

Dark Energy Survey Deep Field photometric redshift performance and training incompleteness assessment

L. Toribio San Cipriano¹, J. De Vicente¹, I. Sevilla-Noarbe¹, W. G. Hartley², J. Myles³, A. Amon^{4,5}, G. M. Bernstein⁶, A. Choi⁷, K. Eckert⁶, R. A. Gruendl^{8,9}, I. Harrison¹⁰, E. Sheldon¹¹, B. Yanny¹², M. Aguena¹³, S. S. Allam¹², O. Alves¹⁴, D. Bacon¹⁵, D. Brooks¹⁶, A. Campos¹⁷, A. Carnero Rosell^{18,13,19}, J. Carretero²⁰, F. J. Castander^{21,22}, C. Conselice^{23,24}, L. N. da Costa¹³, M. E. S. Pereira²⁵, T. M. Davis²⁶, S. Desai²⁷, H. T. Diehl¹², P. Doel¹⁶, I. Ferrero²⁸, J. Frieman^{12,29}, J. García-Bellido³⁰, E. Gaztañaga^{15,21,22}, G. Giannini^{20,29}, S. R. Hinton²⁶, D. L. Hollowood³¹, K. Honscheid^{32,33}, D. J. James³⁴, K. Kuehn^{35,36}, S. Lee³⁷, C. Lidman^{38,39}, J. L. Marshall⁴⁰, J. Mena-Fernández⁴¹, F. Menanteau^{8,9}, R. Miquel^{42,20}, A. Palmese¹⁷, A. Pieres^{13,43}, A. A. Plazas Malagón^{44,45}, A. Roodman^{44,45}, E. Sanchez¹, M. Smith⁴⁶, M. Soares-Santos¹⁴, E. Suchyta⁴⁷, M. E. C. Swanson⁸, G. Tarle¹⁴, M. Vincenzi^{15,46}, N. Weaverdyck^{14,48}, P. Wiseman⁴⁶, and

(DES Collaboration)

(Affiliations can be found after the references)

Received 14 December 2023; accepted 24 February 2024

ABSTRACT

Context. The determination of accurate photometric redshifts (photo-zs) in large imaging galaxy surveys is key for cosmological studies. One of the most common approaches are machine learning techniques. These methods require a spectroscopic or reference sample to train the algorithms. Attention has to be paid to the quality and properties of these samples since they are key factors in the estimation of reliable photo-zs.

Aims. The goal of this work is to calculate the photo-zs for the Y3 DES Deep Fields catalogue using the Directional Neighborhood Fitting (DNF) machine learning algorithm. Moreover, we want to develop techniques to assess the incompleteness of the training sample and metrics to study how incompleteness affects the quality of photometric redshifts. Finally, we are interested in comparing the performance obtained with respect to the EAZY template fitting approach on Y3 DES Deep Fields catalogue.

Methods. We have emulated –at brighter magnitude– the training incompleteness with a spectroscopic sample whose redshifts are known to have a measurable view of the problem. We have used a principal component analysis to graphically assess incompleteness and to relate it with the performance parameters provided by DNF. Finally, we have applied the results about the incompleteness to the photo-z computation on Y3 DES Deep Fields with DNF and estimated its performance.

Results. The photo-zs for the galaxies on DES Deep Fields have been computed with the DNF algorithm and added to the Y3 DES Deep Fields catalogue. They are available at <https://des.ncsa.illinois.edu/releases/y3a2/Y3deepfields>. Some techniques have been developed to evaluate the performance in the absence of "true" redshift and to assess completeness. We have studied the tradeoff on the training sample between highest spectroscopic redshift quality vs. completeness. Some advantages have been found on relaxing the highest quality spectroscopic redshift requirements at fainter magnitudes in favour of completeness. The result achieved by DNF on the Y3 Deep Fields are competitive with the ones provided by EAZY showing notable stability at high redshifts. It should be noted the good results obtained by DNF for the estimation of photo-zs in deep field catalogues and its suitability for the future LSST and Euclid data, which will have similar depths to the Y3 DES Deep Fields.

Key words. dark energy – Galaxies: distances and redshifts – Machine Learning

1. Introduction

The arrival of large photometric galaxy surveys such as Sloan Digital Sky Survey (SDSS, York et al. 2000), Dark Energy Survey (DES, Flaugher et al. 2015), Physics of the Accelerating Universe (PAU, Castander et al. 2012) or the future projects such as Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST, LSST Science Collaboration et al. 2009), and Euclid (Euclid Collaboration et al. 2020), capable of collecting huge amounts of data, are providing invaluable insight about the Universe. One of the crucial elements for cosmological and astrophysical studies is the estimation of accurate redshifts from photometric information which are essential for many cosmological probes as baryon acoustic oscillation (BAO), weak lensing or galaxy clustering. Spectroscopic surveys –measuring the difference in the wavelength of some spectral lines with respect to

their wavelength at rest frame– provide high precision redshifts, but obtaining spectroscopic redshifts of large samples of astronomical objects is very expensive in terms of observing time. Currently, the Dark Energy Spectroscopic Instrument (DESI) project (DESI Collaboration et al. 2016) is capable of measuring thousands of galaxy spectra every night, reducing telescope time. Despite this great advantage long exposure times are still required to obtain good signal-to-noise spectra of faint objects, and photometric data for target selection. An alternative is to measure the fluxes of galaxies with a set of broadband or narrowband filters within an image survey, that is using photometric techniques. These measurements allow us to compute the photometric redshift (also called photo-zs) of a large number of galaxies per image reducing the telescope time at the cost of a lower precision.

The two main approaches to determining photo-zs are *template fitting* and *machine learning* methods. Template methods compare the spectral energy distribution (SED) of each galaxy with that of a set of redshifted rest-frame templates looking for the best match (e.g. Arnouts et al. 1999; Benítez 2000; Bolzonella et al. 2000; Ilbert et al. 2006). Machine learning approaches use reference or training galaxy samples whose spectroscopic redshifts are known in order to learn the relationship between magnitudes, colours and redshifts. With this information, machine learning methods can predict the photometric redshift of a set of target galaxies (e.g. Collister & Lahav 2004; Sadeh et al. 2016; Carrasco Kind & Brunner 2013; De Vicente et al. 2016). Neither method is free of difficulties. Template methods depend on synthetic models and the completeness of the template library used in the fitting, while machine learning methods depend on the quality and variety of the training samples. Specifically, the selection of this spectroscopic training sample is one of the most important decisions to obtain accurate photometric redshift estimations in the machine learning approach. Ideally, the spectroscopic sample should be representative of the whole target galaxy sample, covering the same colour-magnitude space. Unfortunately, the galaxy samples whose photometric redshift is to be determined typically include galaxies with deeper magnitudes that are not included in the spectroscopic sample. Hartley et al. (2020) studied the impact of using incomplete spectroscopic samples in the redshift distribution using the Lima et al. (2008) algorithm. They show that an incomplete spectroscopic training sample could bias the galaxy redshifts. Moreover, the studies of Hildebrandt et al. (2010), Beck et al. (2017), Sánchez et al. (2014), Schmidt et al. (2020), Bonnett et al. (2016) and Abdalla et al. (2011) and Brescia et al. (2021) compare different methods for photo-zs estimation. These works suggest that machine learning methods provide more accurate values of photo-zs than template methods as long as there is a sufficiently adequate sample for training. Outside the magnitude and colour space, template methods seem to perform better than machine learning methods because they can generate synthesized spectra without redshift constraint. Everything seems to indicate that the combination of both template and machine learning is the best option to obtain the best photo-zs accuracy of a sample (e.g. Tanaka et al. 2018; Salvato et al. 2019).

In this work, we study how the incompleteness in the spectroscopic training sample affects Directional Neighborhood Fitting (DNF) photo-zs algorithm (De Vicente et al. 2016) photo-zs, as estimated in the Dark Energy Survey (DES) Year 3 Deep Field sample. DNF is a nearest-neighbour approach for photometric redshift estimation that has become a reference within DES collaboration and included between the five methods to be prioritized in Vera Rubin observatory. To assess the effects of incompleteness, we first derive the relevant parameters to characterize incompleteness, demonstrating how these parameters affect photo-z performance. Then, we show how DNF accounts for incompleteness in the photo-zs errors provided. Finally, we study the incompleteness of the training sample in Y3 DES Deep Fields and compare our results with those obtained from the EAzy template method (Brammer et al. 2008).

The rest of the paper is organised as follows. In Sect. 2, we describe the sample selection and in Sect. 3 the metrics used and the description of DNF algorithm. We carried out an analysis of the effects of incomplete training samples on the estimation of photometric redshift in Sect. 4. In Sect. 5, we estimate photometric redshift for Y3 DES Deep Fields with different training samples. We compare the photo-zs determined by DNF and

EAzy in Sect. 6. Finally, we enumerate the conclusions of this work in Sect. 7.

2. Data

2.1. Spectroscopic sample

We used the spectroscopic sample defined by Gschwend et al. (2018). This sample contains spectroscopic redshifts of galaxies from 34 surveys (see Appendix A) and the photometric information for each of them. The quality of the spectroscopic redshift is flagged by the label FLAG_DES (with FLAG_DES = 4 as certain redshift, FLAG_DES = 3 as probable redshift, FLAG_DES = 2 as possible redshift and FLAG_DES = 1 as unknown redshift). For this work, we only select those objects with spectroscopic redshift determination marked in the catalogue with the best redshift determination, that is: those galaxies with flag of 3 and 4 level (FLAG_DES $\geq$ 3). In addition, we exclude those galaxies with $mag(i) \geq 28$. After these cuts, our spectroscopic sample contains a total of 55,601 galaxies.

2.2. Y3 Deep Fields catalogue

The Y3 DES Deep Fields catalogue¹ used is part of the Dark Energy Survey. The observations were taken using the Dark Energy Camera (DECam, Flaugher et al. 2015) on the Victor M. Blanco 4m telescope at the Cerro Tololo Inter-American Observatory (CTIO) in Chile. DES covered 5000 deg² in *grizY* bands with approximately 10 overlapping dithered exposures in each filter (90 sec in *griz*, 45 sec in *Y*) covering the survey footprint. The Y3 DES Deep Fields catalogue comprises four fields measured with 8 bands (*ugrizJHK_s*) covering an area of ~ 5.88 deg² where the integrated exposure time per pixel is approximately ten times more than the main DES survey area (see details in Hartley et al. 2022). This catalogue contains around 2.8 million galaxies. We have selected those galaxies that have flux measurements in the eight filters and with $mag(i) < 28$, resulting a catalogue that contains around 1.5 million galaxies. We have selected galaxies with $mag(i) < 28$ –still suitable for weak lensing applications– because for higher magnitudes the errors in the photometry are large and the data become unreliable.

3. Metrics and algorithm

3.1. Metrics

This section describes the metrics used in this work to assess the quality of the photo-zs estimates, where z_{spec} , z_{phot} and N represent the spectroscopic redshift, the photometric redshift and the number of objects in the sample, respectively. We define the following metrics to quantify the degree of precision of the photo-zs and its scatter:

- **Bias:** The assessment of the overall photo-zs is determined by the mean bias:

$$\overline{\Delta z} = \frac{1}{N} \sum_{i=1}^N (\Delta z_i),$$

where $\Delta z = z_{spec} - z_{phot}$.

¹ Available at <https://des.ncsa.illinois.edu/releases/y3a2/Y3deepfields>

– Mean Absolute Deviation:

$$\text{MAD}(\Delta z) = \frac{1}{N} \sum_{i=1}^N |\Delta z_i|.$$

– $\sigma_{68}(\Delta z)$: denotes the half width of the central 68% percentile range of galaxies both bias value.

$$\sigma_{68}(\Delta z) = \frac{1}{2}(P_{84} - P_{16}),$$

where P_{16} = 16th percentile of the cumulative distribution and P_{84} = 84th percentile of the cumulative distribution.

– σ_{68} normalised: defined as

$$\sigma_{68}^{\text{Norm}} = \sigma_{68} \left(\frac{|\Delta z|}{1 + z_{\text{spec}}} \right).$$

– Outlier fraction:

$$f = \frac{N_{\text{out}}}{N},$$

where N is the total number of objects and N_{out} the outlier defined by:

$$|\Delta z| \geq 3\sigma,$$

where σ is the standard deviation of the Δz distribution.

3.2. DNF algorithm

Directional Neighborhood Fitting (DNF, De Vicente et al. 2016) is a nearest neighbour algorithm for estimating the redshift of a sample of galaxies. DNF uses the colours and magnitudes/fluxes as a measurement of closeness to a reference sample composed of galaxies whose spectroscopic redshifts are known. DNF provides the main photo- z value and its error estimation along with a secondary value intended for photo- z distribution estimation:

- DNF_Z: the main photo- z estimate determined by the fit of a number of neighbour galaxies to a hyperplane in magnitude space. The process is iterated to remove outliers. In addition the algorithm can provide individual photo- z probability density functions (PDFs).
- DNF_ZSIGMA: indicator of photo- z quality computed from the quadratic sum of the error due to photometry plus the error due to the fit. DNF_ZSIGMA takes the value -99 when DNF does not estimate the photo- z of a galaxy because there is no neighbour galaxy within a given radius.
- DNF_ZN: a secondary photo- z determined by the single nearest neighbour galaxy which is valuable for redshift distribution estimation.

The algorithm provides three alternative metrics for the assessment of closeness: Euclidean, angular and directional. While Euclidean and angular metrics account for magnitude and colour respectively, directional metric integrate both in a unique number. The present work takes advantage of the combination of 5 optical plus 3 near-infrared filters to define non-degenerated colours within the angular metric.

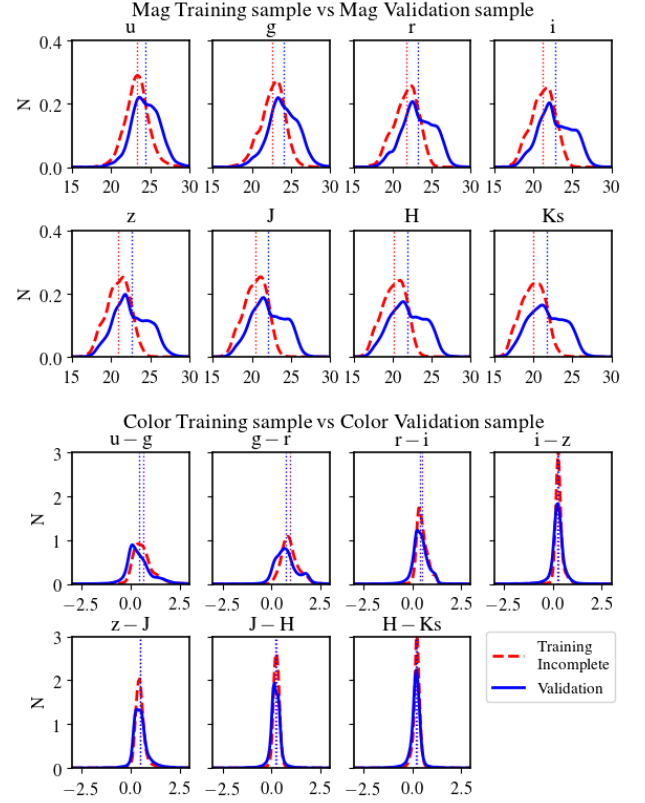


Fig. 1. Magnitude and colour distribution for incomplete training and validation samples. The red dashed lines represent the incomplete training sample and the blue lines the validation sample. The dotted vertical lines are the mean of each distribution. We have not included the curve for the complete training figure because these distributions overlap perfectly with those of the validation sample.

4. Effect of training incompleteness in photometric redshift estimation

We study the effect of using an incomplete spectroscopic training sample to determine the photo- z s with the DNF algorithm. We refer to an incomplete training sample when it does not cover the same range of magnitudes and/or colours as the target sample for which we want to determine its photo- z s.

The spectroscopic sample, in addition to being used to train the algorithm, allows us to study the accuracy and precision of the photo- z estimation. For this purpose, the spectroscopic sample is usually split into two samples: one used to train the algorithm and the other one to validate the photo- z s (known as the training and validation sample, respectively). However, we must be careful when extrapolating the results obtained in the validation sample to the galaxies in the scientific target sample. The scientific sample may well contain galaxies at deeper magnitudes or in a different colour range which are not represented in the training sample and photo- z s may not be correctly estimated.

4.1. Incompleteness emulation with the spectroscopic sample

With the goal to learn how the incompleteness affects photometric redshift performance, we use the spectroscopic sample to emulate, at brighter magnitudes, two different scenarios: a case in which we have completeness of magnitude and colour coverage

from the training to the target sample, and another in which we do not (incomplete case).

We split the spectroscopic catalogue into two sub-samples of equal size (with 27,801 galaxies each) and equal magnitude-colour distribution. We take one of these sub-samples as a validation sample and the other as training sample. We select the galaxies of the training sample in two different ways to emulate the scenarios mentioned. On one hand, we take all galaxies of the training sample (the 27,801 galaxies) to emulate a complete training set, that is a training sample that covers the same colour-magnitude space as the validation sample. On the other hand, we use the training sample to construct an incomplete version. In this second case, we want to emulate, at brighter magnitudes, the incompleteness observed in the Y3 DES Deep Fields catalogue when the training sample is formed by galaxies of the spectroscopic sample with $\text{FLAG_DES}=4$. To achieve this, some high-magnitude galaxies can be manually removed from the training sample until incompleteness is reached. In order to automate this process rather than performing it manually, we have employed the following method. We first calculate Δ_{band} , that is the difference between the mean magnitude of the objects in the spectroscopic sample and in the Y3 DES Deep Fields photometric catalogue, for each band. Then, we subtract Δ_{band} from the magnitudes of every galaxy in the training sample to emulate a similar incompleteness at brighter magnitudes. Applying this magnitude left-shift, we achieve a magnitude incompleteness at the expense of decoupling galaxies from their own redshift. To solve this issue, we use a nearest neighbour algorithm to find, within the shifted sample, real galaxies with similar magnitudes. The algorithm assigns for each left-shifted magnitude a real galaxy from the training sample, many of them repeated. After applying this procedure and dropping out the repeated galaxies, the new training sample, hereafter referred to as the incomplete training sample, is reduced to 5,336 galaxies out of the original 27,801. In this way, we now have a galaxy sample that simulates incompleteness in a magnitude range for which we have information about the spectroscopic redshift, enabling us for the study of the effects of incompleteness.

Figure 1 shows the magnitudes and colour distributions (upper and lower plots, respectively) for the incomplete training sample (red dashed lines) and the validation sample (blue lines). We have not included the comparison to the complete training sample since their distributions overlap perfectly with those of the validation sample by construction.

4.2. Incompleteness assessment

We determine the DNF photometric redshifts for the validation sample with both complete and incomplete training sets defined as in Sect. 4.1. We select objects meeting the conditions $\text{DNF_Z} > 0$, $\text{DNF_ZN} > 0$ and $\text{DNF_ZSIGMA} < 1.0$ to ensure the quality of the sample. The cut-off of DNF_ZSIGMA has been defined taking into account the analysis carried out in Appendix B which studies the possible biases that DNF_ZSIGMA may have as a quality estimator of DNF photo-z. The number of galaxies after these cuts is 26,882 galaxies (96.7% of the sample) using the complete training sample and 22,617 (81.3% of the sample) using the incomplete training sample.

Figure 2 shows the magnitude and colour distributions of the galaxies in the validation sample (blue lines) versus the distributions of their nearest-neighbours galaxies (orange dashed lines) determined from the incomplete training sample. Note that while nearest-neighbour magnitude distributions do not match to the weaker magnitudes in all the filters, the colour distributions

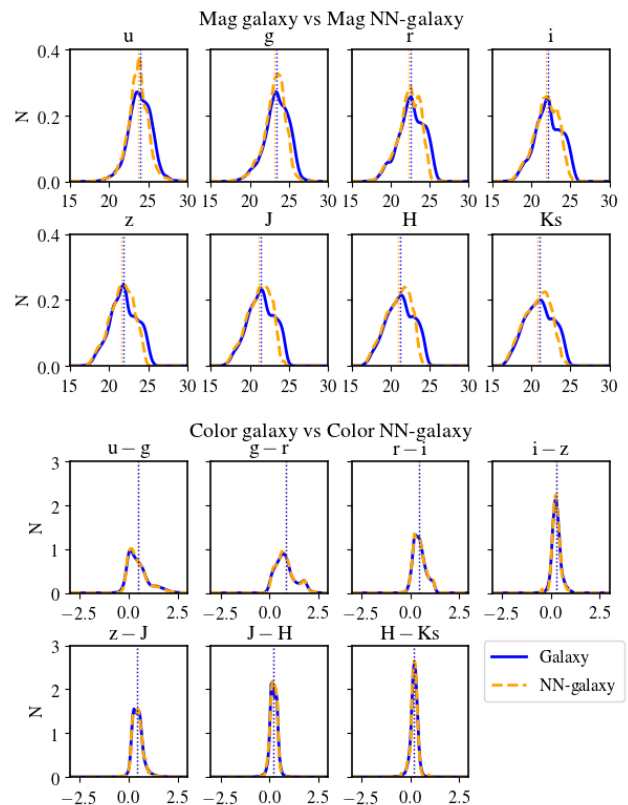


Fig. 2. Magnitude and colour distribution for the nearest neighbour galaxies used for the estimation of photo-z in the incomplete training and the validation sample. The orange dashed lines represent the nearest neighbour galaxies distribution of incomplete training and the blue lines the validation sample. The dotted vertical lines are the mean of each distribution. We have not included the distributions for the complete training sample since they overlap perfectly with those of the validation sample.

are close to being recovered in comparison with those shown in Fig. 1. The matching of the colour distributions between the validation sample and their nearest-neighbour in the incomplete training sample may be a necessary condition to produce a reliable photometric redshift distribution. However, it may not be sufficient due to the possibility of the existence of galaxies with colour combinations not covered by the training sample.

In order to study the effect of incompleteness and how to detect it, we have carried out a principal component analysis (PCA). The PCA has been performed with the magnitudes of the bands $ugrizJHK_s$. The first principal component (PC_1) for this sample represents 92.8% of the variance of the validation sample in magnitude space, while the percentage grows up to 98.1% with the second component (PC_2). We represent the density map of the validation sample for the principal components in the upper panel of Fig. 3. We have also stored the first two eigenvectors obtained for the validation sample to represent in the same basis the training sample. In the bottom panel of Fig. 3 the red dots show the scatter of the incomplete training sample represented using the same eigenvectors of the validation sample. Comparing both panels of Fig. 3, it can be seen that the incomplete training sample does not cover the full validation sample, but this red area is well delimited by the inner bold black line that corresponds to the region for which $\text{DNF_ZSIGMA} < 0.1$. Note that this plot shows the limitations in determining the photo-zs using the DNF algorithm with an incomplete training sample but it

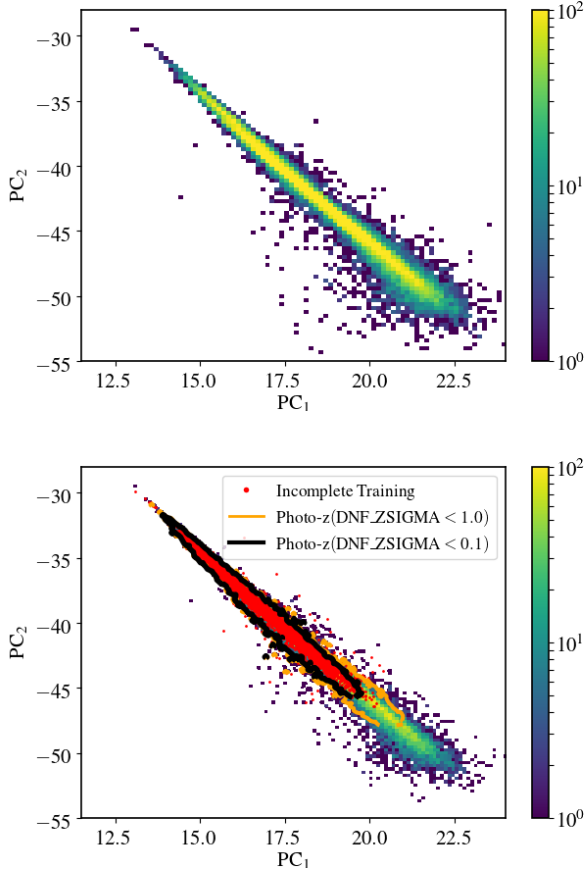


Fig. 3. Density map as a function of the first and second principal component of the galaxies of the validation sample. In the bottom panel, we included the galaxies of the training sample (red dots) and the limit in principal components of the galaxies that DNF provides a value of photo-z with DNF_ZSIGMA < 1.0 (orange line) and DNF_ZSIGMA < 0.1 (bold black line) with this training sample.

Table 1. Summary of metrics.

Training Sample	Num. galaxies	$\overline{\Delta z}$	σ_{68}^{Norm}
Complete	26, 883	0.0104	0.0373
Incomplete	22, 651	-0.0740	0.0500

also shows how DNF_ZSIGMA informs of this fact. The outer orange line corresponds to galaxies with DNF_ZSIGMA < 1.0 (this is 81.3% of the sample). Therefore, we can identify 3 groups of galaxies. Those galaxies covered by the red dots will have precise photo-zs since for these galaxies the training sample covers the full range of principal component. On other hand, DNF tags as unreliable photo-zs those galaxies outside the orange limit. Therefore, we must study the quality of the photo-zs of the galaxies that are located inside the orange limit but are not covered by the training sample (red area).

4.3. Photo-zs performance estimation

We compare the photo-z estimation given by DNF in the validation sample using the incomplete and complete training sam-

ples. Figure 4 shows the comparison between the spectroscopic redshift (z_{spec}) and the photometric redshift (DNF_Z) for incomplete training sample (left panel) and complete training sample (right panel). We can see that the complete training sample not only determines photo-z values for a larger number of galaxies compared to the incomplete case (26, 883 galaxies versus 22, 651), but also presents a lower bias and σ_{68}^{Norm} (see Table 1). In addition to the completeness, the number of galaxies in the training sample is a factor that influences the quality of the photo-zs. In Appendix C we have included the results of calculating the photometric redshift using a complete sample with the same number of galaxies as the incomplete sample. The results show that the completeness allows to calculate more accurate photo-zs than the incomplete case for comparables training sample sizes.

We have studied the behaviour of the photo-zs estimation through the mean absolute deviation, the σ_{68}^{Norm} and outliers. Figure 5 shows the mean absolute deviation (upper panel) and the σ_{68}^{Norm} (bottom panel) of DNF_Z with respect to z_{spec} and DNF_ZN (solid and dashed lines, respectively) as a function of the $mag(i)$ for the complete training sample (blue lines) and incomplete training sample (magenta lines). The solid lines display the mean absolute deviation and σ_{68}^{Norm} of the photo-zs calculated with z_{spec} (which we will refer to as real metric values). It can be readily seen in Fig. 5 that the completeness of the training sample affects the metrics. Then, we cannot assume that the metrics (mean absolute deviation, σ_{68}^{Norm} or those chosen in the study) will have the same behaviour in the validation sample and in the target sample if the training sample shows incompleteness. We must keep in mind that the z_{spec} value of each galaxy is not available when we are calculating the photo-zs for a catalogue so we will not have these measurements to estimate the precision of the photo-zs. Nevertheless, note that DNF_ZN (nearest neighbour photo-z) is able to reproduce z_{spec} distribution for moderate training incompleteness in the same way that colour distributions are well recovered for the case of incomplete training (Fig. 2). In this way, statistical metrics involving z_{spec} are well represented by DNF_ZN. For this reason, we have calculated the mean absolute deviation and the σ_{68}^{Norm} replacing z_{spec} by DNF_ZN (dashed lines). In both plots the behaviour of the mean absolute deviation and the σ_{68}^{Norm} can be considered a good approximation to the real value which changes depending on the training sample. We can take these metrics calculated by DNF_ZN as an upper limit of real ones. Figure 6 shows the outliers as a function of the i band magnitude $mag(i)$ for the complete training case (blue lines) and incomplete training case (magenta lines). The number of outliers is less than 4% in both cases, up until $mag(i) > 24$ where it starts to increase for the incomplete training case. Finally, we complete this study with the behaviour of the photo-z estimation as a function of the spectroscopic redshifts in Appendix D.

5. Photometric redshift Deep Fields catalogue

We want to study the effects of using different training samples on the quality of the photo-zs in the Y3 DES Deep Fields catalogue. For that, we applied the same analysis developed in Sect. 4 using two training samples. The first training sample contains only galaxies with the highest quality of spectroscopic redshift determination (i.e. with FLAG_DES = 4). In this case, the training sample does not contain galaxies with magnitudes as deep as the Y3 DES Deep Fields catalogue. In other words, this training sample is of the highest quality but shows a certain incompleteness with respect to the science sample. The second training sample contains galaxies labelled with spectroscopic redshift

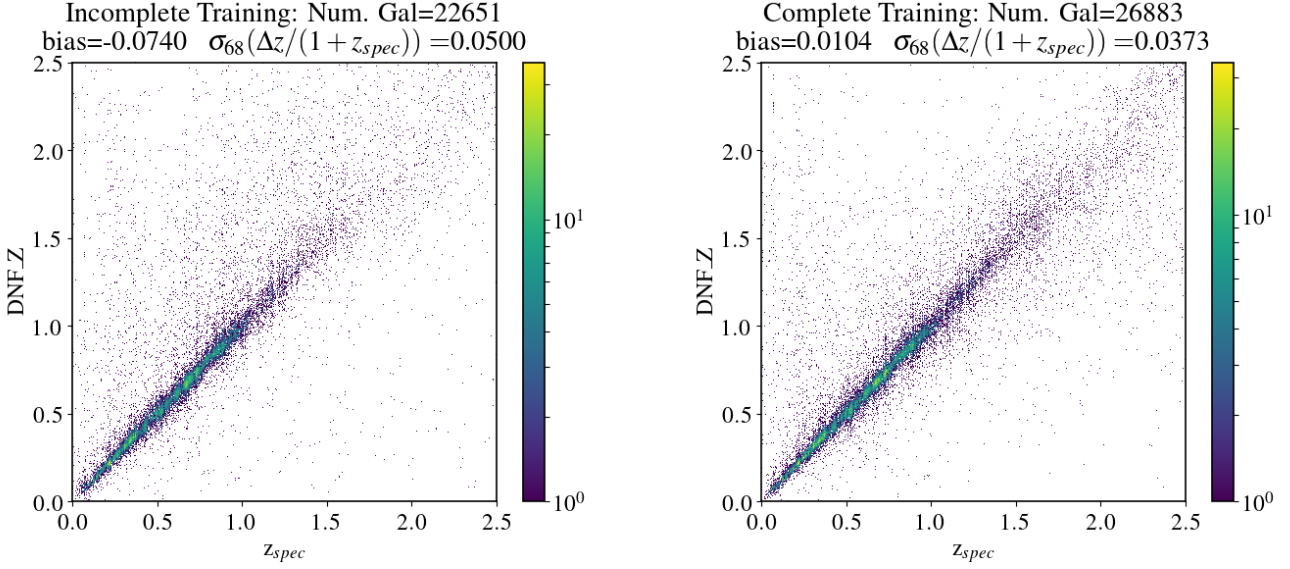


Fig. 4. Scatter plot of spectroscopic redshift z_{spec} and the photo-z DNF_Z for incomplete training (left panel) and complete training (right panel). Note the improvement for $z > 1$ in the complete case.

quality $\text{FLAG_DES} \geq 3$. The inclusion of galaxies whose spectroscopic redshift quality is not optimal but still good in this training sample reduces the problem of incompleteness. In this second case, the training sample reaches the deepest magnitudes of the Y3 DES Deep Fields catalogue. Figure 7 shows the magnitude and colour distributions of both training samples (in the left panels for the incomplete sample and in the right panels for the semi-complete sample). In order to carry out a similar analysis of that performed in Sect. 4, we selected those galaxies with $\text{mag}(i) < 28.0$ and with a positive flux measurement in the eight filters. This sample contains 1,478,705 galaxies.

5.1. Assessment of high quality but incomplete training

For this study, the incomplete training sample is limited to 38,123 galaxies for which their spectroscopic redshift has been determined with very high quality. As we can see in the left panels of Fig. 7, this spectroscopic sample is shallower than the Y3 DES Deep Fields catalogue (red lines and blue lines, respectively). We want to know how this incompleteness affects the photometric redshift calculation. Using DNF and selecting galaxies with $\text{DNF_Z} > 0$, $\text{DNF_ZN} > 0$ and $\text{DNF_ZSIGMA} < 1.0$, we have determined the photometric redshift for 1,254,981 galaxies (84.9%) of the Deep Fields catalogue using this training sample.

The left panel of Fig. 8 shows the density map as a function of the first and second principal component of the Y3 DES Deep Fields catalogue and the galaxies of the training sample (red dots). The orange line is the limit of the photo-zs of this 84.9% of galaxies with the cuts defined above. In addition we overplot another limit (bold black line) using a more stringent cut, namely $\text{DNF_ZSIGMA} < 0.1$ that keeps 441,144 galaxies, that is 29.8%.

5.2. Assessment of medium quality but semi-complete training

The second training sample used to determine the photometric redshift of Y3 DES Deep Fields catalogue contains 55,601 galaxies which have magnitudes as deep as the Y3 DES Deep

Fields catalogue but with a different distribution as shown in the right panels of Fig. 7.

We can see in the right panel of Fig. 8, corresponding to the principal components (the first two eigenvectors represent 93.59% of the sample), that the spectroscopic training sample is located in the area where the density of galaxies is higher. Although it does not cover the entire of principal component area of the field, DNF provides photo-z for almost all galaxies in the sample (1,318,960 galaxies, 91.67%) delimited by the orange line in the figure. We plot another limit with a bold black line that represents galaxies with a more stringent cut of $\text{DNF_ZSIGMA} < 0.1$ as we did before (405,854 galaxies, i.e. 28.2%).

5.3. Performance and comparison of science sets with different training samples

According to the results obtained, based on the cuts defined in Sect. 4, DNF determines photometric redshifts for slightly fewer galaxies when using the incomplete but high quality training sample than in the semi-complete case. The question is how the quality of these photometric redshifts estimates compare. Or in other words, whether it is more important to have high quality spectroscopic redshifts or we can slightly relax that condition to cover the magnitude-colour space as much as possible.

The results of Fig. 9 show the precision of the photo-z estimation by DNF for Y3 DES Deep Fields catalogue defined by mean absolute deviation (left panel) and the $\sigma_{68}^{\text{Norm}}$ (right panel) as a function of the $\text{mag}(i)$. Note that the z_{spec} of each galaxy is not available in Y3 DES Deep Fields catalogue, then to estimate the mean absolute deviation and the $\sigma_{68}^{\text{Norm}}$, we have replaced z_{spec} by DNF_ZN following the analysis done in Sect. 4.3. We can see that the results obtained by the incomplete, high quality training (dashed purple lines) and the semi-complete training (blue lines) samples follow a similar behaviour for $\text{mag}(i) < 24$ although slightly better for the incomplete, high quality training. In this case we obtain a lower error for magnitudes-colour areas covered by the spectroscopic sample.

We have also seen the same behaviour in Sect. 5.1 and 5.2 where the incomplete, high quality training contains more galax-

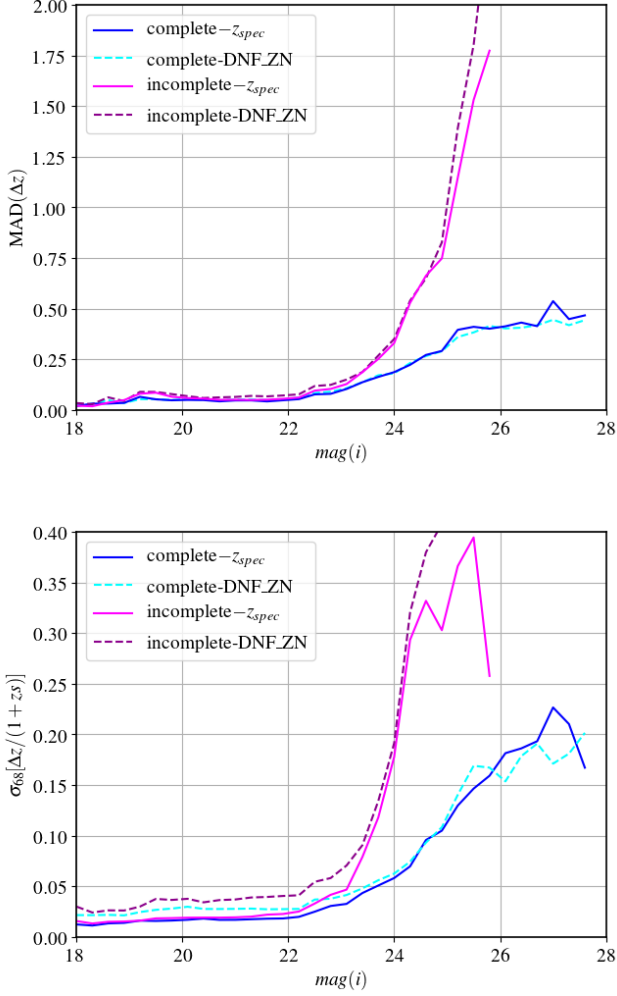


Fig. 5. The precision of the photo-z defined by $MAD(\Delta z)$ (upper panel) and the σ_{68}^{Norm} (bottom panel) as a function of the $mag(i)$ for the complete training sample (blue lines) and incomplete training sample (magenta lines). The solid lines display the metrics calculated with z_{spec} and the dashed lines has been calculated replacing z_{spec} by DNF_ZN.

ies with $DNF_ZSIGMA < 0.1$ even though, globally, the semi-complete training generates more precise photo-zs. For $mag(i) \geq 24$, the semi-complete training sample, formed by galaxies with slightly lower quality spectroscopic redshift, outperforms the photo-zs of the incomplete training sample formed by the highest quality spectroscopic redshift galaxies. The results indicate that completeness plays an important role in determining higher quality photometric redshift values, as expected. But the results also suggest that for specific studies focussed on brighter galaxies, we may be more interested in using only the redshifts of the highest possible quality in our training.

Finally, we studied the absolute median deviation and σ_{68}^{Norm} as a function of the redshift for the two training samples. You can see more details in Appendix D.

6. Comparison between DNF and EAZY

We estimated the photo-zs for the whole deep fields and we added this information to the Y3 DES Deep Fields data².

² Available at des.ncsa.illinois.edu/releases/y3a2/Y3deepfields

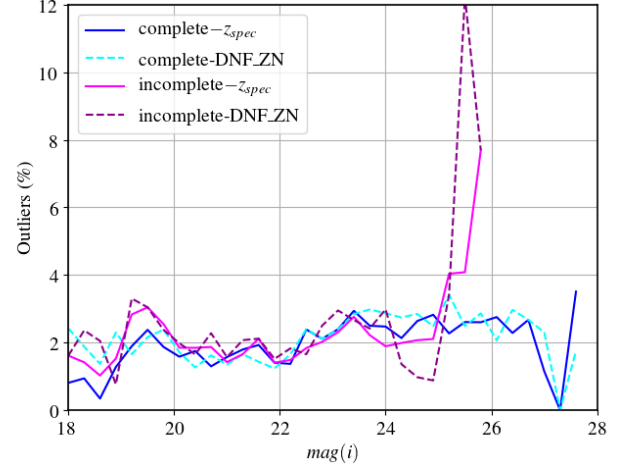


Fig. 6. Outliers as a function of the $mag(i)$ for the complete training (blue lines) and incomplete training (magenta lines). The solid lines display the outliers calculated with z_{spec} and the dashed lines replacing z_{spec} by DNF_ZN.

The training sample used to estimate the photo-zs contains galaxies with spectroscopic redshift information labeled with $FLAG_DES \geq 3$ corresponding to the semi-complete training sample in Sect. 5.2. It is important to note that when computing DNF in this case, we ignore the own spectroscopic redshift for each galaxies in the training sample, to provide a homogeneous comparison of all estimates.

In addition to the DNF photo-zs (De Vicente et al. 2016), the Y3 DES Deep Fields catalogue contains photo-zs determined with the EAZY algorithm (Hartley et al. 2022; Brammer et al. 2008). These two methods approach the photometric redshift problem from different perspectives: EAZY determined the photo-zs by fitting a linear combination of template components while DNF is a machine learning code.

We analysed the photo-zs obtained using both methods. Firstly, we selected from Y3 Deep Fields catalogue those galaxies with spectroscopic redshift information, $mag(i) < 28.0$, flux measurements in the eight filters. This sample contains 55,198 galaxies and covers a large portion of the total sample as we can see in the right panel of Fig. 8. Figure 10 shows the scatter of the photo-zs determined by both methods versus the spectroscopic redshift (on the left DNF and on the right EAZY). The metrics obtained by DNF slightly outperforms those provided by EAZY, justified in part for the completeness of the training sample used in this test. The bias and the σ_{68}^{Norm} are -0.0148 and 0.0519 for EAZY and 0.0065 and 0.0390 for DNF respectively. Both methods give good photo-zs values for $z < 1$. However, for $z \geq 1$ EAZY shows a somewhat biased behaviour. The plot at the bottom shows the photo-z values of DNF (X-axis) vs EAZY (Y-axis). We can see the same bias appeared on right panel, that is EAZY with respect to SPEC_Z. Therefore, this behaviour seems to come from the EAZY estimation. It may be due to the lack of Y-band data. The break is poorly constrained from $z \sim 1$ until the $4000\text{\AA}$ break starts to enter the J-band. The prior tends to favour a lower redshift and so the point estimates are pulled to lower redshift slightly. This would be partially alleviated with the full EAZY PDFs.

In Fig. 11 (left), we compare the photo-z provided by both methods for the Y3 DES Deep Fields catalogue (right panel). We have focussed on galaxies with flux measurements in all

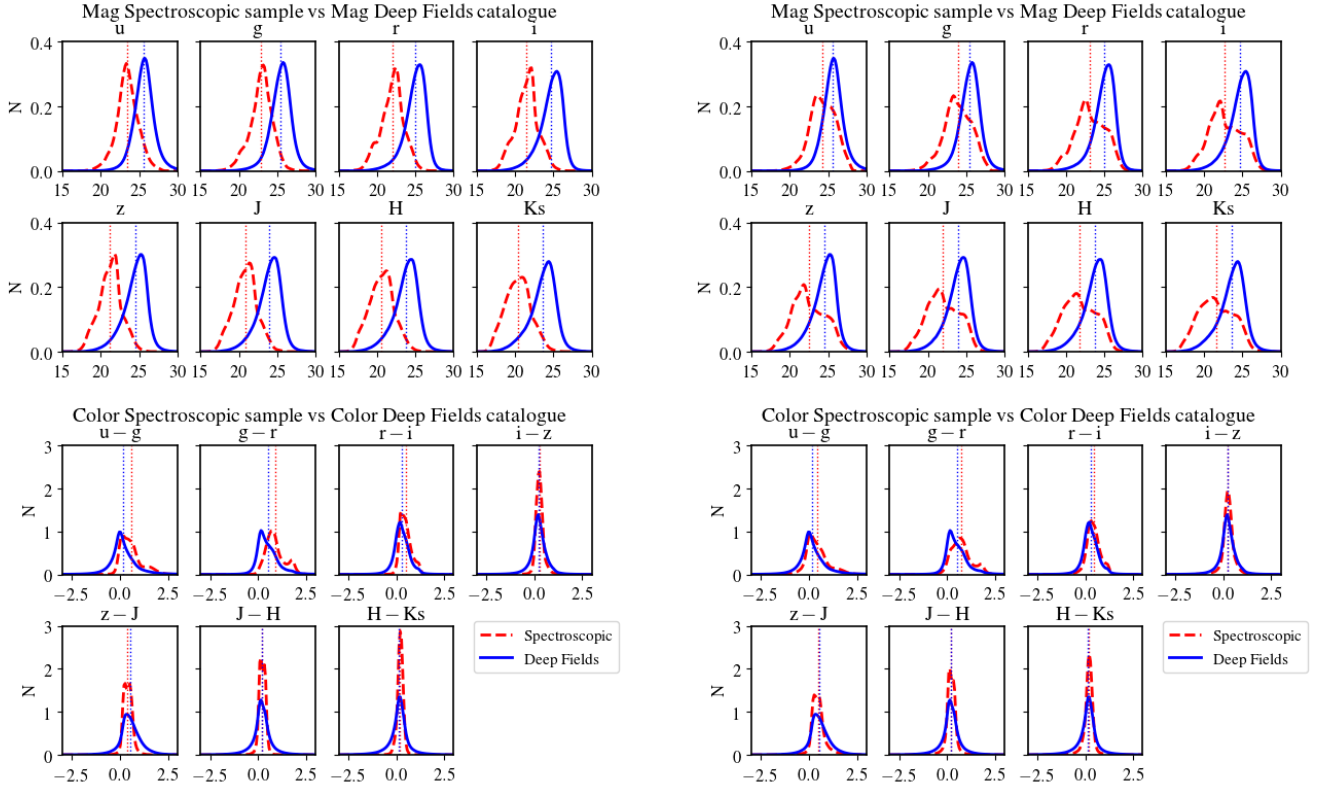


Fig. 7. Magnitude and colour distribution for incomplete training (right panels) and semi-complete training (left panels). The red dashed lines represents the training samples and the blue lines the Y3 DES Deep Fields sample. The dotted lines are the mean of each distribution.

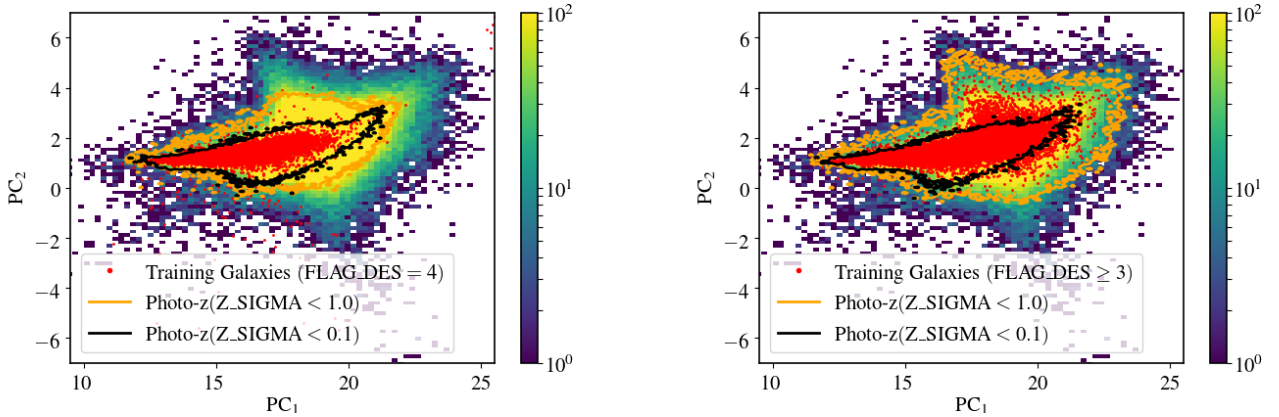


Fig. 8. Density map as a function of the first and second principal component of the galaxies of Y3 DES Deep Fields (density plot in green and yellow colours), the galaxies of the training sample (red dots) and the limit in principal components of the galaxies that DNF provides a value of photoz with $\text{DNF_ZSIGMA} < 1.0$ (orange line) and $\text{DNF_ZSIGMA} < 0.1$ (bold black line) with this training sample. The red blob is less extensive than the green-yellow blob (where the highest density of galaxies is located) when we select only the galaxies with $\text{FLAG_DES} = 4$.

the eight filters and $\text{mag}(i) < 28.0$. It corresponds to a sample of 1,473,381 galaxies. For $z > 1$ we can see a similar behaviour to that observed with spectroscopic redshifts (right panel of Fig. 10). Therefore, this behaviour seems to come from EAZY estimation. On the other hand, there is a cloud of points below the diagonal around $\text{EAZY_Z} \sim 0.5$ which extend along several values of DNF_Z . We can see in Fig. 11 (right) that the cloud can be removed by applying the quality cut $\text{DNF_ZSIGMA} < 0.5$. In general DNF_ZSIGMA allow us to detect galaxies with large errors due to bad photometry, degeneracies or incompleteness.

Determining the best method to be applied to a scientific sample is non-trivial. Salvato et al. (2019) points out that machine learning methods outperforms the template approaches when the training survey is sufficiently complete. However, template methods are more favourable when spectroscopic samples are limited. In the case of DNF and EAZY, the biggest differences appear for $z > 1$, when the completeness of the training sample is poorer. Nevertheless, the photo-zs generated by DNF present better metrics than those of provided by EAZY. According to Salvato et al. (2019), template methods work best for high redshift because of the lack of photometric information to con-

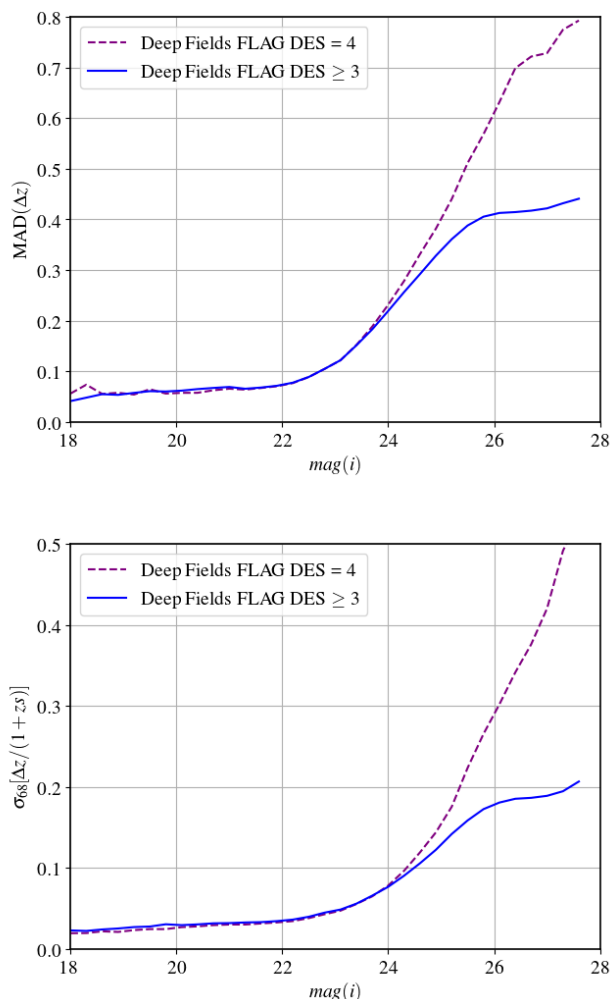


Fig. 9. The precision of the photo- z estimates defined by $\text{MAD}(\Delta z)$ (upper panel) and $\sigma_{68}^{\text{Norm}}$ (bottom panel) as a function of the $\text{mag}(i)$ calculated by DNF_Z and DNF_ZN determined by training sample with only galaxies with $\text{FLAG_DES} = 4$ (incomplete training sample, in purple) and training sample with galaxies with $\text{FLAG_DES} \geq 3$ (semi-complete training sample, in blue).

struct training samples for machine learning methods. In the same sense, the templates are built on physical assumptions that may not be entirely correct or have incomplete coverage in certain areas.

7. Conclusions

This study is an analysis of how the completeness and spectroscopic quality of the training sample affects the photometric redshift determination using the DNF algorithm. The conclusions are the following:

1. We have emulated the problem of an incomplete training sample for DNF with the goal of measuring its effects and to take them into account on the photo- z performance. The principal component analysis provides a graphical method to assess completeness and DNF_ZSIGMA turns out to be a reliable parameter to separate the set of galaxies computed with a complete training.
2. We have analysed the possibility of substituting z_{spec} by DNF_ZN to assess $\Delta(z)$ in the scatter metrics of DNF_Z

($\text{MAD}(\Delta z)$ and $\sigma_{68}^{\text{Norm}}$). The results show that DNF_ZN provides an upper limit of the real values. Using this method the photo- z quality can be estimated when no spectroscopic information is available.

3. We determine the photo- z s of the Y3 DES Deep Fields catalogue using both, a semi-complete training sample with high and medium quality redshift spectroscopy and an incomplete training sample with the highest quality redshift spectroscopy. The obtained results are globally better for the semi-complete sample in spite of its slight relaxation in quality. However, the photo- z improves for that sub-sample where the high-quality incomplete training covers its principal component analysis space. For faint magnitudes it seems better to use a training sample with medium quality spectroscopic redshift covering deeper magnitudes. This result advocates for training completeness at the expenses of sacrificing slightly the quality of the spectroscopic redshifts. The results also suggest that for specific studies focussed on brighter galaxies, we may be more interested in using only redshift of the highest possible quality in our training.
4. We have compared the photometric redshift of Y3 DES Deep Fields catalogue determined with DNF and EAZY. Both methods show a similar behaviour up to $z \sim 1$. For $z > 1$ DNF outperforms EAZY, which shows some bias towards higher redshift.

Acknowledgements. We acknowledge funding support from the Autonomous Community of Madrid through the project TEC2SPACE-CM (S2018/NMT-4291). Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium. Based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2). This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

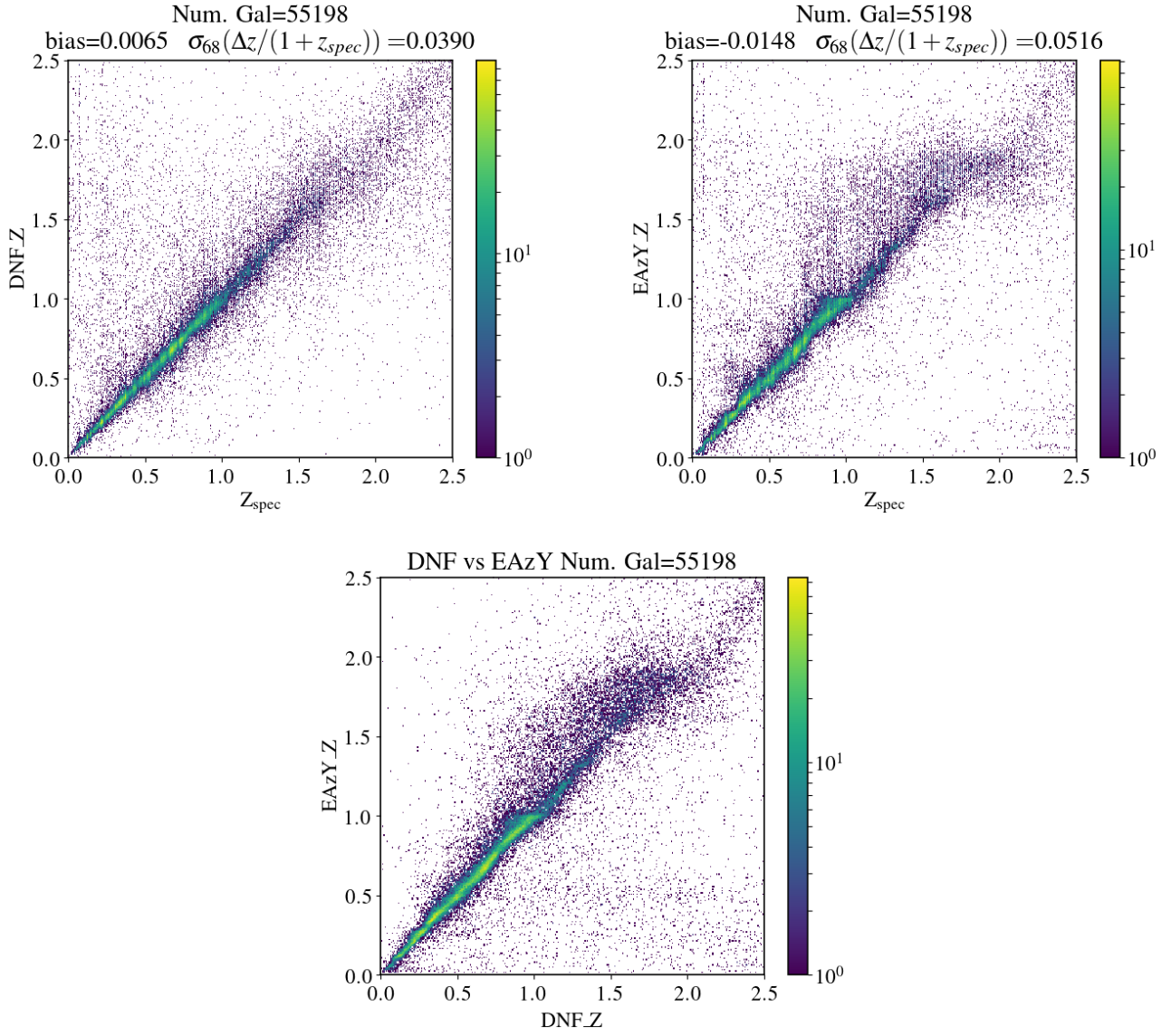


Fig. 10. Scatter plot of spectroscopic sample: SPEC_Z vs DNF_Z (left), SPEC_Z vs EAzy_Z (right) and DNF_Z vs EAzy_Z (bottom)

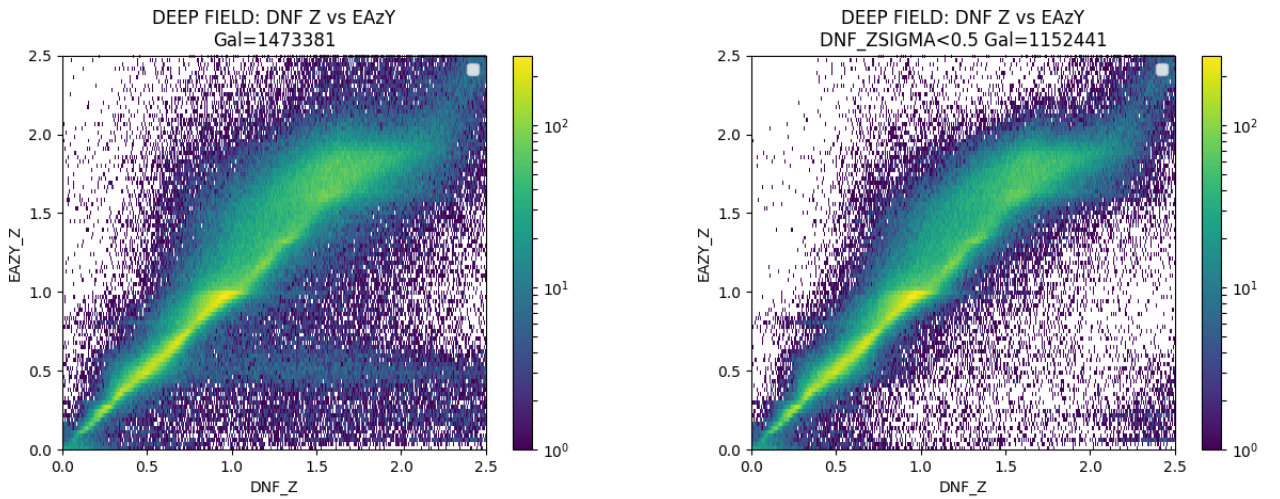


Fig. 11. Y3 DES Deep Fields catalogue. Scatter plot DNF_Z vs EAzy_Z: for $\text{mag} < 28$ (left) and for additional quality cut $\text{DNF_ZSIGMA} < 0.5$ (right)

References

- Abdalla, F. B., Banerji, M., Lahav, O., et al. 2011, *MNRAS*, 417, 1891. doi:10.1111/j.1365-2966.2011.19375.x
- Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, 235, 42. doi:10.3847/1538-4365/aa9e8a
- Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, *MNRAS*, 310, 540. doi:10.1046/j.1365-8711.1999.02978.x
- Bayliss, M. B., Ruel, J., Stubbs, C. W., et al. 2016, *ApJS*, 227, 3. doi:10.3847/0067-0049/227/1/3
- Bazin, G., Ruhlmann-Kleider, V., Palanque-Delabrouille, N., et al. 2011, *A&A*, 534, A43. doi:10.1051/0004-6361/201116898
- Beck, R., Lin, C.-A., Ishida, E. E. O., et al. 2017, *MNRAS*, 468, 4323. doi:10.1093/mnras/stx687
- Benítez, N. 2000, *ApJ*, 536, 571. doi:10.1086/308947
- Blake, C., Amon, A., Childress, M., et al. 2016, *MNRAS*, 462, 4240. doi:10.1093/mnras/stw1990
- Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, *A&A*, 363, 476. doi:10.48550/arXiv.astro-ph/0003380
- Bonnett, C., Troxel, M. A., Hartley, W., et al. 2016, *Phys. Rev. D*, 94, 042005. doi:10.1103/PhysRevD.94.042005
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503. doi:10.1086/591786
- Brescia, M., Cavuoti, S., Razim, O., et al. 2021, *Frontiers in Astronomy and Space Sciences*, 8, 70. doi:10.3389/fspas.2021.658229
- Carrasco Kind, M. & Brunner, R. J. 2013, *MNRAS*, 432, 1483. doi:10.1093/mnras/stt574
- Castander, F. J., Ballester, O., Bauer, A., et al. 2012, *Proc. SPIE*, 8446, 84466D. doi:10.1117/12.926234
- Cavuoti, S., Brescia, M., Longo, G., et al. 2012, *A&A*, 546, A13. doi:10.1051/0004-6361/201219755
- Cavuoti, S., Tortora, C., Brescia, M., et al. 2017, *MNRAS*, 466, 2039. doi:10.1093/mnras/stw3208
- Childress, M. J., Lidman, C., Davis, T. M., et al. 2017, *MNRAS*, 472, 273. doi:10.1093/mnras/stx1872
- Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, *ApJ*, 741, 8. doi:10.1088/0004-637X/741/1/8
- Colless, M., Dalton, G., Maddox, S., et al. 2001, *MNRAS*, 328, 1039. doi:10.1046/j.1365-8711.2001.04902.x
- Collister, A. A. & Lahav, O. 2004, *PASP*, 116, 345. doi:10.1086/383254
- Blanton, M. R., et al. 2013, *ApJ*, 767, 118. doi:10.1088/0004-637X/767/2/118
- Cool, R. J., Moustakas, J., Blanton, M. R., et al. 2013, *ApJ*, 767, 118. doi:10.1088/0004-637X/767/2/118
- Cooper, M. C., Yan, R., Dickinson, M., et al. 2012, *MNRAS*, 425, 2116. doi:10.1111/j.1365-2966.2012.21524.x
- Davis, C., Gatti, M., Vielzeuf, P., et al. 2017, *arXiv:1710.02517*. doi:10.48550/arXiv.1710.02517
- Davis, M., Faber, S. M., Newman, J., et al. 2003, *Proc. SPIE*, 4834, 161. doi:10.1117/12.457897
- De Vicente, J., Sánchez, E., & Sevilla-Noarbe, I. 2016, *MNRAS*, 459, 3078. doi:10.1093/mnras/stw857
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, *arXiv:1611.00036*. doi:10.48550/arXiv.1611.00036
- Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, *MNRAS*, 413, 971. doi:10.1111/j.1365-2966.2010.18188.x
- Euclid Collaboration, Desprez, G., Paltani, S., et al. 2020, *A&A*, 644, A31. doi:10.1051/0004-6361/202039403
- Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, *AJ*, 150, 150. doi:10.1088/0004-6256/150/5/150
- Garilli, B., Le Fèvre, O., Guzzo, L., et al. 2008, *A&A*, 486, 683. doi:10.1051/0004-6361/20078878
- Garilli, B., Guzzo, L., Scodreggio, M., et al. 2014, *A&A*, 562, A23. doi:10.1051/0004-6361/201322790
- Geha, M., Wechsler, R. H., Mao, Y.-Y., et al. 2017, *ApJ*, 847, 4. doi:10.3847/1538-4357/aa8626
- Gschwend, J., Rossel, A. C., Ogando, R. L. C., et al. 2018, *Astronomy and Computing*, 25, 58. doi:10.1016/j.ascom.2018.08.008
- Hartley, W. G., Chang, C., Samani, S., et al. 2020, *MNRAS*, 496, 4769. doi:10.1093/mnras/staa1812
- Hartley, W. G., Choi, A., Amon, A., et al. 2022, *MNRAS*, 509, 3547. doi:10.1093/mnras/stab3055
- Hildebrandt, H., Arnouts, S., Capak, P., et al. 2010, *A&A*, 523, A31. doi:10.1051/0004-6361/201014885
- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *A&A*, 457, 841. doi:10.1051/0004-6361/20065138
- Jones, D. H., Read, M. A., Saunders, W., et al. 2009, *MNRAS*, 399, 683. doi:10.1111/j.1365-2966.2009.15338.x
- Kaiser, N., Burgett, W., Chambers, K., et al. 2010, *Proc. SPIE*, 7733, 77330E. doi:10.1117/12.859188
- Le Fèvre, O., Vettolani, G., Paltani, S., et al. 2004, *A&A*, 428, 1043. doi:10.1051/0004-6361:20048072
- Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, *A&A*, 439, 845. doi:10.1051/0004-6361:20041960
- Lidman, C., Ardila, F., Owers, M., et al. 2016, *PASA*, 33, e001. doi:10.1017/pasa.2015.52
- Lidman, C., Ruhlmann-Kleider, V., Sullivan, M., et al. 2013, *PASA*, 30, e001. doi:10.1017/pasa.2012.001
- Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, *ApJS*, 184, 218. doi:10.1088/0067-0049/184/2/218
- Lima, M., Cunha, C. E., Oyaizu, H., et al. 2008, *MNRAS*, 390, 118. doi:10.1111/j.1365-2966.2008.13510.x
- LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, *arXiv:0912.0201*. doi:10.48550/arXiv.0912.0201
- Mao, M. Y., Sharp, R., Norris, R. P., et al. 2012, *MNRAS*, 426, 3334. doi:10.1111/j.1365-2966.2012.21913.x
- Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2017, *ApJ*, 841, 111. doi:10.3847/1538-4357/aa6f08
- Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, *ApJS*, 225, 27. doi:10.3847/0067-0049/225/2/27
- Muzzin, A., Wilson, G., Yee, H. K. C., et al. 2012, *ApJ*, 746, 188. doi:10.1088/0004-637X/746/2/188
- Nanayakkara, T., Glazebrook, K., Kacprzak, G. G., et al. 2016, *ApJ*, 828, 21. doi:10.3847/0004-637X/828/1/21
- Nord, B., Buckley-Geer, E., Lin, H., et al. 2016, *ApJ*, 827, 51. doi:10.3847/0004-637X/827/1/51
- Parkinson, D., Riemer-Sørensen, S., Blake, C., et al. 2012, *Phys. Rev. D*, 86, 103518. doi:10.1103/PhysRevD.86.103518
- Rest, A., Scolnic, D., Foley, R. J., et al. 2014, *ApJ*, 795, 44. doi:10.1088/0004-637X/795/1/44
- Sadeh, I., Abdalla, F. B., & Lahav, O. 2016, *PASP*, 128, 104502. doi:10.1088/1538-3873/128/968/104502
- Salvato, M., Ilbert, O., & Hoyle, B. 2019, *Nature Astronomy*, 3, 212. doi:10.1038/s41550-018-0478-0
- Schmidt, S. J., Malz, A. I., Soo, J. Y. H., et al. 2020, *MNRAS*, 499, 1587. doi:10.1093/mnras/staa2799
- Scolnic, D., Rest, A., Riess, A., et al. 2014, *ApJ*, 795, 45. doi:10.1088/0004-637X/795/1/45
- Silverman, J. D., Kashino, D., Sanders, D., et al. 2015, *ApJS*, 220, 12. doi:10.1088/0067-0049/220/1/12
- Stalin, C. S., Petitjean, P., Srianand, R., et al. 2010, *MNRAS*, 401, 294. doi:10.1111/j.1365-2966.2009.15636.x
- Sullivan, M., Conley, A., Howell, D. A., et al. 2011, *VizieR Online Data Catalog*, J/MNRAS/406/782
- Sánchez, C., Carrasco Kind, M., Lin, H., et al. 2014, *MNRAS*, 445, 1482. doi:10.1093/mnras/stu1836
- Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, *PASJ*, 70, S9. doi:10.1093/pasj/psx077
- Tasca, L. A. M., Le Fèvre, O., Ribeiro, B., et al. 2017, *A&A*, 600, A110. doi:10.1051/0004-6361/201527963
- Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, *ApJ*, 812, 114. doi:10.1088/0004-637X/812/2/114
- York, D. G., Adelman, J., Anderson, J. E., et al. 2000, *AJ*, 120, 1579. doi:10.1086/301513
- Yuan, F., Lidman, C., Davis, T. M., et al. 2015, *MNRAS*, 452, 3047. doi:10.1093/mnras/stv1507

¹ Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain e-mail: laura.toribio@ciemat.es

² Department of Astronomy, University of Geneva, ch. d'Écogia 16, CH-1290 Versoix, Switzerland

³ Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA

⁴ Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

⁵ Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

⁶ Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

⁷ NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA

⁸ Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

- ⁹ Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA
- ¹⁰ School of Physics and Astronomy, Cardiff University, CF24 3AA, UK
- ¹¹ Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA
- ¹² Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA
- ¹³ Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
- ¹⁴ Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
- ¹⁵ Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK
- ¹⁶ Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
- ¹⁷ Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15312, USA
- ¹⁸ Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain
- ¹⁹ Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain
- ²⁰ Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain
- ²¹ Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain
- ²² Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
- ²³ Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester, M13 9PL, UK
- ²⁴ University of Nottingham, School of Physics and Astronomy, Nottingham NG7 2RD, UK
- ²⁵ Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany
- ²⁶ School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia
- ²⁷ Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India
- ²⁸ Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, NO-0315 Oslo, Norway
- ²⁹ Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
- ³⁰ Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, 28049 Madrid, Spain
- ³¹ Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA
- ³² Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
- ³³ Department of Physics, The Ohio State University, Columbus, OH 43210, USA
- ³⁴ Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
- ³⁵ Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia
- ³⁶ Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA
- ³⁷ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
- ³⁸ Centre for Gravitational Astrophysics, College of Science, The Australian National University, ACT 2601, Australia
- ³⁹ The Research School of Astronomy and Astrophysics, Australian National University, ACT 2601, Australia
- ⁴⁰ George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA
- ⁴¹ LPSC Grenoble - 53, Avenue des Martyrs 38026 Grenoble, France
- ⁴² Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain
- ⁴³ Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
- ⁴⁴ Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA
- ⁴⁵ SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
- ⁴⁶ School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK
- ⁴⁷ Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831
- ⁴⁸ Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

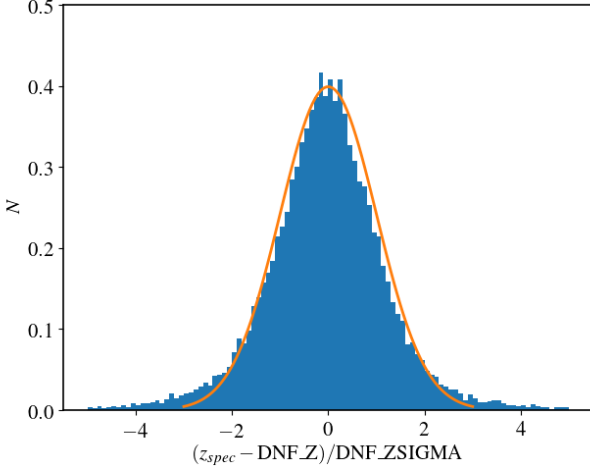


Fig. B.1. Comparison of *pull* distribution (blue) with a standard Gaussian distribution (orange line).

Appendix A: Spectroscopic data

We have listed in Table A.1 the 34 spectroscopic samples compiled by Gschwend et al. (2018) to create the spectroscopic sample used in this work.

Appendix B: DNF_ZSIGMA as an indicator of the quality of photo-z

DNF_ZSIGMA is the indicator of the quality of each photo-z provided by DNF. These values are computed from the quadratic sum of the error due to the photometry plus the error due to the fit. In this Appendix, we analyse the DNF_ZSIGMA values. For this purpose, we have calculated the *pull* defined as follow:

$$pull = \frac{z_{spec} - DNF_Z}{DNF_ZSIGMA},$$

where z_{spec} is the spectroscopic redshift and DNF_Z the photometric redshift.

Figure B.2 compares the *pull* distribution (blue) with a standard Gaussian distribution with mean zero and unit width (orange line) for the values obtained from the complete sample. The *pull* together with the central limit theorem allows us to analyse the possible dispersion and bias in the DNF_ZSIGMA values comparing the *pull* distribution with a standard Gaussian. The results obtained from the *pull* using the complete training sample fit to the Gaussian distribution. The *pull* distribution is slightly narrower in the centre and with larger wings. These differences are showing that DNF_ZSIGMA overestimates the errors for photo-zs with small errors and underestimates for large errors.

Appendix C: Effect of training sample size on Photometric Redshift

In addition to incompleteness, the number of galaxies in the training sample is also a factor that must be taken into account to determine the quality of the photometric redshift. We wanted to check what the results would be if the complete sample had the same number of galaxies as our incomplete sample, that is,

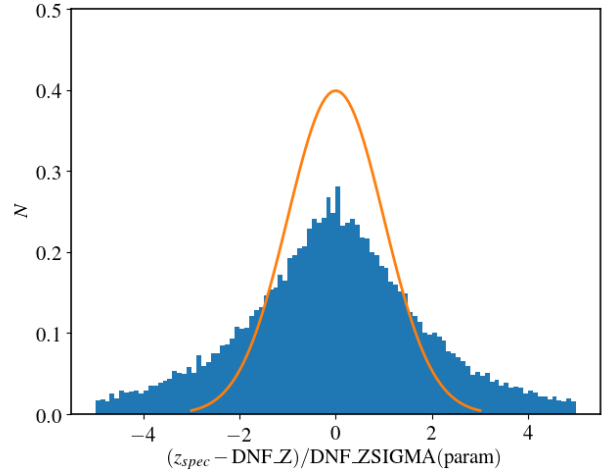
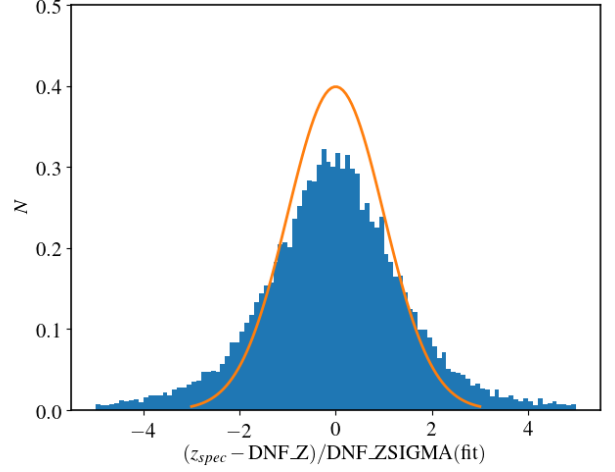


Fig. B.2. Comparison of *pull* distribution (blue) with a standard Gaussian distribution (orange line) for the error due to the fit (upper panel) and to the photometry (bottom panel).

5,336 galaxies. In this appendix, the Fig. C.1 shows the comparison between the spectroscopic redshift (z_{spec}) and the photometric redshift (DNF_Z) for a training sample that is complete but consists of 5,336 galaxies. The number of galaxies that DNF has calculated photometric redshifts for with the same cuts defined in 4.2 is 26,608 galaxies (95.7% of the sample). This value is very close to the case of the complete sample with 27,801 galaxies (96.8%) and considerably improves the result of the incomplete sample (81.3%). On the other hand, the results found are intermediate values between the incomplete and complete cases for bias and σ_{68}^{Norm} .

Appendix D: Quality metrics as a function of the redshift

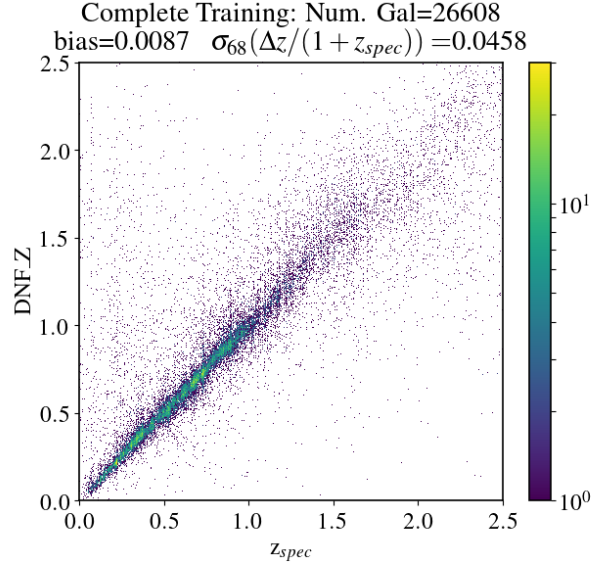
We studied the behaviour of the photo-zs estimation as a function of the spectroscopic redshift for the complete and incomplete spectroscopic training samples defined in Sect. 4. The Fig. D.1 and D.2 show the behaviour of the absolute median deviation and the σ_{68}^{Norm} as a function of z_{spec} for the complete training sample (blue lines) and incomplete training sample (magenta lines). We have also calculated the mean absolute deviation and the σ_{68}^{Norm} replacing z_{spec} by DNF_ZN (dashed lines). As in Fig. 5 of Sect.

Table A.1. Spectroscopic samples used in Gschwend et al. (2018).

Survey	Ref.
PRIMUS	Coil et al. (2011); Cool et al. (2013)
SDSS DR14	Abolfathi et al. (2018)
DES AAOmega	Yuan et al. (2015); Childress et al. (2017)
VIPERS	Garilli et al. (2014) and http://vipers.inaf.it/rel-pdr1.html
WiggleZ	Parkinson et al. (2012) and http://wigglez.swin.edu.au/site/
VVDS	Garilli et al. (2008); Le Fèvre et al. (2004)
zCOSMOS	Lilly et al. (2009)
3D-HST	Momcheva et al. (2016) and http://3dhst.research.yale.edu/Data.php
DEEP2	Davis et al. (2003, 2017) and http://deep.ps.uci.edu/DR4/home.html
2dF	Colless et al. (2001) and http://www.2dfgrs.net/
GAMA	Driver et al. (2011)
ACES	Cooper et al. (2012) and http://mur.ps.uci.edu/cooper/ACES/zcatalog.html
6dF	Jones et al. (2009) and http://www.6dfgs.net/
DES IMACS	Nord et al. (2016)
SAGA	Geha et al. (2017) and http://sagasurvey.org/
NOAO OzDES	Yuan et al. (2015); Childress et al. (2017)
XXL AAOmega	Lidman et al. (2016) and http://cosmosdb.iasf-milano.inaf.it/XXL/
SPT GMOS	Bayliss et al. (2016)
UDS	http://www.nottingham.ac.uk/astronomy/UDS/UDSz/
SNLS FORS	Bazin et al. (2011)
ATLAS	Mao et al. (2012)
Pan-STARRS	Rest et al. (2014); Scolnic et al. (2014); Kaiser et al. (2010)
C3R2	Masters et al. (2017)
SpARCS	Muzzin et al. (2012)
SNVETO	http://www.ast.cam.ac.uk/fo250/Research/SNveto/
FMOS-COSMOS	Silverman et al. (2015) and http://member.ipmu.jp/fmos-cosmos/FC_catalogs.html
SNLS AAOmega	Lidman et al. (2013); Yuan et al. (2015); Childress et al. (2017) and http://apm5.ast.cam.ac.uk/arc-bin/wdb/aat_database/observation_log/make
CDB	Sullivan et al. (2011)
VUDS	Tasca et al. (2017) and http://cesam.lam.fr/vuds/DR1/
ZFIRE	Nanayakkara et al. (2016) and http://zfire.swinburne.edu.au/data.html
MOSFIRE	http://mosdef.astro.berkeley.edu
2dFlenS	Blake et al. (2016) and http://2dfpens.swin.edu.au/http://2dfpens.swin.edu.au/
GLASS	Treu et al. (2015) and https://archive.stsci.edu/prepds/glass/
XMM-LSS	Stalin et al. (2010)

4.3, in both plots the behaviour of the mean absolute deviation and the σ_{68}^{Norm} can be considered a good approximation to the real value which changes depending on the training sample. The high errors that can be observed for z_{spec} close to zero are due to stars wrongly classified in the validation sample.

In addition, we studied the behaviour of the photo-zs estimation as a function of the redshift for the galaxies of Y3 Deep Field catalogue using the incomplete and semi-incomplete training sample defined in Sect. 5. In this case, as we lack information on the spectroscopic redshift, we have replaced z_{spec} by DNF_Z. The results of Fig. D.3 and D.4 show that $MAD(\Delta z)$ and σ_{68}^{Norm} get worse for higher redshift. Both training samples have similar results for $z < 1.4$. After this value, semi-incomplete training sample works better than incomplete training sample.

**Fig. C.1.** Scatter plot of spectroscopic redshift z_{spec} and the photo-z DNF_Z for complete training of 5,336 galaxies.

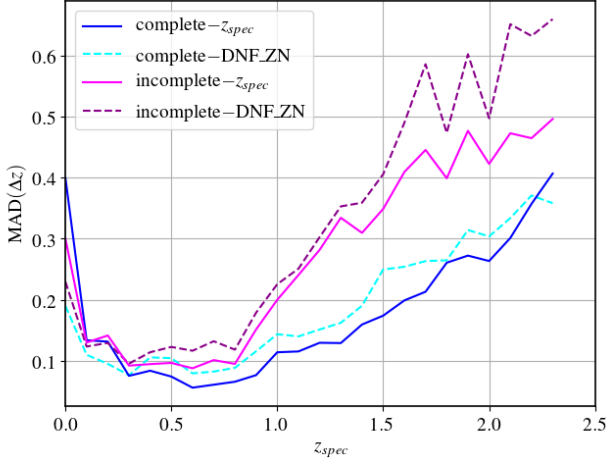


Fig. D.1. $MAD(\Delta z)$ as a function of the z_{spec} for the complete training (blue lines) and the incomplete training (magenta lines). The solid lines display the metrics calculated with z_{spec} and the dashed lines replacing z_{spec} by DNF_ZN.

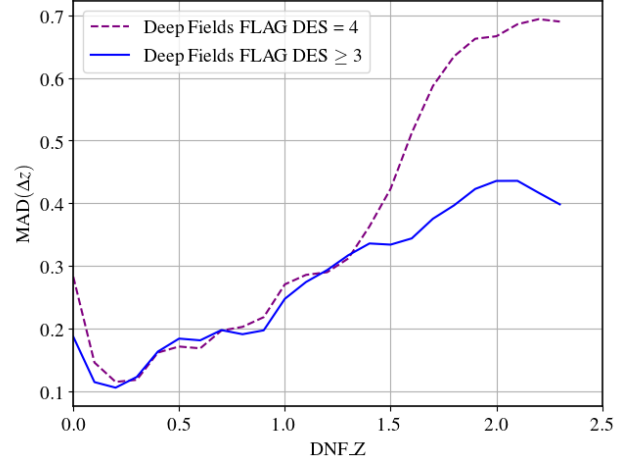


Fig. D.3. The precision of the photo-z estimates defined by the absolute median deviation as a function of the z_{spec} calculated by DNF_Z and DNF_ZN determined by training sample with only galaxies with FLAG_DES = 4 (incomplete training sample) and training sample with galaxies with FLAG_DES ≤ 3 (semi-complete training sample), in purple and blue, respectively.

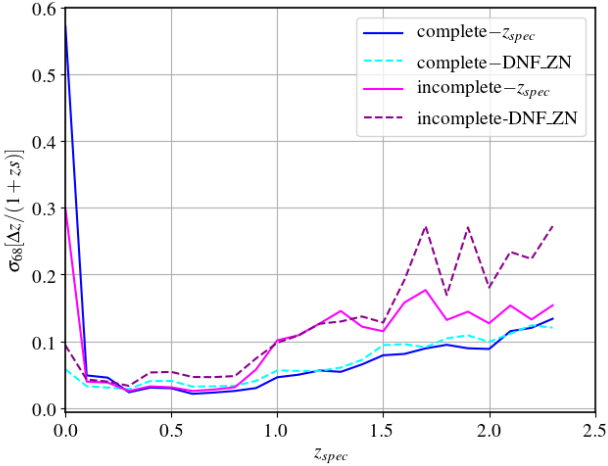


Fig. D.2. σ_{68}^{Norm} as a function of the z_{spec} for the complete training (blue lines) and the incomplete training (magenta lines). The solid lines display the metrics calculated with z_{spec} and the dashed lines replacing z_{spec} by DNF_ZN.

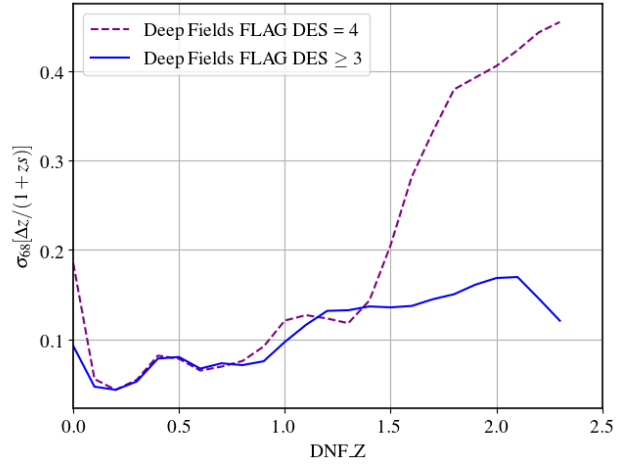


Fig. D.4. The precision of the photo-z estimates defined by the σ_{68}^{Norm} as a function of the z_{spec} calculated by DNF_Z and DNF_ZN determined by training sample with only galaxies with FLAG_DES = 4 (incomplete training sample) and training sample with galaxies with FLAG_DES ≤ 3 (semi-complete training sample), in purple and blue, respectively.