

The WaZP DC2 Run: ~~Performance and Comparison to redMaPPer~~

ANOTHER DAY



Rance Solomon
Postdoc @ LAPP
solomon@lapp.in2p3.fr



LSST-France Meeting
Grenoble, France
June 2023

In collaboration with:

Michel Agüena (APC)
Christophe Benoist (OCA, LineA)
Dominique Boutigny (LAPP)
Luiz da Costa (LineA)
Thibault Guillemin (LAPP)



In this talk:

- WaZP introduction
- the WaZP DC2 run with FlexZBoost
- FlexZBoost issues at high magnitude
- status update on Python WaZP

Related projects:

[DC2 Project](#): The WaZP cluster catalog on DC2 data

Create and evaluate the properties of the WaZP cluster catalog on DC2 truth and image sim

DESC Project: [[248](#)]

Comparison of cluster finder algorithms on DC2 data

What have we learned in the past 2 days?

- 1) some stuff that does not provide a good segway into this talk
- 2) **Marina: galaxy clusters provide a strong probe into cosmology**
- 3) bad things happen to those who don't survey 🦖 + 🐒
- 4) more unrelated things
- 5) **Nathan: to do cluster cosmology we need well tested cluster finding algorithms**
- 6) other stuff that happened between Nathan and now

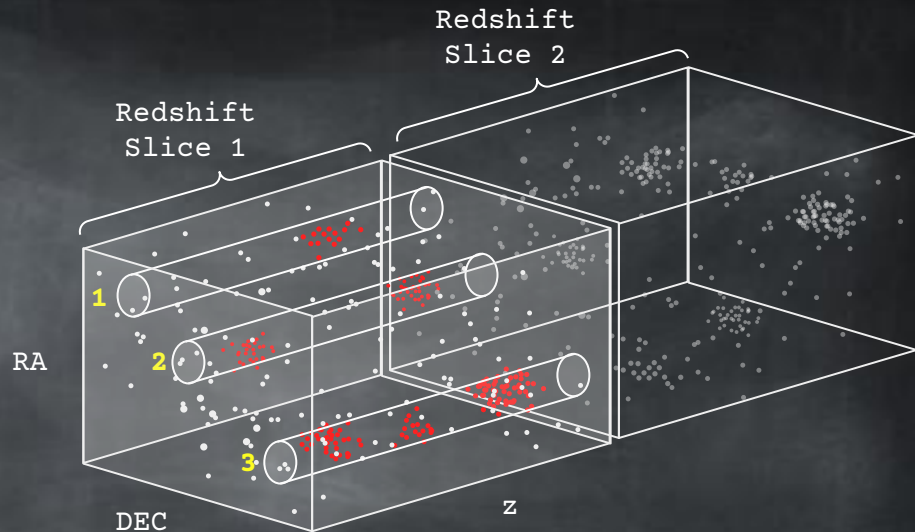
WaZP (written by Christophe Benoist – LIneA)

WaZP: a spatial density based cluster finder

How they do it:

- Redshift space is cut into thin slices.
- On each slice a cluster-scale wavelet kernel is applied to create density maps.
- Density maps are cut into densely packed cylinders ($r \approx 300\text{kpc}$).
- Over-densities are then taken as clusters.
- Each cluster member is assigned a membership probability based on distance from cluster center.

Optimized for photometric redshifts making it a good candidate for a DESC clustering algorithm.



Cylinder 1:



Cylinder 2:



Cylinder 3:

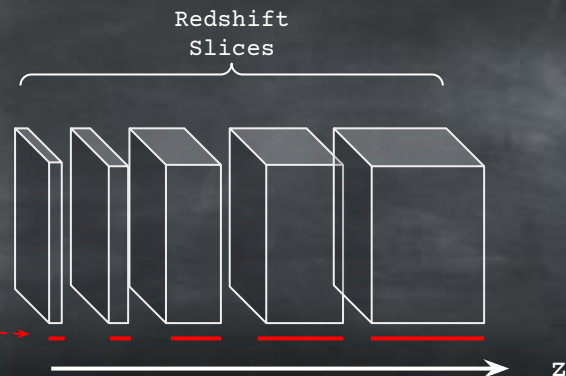
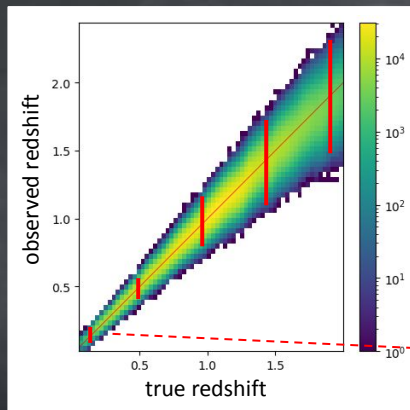
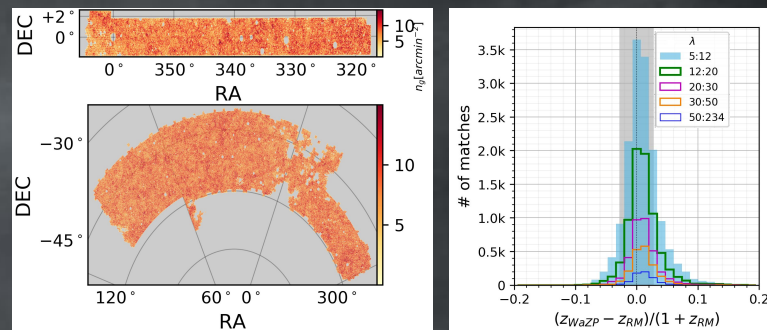


WaZP (written by Christophe Benoist – LineA)

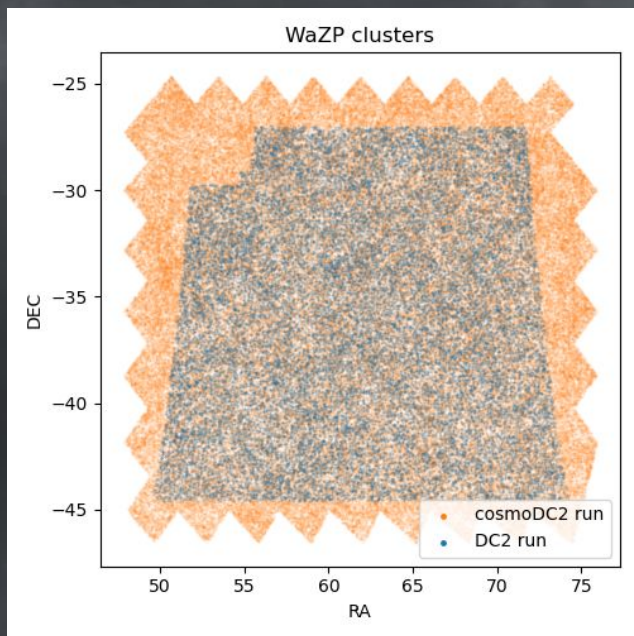
M. Aguena, et. al. ([arxiv:2008.08711](https://arxiv.org/abs/2008.08711))

FORTTRAN version:

- Performed well on DES Y1 data, on par with redMaPPer.
- Version used to produce WaZP related GCR catalogs.
- Ran on cosmoDC2 in ~ 3 hours on 1200 cores.
- Also used for WaZP DC2 run.
- Exclusively ran at LineA (i.e. not available on CC-in2p3).
- A key step is to model growth in galaxy redshift scatter to adjust slice width accordingly.
 - Otherwise, high z clusters will be biased towards low richness.



WaZP Catalog Status



Catalogs now in GCR:

cosmoDC2_v1.1.4_wazp_v1.0_truez

cosmoDC2_v1.1.4_wazp_v1.0_flexzboost_v1

Produced but not yet in GCR:

WaZP run on DC2 with FlexZBoost

FlexZBoost reminder: ([github](#))

- converts mean point redshift estimates to pdzs
- based on FlexCoDE, uses XGBoost to perform regression

WaZP Performance on DC2

At catalog level

WaZP DC2 catalogs:

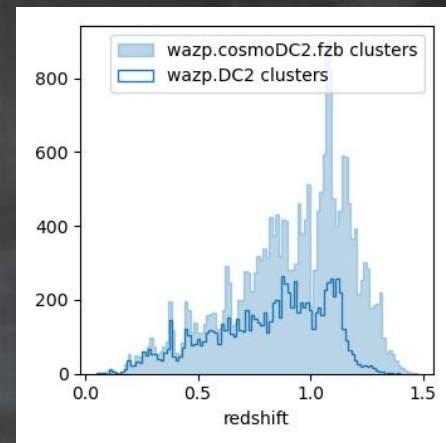
- /sps/lsst/groups/clusters/dc2/cosmoDC2_v1.1.4/run_2.2i/dr6/wazp/7081/

DC2 truth catalog:

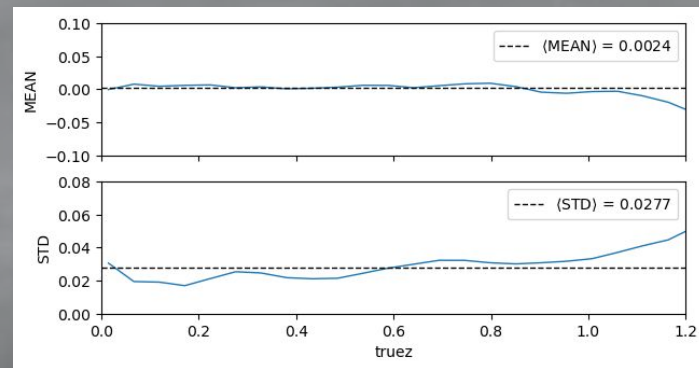
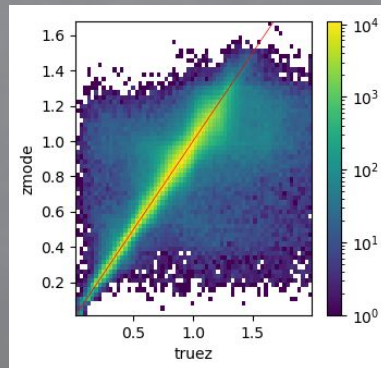
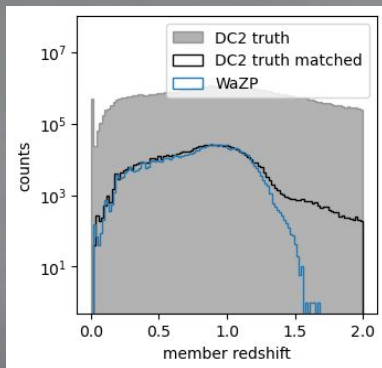
- /sps/lsst/groups/desc/PZ/PhotoZDC2/run2.2i_dr6_v2/CATALOGS_W_TRUTH/ \ Run2.2i_dr6_dereddened_tract_{tract}_withtruez.hdf5
- /sps/lsst/groups/desc/PZ/PhotoZDC2/run2.2i_dr6_v2/FLEXZBOOST/PDF_FILES/

applied cuts:

- $NGALS > 20$
- $m_z < m_z^* + 2$



Galaxy pz from FlexZBoost



WaZP Performance on DC2

At matched level

```
## Finding unique matches of catalog 1
Unique Matches (cosmoDC2_v1)
+ 34,413/471,965 objects matched.
```

```
## Finding unique matches of catalog 2
Unique Matches (wazp.DC2.v0)
+ 34,413/37,993 objects matched.
Cross Matches (cosmoDC2_v1)
+ 34,411/471,965 objects matched.
Cross Matches (wazp.DC2.v0)
+ 34,411/37,993 objects matched.
```

```
cosmoDC2_v1
* CLEvar used in matching: 0.13.7
* Total objects: 471,965
* multiple (self): 62,495
* multiple (other): 62,495
* unique (self): 34,413
* unique (other): 34,413
* cross: 34,411
```

```
wazp.DC2.v0 ← NGALS > 0
* CLEvar used in matching: 0.13.7
* Total objects: 37,993
* multiple (self): 35,174
* multiple (other): 35,174
* unique (self): 34,413
* unique (other): 34,413
* cross: 34,411
```

$$\frac{34\,411}{37\,993} = 0.906$$

```
## Finding unique matches of catalog 1
Unique Matches (cosmoDC2_v1)
+ 9,501/471,965 objects matched.
```

```
## Finding unique matches of catalog 2
Unique Matches (wazp.DC2.v1)
+ 9,501/9,658 objects matched.
Cross Matches (cosmoDC2_v1)
+ 9,501/471,965 objects matched.
Cross Matches (wazp.DC2.v1)
+ 9,501/9,658 objects matched.
```

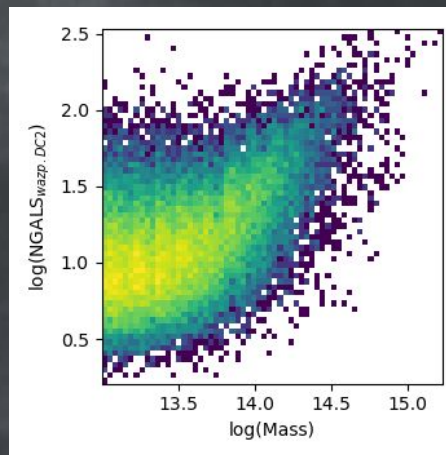
```
cosmoDC2_v1
* CLEvar used in matching: 0.13.7
* Total objects: 471,965
* multiple (self): 22,500
* multiple (other): 22,500
* unique (self): 9,501
* unique (other): 9,501
* cross: 9,501
```

```
wazp.DC2.v1 ← NGALS > 20
* CLEvar used in matching: 0.13.7
* Total objects: 9,658
* multiple (self): 9,501
* multiple (other): 9,501
* unique (self): 9,501
* unique (other): 9,501
* cross: 9,501
```

$$\frac{9\,501}{9\,658} = 0.984$$

CLEvar ([github](#)):

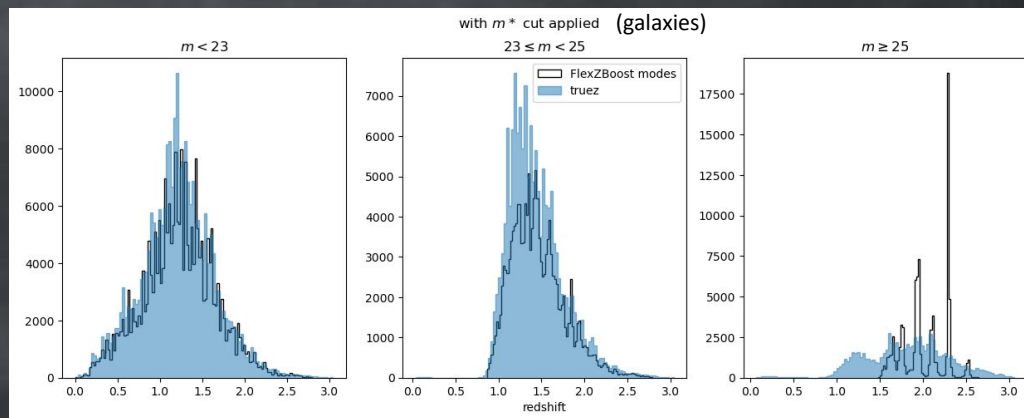
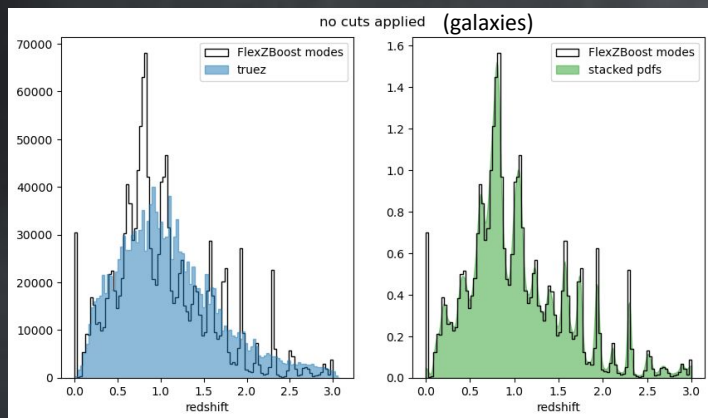
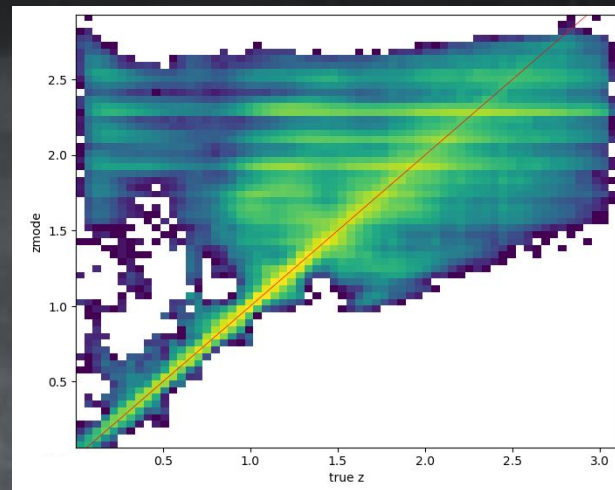
library to validate cluster detection, written by Michel Aguena



Trouble with FlexZBoost runs

FlexZBoost has an issue in redshift recovery for dim objects.

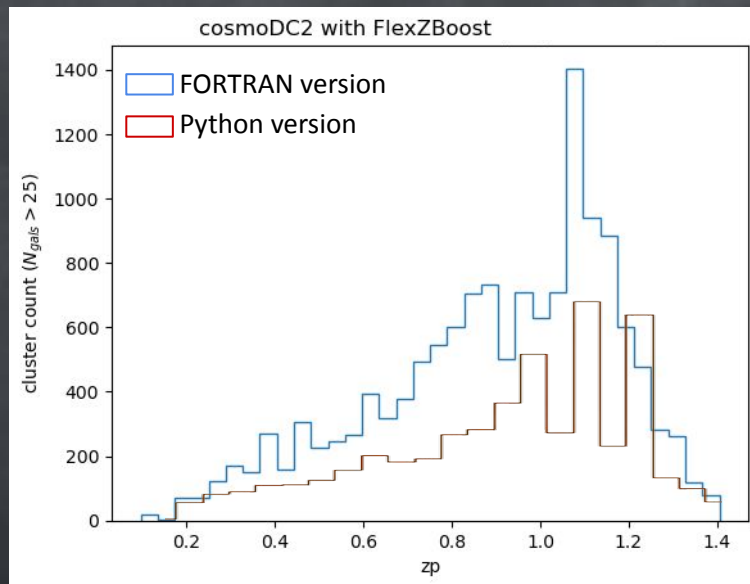
- FlexZBoost was trained on $i < 25.3$ objects.
- Dimmer objects are having the FlexZBoost basis functions imprinted on them.
- In discussion with pz working group for possible solutions.



WaZP Python ([github](#))

Python version:

- In development now.
- Will be executable on CC-in2p3.
- Ran on DC2 in ~2 hours (18 threads, 36 tiles)
- Some current issues in estimation of cluster redshifts – working on it!



Summary

- WaZP cluster finder does not depend on astrophysics and is optimized for photometric redshifts.
- FORTRAN version of WaZP has been run on cosmoDC2, cosmoDC2.fzb, and DC2 catalogs
- Understanding how the scatter of redshift evolves is important to the cluster finder.
- WaZP with FlexZBoost catalogs display peaks in redshift distribution for dim objects.
- Python version of WaZP is around the corner!

Merci beaucoup!

Backup:

Effect of dim FlexZBoost galaxies on cluster count. (produced by Thibault Guillemin)

