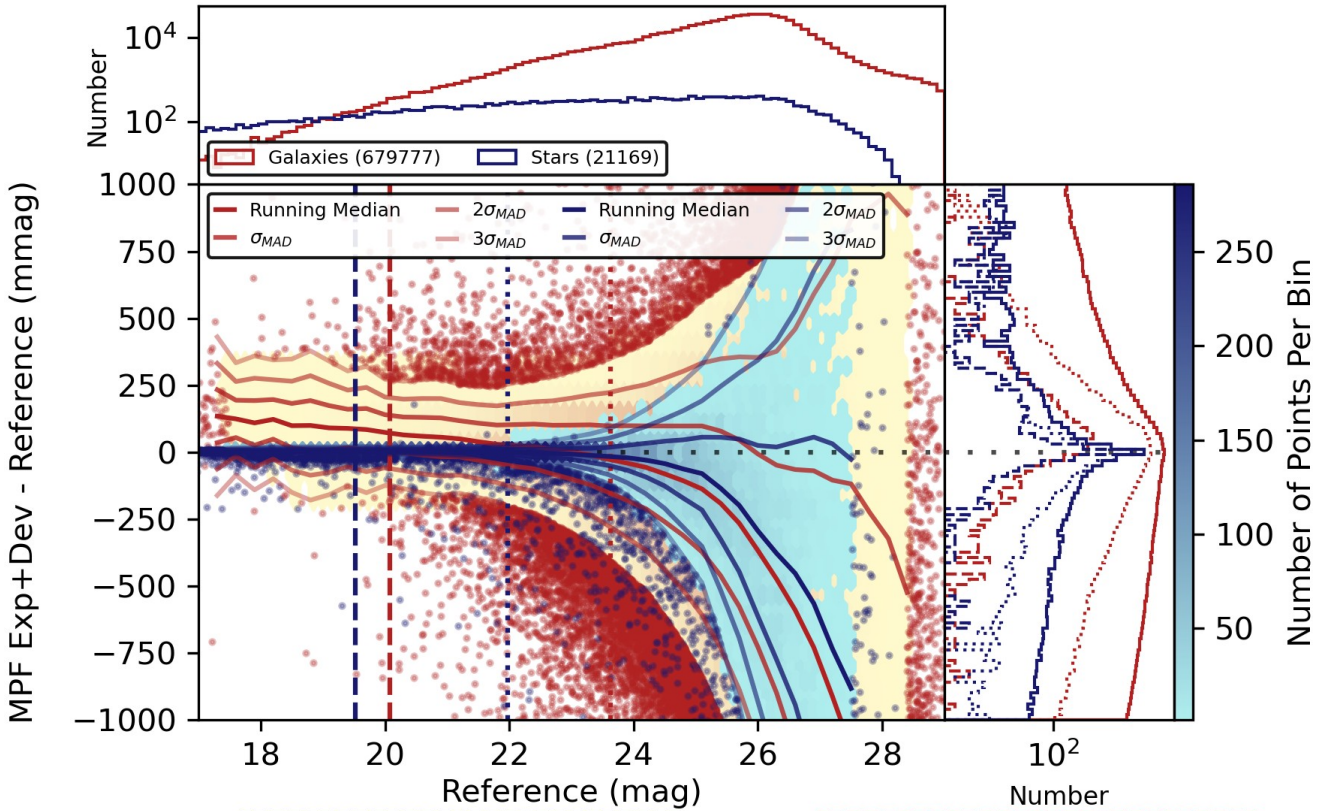
Maybe needs better init.

matchedRefMagDiff

u/dtaranu/tickets/DM-42157-08/fit_ugrizy_match_expdev/20240227T101226Z

PhotoCalib: None, Astrometry: None

Table: matched_truth_summary_objectTable_tract_multiprofit, Tract: 3828, Bands: r, S/N(r) > 10.0



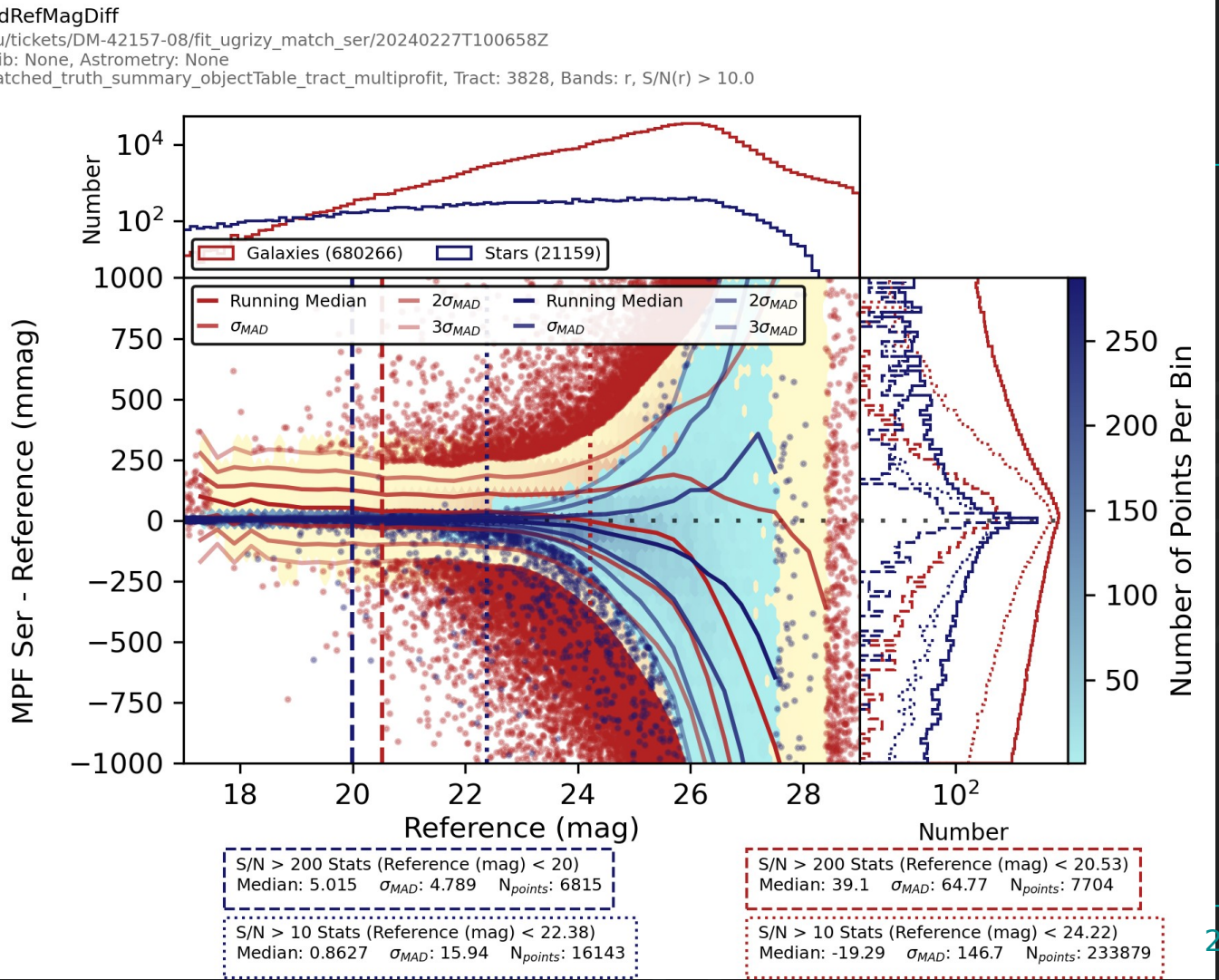
S/N > 200 Stats (Reference (mag) < 19.52)
Median: 4.856 σ_{MAD}: 5.191 N_{points}: 5306

S/N > 10 Stats (Reference (mag) < 21.97)
Median: 0.698 σ_{MAD}: 15.15 N_{points}: 14244

S/N > 200 Stats (Reference (mag) < 20.08)
Median: 60.9 σ_{MAD}: 65.84 N_{points}: 4509

S/N > 10 Stats (Reference (mag) < 23.62)
Median: -40.14 σ_{MAD}: 152.1 N_{points}: 150676

Free centroid Sersic again, for reference

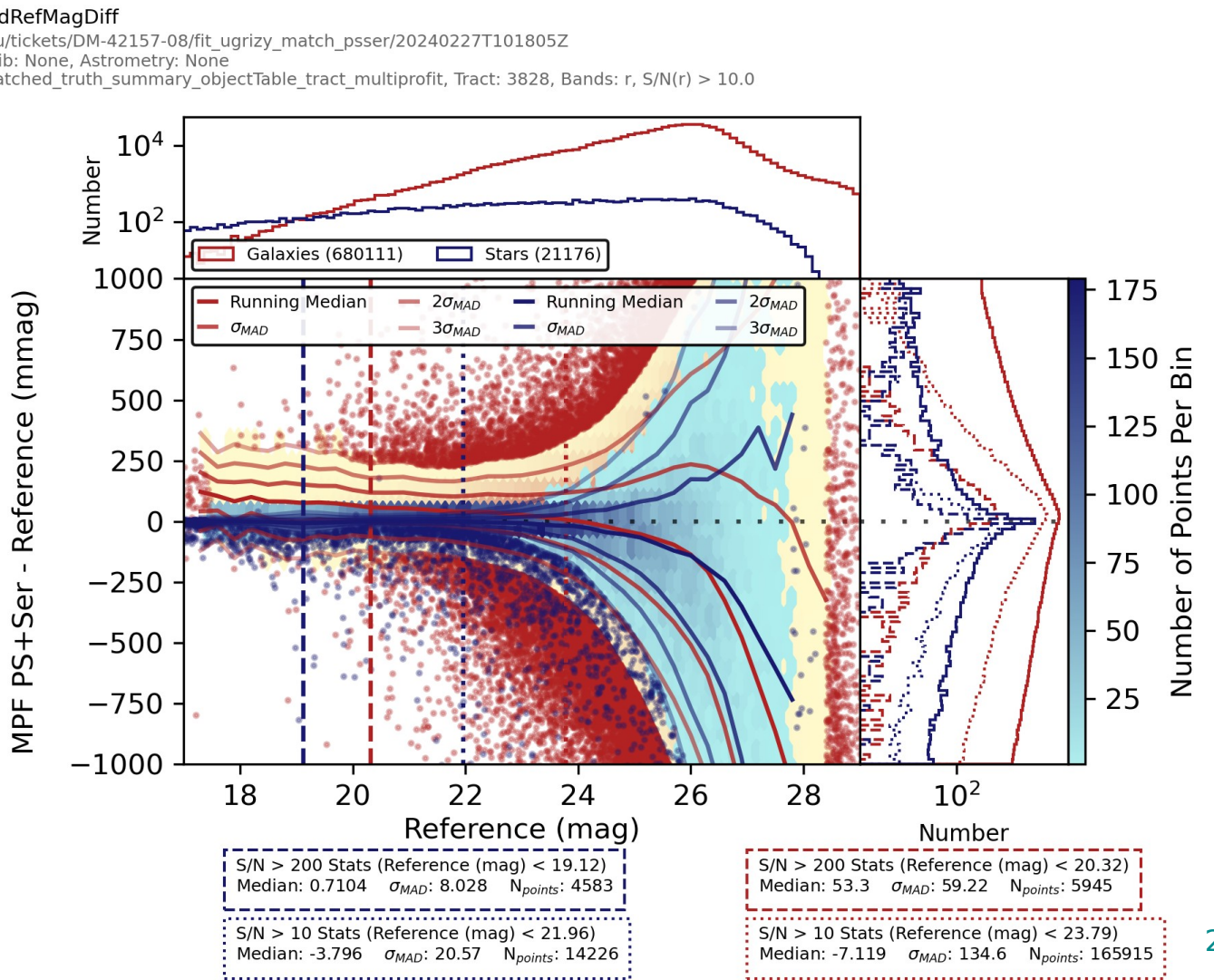


VERA C. RUBIN OBSERVATORY

Free centroid PS+Ser

Galaxy bias worse, scatter similar.

Stars worse (PS too small)



Performance

It’s actually pretty good now!

Tract 3828, patch 24:
CModel sum = 1749s (sum ugrizy, initial/exp/dev)
Free cen Sersic sum = 1026s

632s in fitting routine, 435s of that spent evaluating models

(i.e. there is room to optimize *some* overhead, but not all)

PSF fitting still very slow – is it included in CModel times?

Performance – PSF fitting

PSF fitting:

659s u-band, 1604s i-band, 1130s z-band total times per patch

(yes, spending more time PSF fitting in one band than per object...)

Only 12-27% of time spent actually fitting, and 45% of that in model eval

i.e. overhead dominates, probably coadd PSF eval (save us, cell coadds)
scipy optimizer could do better (on objects, 69% of time in model eval)

Performance thoughts

MultiProFit *should* be faster, all else equal (i.e. doing CModel fits)

It does analytic Jacobian – meas_modelfit does finite diff

I suspect meas_modelfit's optimizer is better than scipy defaults.

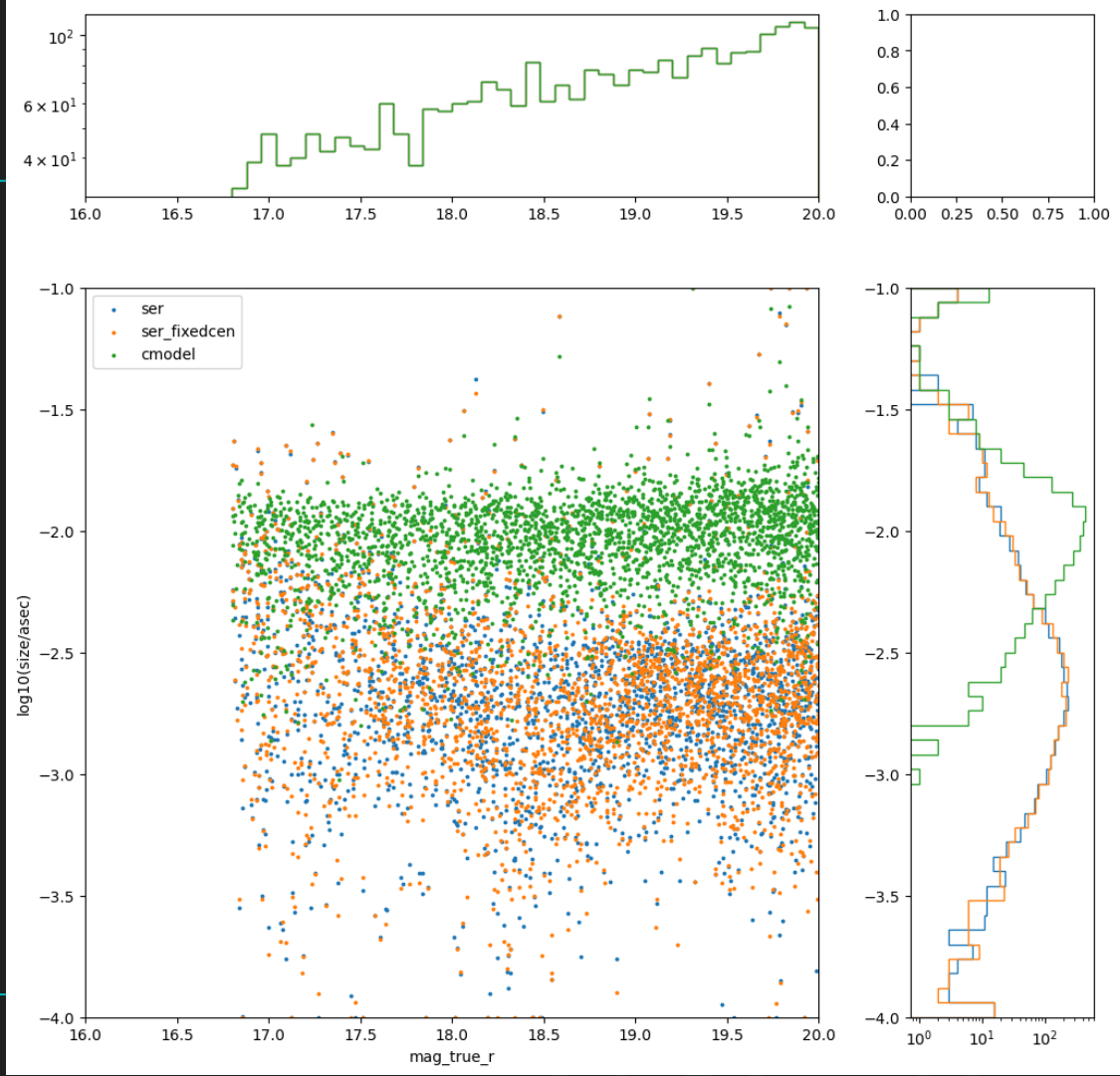
meas_modelfit supports Gauss-Hermite PSF (with skew & kurtosis)
probably not worthwhile; higher-order terms usually near zero

Hermite PSF might explain better bright star photometry.
Adding more Gaussians might be better, except MPF can't do >2 now

CModel has ~0.01” sizes for stars (why?)

MPF has optional PSF shrink param, 0.01 pix = 0.002” (log10=-2.7), right where the mode is. Intended to absorb PSF model erros.

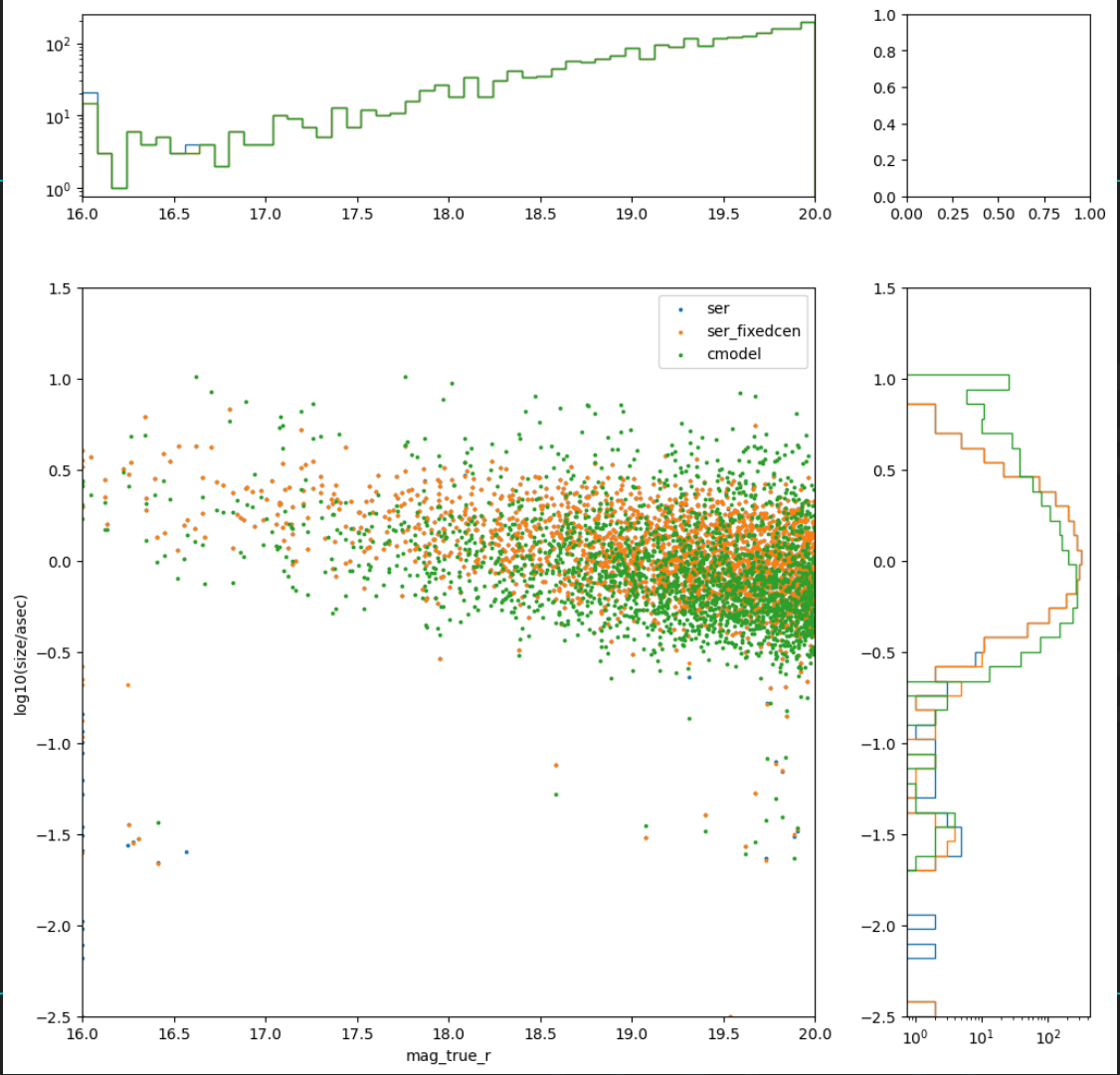
(stars, true r < 20)



Galaxy sizes are ok for good matches.

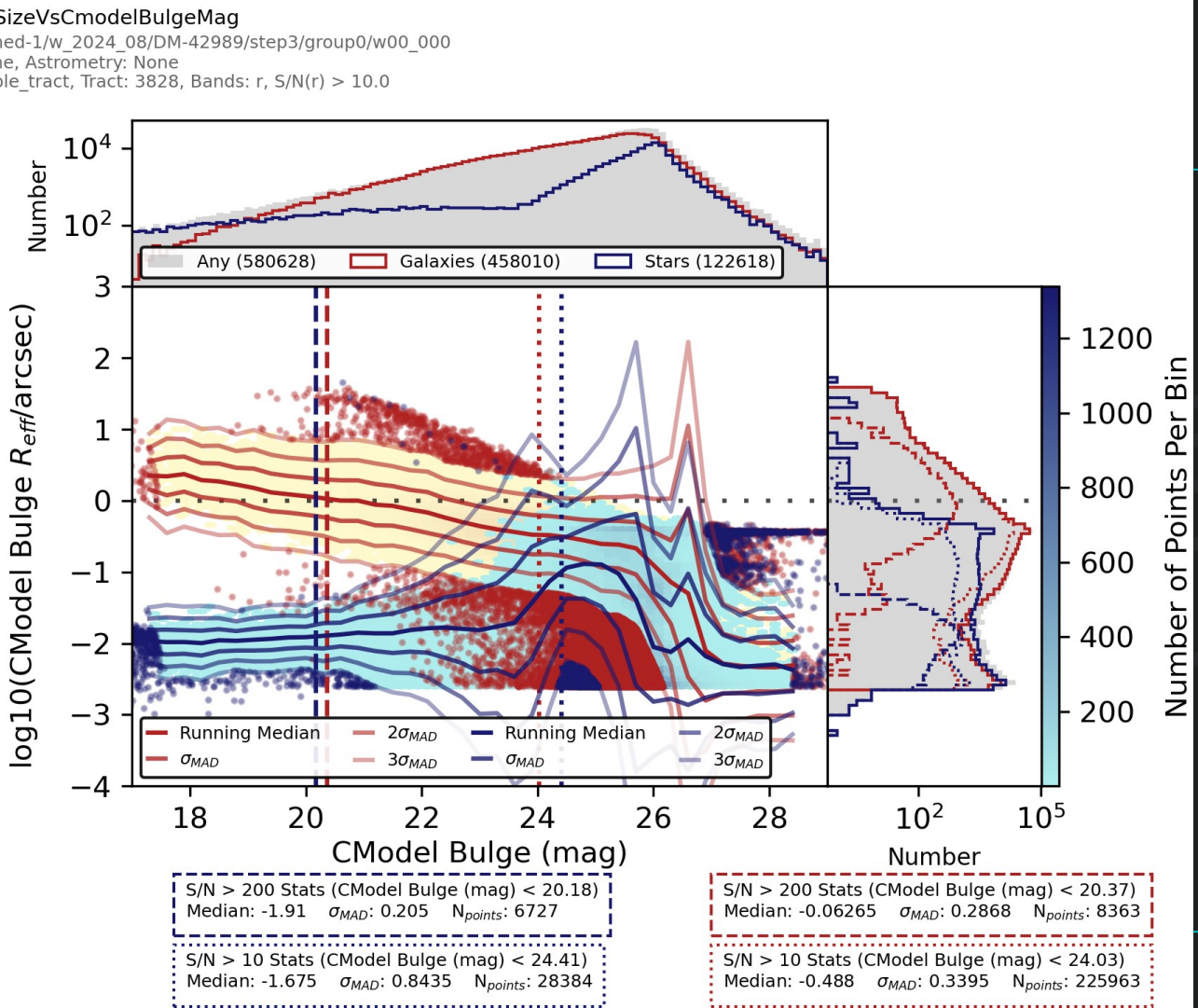
I did a flux-weighted average of exp/dev sizes for CModel, which isn't exactly equivalent to Sersic r_eff. Still ok match.

(galaxies, true r < 20)



There are atools plots of sizes (objectTable extended)

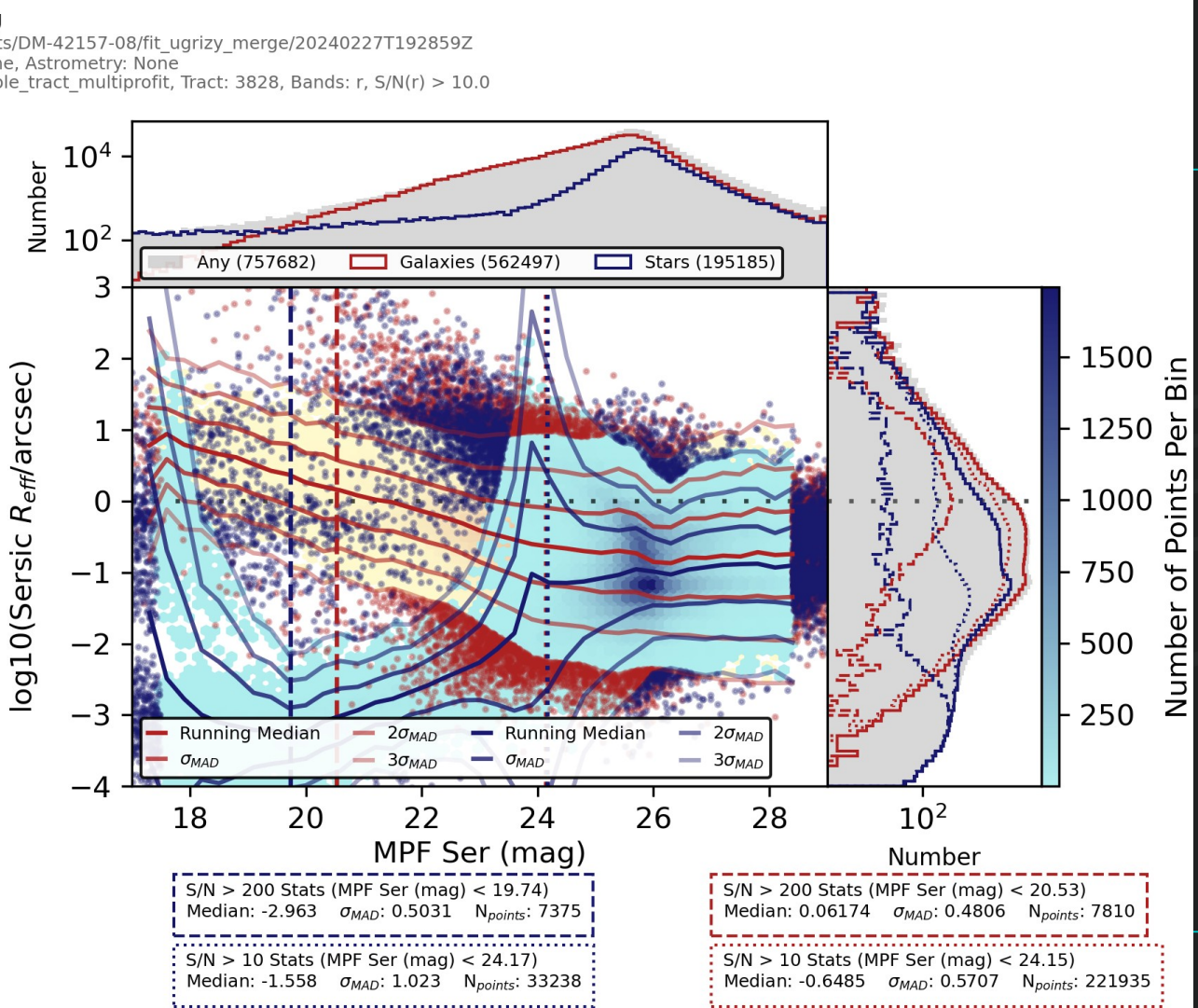
Keeping an eye out for “super spreaders” ...they still exist...



The MPF plot looks wild b/c I forgot flags.

Outliers are mostly false positive detections?

Still super spreaders.



Colors

I promise this is the last category of matched difference plots.

Colors usually have smaller scatter than mags – of course, b/c bias should correlate with no color gradients, and maybe with too.

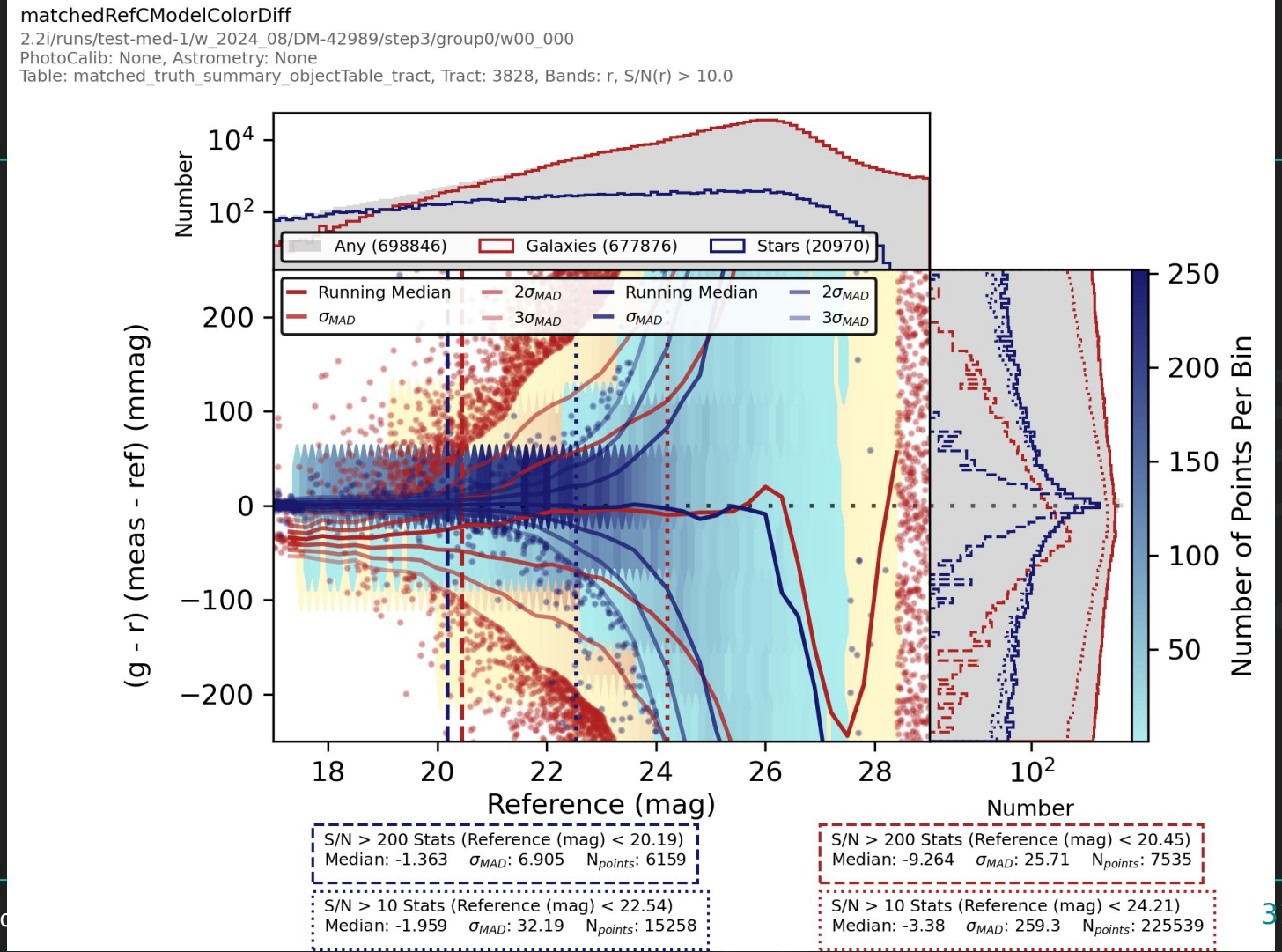
They'd better be good to pass to photo-z codes.

GaaP thought to be better – have we actually shown it in DC2?
(I have not checked yet)

CModel colors are biased.

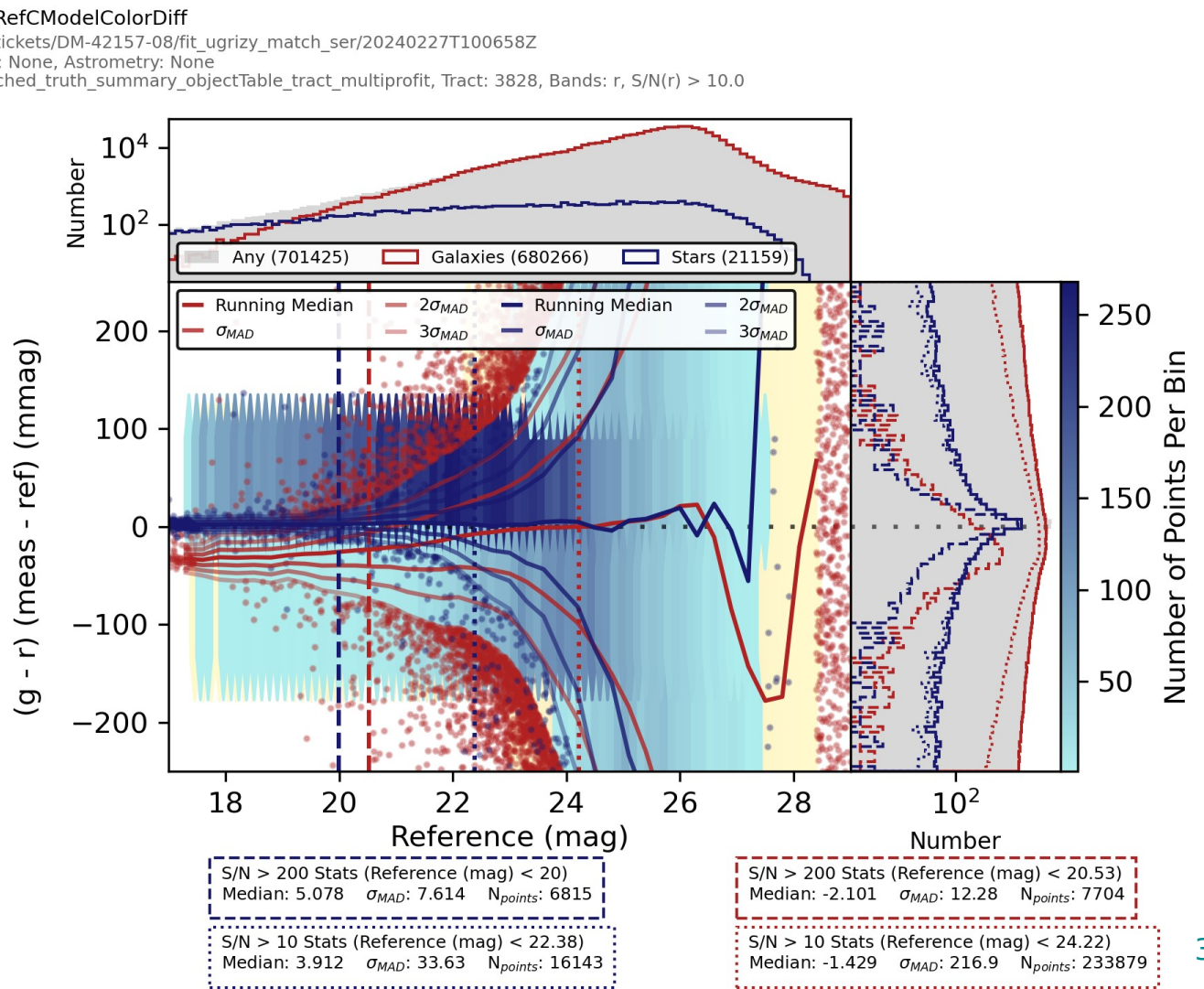
Bias goes up with brightness.

Is 10-20 mmag ok? Maybe.



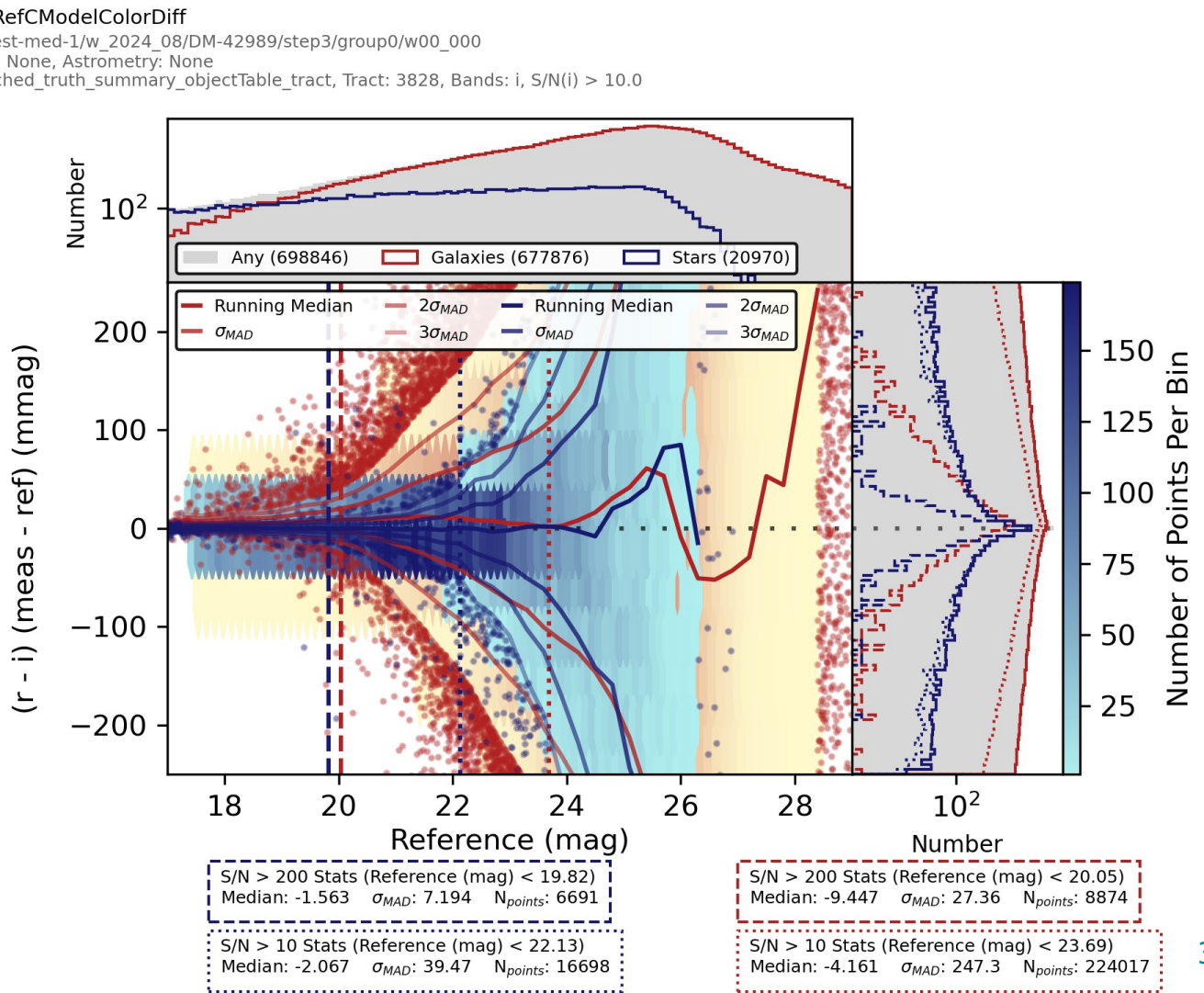
Sersic colors are biased, same as CModel.

Scatter is improved.



r-i not as biased.

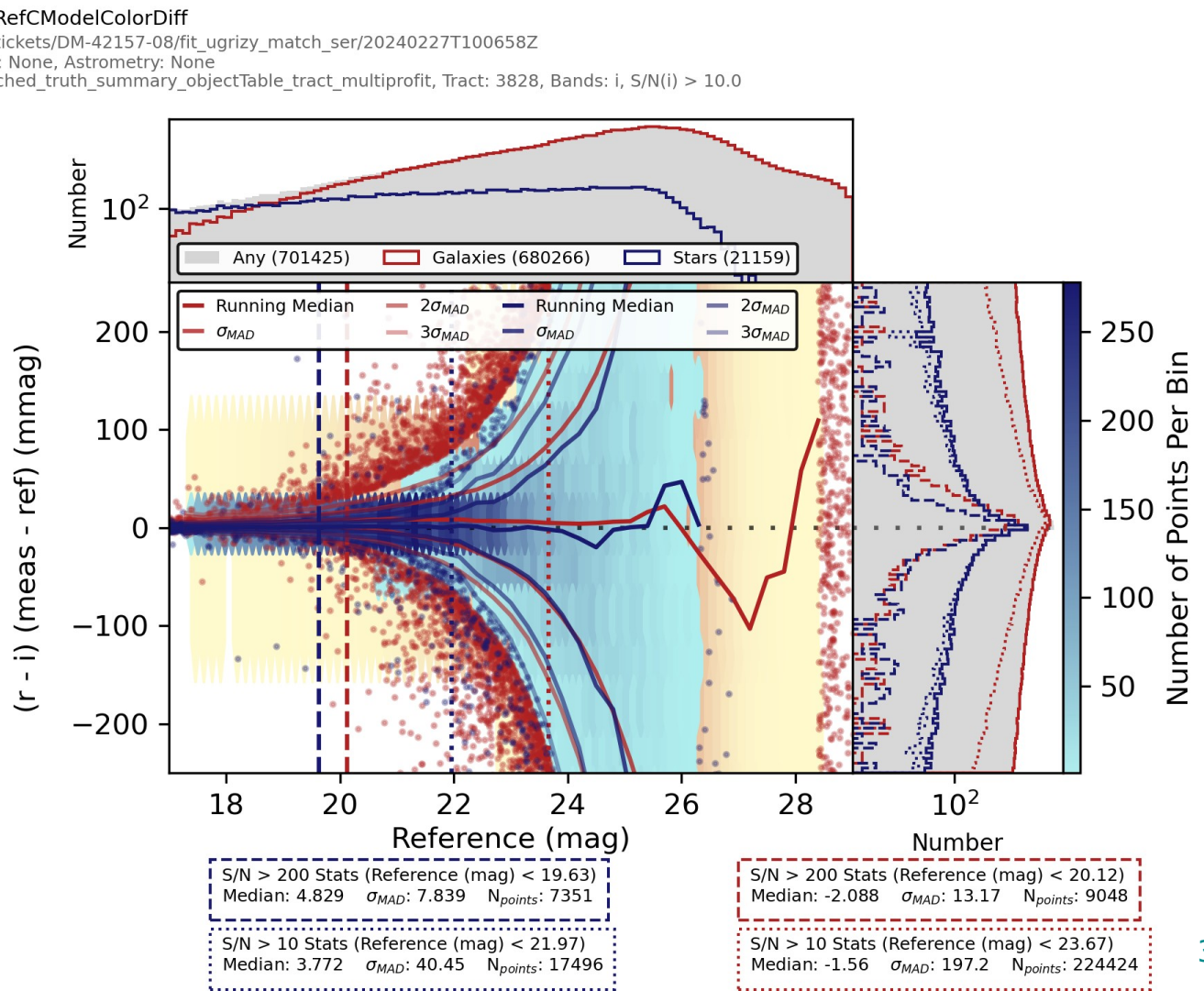
Might be worse if diff colors have diff bias...



Sersic r-i colors better than Cmodel.

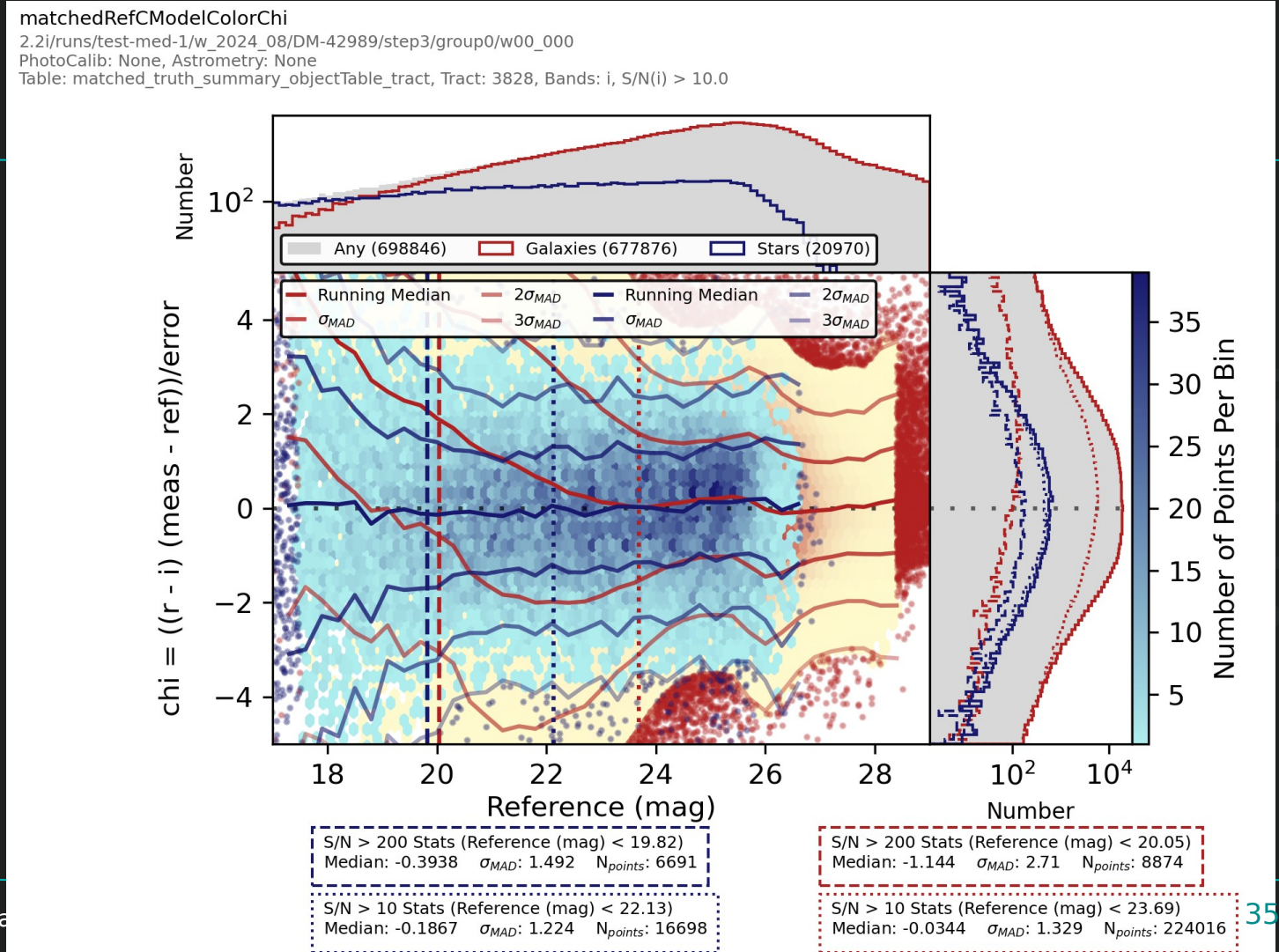
Actually, galaxy colors

Accuracy pretty close to stars...



CModel r-i errors not so bad.

Still too small for galaxies.



Sersic r-i errors actually sort of okay.

If the bias wasn't there, I'd be satisfied.

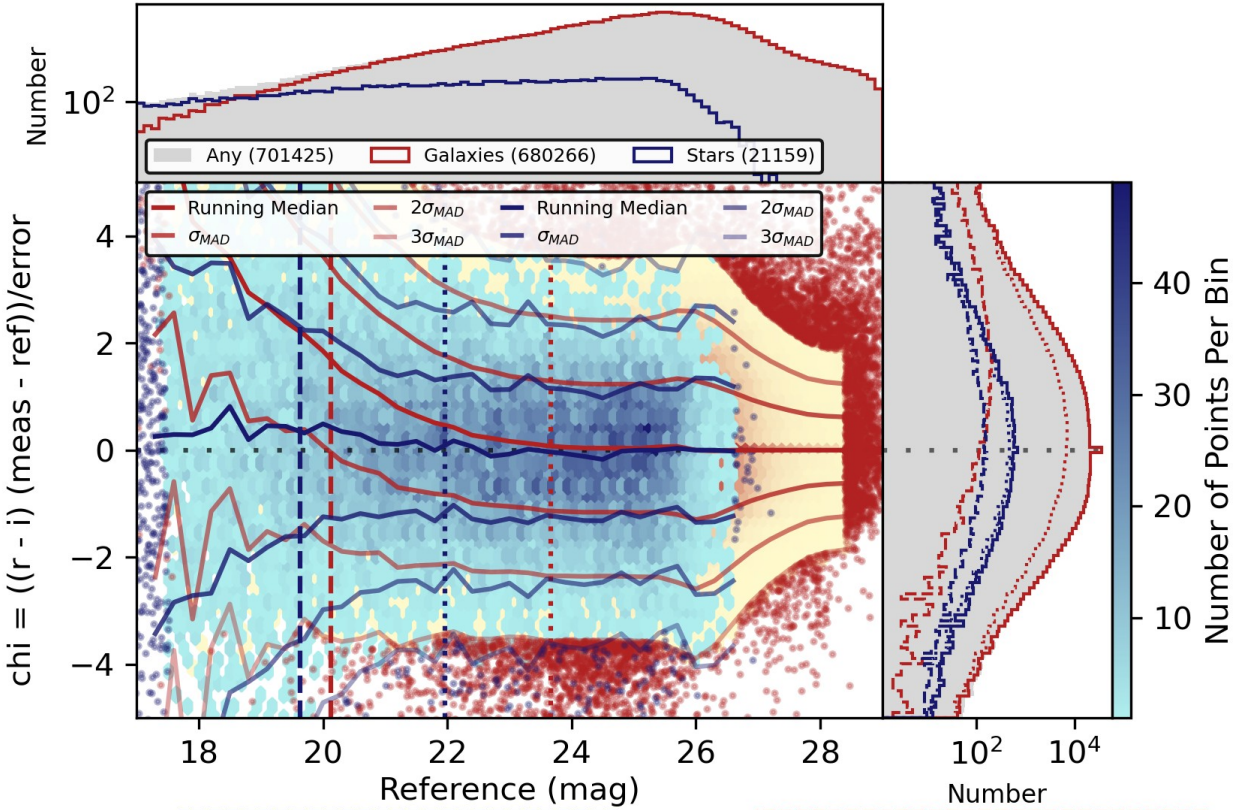
Low S/N errs too big!

matchedRefCModelColorChi

u/dtaranu/tickets/DM-42157-08/fit_ugrizy_match_ser/20240227T100658Z

PhotoCalib: None, Astrometry: None

Table: matched_truth_summary_objectTable_tract_multiprofit, Tract: 3828, Bands: i, S/N(i) > 10.0



Conclusions?

Next step: file RFC to add (meas_extensions_)multiprofit and dependencies to lsst_distrib

Test on HSC (not expecting surprises, but...)

Add to ci_imsim & ideally test-med-1 reprocessing

Consider merging columns into objectTable_tract

(caveat with outright replacing CModel immediately – single Sersic won't have bulge/disk fluxes. ExpDev would be ideal if it outperformed Ser...

Blend Inspect

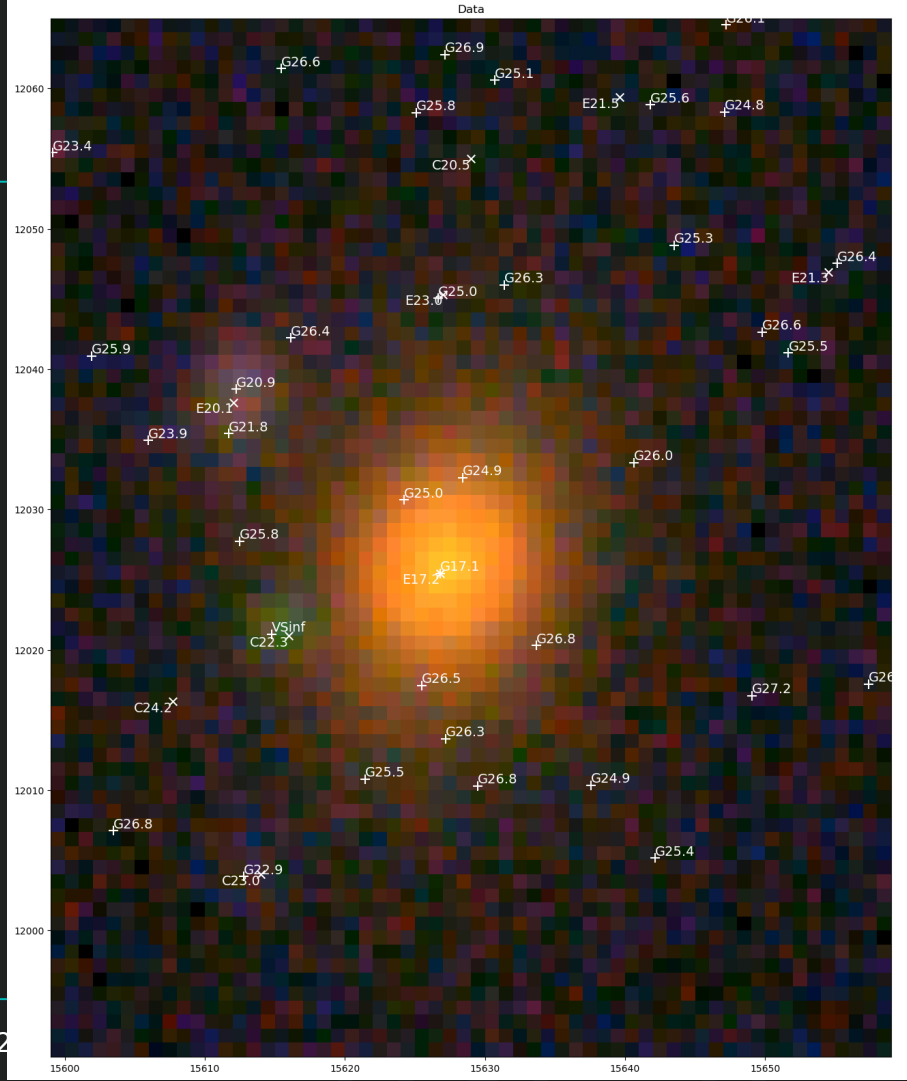
A task that loads a MPF fit catalog and its inputs and rebuilds models for child objects to inspect residuals.

This is one of the few processed blends in ci_imsim.

x: true mag (VS = variable *)

+: meas mag, Cmodel?

(C/E = compact/extended)



Blend Inspect

https://github.com/lsst-dm/multiprofit_validation/blob/tickets/DM-42270/notebooks/blend_inspect_ci_imsim_w_2024_05.ipynb

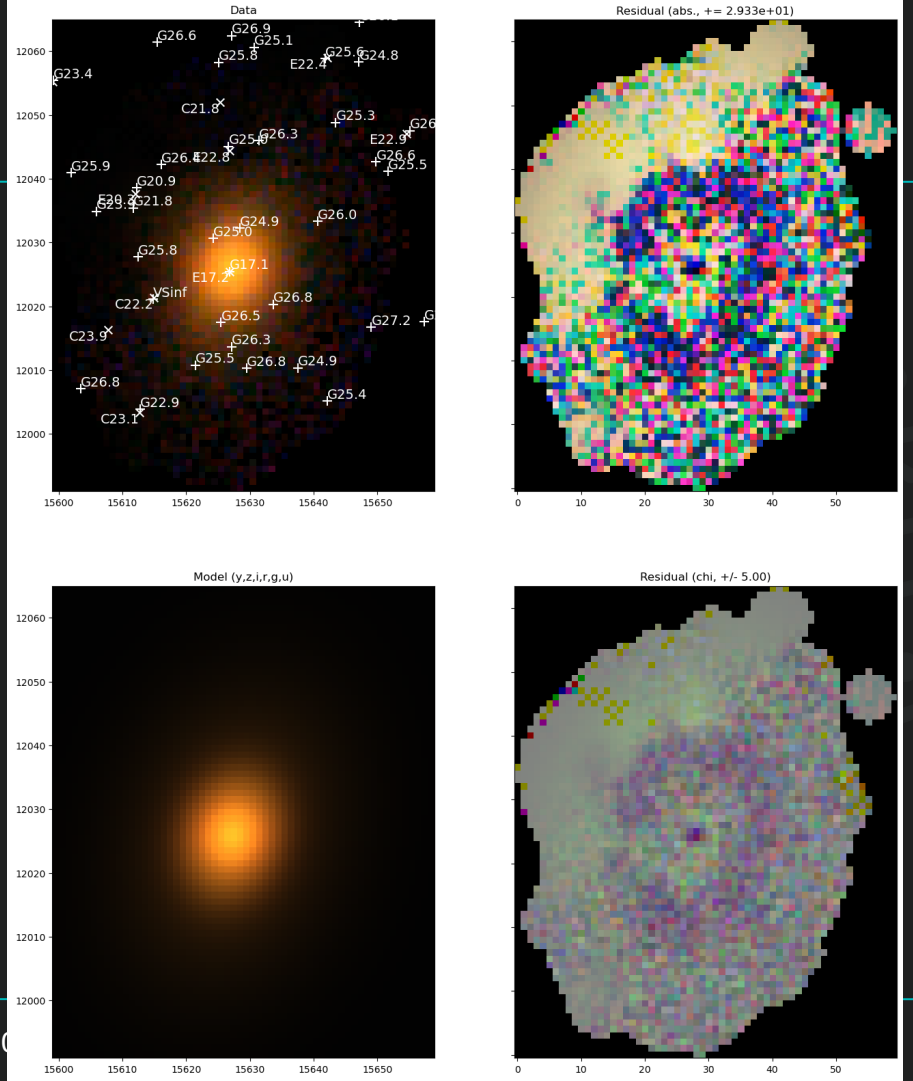
If you want to follow along (no spaces)

Sersic fit #1

The brightest object in the blend – it's a bright galaxy. Must have satellites.

Residuals suggest non-Sersic profile and/or color gradients

(I should plot the structural parameters too, and get the true ones... some day)

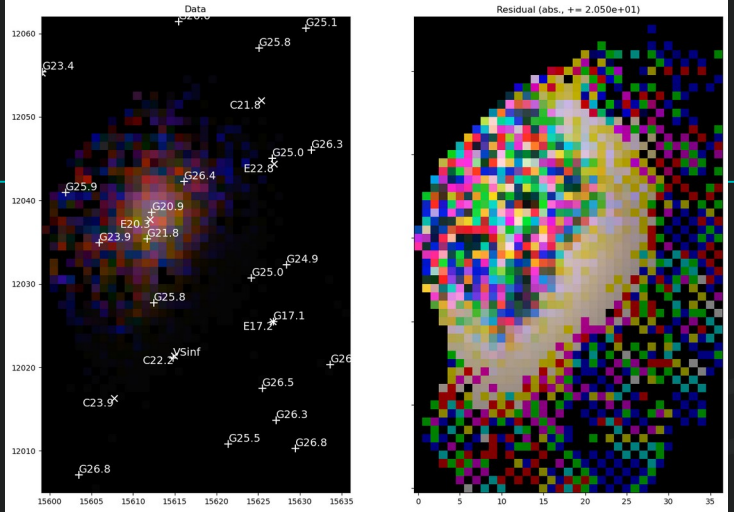


Sersic fit #2

Two blended real galaxies

Probably not possible to deblend at ci depths...

Remind me to check full DC2

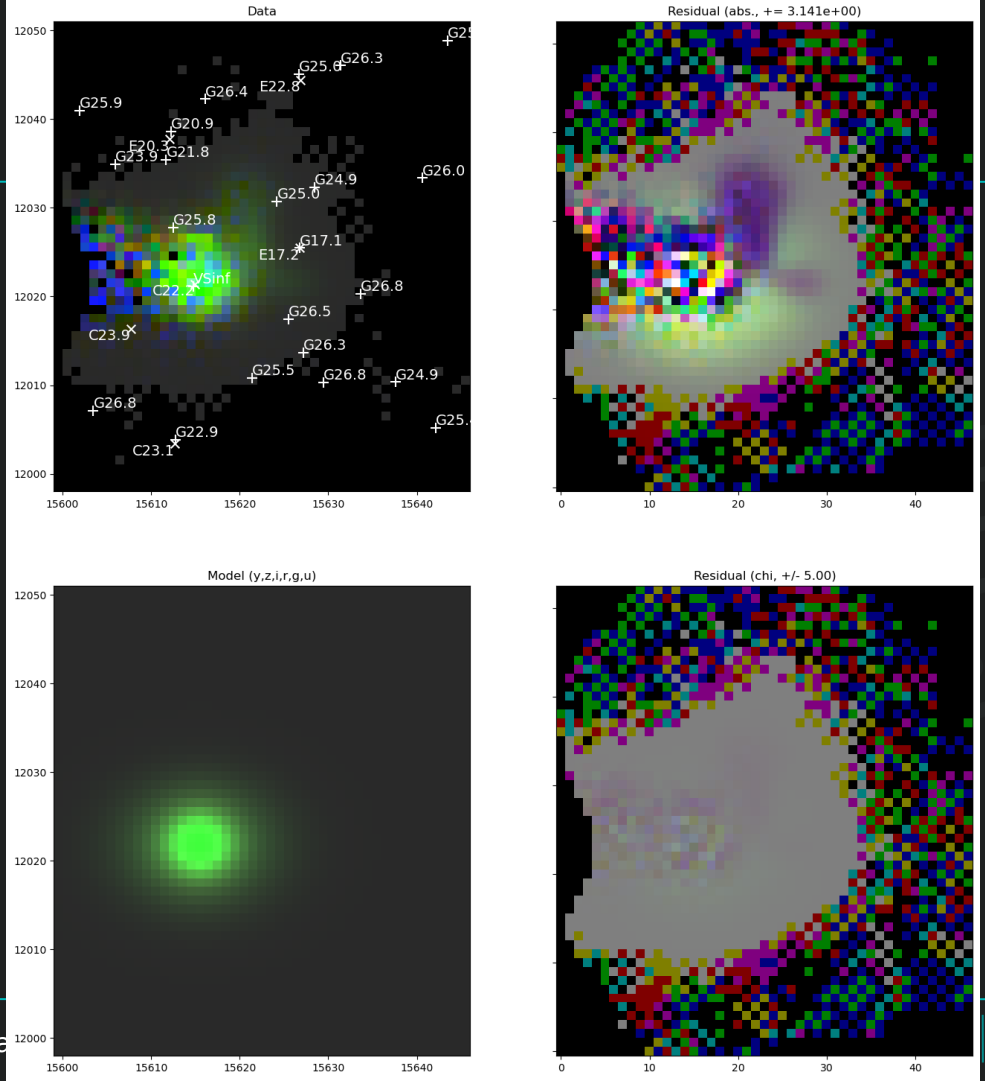


Sersic fit #3

This must be a supernova.

What are the odds?

At any rate, it’s a variable source. Most very green things are… hopefully.



Sersic #4

This isn’t anything. Probably just shredding of the brightest object.

It won’t match to anything and it will probably be large.

make_lupton_rgb kwargs could be improved here.