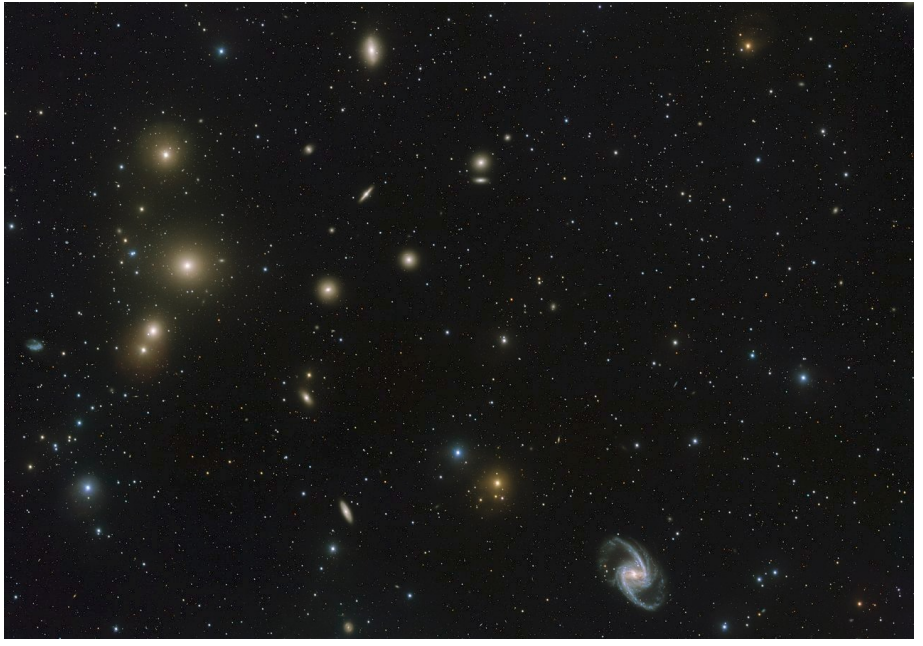
Towards a photometric-selected catalog of Fornax Cluster galaxies

Laerte Sodré Jr. - IAG/USP



thanks to Erik, Juliana , Isabela, Analía, Paulo, Claudia, Clécio, Fábio, Amanda, Gustavo, ...

FAPESP CNPq

What galaxies in the FC direction are indeed cluster members?

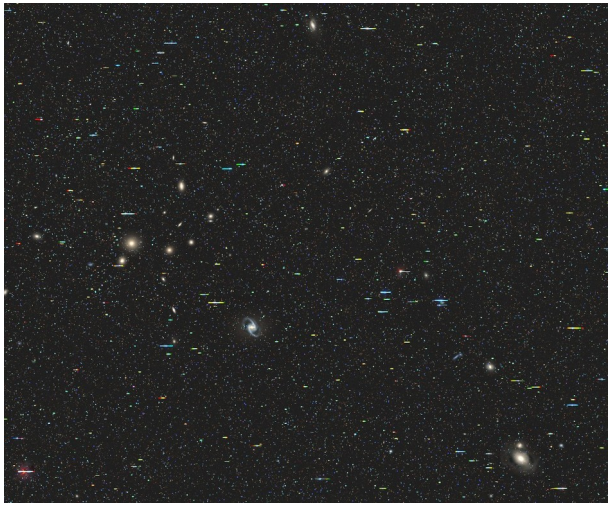
- **only 253 unique identifications with z_{spec} in Maddox et al. (2019)**

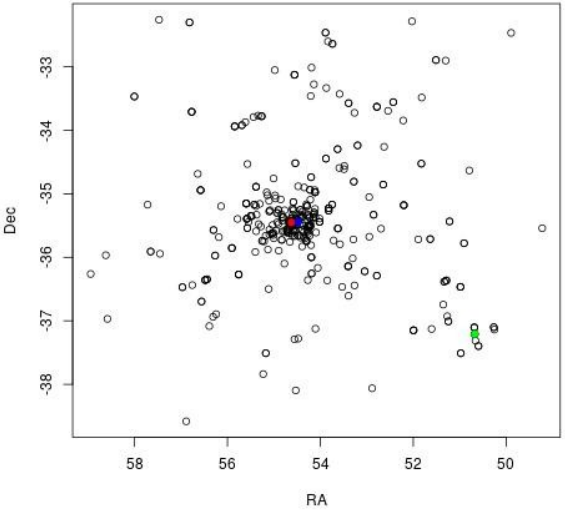
- **photometric redshifts don't work: z_{fornax} ~ 0.005**

  typical photo-z errors, ~0.020

  (Erik V.R.L+2021, arXiv:2110.13901)

- ***how to obtain a "reliable sample" from S-PLUS photometry?***





FC membership through classification

- **approach: two-class classification:**

  class 0: z > zlim;

  class 1: zmin < z < zlim zmin=0.002 zlim = 0.0075

- **input for classification: 11 colors (wrt to r_petro) , r_petro, MU_MAX, R20**

  r_petro < 21.3 maximum error in gri = 0.1

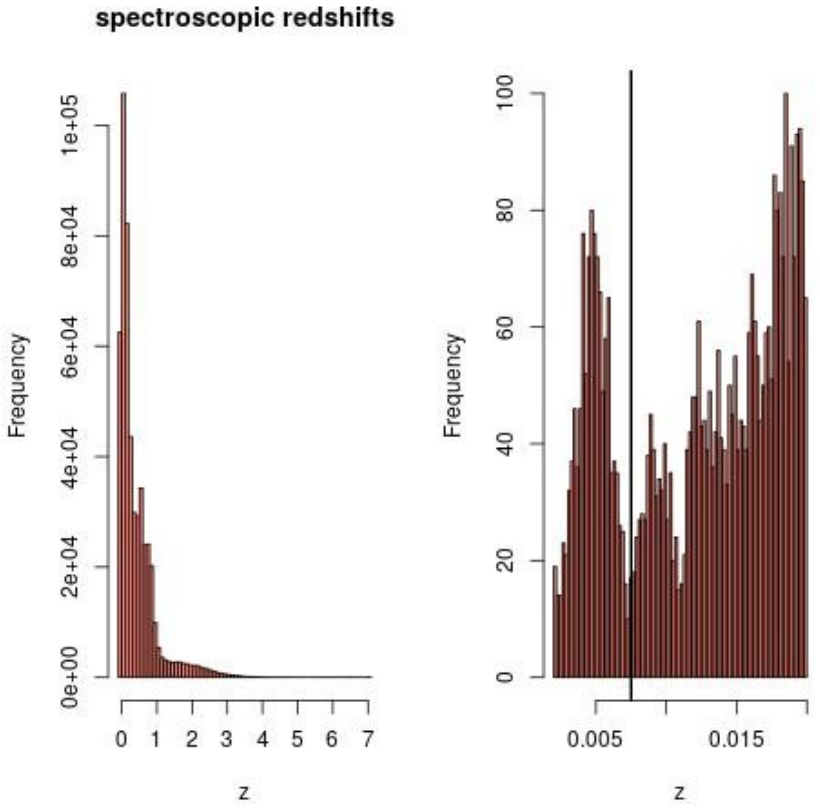
- **training data: iDR3_Petro_SpecZ_V6_CCM89.csv**

  512607 spectroscopic redshifts

  105058 with complete input data

- **437 objects in class 1 and 105058 objects in class 2**

  only 437 objects in class 1!



FC membership through classification

- *only 437 objects in class 1!*
- **classification requires balanced samples and**

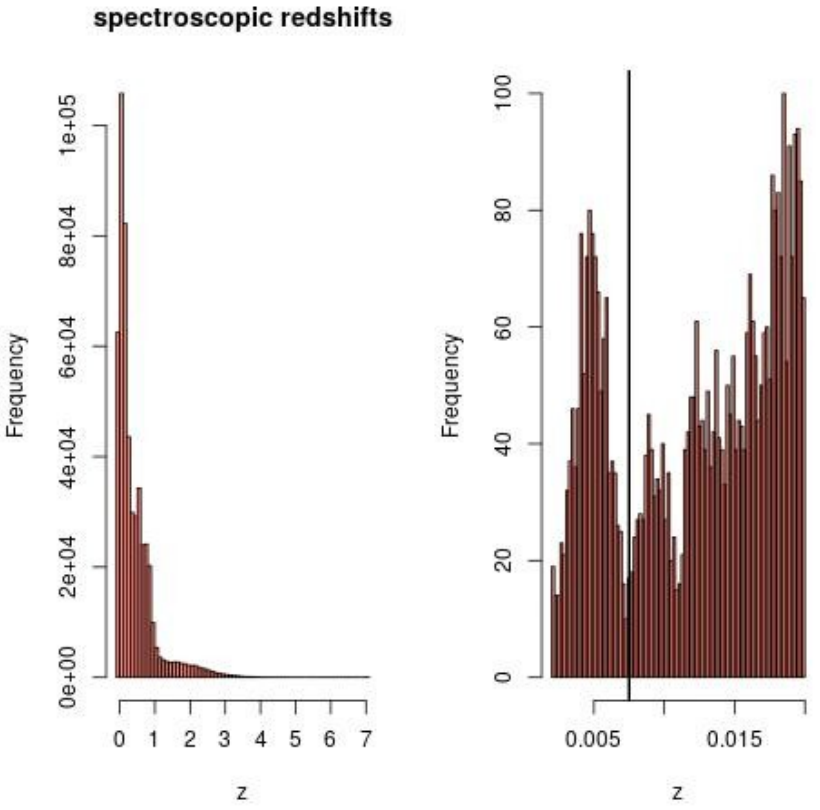
  machine learning likes high amount of data

- **data augmentation:**

  1. create a sample 30 larger of class 1 objects by sampling from input data errors (and assuming that the MU_MAX error = r_petro error, and neglecting R20 errors): 437 x 30 = 13110

  2. sample randomly the same number of class 0 objects

  3. repeat this 100 times, producing a classification each time



Deep Learning model

- *only 437 objects in class 1!*
- classification requires balanced samples and

  machine learning likes high amount of data

- data augmentation:

  1. create a sample 30 larger of class 1 objects by sampling from input data errors (and assuming that the MU_MAX error = r_petro error, and neglecting R20 errors): 437 x 30 = 13110

  2. sample randomly the same number of class 0 objects

  3. repeat this 100 times, producing a classification each time

- training, validation, test:

  80%, 10%, 10%

  each run: 13110 x 2 = 26220 objects

- algorithm:

  densely connected net: 11:128:128d:64d:16:1

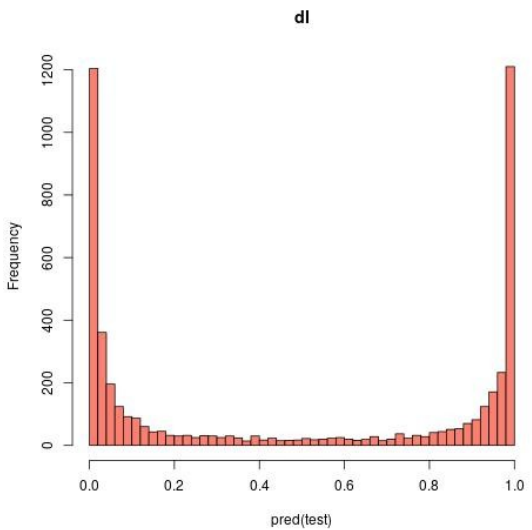
```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense (Dense)                (None, 128)               1920
_________________________________________________________________
dense_1 (Dense)              (None, 128)               16512
_________________________________________________________________
dropout (Dropout)            (None, 128)               0
_________________________________________________________________
dense_2 (Dense)              (None, 64)                8256
_________________________________________________________________
dropout_1 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_3 (Dense)              (None, 16)                1040
_________________________________________________________________
dense_4 (Dense)              (None, 1)                 17
=================================================================
Total params: 27,745
Trainable params: 27,745
Non-trainable params: 0
```

Deep Learning model

- **Training:**
- **example of a confusion matrix**

	T0	T1
C0	1289	19
C1	76	1293

median accuracy: 96.7%

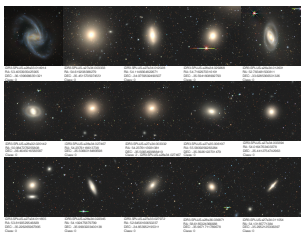


- **Fornax sample: Fornax_Laerte_All_02.10.csv**
- **selection: complete data,**

  r_petro < 21.3, Prob_Gal > 0.8

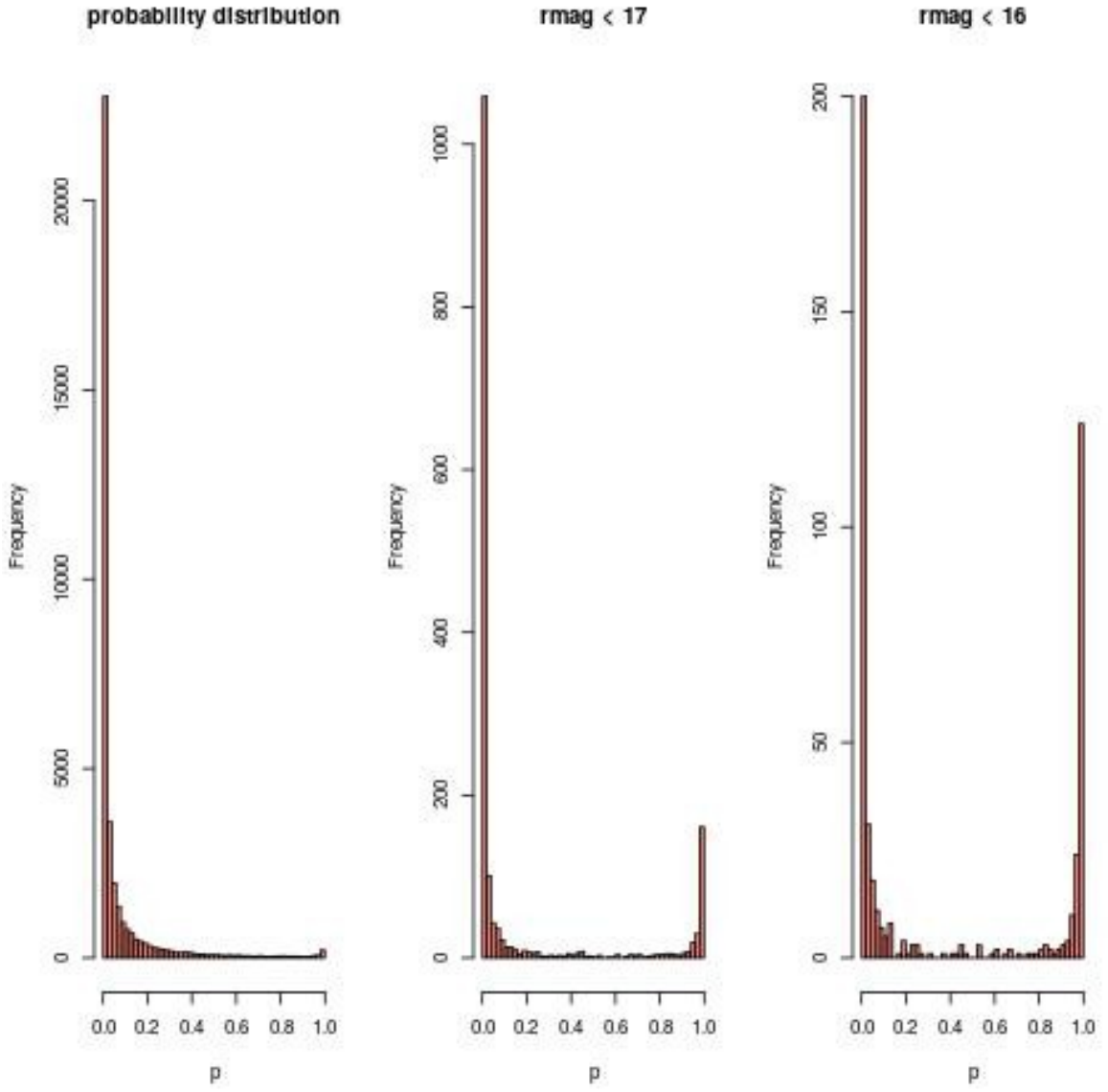
  37278 objects
- **for each simulation we computed the probabilities of class = 1, and the final probability for each object is the mean of 100 simulations**
- **To reduce contamination we consider as cluster members objects with p > 0.7**

  sample size: (~) 842 objects

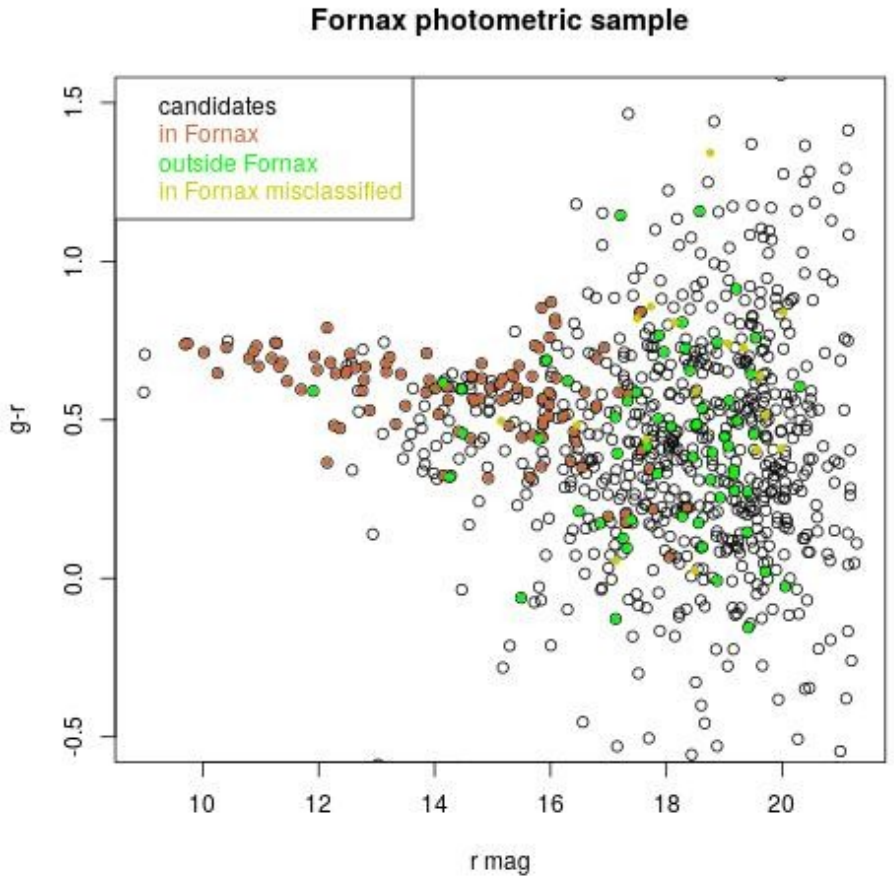


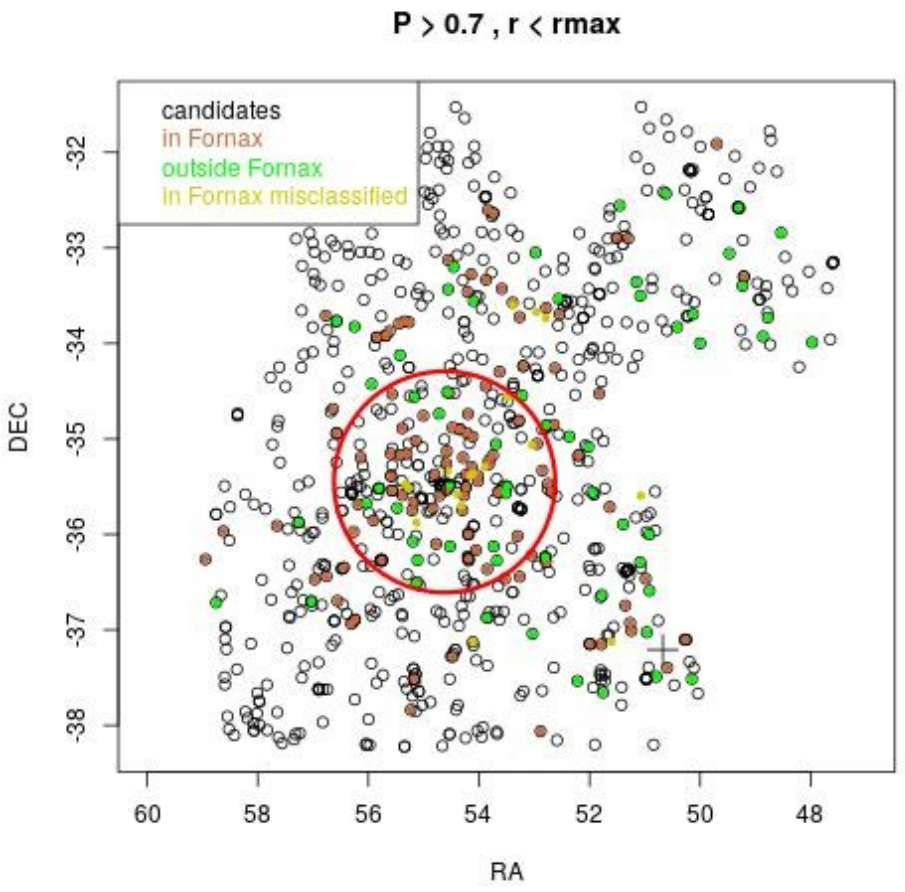
***cleaning the sample with legacy images*: contaminants, noise, multiple detections - 573 candidates**

Fornax sample

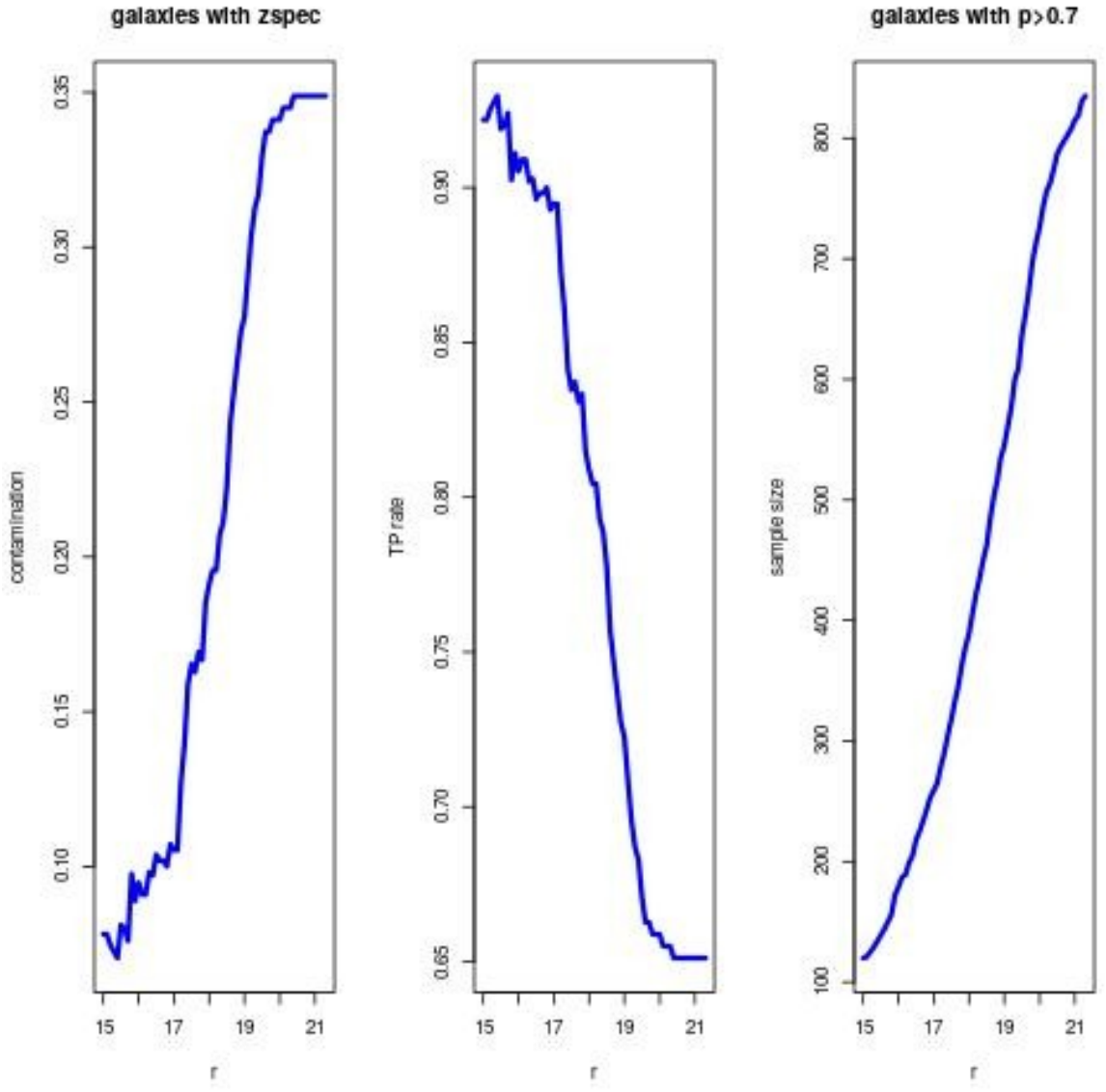


Evaluating the results with galaxies with zspec

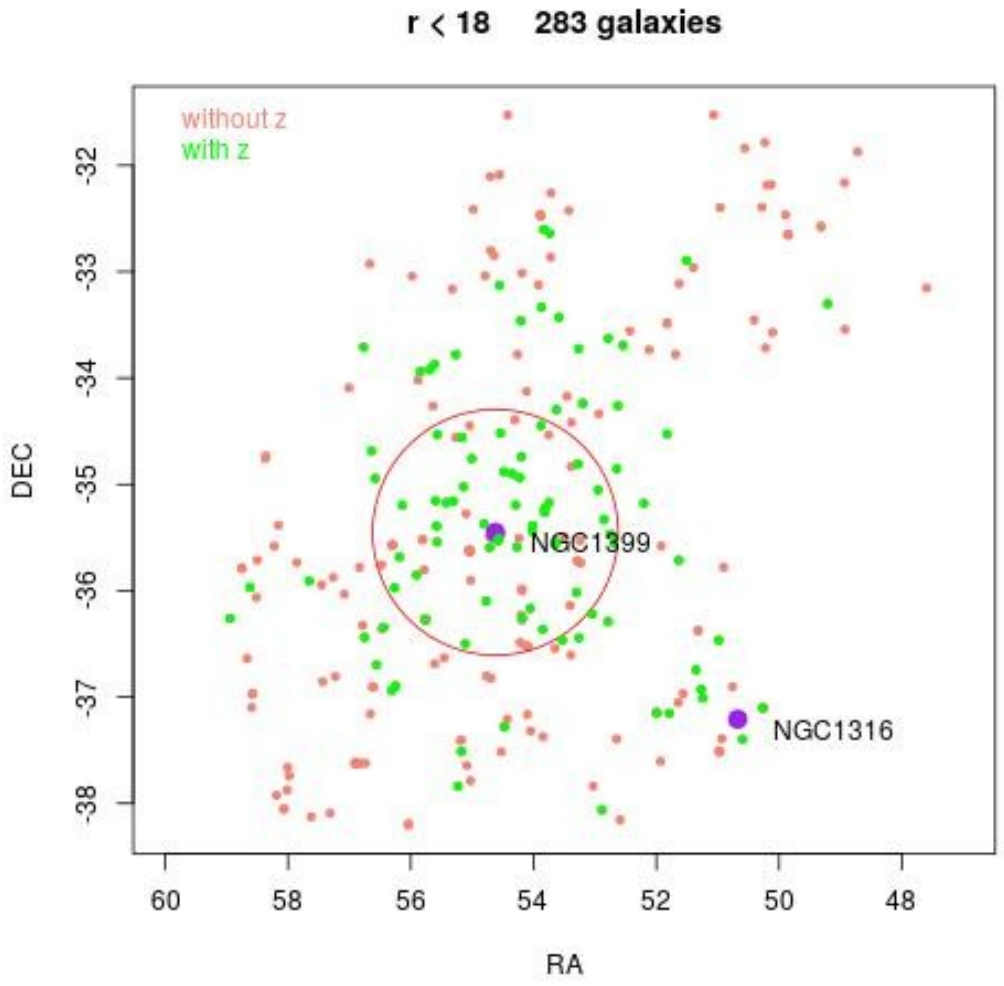


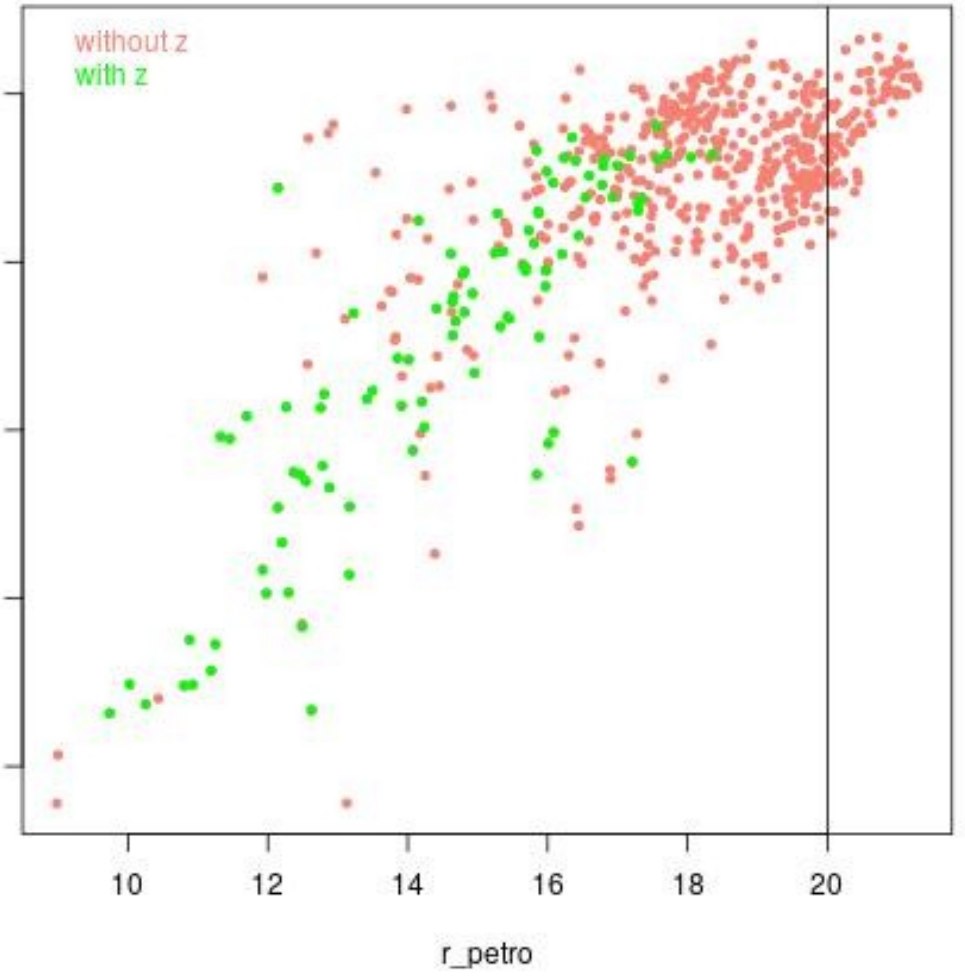


Fornax sample



cleaned photometric sample



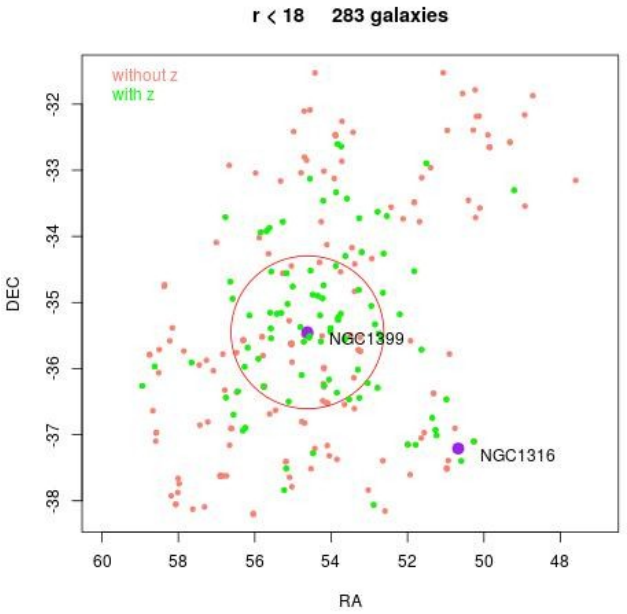


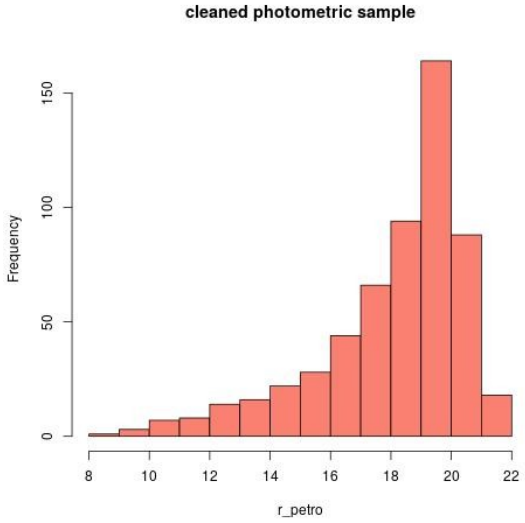
summary

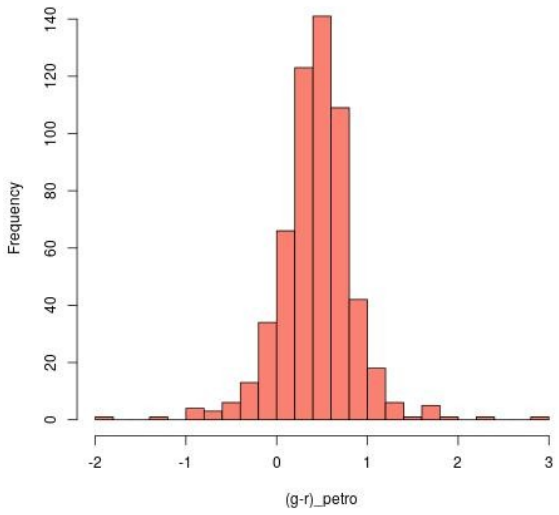
- it seems that for rpetro < 18 we can obtain a catalogue with high purity and low contamination

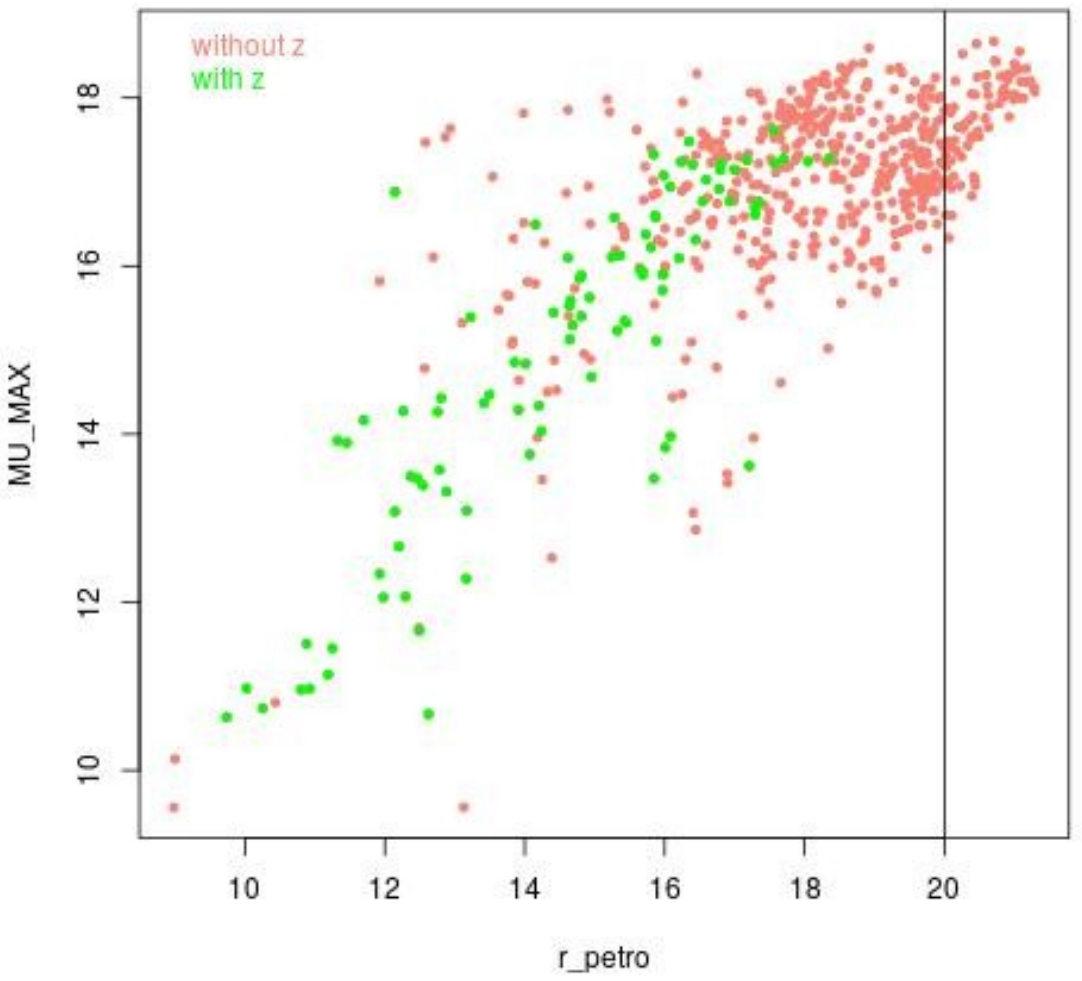
next steps:

- a more robust DL algorithm
- more inputs?
- revision of the radial velocities in Fornax
- spectroscopic follow-up?
- physical analysis!




cleaned photometric sample
Frequency
0
50
100
150
8
10
12
14
16
18
20
22
r_petro


Frequency
0
20
40
60
80
100
120
140
-2
-1
0
1
2
3
(g-r)_petro


without z
with z
MU_MAX
10
12
14
16
18
10
12
14
16
18
20
r_petro