Photometric Membership Estimation for S-PLUS Galaxy Clusters

Paulo Lopes (Observatório do Valongo / UFRJ)
André Ribeiro (LATO / UESC)
Eduardo Cypriano (IAG / USP)
Lia Doubrawa (IAG / USP)
Douglas Bambrila (OV/UFRJ)
Cláudia M. de Oliveira (IAG / USP)

15th S-PLUS meeting – June 3, 2021

1

Cluster membership

Why is it important?

WHY IS IT IMPORTANT?

When a spectroscopic follow-up is poor or absent, galaxy cluster population studies have been traditionally done with statistical corrections (or RS, photo-z dist, radial and lum profiles).

They may be affected by systematic effects, potentially biasing the results.

Only valid for combined samples, preventing investigation on individual galaxies.

Only a few works have previously addressed the issue of robust membership assignment from photometric information (e.g., George+11, Rozo+15, Castignani & Benoist 2016).

WHY IS IT IMPORTANT?

Current deep large area cluster catalogs already provide many more cluster candidates than is possible to probe spectroscopically.

This situation will only get worst with new surveys (e.g., S-PLUS, J-PAS, EUCLID, LSST and eROSITA).

WHY IS IT IMPORTANT?

CMDs (HFF data) of MACS J0416 (part of the BUFFALO project). They identify cluster members as galaxies with $0.35 < z_{phot} < 0.44$, and those with z_{spec} satisfying $|z_{cl} - z_{spec}| < 0.0104$, with $z_{cl} = 0.3979$. They also include galaxies within 3σ of the cluster red-sequence.

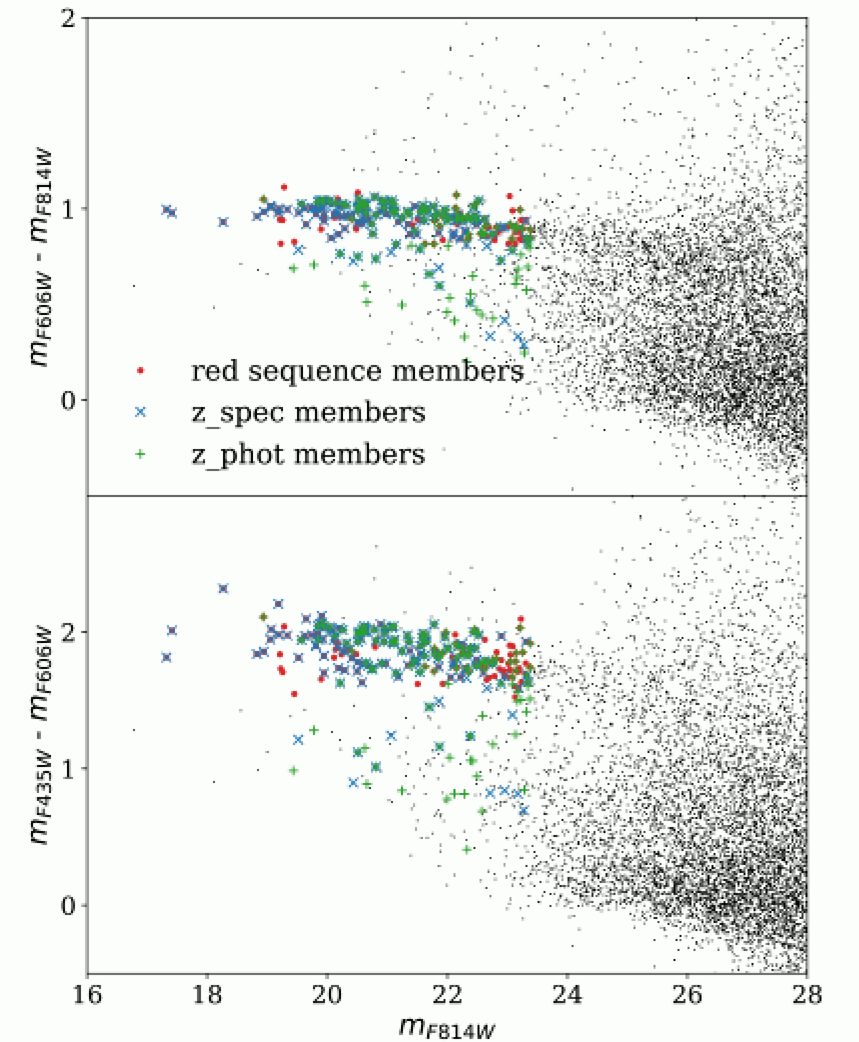


Figure 1. Colour-magnitude diagrams for all galaxies within the cluster and the parallel fields. Red sequence galaxies are selected according to a gaussian fit, including all the galaxies within 3σ, and are marked as red dots. Galaxies selected according to their photometric and spectroscopic redshifts are marked as blue crosses and green pluses, respectively.

Gonzalez+19

Problem:

Estimate membership only from photometric parameters.

This work:

Develop a new method based on a machine learning approach, relying on simple photometric parameters (colours, magnitude, concentration and radius), and environmental properties (radial offset and local density).

First work to address this issue with a machine learning method (Lopes & Ribeiro 2020).

2

Cluster samples

and data

CLUSTER SAMPLES

30 clusters with $z_{phot} \leq 0.045$ (complete to M* + 3 in the r-band).

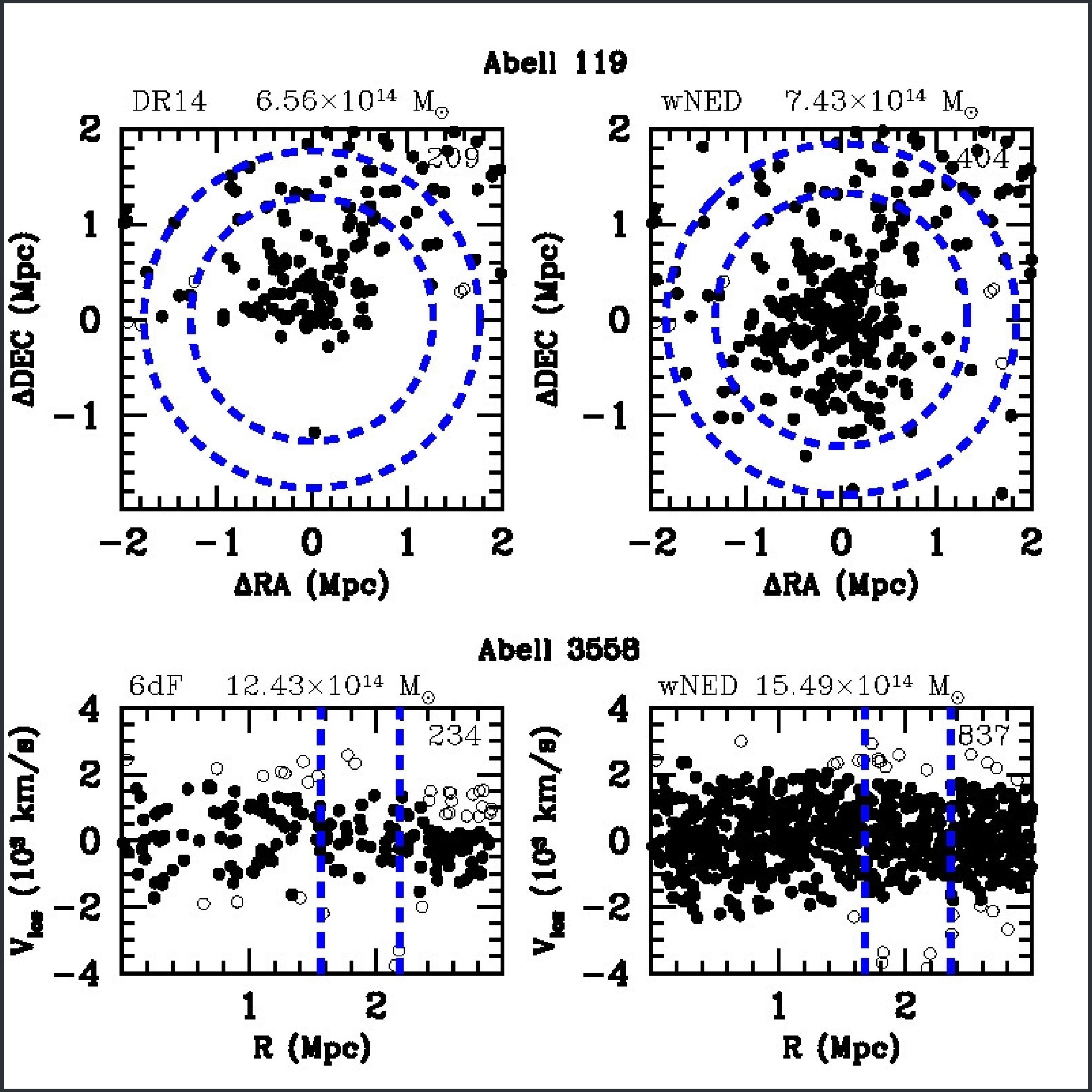
Selected in the optical, X-rays or through the SZ effect (Lopes+09, 18).

DATA

Photometry and spectra: Sloan Digital Sky Survey (SDSS);

Surveys' completeness affected by issues like fiber collision: we have also gathered additional redshifts from the NASA/IPAC Extragalactic Database (NED).

ADDING NED REDSHIFTS



Lopes+18

SPECTROSCOPIC CLUSTER MEMBERSHIP ASSIGNMENT

Initially we have all galaxies within ~4 Mpc and $|\Delta_v| \leq 4000$ km/s.

Run the shifting gapper procedure to select members and interlopers.

For the photometric classification we want all galaxies along the line of sight.

The remaining galaxies with $|\Delta_v| > 4000$ km/s are labeled as interlopers.

Stacked sample with 15802 galaxies, being 4420 within R200.

SPECTROSCOPIC CLUSTER MEMBERSHIP ASSIGNMENT

That is our main sample, split in half for training and validation.

We consider at all times only galaxy and cluster parameters derived from photometry.

3 Machine learning methods

Completeness (C) and Purity (P)

Completeness (true positive rate) is defined as the fraction of true members that are classified as members.

Purity (positive predictive value) is the fraction of true members among the objects classified as members.

$$C = \frac{N_{\mathrm{true}} - N_{\mathrm{missed}}}{N_{\mathrm{true}}}$$

$$P = \frac{N_{\mathrm{selected}} - N_{\mathrm{interlopers}}}{N_{\mathrm{selected}}}$$

$$N_{\mathrm{true}} = \frac{P}{C} N_{\mathrm{selected}}.$$

Completeness (C) and Purity (P)

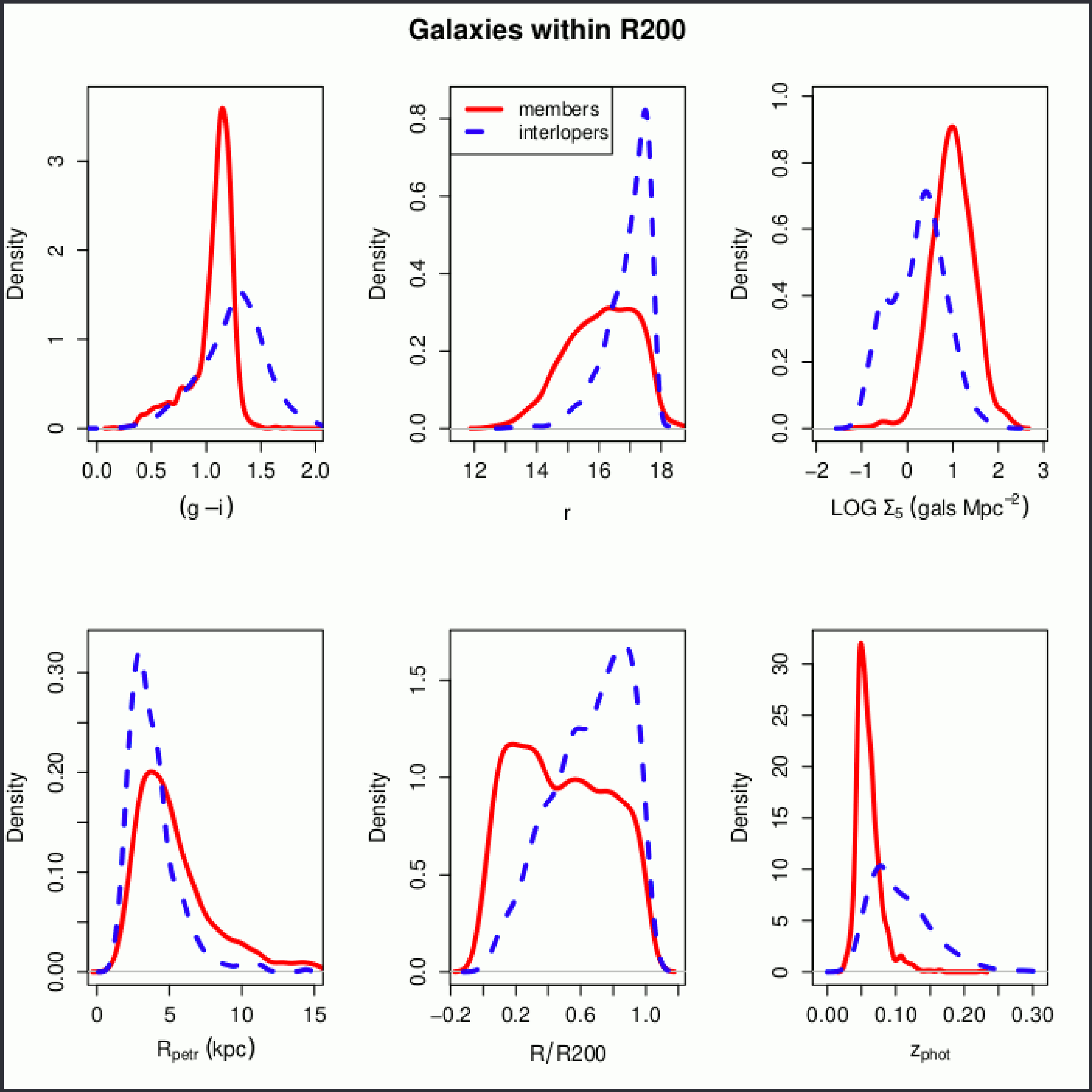
$N_{selected}$: number of galaxies photometrically classified as members.

N_{true}: the true number of members.

$N_{interlopers}$: number of objects wrongly classified as members.

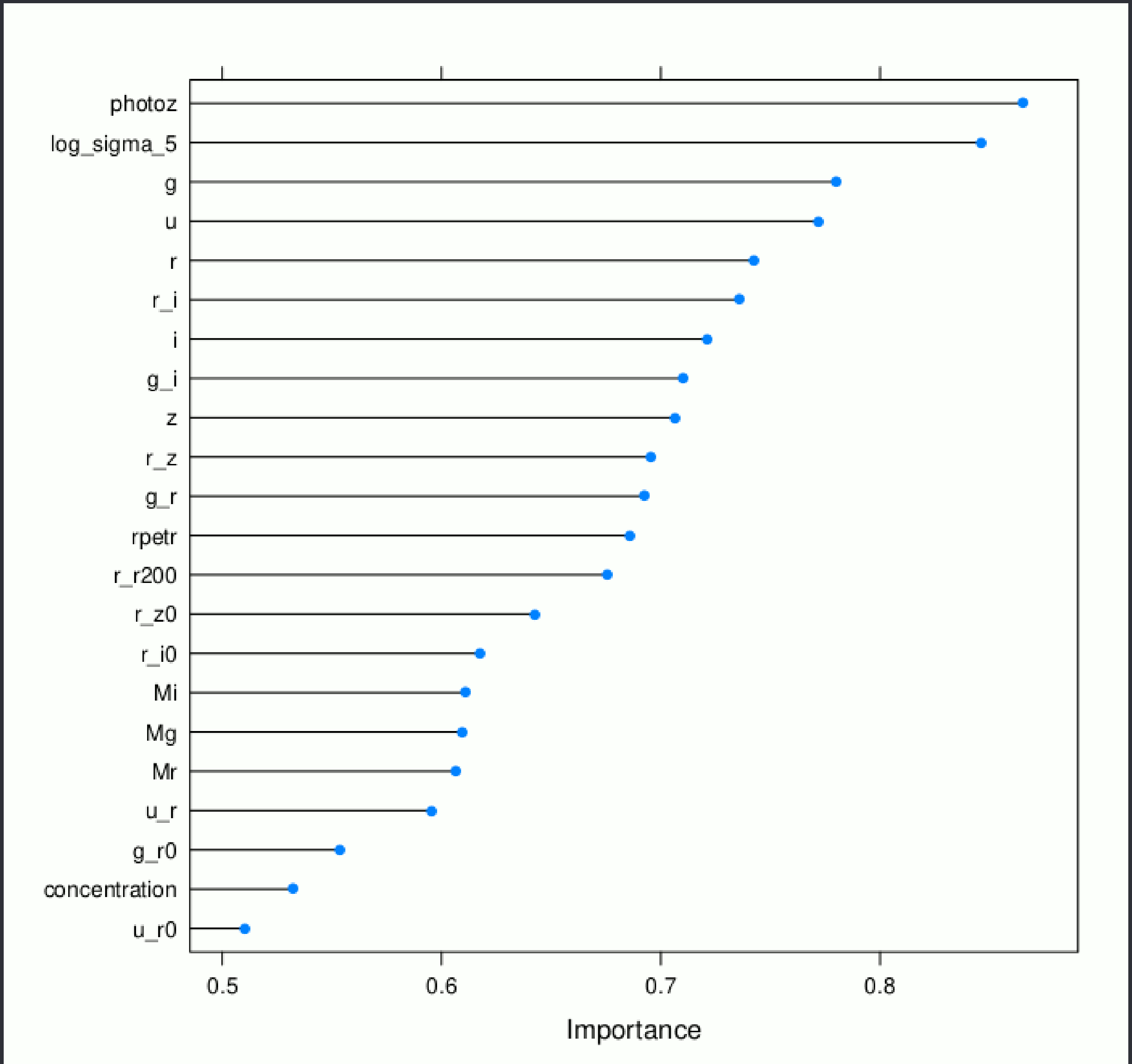
N_{missed}: number of true members that were not photometrically classified as so.

Features



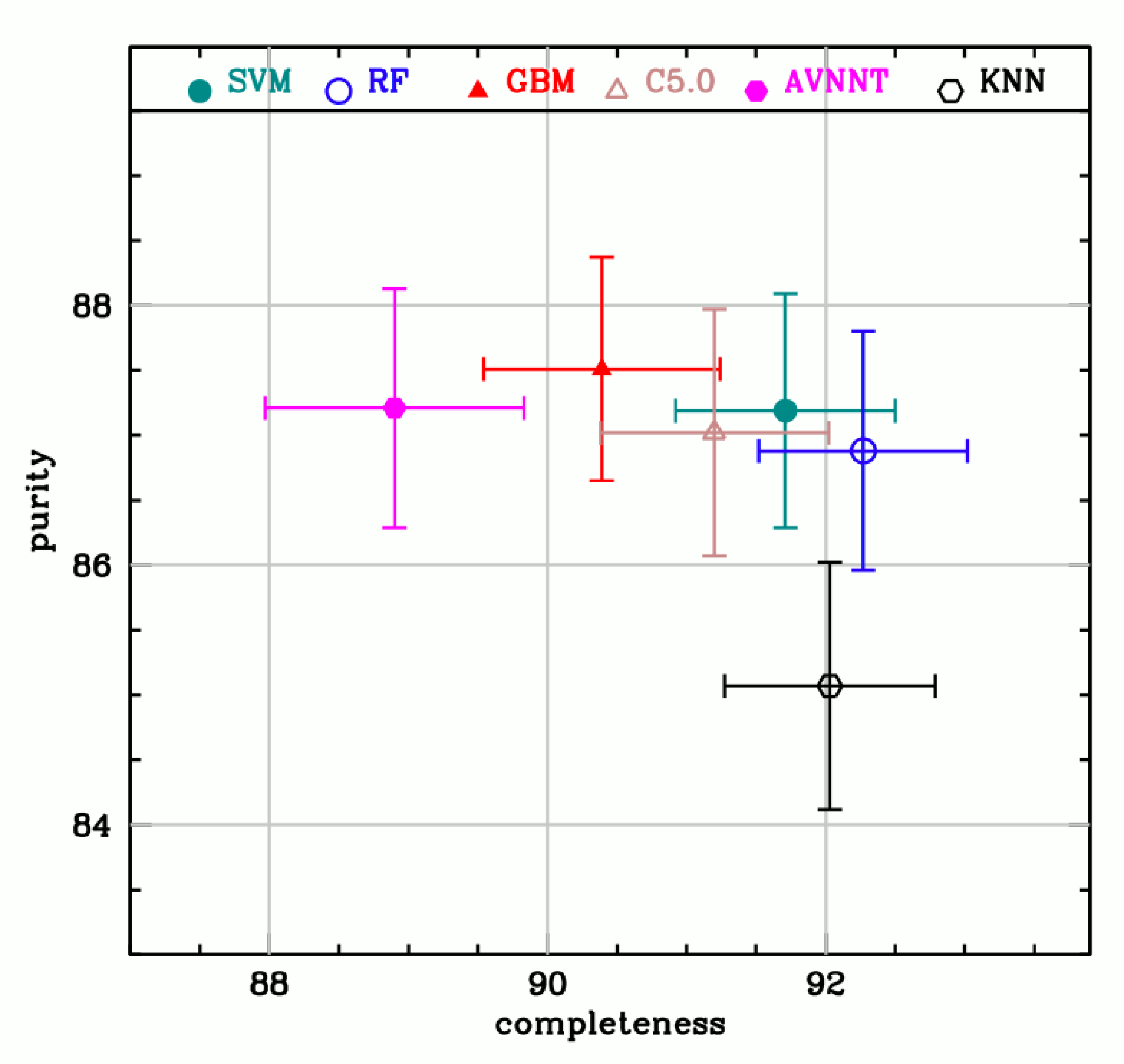
Distribution of six photometric parameters of members (red) and interlopers (blue) of the ESZ+X-ray sample.

Feature Selection



Importance of different variables available in the data set.

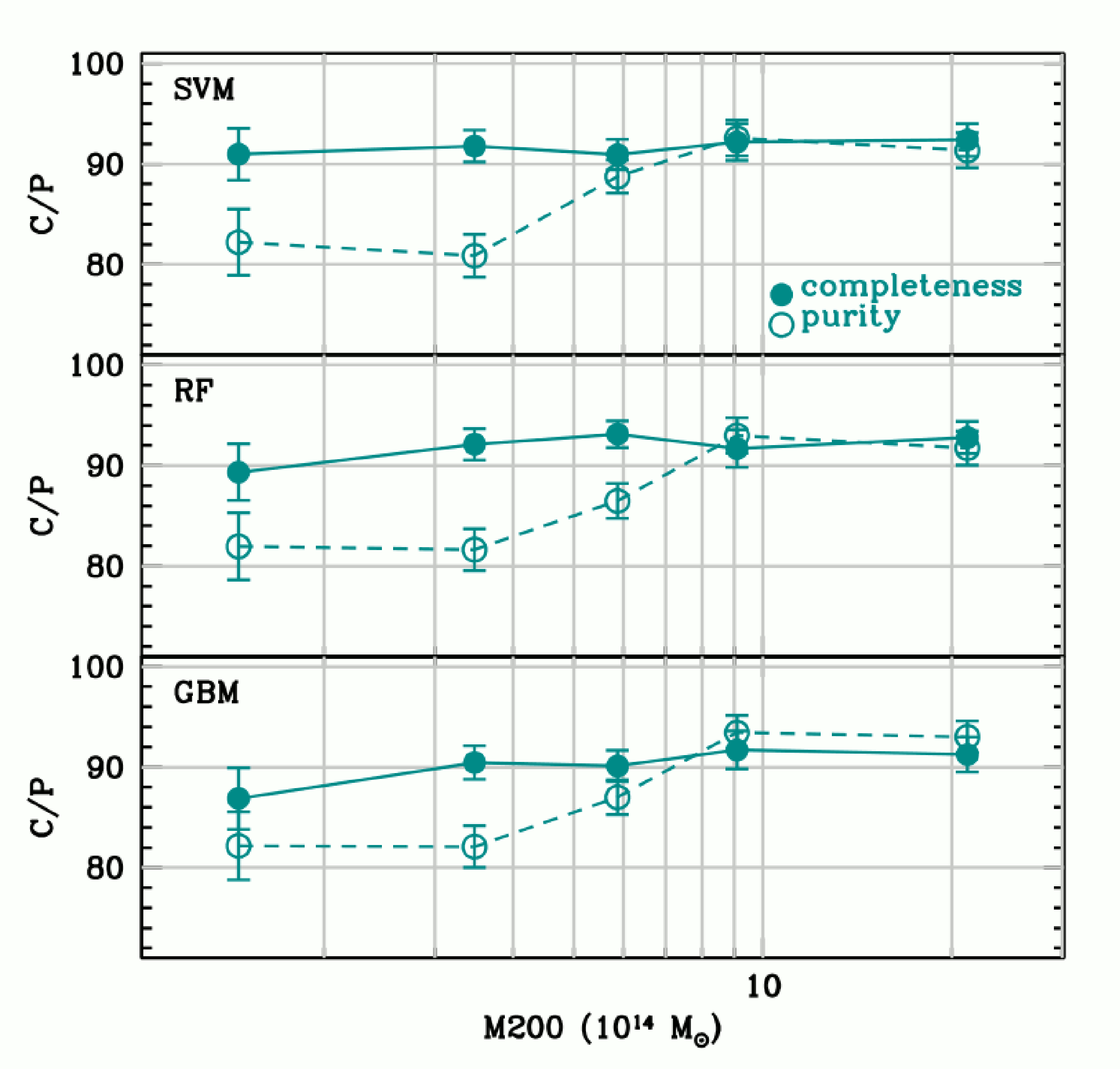
Model Selection and Tune



Purity vs completeness obtained with the the six best models tested.

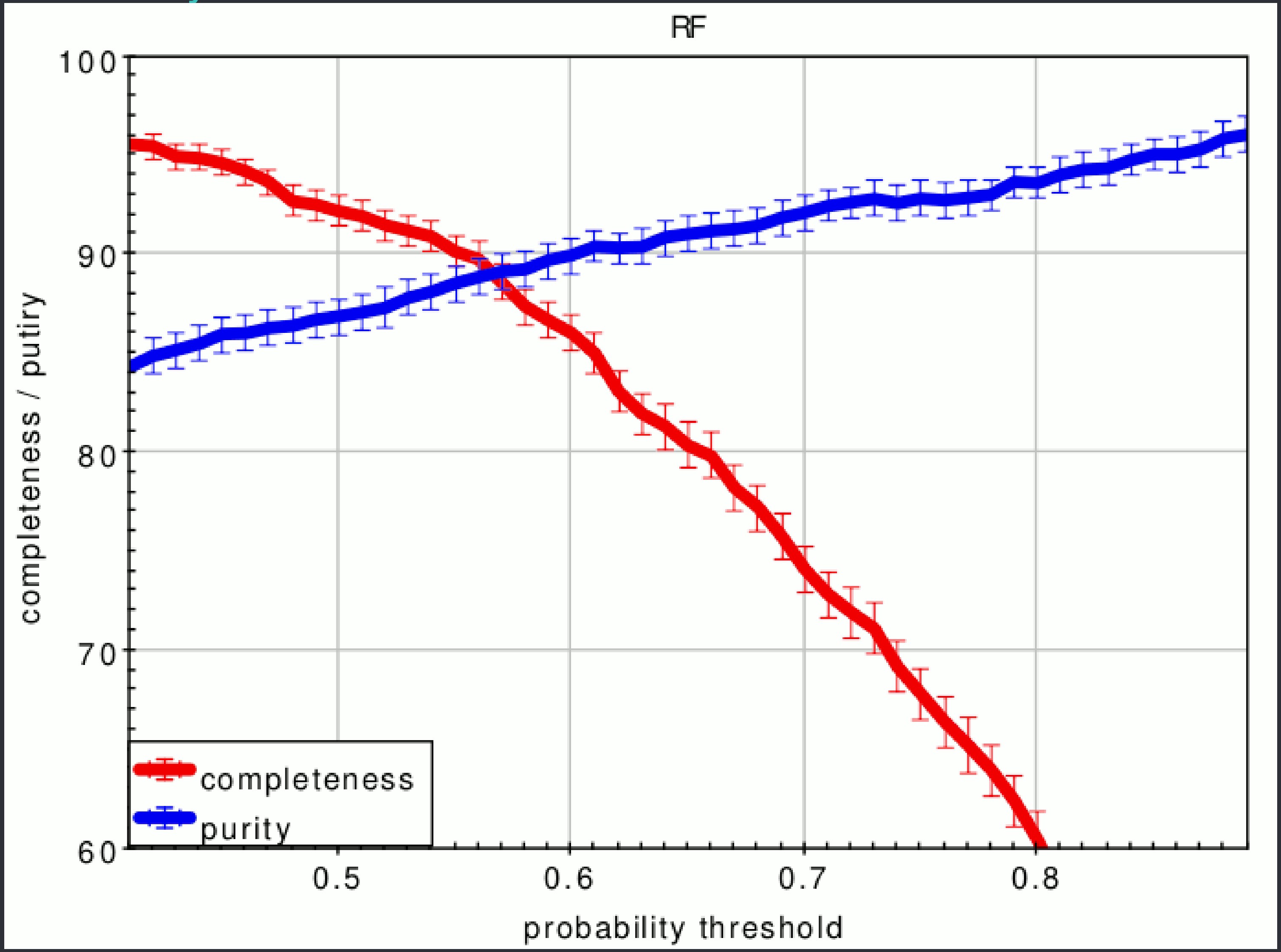
4 Results

Cluster mass dependence



C and P as a function of cluster mass.

Probability threshold

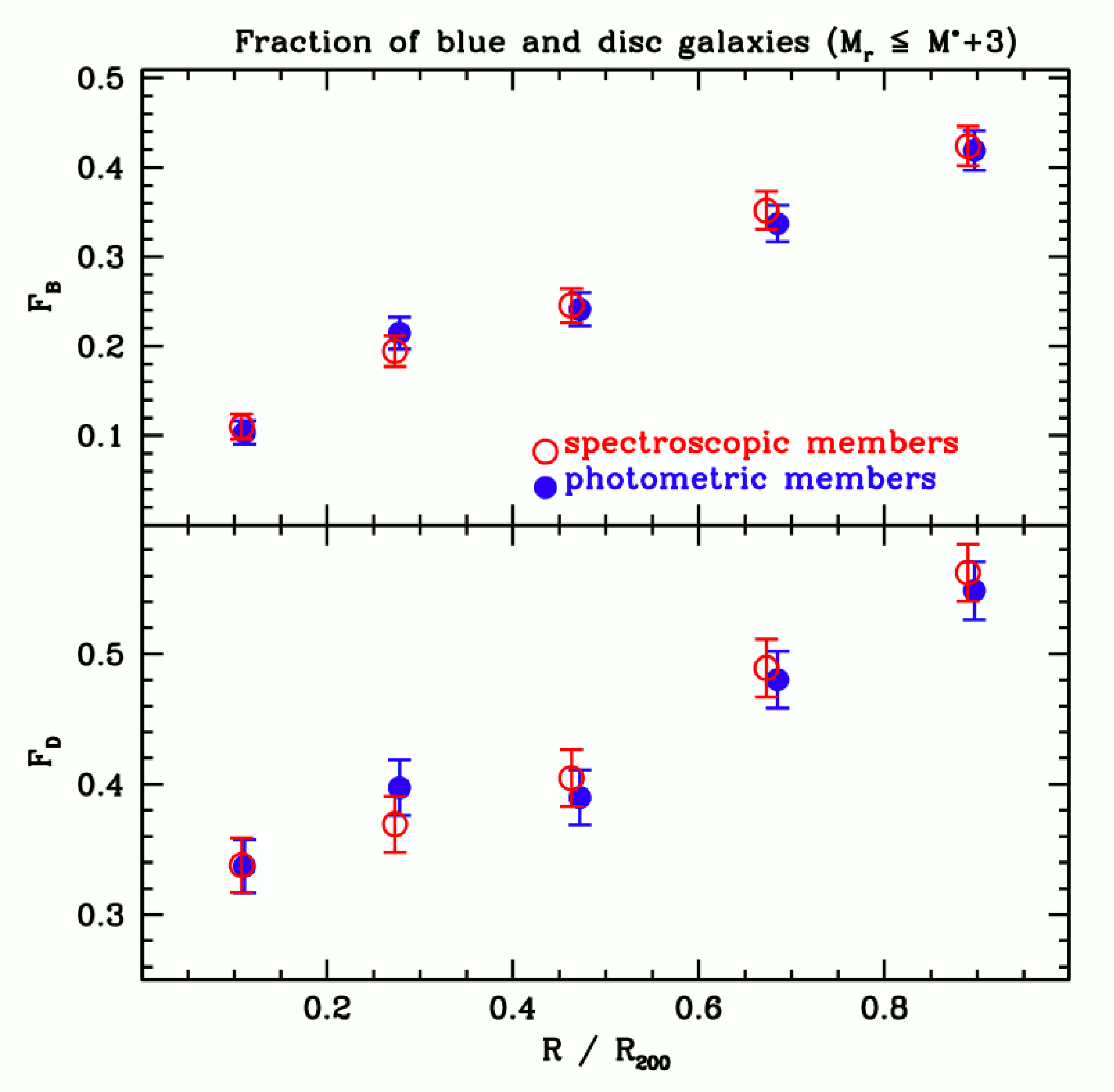


C and P as a function of the probability threshold.

5 Applications

Colour and morphology environmental variation

Variation of the fractions of blue (top) and disc (bottom) cluster galaxies as a function of clustercentric distance.



Many other applications

Without relying on a background correction one can investigate, for instance:

Cluster luminosity function.

Spatial segregation of blue and red galaxies.

Properties of transitional galaxies.

ICL

AGN fraction in clusters.

Scaling relations and mass calibration.

Lensing (after carefull separation of cluster and background galaxies)

Application to S-PLUS

6792 Pzwav groups and clusters (SN > 3.0) selected in the S82 (Werner et al., in prep.)

~ 93k spectroscopic redshifts in the S82 (Molino+20)

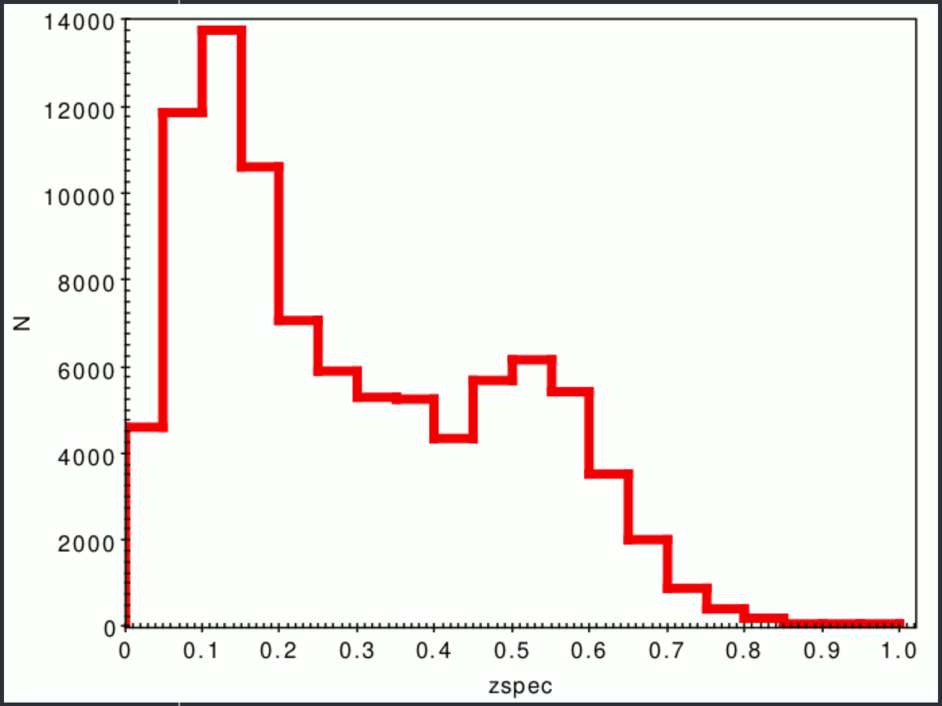
Gap technique within 0.5 Mpc/h. Obtained spectroscopic redshifts for 619 systems

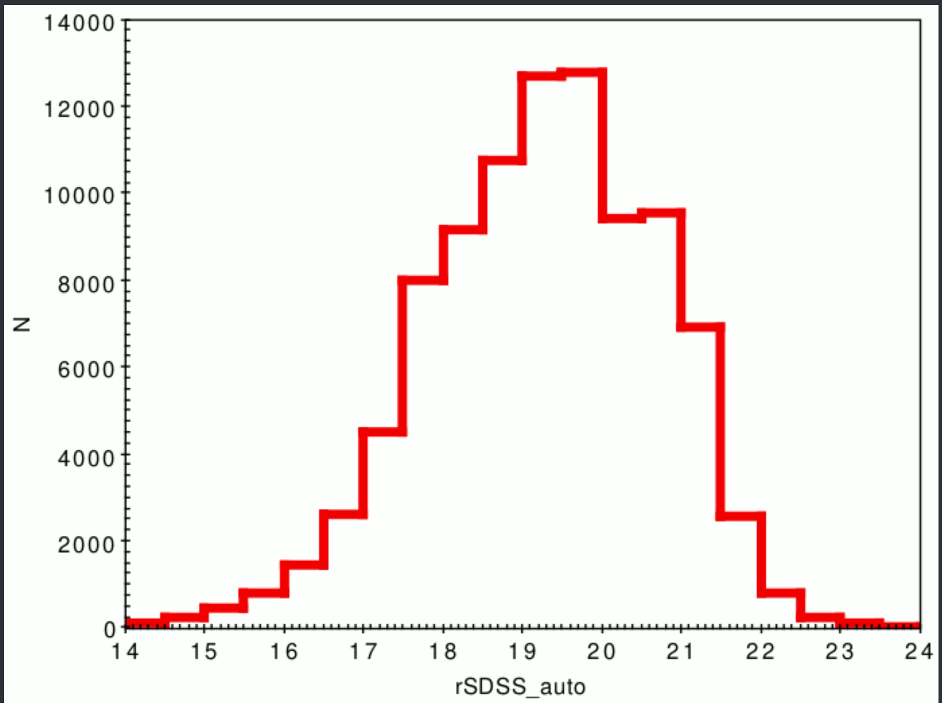
Shifting gapper technique within 2.5 Mpc/h: spectroscopic selected members for 161 groups, velocity dispersion, R200 and M200 estimates.

145 systems with $z < 0.2$

Preliminary results with DR1

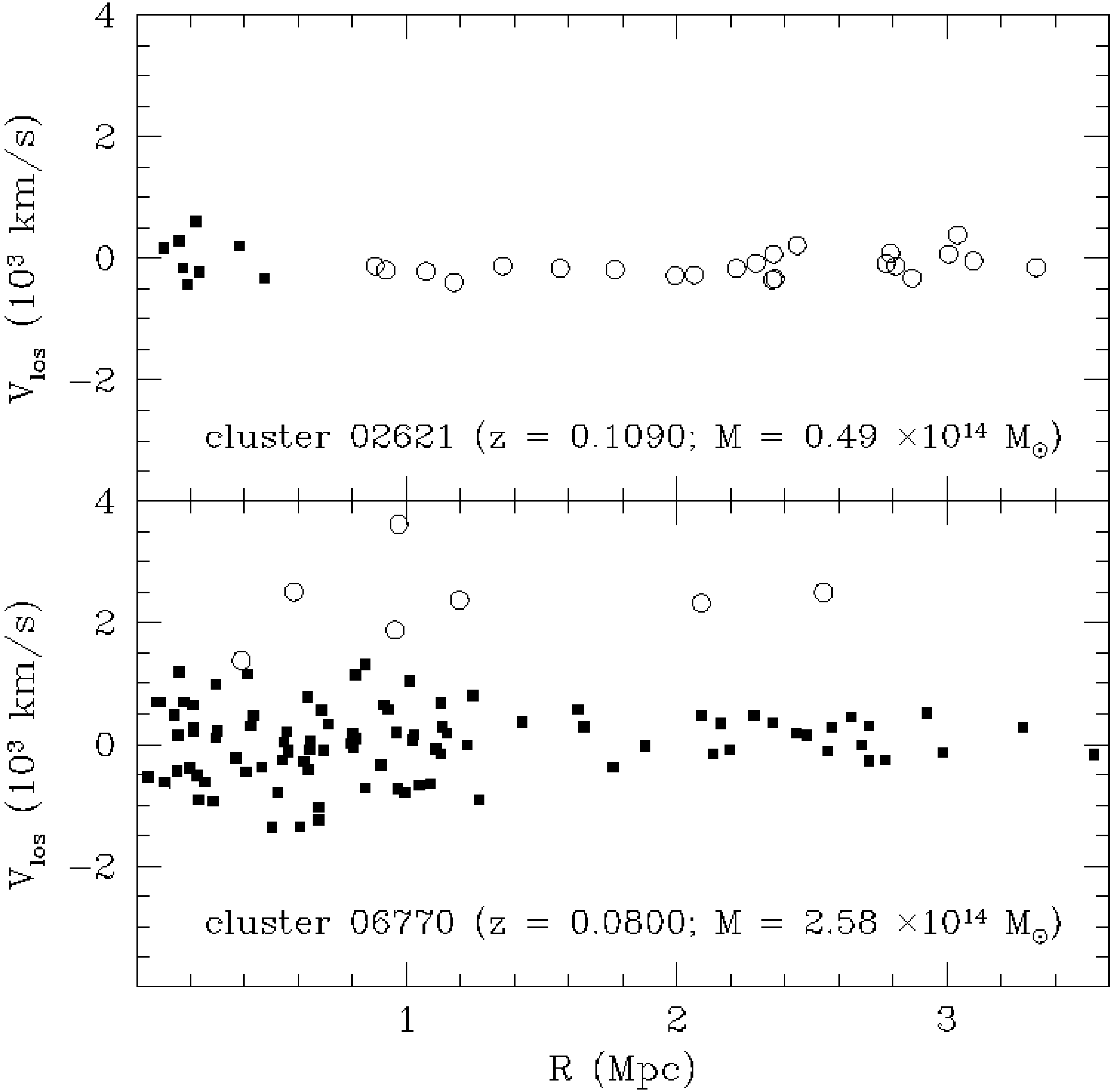
Application to S-PLUS





Spectroscopic redshifts in the S82 (Molino+20)

Thanks to Maria Luísa Buzzo and Fábio Herpich

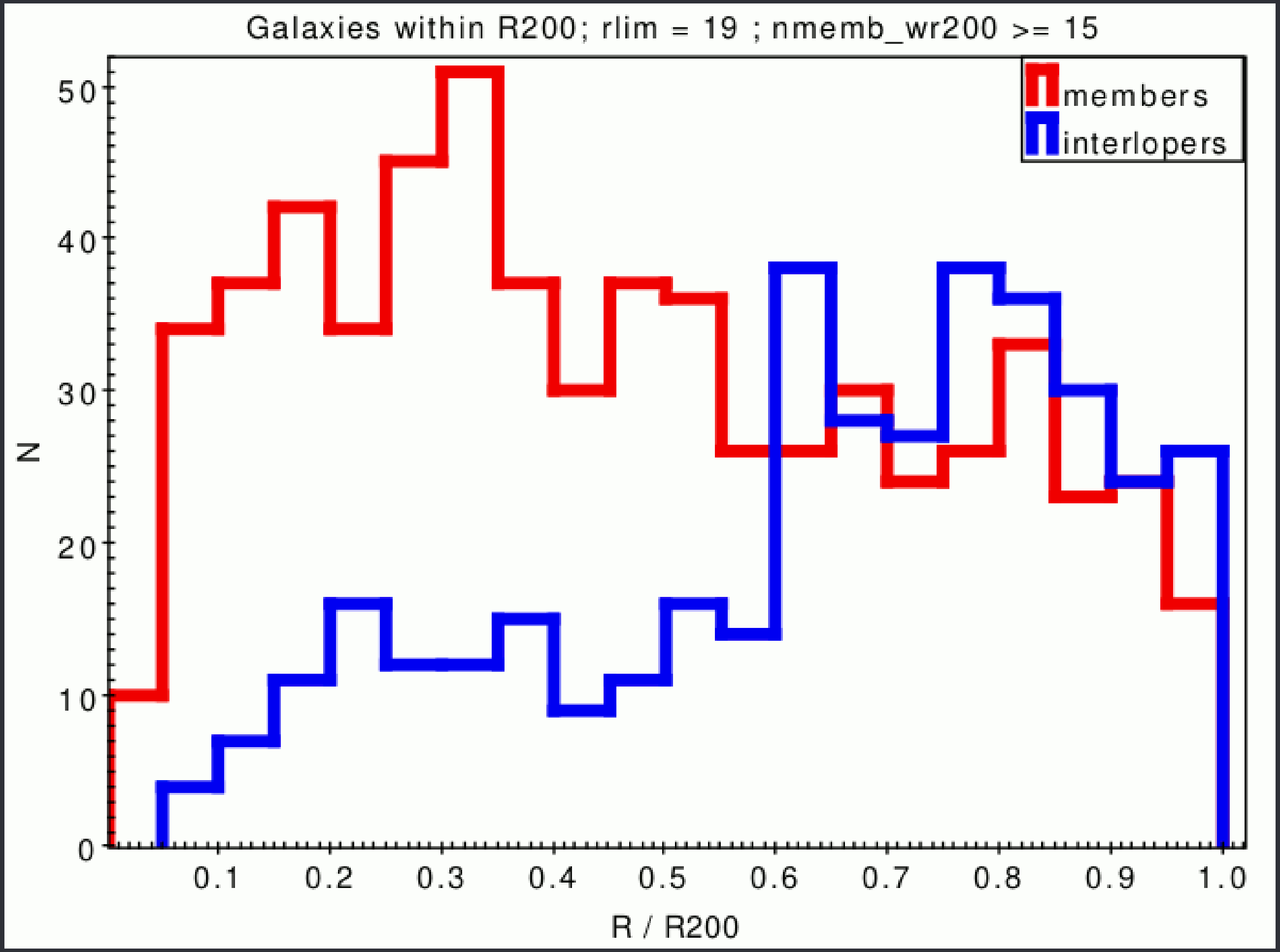


Also check the work of Camila Cid Parada (14th meeting)

Application to S-PLUS

28 groups with N200 > 15 (most z < 0.15; 996 galaxies)

C ~ 94% and P ~84%



6 Conclusions

CONCLUSIONS AND FUTURE WORK

Results

Cheap and reliable assessment of membership in clusters.

ML approach using only photometric information.

Completeness and purity of ~90% (SDSS and HST data).

Many astrophysical and cosmological applications.

Future work with S-PLUS

Prove it works for low mass systems.

Apply to iDR3 (VAC?)

Possibly include massive clusters outside the S82

Consider improved z_{phot}, from S-PLUS.

Go beyond R200 (2x ?)

Filaments between merging clusters?