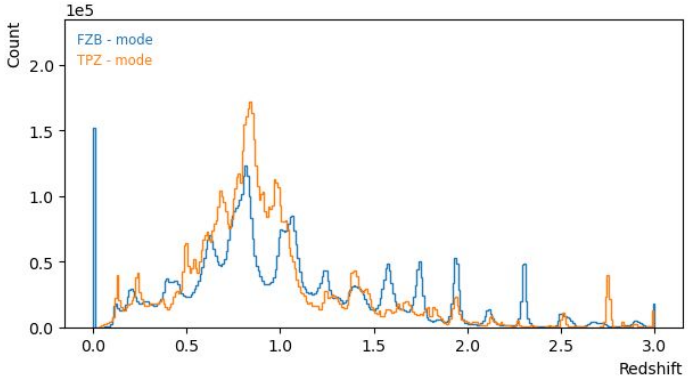
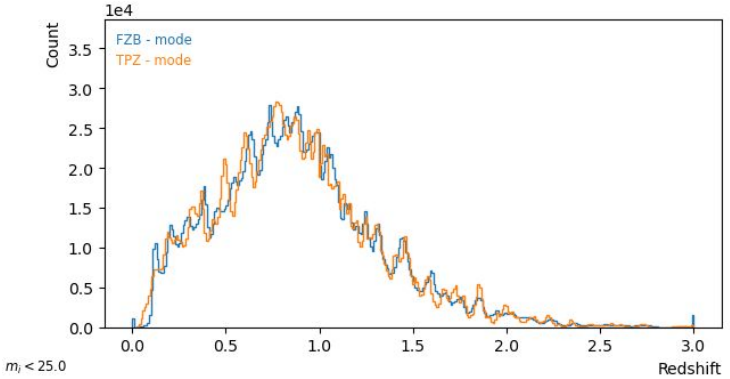
Effect of Photo-z Performance in Clusters

Rance Solomon (LAPP)

LSST-France meeting
Marseille
June 2024

From FZB to TPZ





➔ FlexZBoost photo-z algorithm displayed strong peaked behavior

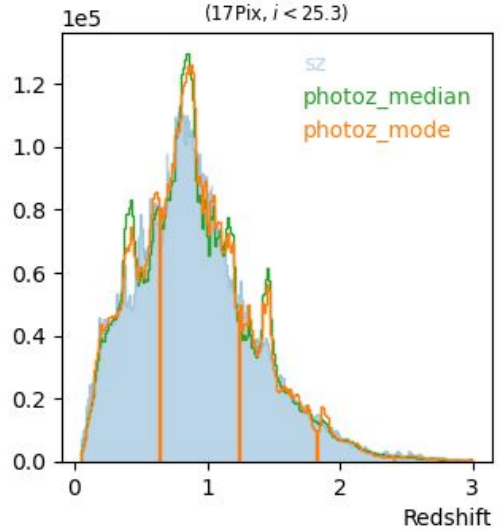
➔ TPZ (and several others) displayed similar systematics

➔ And reducing to represented sample still displayed significant peaked behavior

From FZB to TPZ

Work began with Sam Schmidt of the PZ WG to build a minimal working (i.e. not peaked) photo-z catalog.

- ➔ TPZ is simple Decision Tree based regresser – good for trouble shooting
- ➔ found that custom built decision tree was causing peaked structures
- ➔ switching decision trees also decreased running time by ~1000 fold



Final validation and performance tests are being done before adding to GCR, but currently available to play with at CC-in2p3:

```
/sps/lsst/groups/photoz/TPZ/estimated/validation_cosmodc2_v1.1.4/training_500k/
```

Project Purpose (DESC project [[353]])

Goal: Can we build a training set that is optimized for clusters?

Practical details:

- ➔ base galaxy catalog: cosmoDC2.small
- ➔ photo-z algorithm: TPZ (Trees for Photo-Z)
- ➔ cluster finder: WaZP

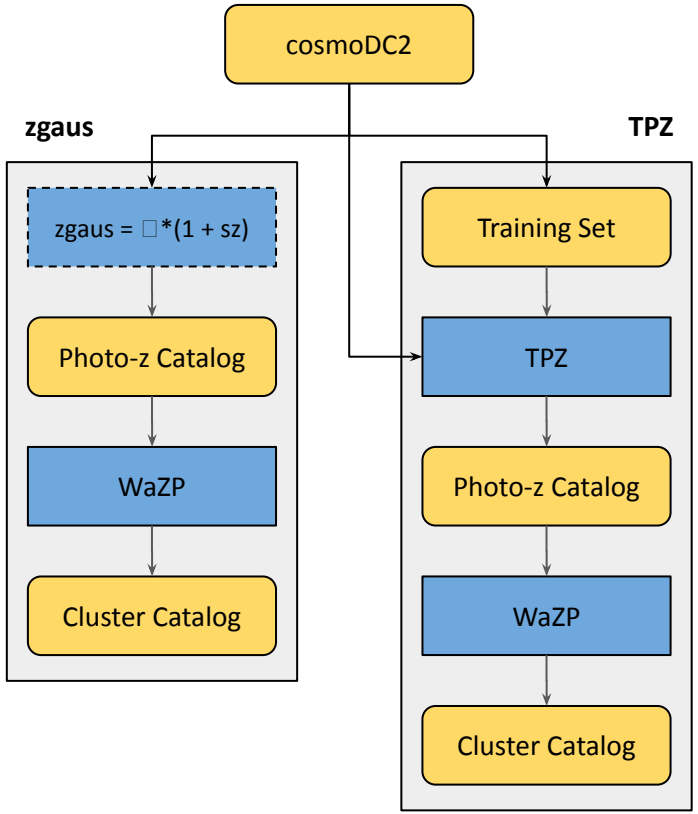
Study consists of:

Two mock Gaussian cases:

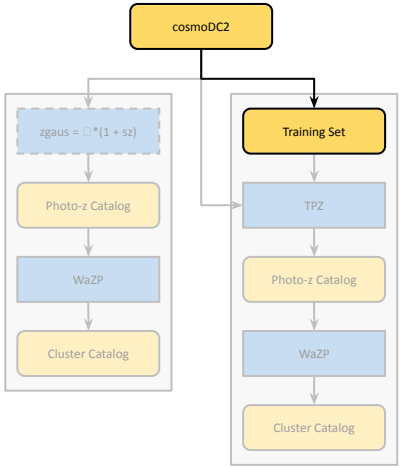
1. zgaus = 0.01*(1 + sz)
2. zgaus = 0.03*(1 + sz)

Four TPZ cases:

3. T30k – 30k random galaxies
4. T500k – 500k random galaxies (control set)
5. T500k.mbsOnly – 500k random *members* ($M > 10^{13.5} M_{\odot}$)



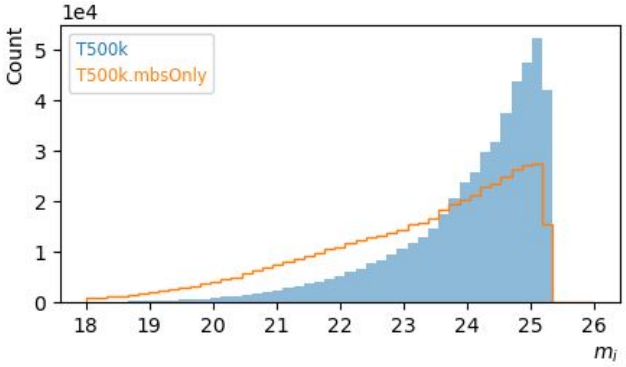
Building TPZ training sets

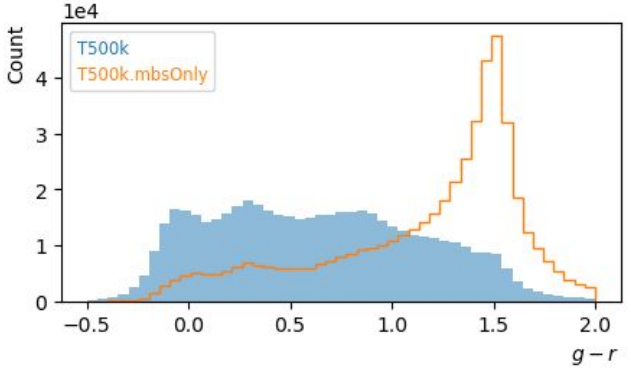


Random selection in RA-DEC-z space from across full cosmoDC2 catalog.

Training set requirements:

- ➔ mag_i < 25.3
- ➔ pass at least 3 of the LSST 10-year mag depths



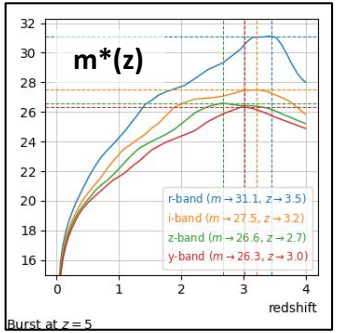


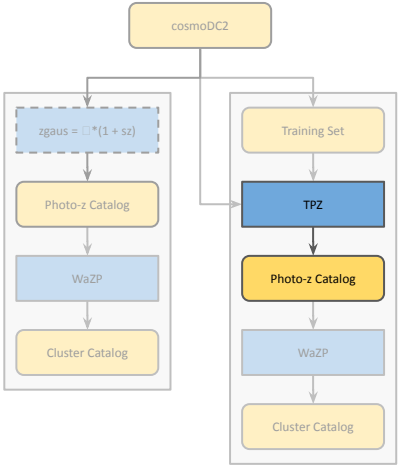
- ➔ basic training set dominated by high mag
- ➔ mbsOnly training set has brighter population
- ➔ mbsOnly training set is much more red

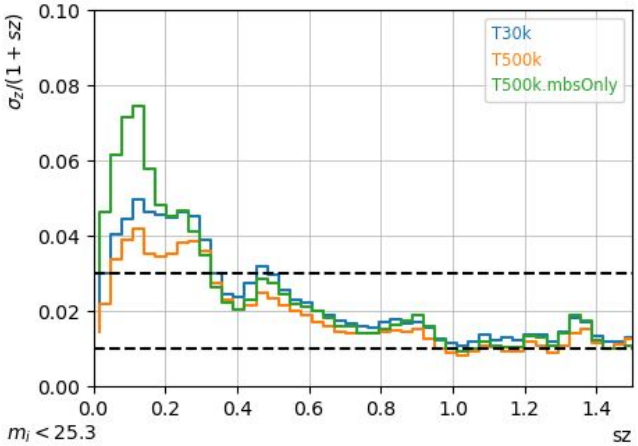
TPZ performance

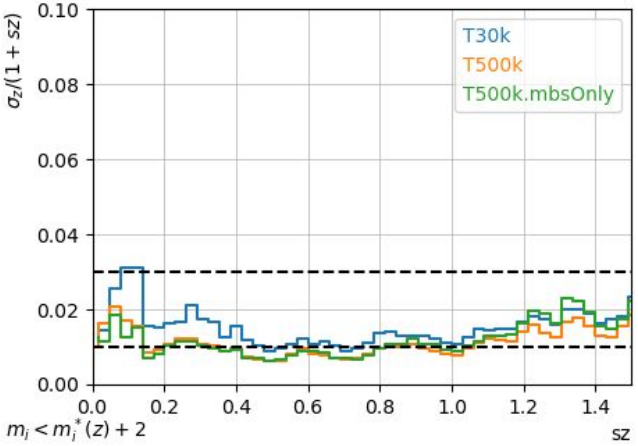
Training set has $m_i < 25.3$ but WaZP and AMICO apply additional m* cuts

- ➔ T500k.mbsOnly worse in low z – but improves significantly with m* cut
- ➔ training sets show little difference at high z

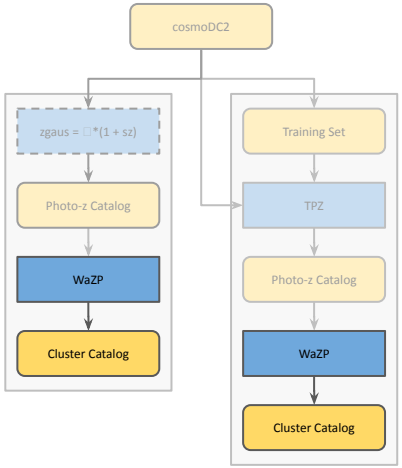




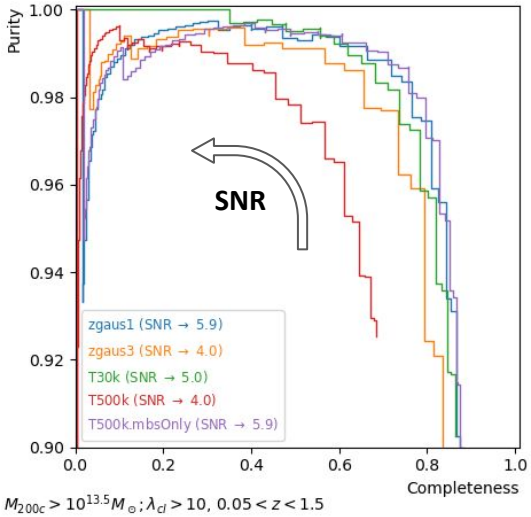


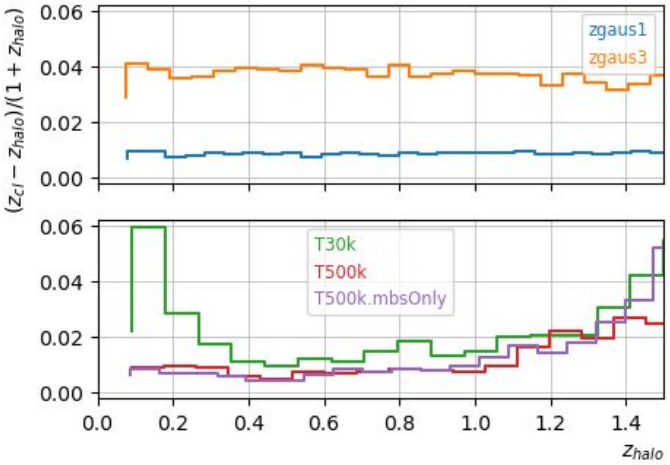


WaZP Detections

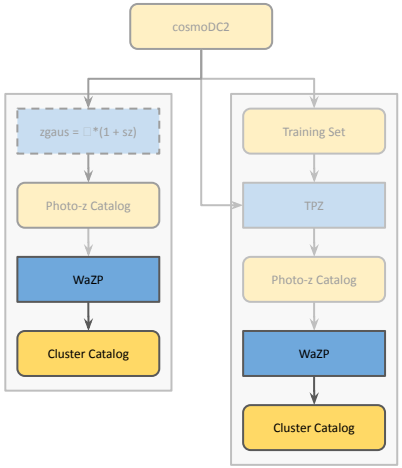


- ➔ T500k.mbsOnly performs as good as 1% zgaus
- ➔ T30k has better purity-completness than T500k but worse redshift resolution

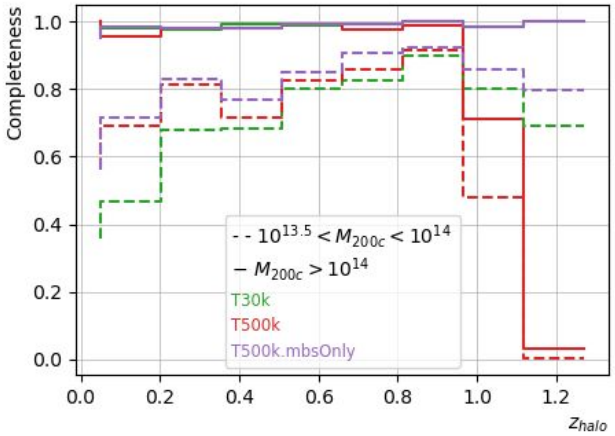


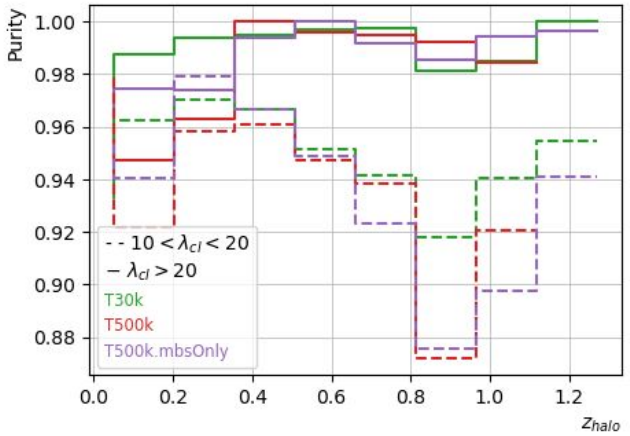


WaZP Detections



- ➔ T500k completeness drops off above z_{halo} = 1.1
  - needs to be looked into and may explain underperformance
- ➔ T30k has better purity in low richness clusters
  - poor photo-z may be acting as additional "filter" – washing out possible false detections



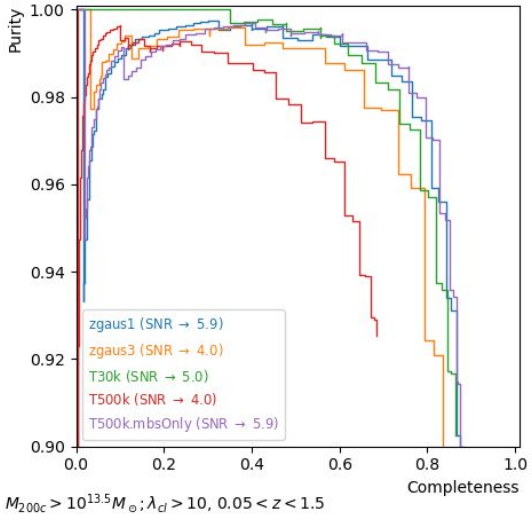


Conclusions

- ➔ T500k.mbsOnly run shows best performance – possibly using redder training set can reproduce this
- ➔ T500k run has issue at high redshift that needs to be revisited

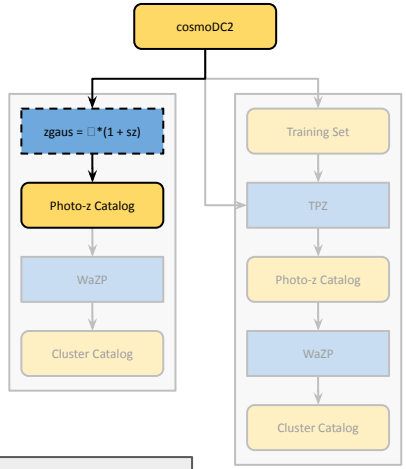
Future work:

- ➔ brighter training set
- ➔ redder/bluer training set

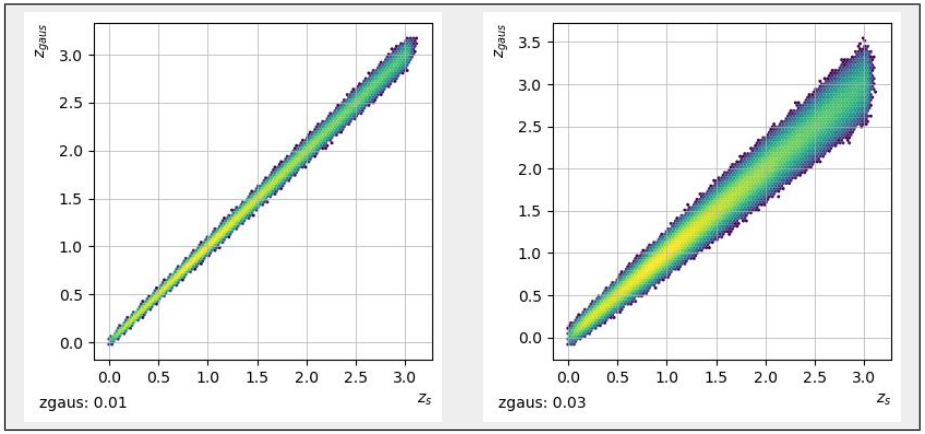


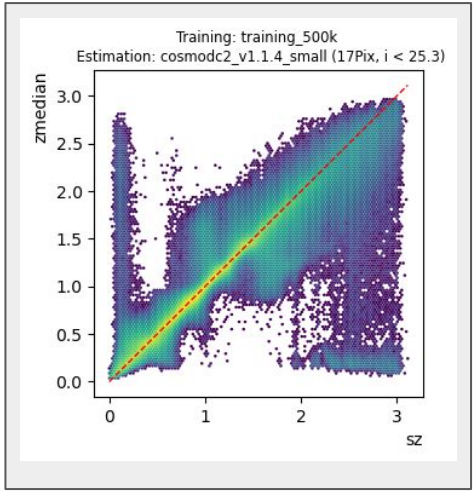
Backup

Building zgaus



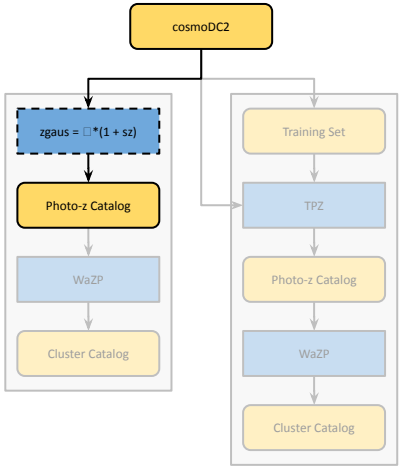
- ➔ zgaus provides “ideal” pz-sz plane – compare to TPZ
- ➔ PZ WG aims for 1% scatter in point estimates for cosmoDC2 simulation
- ➔ 3% scatter represents “worse case scenario”





Note: colormaps in log scale

Building zgaus



Caution:

- zgaus assumes photo-z performance is mag independent
- but photo-z's will perform better at lower mag
- since cluster members are generally brighter, zgaus is not *ideal* but is *simple*