Effects of incompleteness in the training sample for photo-z estimation by DNF algorithm

Laura Toribio San Cipriano | laura.toribio@ciemat.es

J. De Vicente, E. Sánchez, I. Sevilla-Noarbe, J. Asorey, D. Cid & J. Mena-Fernández

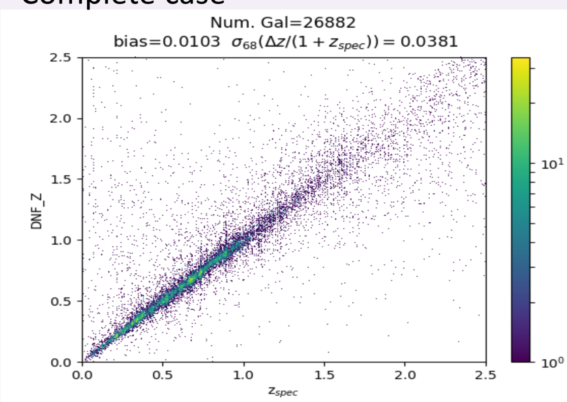
One of the crucial keys in cosmological studies is the estimation of an **accurate redshift** for a large number of galaxies. Sometimes, the spectroscopic sample used as training sample for ML approaches doesn't cover the same magnitude and color space as the target sample. This issue raise doubts about the confidence of the photometric redshift provided by the algorithms.

We present the effect of using complete or incomplete spectroscopic training sample to determine photo-zs by DNF algorithm.

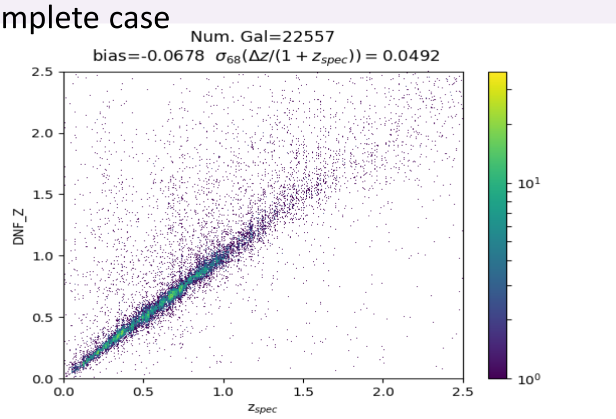
We provide a new method for determining the level of confidence in the photo-z values and the incompleteness assessment of the results.

How does incompleteness affect photo-zs?

Complete case



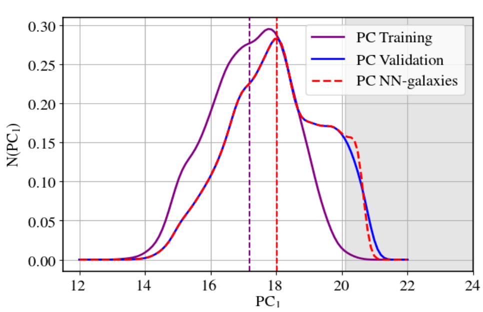
Incomplete case



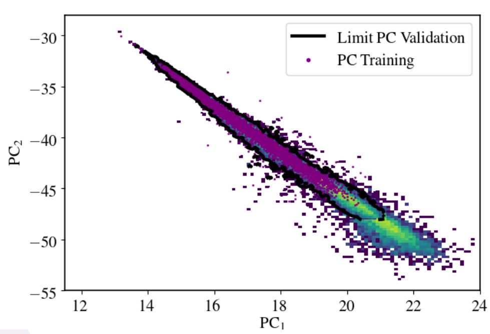
We have used the galaxies Y3 Dark Energy Survey (DES) of the public Deep Fields catalogue with information in 8 bands (ugrizJHKs).

Detecting incompleteness

We have performed a **principal component analysis** to study the completness of the training sample.



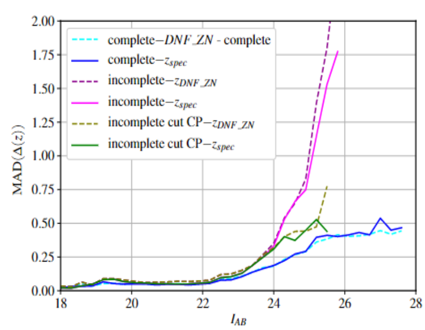
This method makes it possible to detect galaxies whose photo-zs are affected by incompleteness.

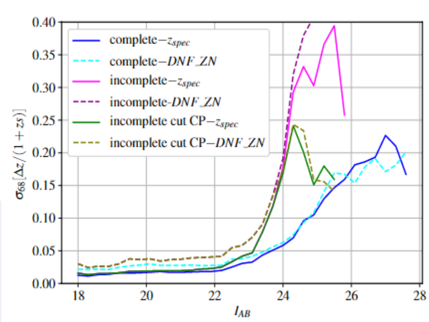


DNF provides accurate photo-zs within the limit (black line) determined by quality cuts.

Measuring incompleteness

Normally, it is assumed that the metrics obtained in validation will have the same behaviour in the target sample. However, it is easy to see how **the training sample affects the metrics.**





To evaluate DNF_Z, we have substituted zspec by DNF_ZN when spectroscopic redshift is not available. The behaviour of these metrics can be considered as **a good approximation of the real value** which changes depending on the training sample. We can take these metrics computed by DNF_ZN as **an upper limit** of the reality.

Funding support for this work was provided by the Autonomous Community of Madrid through the project TEC2SPACEM (S2018/NMT-4291) and the Project Física Fundamental y Cosmología con cartografiados extragálacticos (PGC2018-094773-B-C33).

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas