



HAL
open science

Photometric redshift estimation for Rubin observatory data preview 1 with redshift assessment infrastructure layers (RAIL)

T. Zhang, E. Charles, J F Crenshaw, S J Schmidt, P. Adari, J. Gschwend, S. Mau, B. Andrews, E. Aubourg, Y. Bains, et al.

► To cite this version:

T. Zhang, E. Charles, J F Crenshaw, S J Schmidt, P. Adari, et al.. Photometric redshift estimation for Rubin observatory data preview 1 with redshift assessment infrastructure layers (RAIL). 2025. <hal-05343786>

HAL Id: hal-05343786

<https://hal.science/hal-05343786v1>

Preprint submitted on 3 Nov 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Photometric Redshift Estimation for Rubin Observatory Data Preview 1 with Redshift Assessment Infrastructure Layers (RAIL)

T. Zhang^{1,†}, E. Charles^{2,3}, J.F. Crenshaw^{4,5}, S.J. Schmidt⁶, P. Adari^{2,7}, J. Gschwend⁸, S. Mau^{2,3}, B. Andrews¹, E. Aubourg⁹, Y. Bains¹⁰, K. Bechtol¹¹, A. Boucaud⁹, D. Boutigny¹², P. Burchat^{2,3}, J. Chevalier¹³, J. Chiang², H.-F. Chiang², D. Clowe¹⁴, J. Cohen-Tanugi¹⁵, C. Combet¹⁶, A. Connolly⁴, S. Dagoret-Campagne¹³, P.N. Daly¹⁷, F. Daruich¹⁸, G. Daubard¹⁹, J. De Vicente²⁰, H. Drass¹⁸, K. Fanning², E. Gawiser²¹, M. Graham^{4,5}, L.P. Guy²², Q. Hang²³, P. Ingraham¹⁸, O. Ilbert²⁴, M. Jarvis²⁵, M.J. Jee^{26,6}, T. Jenness¹⁸, A. Johnson², C. Juramy-Gilles¹⁹, S.M. Kahn²⁷, J.B. Kalmbach^{2,3,5}, Y. Kang^{18,2}, A. Kannawadi²⁸, L.S. Kelvin²⁹, S. Liang², O. Lynn³⁰, N.B. Lust²⁹, M. Lutfi¹⁸, A. Malz³⁰, R. Mandelbaum³⁰, S. Marshall², J. Meyers², M. Migliore⁹, M. Moniez¹³, J. Neveu¹⁹, J.A. Newman¹, E. Nourbakhsh²⁹, D. Oldag^{4,5}, H. Park²⁸, S. Pelesky³⁰, A.A. Plazas Malagón^{2,3}, B. Quint¹⁸, M. Rahman³¹, A. Rasmussen², K. Reil², W. Roby³², A. Roodman², C. Roucelle⁹, M. Salvato³³, B. Sánchez³⁴, D. Sanmartin¹⁸, R.H. Schindler^{3,2}, J. Scora³¹, J. Sebag¹⁸, N. Sedaghat⁴, I. Sevilla-Noarbe²⁰, R. Shirley³³, A. Shugart¹⁸, R. Solomon¹², D. Taranu²⁹, G. Thayer², L. Toribio San Cipriano²⁰, E. Urbach³⁵, Y. Utsumi², W. van Reeve¹⁸, A. von der Linden⁷, C.W. Walter²⁸, W.M. Wood-Vasey¹, J. Zuntz³⁶, LSST Dark Energy Science Collaboration

(Affiliations can be found after the references)

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present the first systematic analysis of photometric redshifts (photo- z) estimated from the Rubin Observatory Data Preview 1 (DP1) data taken with the Legacy Survey of Space and Time (LSST) Commissioning Camera. Employing the Redshift Assessment Infrastructure Layers (RAIL) framework, we apply eight photo- z algorithms to the DP1 photometry, using deep *ugrizy* coverage in the Extended Chandra Deep Field South (ECDFS) field and *griz* data in the Rubin_SV_38_7 field. In the ECDFS field, we construct a reference catalog from spectroscopic redshift (spec- z), grism redshift (grism- z), and multiband photo- z for training and validating photo- z . Performance metrics of the photo- z are evaluated using spec- z s from ECDFS and Dark Energy Spectroscopic Instrument Data Release 1 samples. Across the algorithms, we achieve per-galaxy photo- z scatter of $\sigma_{\text{NMAD}} \sim 0.03$ and outlier fractions around 10% in the 6-band data, with performance degrading at faint magnitudes and $z > 1.2$. The overall bias and scatter of our machine-learning based photo- z s satisfy the LSST Y1 requirement. We also use our photo- z to infer the ensemble redshift distribution $n(z)$. We study the photo- z improvement by including near-infrared photometry from the Euclid mission, and find that Euclid photometry improves photo- z at $z > 1.2$. Our results validate the RAIL pipeline for Rubin photo- z production and demonstrate promising initial performance.

Key words: galaxies: distances and redshifts – methods: statistical

1 INTRODUCTION

The NSF–DOE Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezic et al. 2019) is a 10-year survey that will repeatedly image the southern sky in six optical bands (*ugrizy*). During its operation LSST will produce an unprecedentedly large dataset of billions of galaxies, providing a statistical foundation to probe the nature of dark energy through multiple cosmological probes, including weak gravitational lensing (Kilbinger 2015), large-scale structure

(Van Waerbeke et al. 2000), and supernovae (Scolnic et al. 2018; Abbott et al. 2019).

Rubin Observatory’s Data Preview 1 (DP1; NSF-DOE Vera C. Rubin Observatory 2025c) provides real observations from the LSST Commissioning Camera (LSSTComCam; SLAC National Accelerator Laboratory & NSF-DOE Vera C. Rubin Observatory 2024) mounted on the 8.4-meter Simonyi Survey Telescope, with multiband imaging across selected fields. These datasets (NSF-DOE Vera C. Rubin Observatory 2025a) are the first real data released by the Rubin Observatory and are a testbed to validate and understand the data processing pipelines.

† tq.zhang@pitt.edu

Photometric redshifts (photo- z) with well-controlled systematic biases are essential to measuring and modeling the aforementioned probes, and therefore are critical to achieving the science goals of the LSST Dark Energy Science Collaboration (DESC; [Graham et al. 2022](#); [Newman & Gruen 2022](#)). Precise and accurate photo- z estimates for the galaxies are required to select the galaxies for tracing the large-scale structure, dividing lens and source galaxies into tomographic bins. Photo- z is also used to infer the redshift distribution of galaxy ensembles and calibrate standard candles in the absence of spectroscopic coverage.

In this paper, we present the photo- z of galaxies in Rubin Observatory’s DP1 dataset, produced using the Redshift Assessment Infrastructure Layers (RAIL; [The RAIL Team et al. 2025](#)). We use a compilation of spectroscopic, grism and deep multiband photo- z galaxies to construct the reference sample; we use the reference sample to train and test machine-learning photo- z algorithms, and calibrate the template-fitting algorithms in RAIL; we evaluate the the photo- z algorithms performance on the test set and validation set, made by cross-matching DP1 galaxies to the Dark Energy Spectroscopic Instrument (DESI) Data Release 1 DR1 galaxies ([Collaboration et al. 2025](#)); we study the improvement by including near-infrared (NIR) photometry from the Euclid mission ([Euclid Collaboration et al. 2025](#)); we produce per-galaxy photo- z for DP1 galaxies that pass certain quality flags. This work is an extension of the initial study of DP1 photo- z by [Charles et al. \(2025\)](#). The photo- z in this work are essential to other DESC early science projects, e.g., a cluster lensing analysis of the Abell 360 cluster (The LSST Dark Energy Science Collaboration et al. *in prep.*).

The paper is organized as follows: Section 2 describes the DP1 datasets and reference redshift samples; Section 3 describes the photometric redshift estimation methods and our bookkeeping software; Section 4 presents the photo- z performance results; and Section 5 summarizes our findings and outlines future studies.

2 DATA

2.1 Rubin DP1

The Rubin Data Preview 1 (DP1) includes $\sim 15\text{deg}^2$ of multiband optical imaging $ugrizy$ across six selected fields ([NSF-DOE Vera C. Rubin Observatory 2025c](#)), including the Extended Chandra Deep Field South (ECDFS), Euclid Deep Field South (EDFS), a low galactic latitude field Rubin_SV_95_-25 (SV_95_-25) fields, and a low ecliptic latitude field Rubin_SV_38_7 (SV_38_7) field with $griz$ photometry. The photometric catalogs consist magnitudes using a variety of flux measurement methods, including the 1'' Gaussian Aperture (Gaap1p0) fluxes, the 3'' Gaussian Aperture (Gaap3p0) fluxes, the CModel fluxes, the Kron fluxes ([Kron 1980](#)), the PSF aperture fluxes, and the Sérsic aperture fluxes from the DP1 coadd object catalog ([NSF-DOE Vera C. Rubin Observatory 2025b](#)). We estimate photo- z in these four fields with the Gaap1p0 photometry. The stacked PDF of our photo- z is shown in Appendix A.

To ensure the quality photometry necessary for photometric redshift estimation, we include only galaxies with i -band PSF flux signal-to-noise (SNR) ratio exceeding 5. Furthermore, we selected extended sources by requiring `g_extendedness` $>$ 0.5 and `r_extendedness` $>$ 0.5, which selects against point-source objects. The 5σ limiting magnitudes of Gaap1p0 of the ECDFS field are [26.4, 27.8, 27.1, 26.7, 25.8, 24.6].

In the ECDFS, EDFs, and SV_95_-25 fields, this selection yielded galaxies with six-band photometry, resulting in a sample of approximately 375,000 galaxies. In the SV_38_7 field, where observations

are limited to four bands ($griz$), the same selection produced a sample of about 169,000 galaxies. For reference, the full DP1 object catalog contains roughly 2 million objects before applying these selection criteria.

Milky Way dust extinction reduces the flux disproportionately more at bluer wavelengths, and needs to be corrected (dereddened). We use the extinction maps of [Schlafly & Finkbeiner \(2011\)](#) to get $E(B - V)(\alpha, \delta)$ values, and use attenuation coefficient $a_{b,V} = [4.81, 3.64, 2.70, 2.06, 1.58, 1.31]$ in the u, g, r, i, z, y bands. The dereddened magnitudes are

$$\text{mag}_{b,\text{deredden}} = \text{mag}_b - a_{b,V} E(B - V)(\alpha, \delta), \quad (1)$$

where $\text{mag}(\text{mag}_{\text{deredden}})$ are the observed and dereddened magnitudes in the corresponding band b . α, δ represent the right ascension and declination.

2.2 Reference Sample

To construct the photometric redshift training and testing sets, we assembled a reference sample in the Extended Chandra Deep Field South (ECDFS) by collecting galaxies with spec- z s, grism- z s, and high-quality multiband photo- z from multiple surveys listed in Table 1.

Confidence, which takes values between 0.0 and 1.0, is loosely defined as the probability that an individual redshift estimate is correct. Most of the spectroscopic sets provide these estimates for their redshifts. For the few that do not, we assigned the confidence as 0.95. For the grism and multiband photo- z surveys, we set the confidence equal to $1 - f_{\text{out}}$, where f_{out} is the reported outlier rate of these catalogs. To facilitate custom quality cuts, the catalog contains flags indicating whether each redshift originates from spectroscopy, grism, or multiband photo- z , as well as confidence values.

The component redshift catalogs were combined into a single reference catalog. When combining the component redshift catalogs, sources within 0.75'' were identified as duplicates. For these sources, only the highest quality redshift is kept, i.e., spec- z s are preferred over grism- z s, which are preferred over photo- z , and higher confidence values are preferred for redshifts of the same type.

Note that redshifts from grism and photo- z surveys have larger scatter and bias than spectroscopic surveys; however, these characteristics are not captured by the confidence parameter. We also note the redshift quality in some grism surveys, e.g., PRIMUS, is similar to multi-band photo- z due to limited wavelength resolution ([Cool et al. 2013](#)). We encourage users of our catalog to investigate the details of each component survey that comprises our reference catalog and apply their own quality cuts as suit their needs.

2.2.1 Training and Testing Sets

We cross-matched the aforementioned reference galaxies to the ECDFS DP1 object catalog using a 0.75'' radius. To ensure sufficiently high quality photometry, we define the training set as all cross-matched objects with detections in all observing bands and with an i -band CModel flux SNR ratio exceeding 20. We note that galaxy model fluxes have underestimated errors, so the true SNR of the cut is likely lower than 20. We also note that requiring detections in all 6 bands will likely eliminate all “dropout” galaxies in the catalog, e.g., Lyman-break galaxies ([Giavalisco 2002](#)).

We also require the “confidence” to be greater than 0.9. We then split the selected sample, with 70% used for training and the remaining 30% reserved for a test set. These selections result in sets of 6778

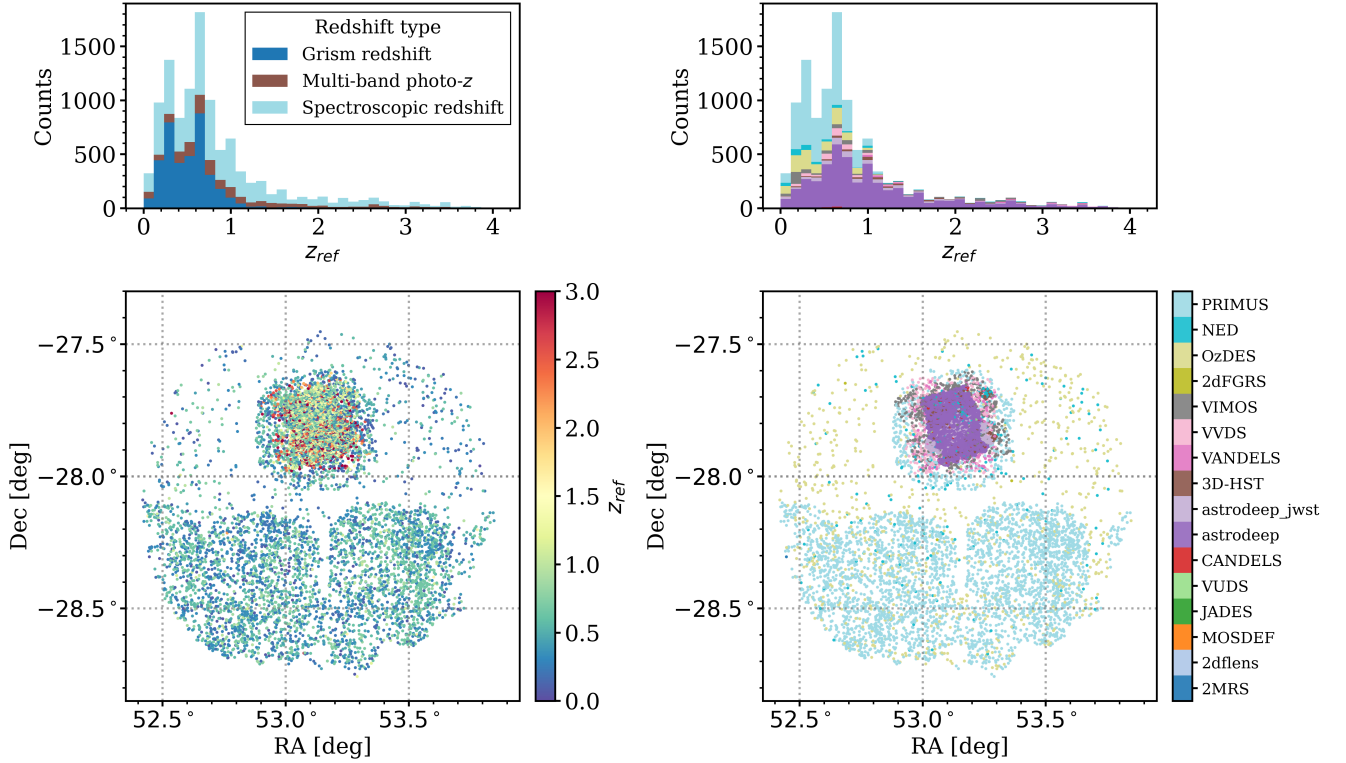


Figure 1. Training set in the ECDFS field. Top left: redshift distribution by methods to obtain redshift. Top right: redshift distribution by surveys. Bottom left: Scatter plot of the ECDFS reference catalog color-coded by redshift. Bottom right: scatter plot of the ECDFS reference catalog color-coded by the survey name. The 1-D distributions are normalized. Our reference galaxy catalog is constructed by a wide range of redshifts, and spans a wide range of redshifts.

training galaxies and 2905 test galaxies. The average confidence of the selected sample is 0.97, indicating approximately $\sim 3\%$ redshift outliers in the training set. Note that we tolerate this relatively high level of redshift outliers due to the limited sample size for the training galaxies.

In Fig. 2, we show the color-magnitude-redshift space of the training galaxies, and compare that to the color-magnitude space of all ECDFS galaxies with i -band $\text{SNR} > 5$. With the combination of spec- z , grism- z and multiband photo- z , the training set covers most color space of the ECDFS sample. Noticeably, the magnitude of the training galaxies only goes down to $i_{\text{mag}} < 24.6$, compared to $i_{\text{mag}} < 26.7$ for the object catalog.

We note that we are training our photo- z algorithms on training galaxies with higher significance than the object catalog. Only high-significance photometry is taken to reduce noise in the training process. Future study is required to determine the optimal selection criteria for maximizing performance on a faint galaxy catalog. Fig. 1 provides an overview of the ECDFS training and testing sets.

We note that the PRIMUS survey is dominating the low- z range, and the astrodeep sample is dominating the high- z range. Potentially, the selection criteria of these specific surveys might be imprinted onto our photo- z .

2.2.2 DESI DR1

For an independent validation of the four-band photometric redshift estimates, we cross-matched the SV_38_7 DP1 object catalog with galaxies from the DESI Bright Galaxy Sample (BGS), Lumi-

nous Red Galaxy (LRG), and Emission Line Galaxy (ELG) samples from the DESI DR1 spectroscopic catalog (DESI Collaboration et al. 2025). This cross-match produced 2,728 matched objects across $z = 0$ to 1.6. We used these galaxies as a validation set to assess the performance of the four-band photo- z estimates in a field with shallower depth and fewer filters. The i -band magnitudes and matched spec- z s of the galaxies are shown in Fig. 3.

Matching was performed purely using spatial coordinates. If a match was found in multiple subcatalogs, the one with the highest weight was selected. The BGS sample, which provides redshifts for $z \leq 0.5$, contains no sources with declination $\delta \geq 7^\circ$, and covers only half of the SV_38_7 field, limiting our ability to assess the low-redshift photo- z . Furthermore, in the matched catalog, the maximum 5σ i -band magnitude depth is 23.8, with a magnitude distribution peak at $i = 22.9$ —both approximately 1.5 magnitudes shallower than the photometry for the full field. Despite these limitations, this sample provides an independent spectroscopic validation set and a test to apply the model trained on deep-field training catalogs to a wide field photometry.

2.2.3 Euclid Crossmatch

To obtain joint optical and NIR photometry for galaxies in the ECDFS field, we positionally cross-matched the DP1 catalog with the Euclid Quick Data Release (Q1) (Euclid Collaboration et al. 2025) catalog. The Euclid catalog was filtered to retain only sources with a SNR ratio (SNR) greater than 5 in the VIS band and with no quality flags raised. For each matched source, we retained Euclid coordinates (α ,

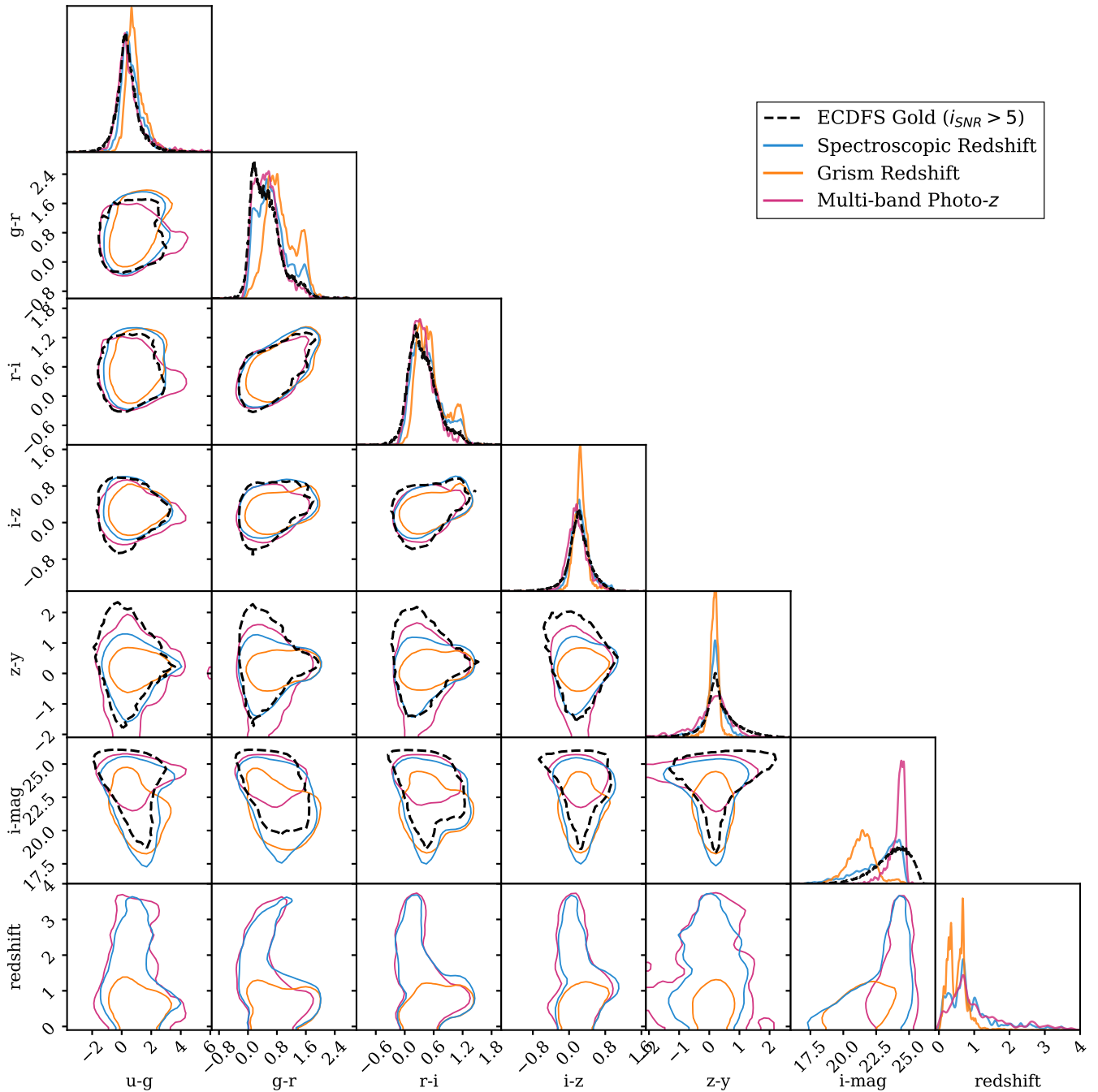


Figure 2. Corner plot of the training set color-magnitude-redshift distributions in the ECDFS field. Each panel shows 95% confidence contours among redshift, adjacent-band colors ($u - g$, $g - r$, $r - i$, $i - z$, $z - y$), and the i -band magnitude. The blue contours represent galaxies with spec-zs, the green contours represent grism-zs, the red contours represent the multiband photo-z, while the black contours show all galaxies in the ECDFS field with i -band SNR greater than 5. The spec-z, grism-z and photo-z samples compose 53%, 34%, and 13% of the entire training/testing galaxies, respectively. Our reference sample covers a reasonable color space of our ECDFS samples; however, in some color spaces, we still lack reference galaxies.

δ), the point-spread function (PSF) magnitudes in the VIS band as well as uniform aperture magnitudes in the Y, J, and H NIR bands. All Euclid fluxes were converted to magnitudes using a zero point magnitude of 23.9. Sources with flagged measurements in any of the Y, J, or H bands ($\text{flag_band} = 1$) were excluded from those bands by setting their magnitudes and uncertainties to NaN.

The 5σ limit magnitudes of the Y, J, H bands from Euclid Q1 are around 24.0, and the VIS band magnitude limit is 26. The Euclid infrared magnitude limit is about 2 mag shallower than the $ugriz$ band of ECDFS. However, since we are applying an SNR cut above 20 to the i -band for the training and testing galaxies, most training and testing galaxies have a detection in the Euclid YJH bands.

To study the benefit of including NIR photometry from Euclid, we

Table 1. Component surveys (names, types, confidence, incremental number of objects (total cross-match without duplicates and confidence cuts), and reference) of the redshift reference sample. Redshift type: s = spec-z, g = grism-z, p = multiband photo-z.

* Note: multiband photo-z redshifts and grism and spec-zs with confidence < 0.90 were not used in any training sets employed in this note.

**Original datasets in this catalog are Momcheva et al. (2016a); Cooper et al. (2012); Pentericci et al. (2018b); Wolf et al. (2004); Grazian et al. (2006); Straughn et al. (2009, 2008); Norman et al. (2002); Wuyts et al. (2008); Croom et al. (2001); van der Wel et al. (2004); Daddi et al. (2004); Kriek et al. (2007); Pasquali et al. (2006); Malhotra et al. (2005); Hathi et al. (2008); Rhoads et al. (2009); Mignoli et al. (2005); Doherty et al. (2005); Le Fèvre et al. (2004); Kriek et al. (2015a); Morris et al. (2015); Popesso et al. (2009); Ravikumar et al. (2007); Szokoly et al. (2004); Tomida et al. (2013); Trump et al. (2015); Pentericci et al. (2018a); McLure et al. (2018); Balestra et al. (2010a); Vanzella et al. (2009); Le Fèvre et al. (2015, 2013b); Vanzella et al. (2008); Wuyts et al. (2009); Silverman et al. (2010)

Survey	Type	Conf	#	Ref.
2dFGRS	s	1.00	3 (8)	Colless et al. (2001)
		0.99	4	
		0.90	1	
2df lens	s	1.00	1	Blake et al. (2016)
2MRS	s	0.95	1 (6)	Huchra et al. (2012)
6dFGRS	s	0.98	2	Jones et al. (2009)
3D-HST	g	0.99	5 (3803)	Momcheva et al. (2016b)
		0.95	277	
ASTRODEEP	s	1.00	4165 (13861)	Merlin et al. (2021) **
		p*	0.97	
ASTRODEEP-JWST	s	1.00	594 (6303)	Merlin et al. (2024)
		p*	0.92	
CANDELS	s	1.00	53 (13447)	Kodra et al. (2023)
		p*	0.93	
JADES	s	0.99	11 (318)	D'Eugenio et al. (2025)
		0.95	34	
		0.90	24	
MOSDEF	s	0.99	9 (45)	Kriek et al. (2015b)
NED	s	0.95	847 (2956)	Helou et al. (1991)
OzDES	s	0.99	897 (910)	Lidman et al. (2020)
PRIMUS	g*	0.92	3653 (6263)	Cool et al. (2013)
		0.85	1687	
VANDELS	s	1.00	196 (414)	Garilli et al. (2021)
VIMOS	s	1.00	499 (1343)	Balestra et al. (2010b)
		0.95	43	
VUDS	s	1.00	9 (150)	Tasca et al. (2017)
		0.95	9	
		0.80	3	
VVDS	s	1.00	101 (656)	Le Fèvre et al. (2013a)
		0.95	193	
Totals	s		7699	
	g		5622	
	p		9301	
	all		22622	

created a training and testing set with the Euclid Y, J, and H photometry in addition to the DPI u, g, r, i, z, y photometry. We dropped all non-detections in Y, J, or H bands in addition to the selection made in Section 2.2.1. This results in 5010 training set galaxies and 2158 testing set galaxies, a $\sim 26\%$ decrease compared to the count of previously selected DPI training and testing galaxy counts.

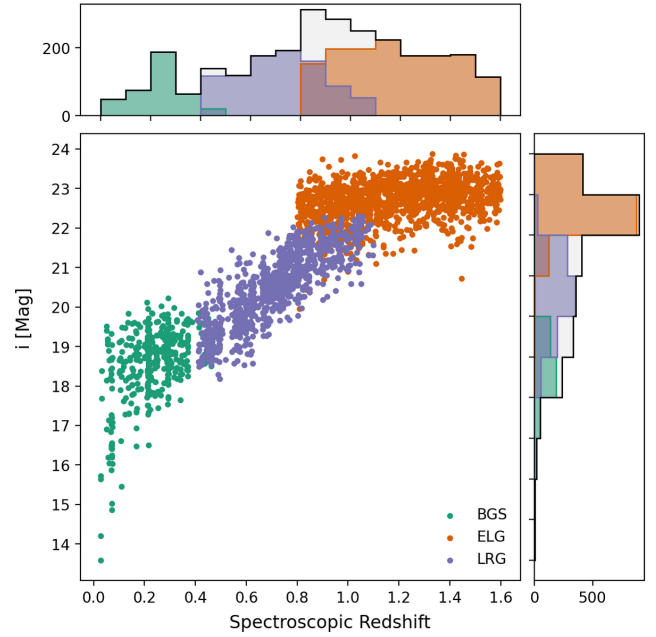


Figure 3. Redshift and i -magnitude distribution for matched DESI objects. For the scatter plot and histograms, the BGS sample is shown in green, ELG in orange, and LRG in purple. The distribution of the total sample is shown in black in the outer histograms. The i -mag distribution peaks at $i = 23$ and the redshift distribution at $z = 0.8$.

3 METHODOLOGY

In this section, we briefly describe the eight photo- z algorithms used in this work. We refer readers to The RAIL Team et al. (2025) and <https://rail-hub.readthedocs.io/> for details about RAIL and these algorithms.

3.1 Template-Fitting Photo- z Algorithms

Template-fitting photo- z algorithms estimate redshifts by comparing observed galaxy photometry to a library of spectral energy distribution (SED) templates spanning a range of galaxy types and redshifts. In this study, we deploy two widely used template-fitting codes: BPZ (Benítez 2000; Coe et al. 2006) and LePhare (Arnouts et al. 1999; Ilbert et al. 2009; Ilbert et al. 2024). Both codes compute synthetic fluxes by redshifting the templates and convolving them with the Rubin filter responses to produce model fluxes, then calculating χ^2 values by comparing these model fluxes to the observed photometry and uncertainties for each template at each position on a sample redshift grid. These χ^2 values are converted to likelihoods, and Bayesian priors are applied to incorporate expected redshift distributions as a function of magnitude and/or galaxy type. A one-dimensional redshift posterior probability is computed by marginalizing (summing) over the template SEDs to produce the final PDF. The best-fit redshift is identified as the mode of the posterior probability, and the best-fit SED is defined as the SED with the maximum contribution to the posterior at that single fixed redshift, and is not valid at other redshifts, and thus, not a precise representation of the galaxy-type vs. redshift degeneracy.

Both codes use a base SED set described in Ilbert et al. (2009) (also included in the base LePhare distribution), consisting of Elliptical and Spiral templates from Polletta et al. (2007) as well as bluer

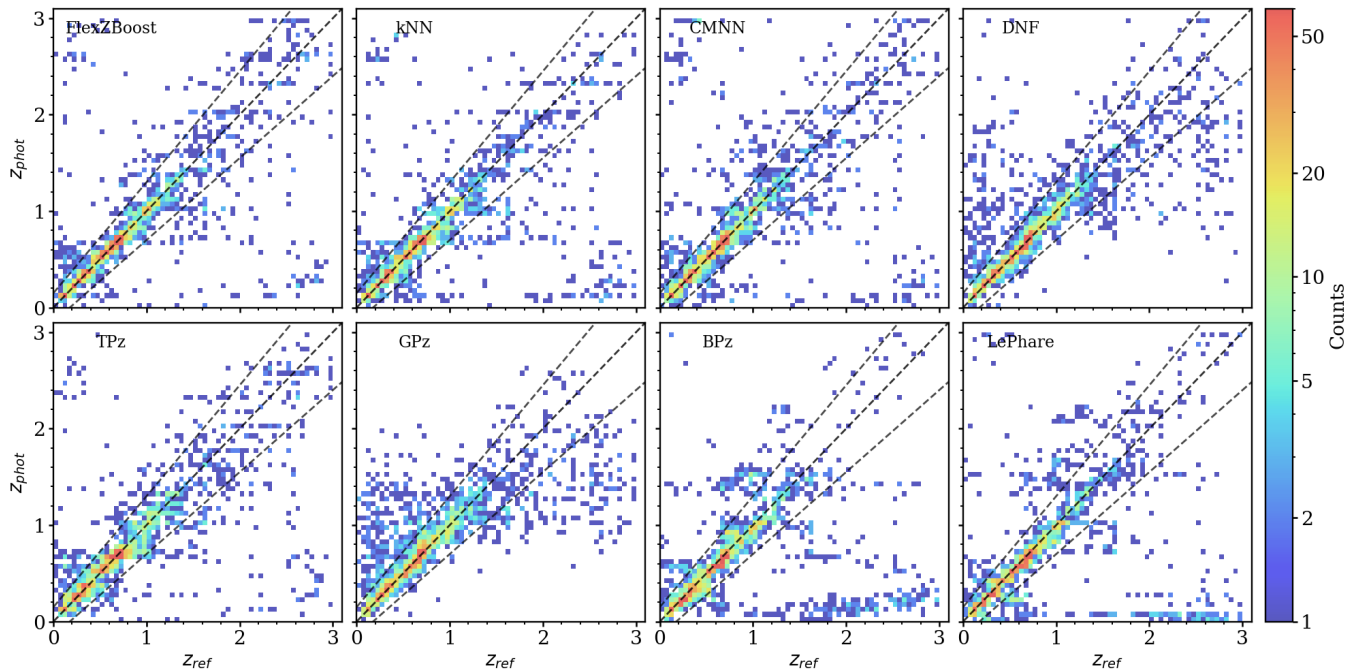


Figure 4. Two-dimensional histograms comparing the mode of the photometric redshifts (z_{phot}) to reference redshifts (z_{ref}) for eight photo- z algorithms applied to the DP1 test sample. Each panel shows one algorithm: FlexZBoost, kNN, CMNN, DNF, TPz, GPz, BPz, and LePhare. The grey dashed lines indicate the identity line ($z_{\text{phot}} = z_{\text{ref}}$) and the $|\Delta z| = 0.15$ lines. The color scale represents the number of objects in each bin on a logarithmic scale. These plots illustrate the overall agreement and outlier behavior of each algorithm across the redshift range $0 < z < 3$.

starburst SEDs generated from Bruzual & Charlot (2003) models. These base templates contain no internal dust extinction. The LePhare code is designed to add a grid of $E(B-V)$ values to each template with values 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5. BPz lacks this capability, and so new templates are created explicitly by adding extinction from Noll et al. (2009) (Calzetti-like dust with a UV bump) to each SED bluer than the Sb type (i. e. no dust is added to the Elliptical, S0, and Sa models, but dust is added to all other SEDs). $E(B-V)$ extinctions with values 0.1, 0.2, 0.3, 0.4, 0.5 are used to generate the final set of 136 SEDs for BPz.

As template-fitting methods can compute model fluxes at any redshift (as long as the SED model spans the wavelengths of the filter curves at that redshift), they can extrapolate beyond the limitations of sparse or missing redshifts in reference training sets, and thus have the potential to be more robust at higher redshifts. However, the redshift predictions are limited by the agreement between the assumed SED models and the real Universe: any mismatch in galaxy evolution as a function of lookback time, luminosity, and other variables will lead to biases in the redshift predictions. So, careful construction of template sets and calibration using deep multiwavelength samples will be critical to the performance of template-based methods.

3.2 Machine-Learning Photo- z Algorithms

Machine-learning photo- z algorithms predict photometric redshifts by learning empirical mappings from multiband photometry to the reference redshift, modeling the complex color–redshift relationships. They can be less prone to systematic bias in the magnitude measurement because they are learning the magnitude-redshift mapping empirically. However, machine-learning methods tend to predict redshift based on training set information on the target set by acting

as implicit priors on the color–redshift relation, and can generate biased results if the training galaxies are not representative of the full galaxy catalog.

The methods employed in this work are:

- TPz (Trees for photo- z) uses random forests to perform regression of redshift as a function of multiband photometry (Carrasco Kind & Brunner 2013)¹.
- FlexZBoost uses boosted decision trees to model the non-linear mapping from colors to redshift (Izbicki & Lee 2017)².
- kNN (K-Nearest Neighbors) predicts the redshift of a galaxy based on the average redshift of its closest neighbors in color-magnitude space.³
- CMNN (Color-Matched Nearest Neighbors) improves upon basic kNN by weighting neighbors according to Mahalanobis distances in color space (Graham et al. 2018)⁴.
- GPz (Gaussian Processes for photo- z) models the redshift-color relation with Gaussian processes (Almosallam et al. 2016)⁵.
- DNF (Directional Neighborhood Fitting) fits local linear models around each galaxy using its nearest neighbors (De Vicente et al. 2016)⁶.

¹ http://github.com/LSSTDESC/rail_tpz

² http://github.com/LSSTDESC/rail_flexzboost

³ http://github.com/LSSTDESC/rail_sklearn

⁴ http://github.com/LSSTDESC/rail_cmnn

⁵ http://github.com/LSSTDESC/rail_gpz_v1

⁶ http://github.com/LSSTDESC/rail_dnf

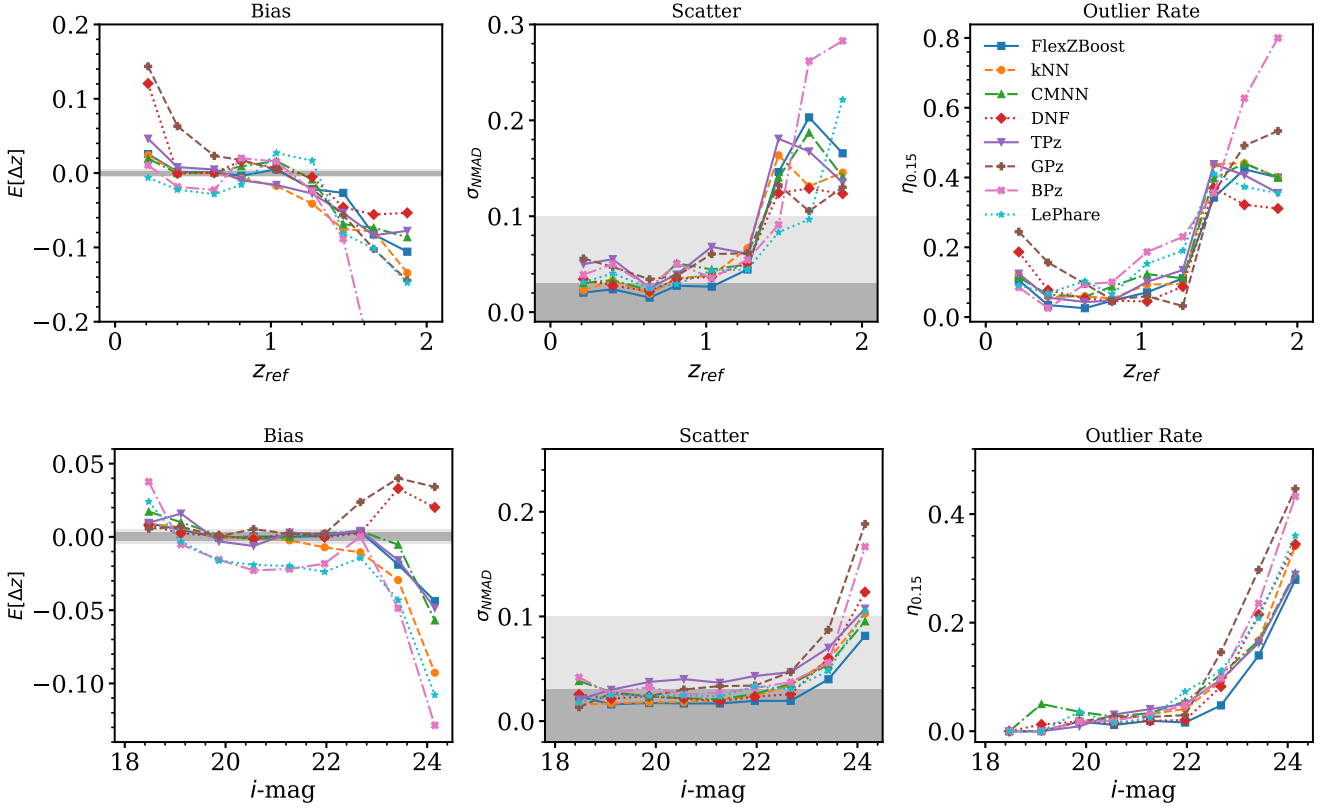


Figure 5. Photo- z performance metrics as a function of redshift (top row) and i -band magnitude (bottom row) for eight algorithms. Each panel shows three metrics: the photo- z bias $\mathbb{E}[\Delta z]$, scatter σ_{NMAD} , and outlier rate $\eta_{0.15}$. The grey shaded regions show the LSST Y1 and Y10 requirements on the mean and scatter of photo- z (The LSST Dark Energy Science Collaboration et al. 2018). The photo- z of all algorithms starts to deteriorate when $z > 1.2$ and $i\text{-mag} > 23$. We notice that some algorithms display significant bias at low and high redshift, and template-fitting methods have more bias across the magnitude range than empirical methods. We do not show statistics beyond $z_{\text{ref}} = 2$ because of the limited number of reference galaxies.

3.3 Workflow Management and Bookkeeping

We developed a workflow management and bookkeeping software system, `RAIL_projects`⁷ for RAIL pipelines. The workflow manager produces configuration files for `ceci`⁸ pipelines based on “flavors” that specify the input datasets, algorithms and corresponding parameters, and selection criteria. `RAIL_projects` enables efficient comparisons between algorithms, parameters, and selection criteria, which can be highly useful in future LSST analyses and photo- z analyses in other surveys. The configuration of the pipelines are stored in `RAIL_project_config`⁹.

In this work, there are three flavors of photo- z models being trained:

(i) **dp1_6band**: models trained on the *ugrizy* Gaap1p0 magnitude of the training and testing galaxy sets in ECDFS. These models are then used to produce photo- z of the objects in the ECDFS, EDFS and SV_95_-25 fields. This configuration can be found in `dp1/dp1_v4.yaml`.

(ii) **dp1_4band**: models trained on the *griz* Gaap1p0 magnitude

of the training and testing galaxies in ECDFS. These models are used to produce photo- z of the object in the SV_38_7 field.

(iii) **dp1_paper_euclid_nir**: only FlexZBoost and BPz are applied in this flavor, using the DP1 *ugrizy* and Euclid *YJH* magnitude. The model is only applied to the test galaxies in ECDFS.

3.4 Metrics

In this study, we use simple metrics based on the mode of the photo- z and the reference redshift to evaluate the performance of the photo- z algorithms. We compute the photo- z deviation by

$$\Delta z = \frac{z_{\text{mode}} - z_{\text{ref}}}{1 + z_{\text{ref}}}. \quad (2)$$

Here z_{mode} is the redshift mode, i.e., the redshift that corresponds to the maximum of the PDF. z_{ref} is the reference redshift, sometimes referred to as the “true redshift”.

For a given set of galaxies, we compute the bias as the mean of Δz , the scatter as the bias-corrected normalized median absolute deviation (NMAD) σ_{NMAD} of Δz ,

$$\sigma_{\text{NMAD}} = 1.48 \times \text{median}(|\Delta z - \text{median}(\Delta z)|), \quad (3)$$

and the catastrophic outlier rate $\eta_{0.15}$ by the fraction of galaxies with $\Delta z > 0.15$.

Photo- z are often used to assign both lens and source galaxies to

⁷ https://github.com/LSSTDESC/rail_projects

⁸ <https://github.com/LSSTDESC/ceci>

⁹ https://github.com/LSSTDESC/rail_project_config

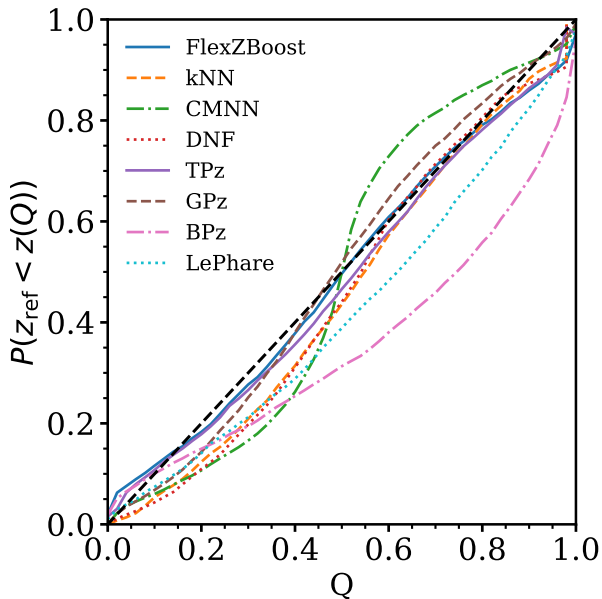


Figure 6. The PIT-QQ plot for eight photo- z estimators. The curves show the empirical CDF of $\{z_i\}$ as a function of Q , i.e., $P(z_{\text{ref}} < z(Q))$, where $z(Q)$ is the Q -th posterior quantile. The black dashed line denotes perfect calibration (uniform PIT).

tomographic bins in joint analyses of galaxy clustering and weak lensing (e.g., Abbott et al. 2022). Therefore, we compute a “binning accuracy” as a metric for binning performance, defined as the probability that the bin assigned based on the photo- z mode is the same as that assigned using z_{ref} . The binning strategy we use is to split all galaxies into five bins with equal number counts. We note that the binning accuracy should be taken as a qualitative metric to assess the accuracy of binning for the algorithm, rather than as a metric for assessing the impact on cosmology analysis.

We use the Probability Integral Transform quantile–quantile (PIT-QQ) plot as a diagnostic tool to evaluate the coverage of the photo- z PDFs. By comparing the empirical cumulative density function (CDF) at quantile Q , i.e., $P(z_{\text{ref}} < z(Q))$ against the diagonal line, the plot shows whether the predicted PDFs are well calibrated. Deviations from the diagonal line indicate systematic issues such as overconfident, underconfident error bars, biased, or skewed uncertainties.

4 RESULTS

4.1 Photo- z Optimization

In this section, we briefly describe the effort towards achieving an optimized photo- z result in this work. We note that we do not consider the results here to be the fully optimized photo- z performance for future Rubin data.

As an initial data exploration, we train and test all eight photo- z algorithms with default settings using different types of photometry. We observe that the Gaap1p0, the CModel, and the Sersic flux are generally better performing than other types of flux measurements in terms of scatter and outlier rate metrics, and they are comparable to each other within minor fluctuations between algorithms. As a

result, we choose Gaap1p0 as flux measure employed for the rest of the work.

We adjust the values of the hyperparameters of the machine-learning methods and adjust the template collections for the template-fitting algorithms to optimize the performance of the photo- z methods. Here are the notable hyperparameter changes compared to the default parameters in the RAIL v1.2¹⁰:

- Common settings: PDF and point estimates are calculated between $z = 0$ and 3.
- TPz: number of random bootstrap samples is set to 10 (default is 8), number of trees in the random forest is set to 10 (default is 5), the minimum number in terminal leaf is set to 2 (default is 5).
- kNN: maximum number of neighbors is set to 10 (default is 7).
- GPz: use the variable diagonal covariance mode. The number of training iterations is set to 1000 (default is 200, and does not converge).
- FlexZBoost: maximum number of basis functions is set to 50 (default is 35).
- BPz: use the base templates from LePhare with a set of dust absorption models applied, resulting in a larger range of dusty model SEDs than standard BPZ runs (See Section 3.1). Use a Hubble Deep Field North (HDFN) like prior from Benítez (2000), with no zero-point adjustments.
- LePhare: enable magnitude zero point adjustment (default is off). Remove a maximum of two bands when χ^2 of the fit exceeds 300 (default is 500).

4.2 Per-galaxy Photo- z

We evaluate the per-galaxy photo- z performance of the eight algorithms on our test set, which has the same population as the training set. Fig. 4 shows that most algorithms achieve good agreement with the spectroscopic redshifts for bright, low-to-moderate redshift galaxies, as seen by the concentration of points along with the identity line. However, systematic biases and increased scatter become apparent at higher redshifts ($z \gtrsim 1.2$), where training data are sparse and photometric uncertainties are high.

The patterns of outlier distribution vary across algorithms. Particularly, for the template-fitting algorithms, we notice that there are galaxies with $1.5 < z_{\text{ref}} < 3$ that are given $z_{\text{phot}} < 0.5$. We suspect this can be caused if the Lyman break of the galaxy are misidentified as a Balmer break by the algorithm.

We summarize the overall performance metrics in Table 2. Overall, the machine-learning based algorithms have bias less than 0.005, which is LSST Y1 requirement (The LSST Dark Energy Science Collaboration et al. 2018). Some algorithms, such as FlexZBoost and DNF, and kNN exhibit tighter scatter, while others, like GPz, show larger overall dispersion. All algorithms has scatter less than 0.1, the LSST Y1 requirement. We also report the catastrophic outlier rates defined as the fraction of galaxies with $|\Delta z| > 0.15$. In terms of binning accuracy, FlexZBoost scores the highest accuracy of 78.1%, while other methods are in the range of 60% to 70%.

In Fig. 5, we show the bias, scatter, and outlier rate of the algorithms as functions of redshift and i -band magnitude. Generally, we observe that the photo- z performance is better at intermediate redshift ($0.3 < z < 1.2$), and brighter magnitude ($i\text{-mag} < 22.5$). We also observe a wide range of performance across different methods

¹⁰ The full set of hyperparameter change can be found in the “dp1_paper” flavor in https://github.com/LSSTDESC/rail_project_config/blob/main/dp1/dp1.yaml

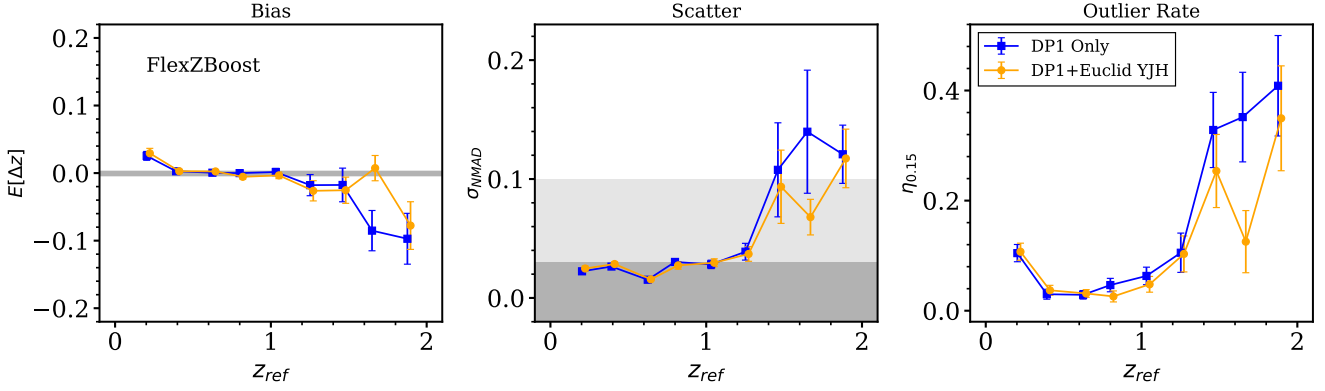


Figure 7. The same photometric redshift performance metrics as Fig. 5 of the FlexZBoost algorithm as a function of redshift, comparing LSST ComCam-only photometry (blue) to the combination of LSST ComCam and Euclid NIR YJH bands (orange). The inclusion of Euclid NIR data notably reduces bias, scatter, and outlier rates at redshift $z > 1.5$. The error bars are computed by bootstrapping the testing set. The grey shaded regions show the LSST Y1 and Y10 requirements on the mean and scatter of photo- z .

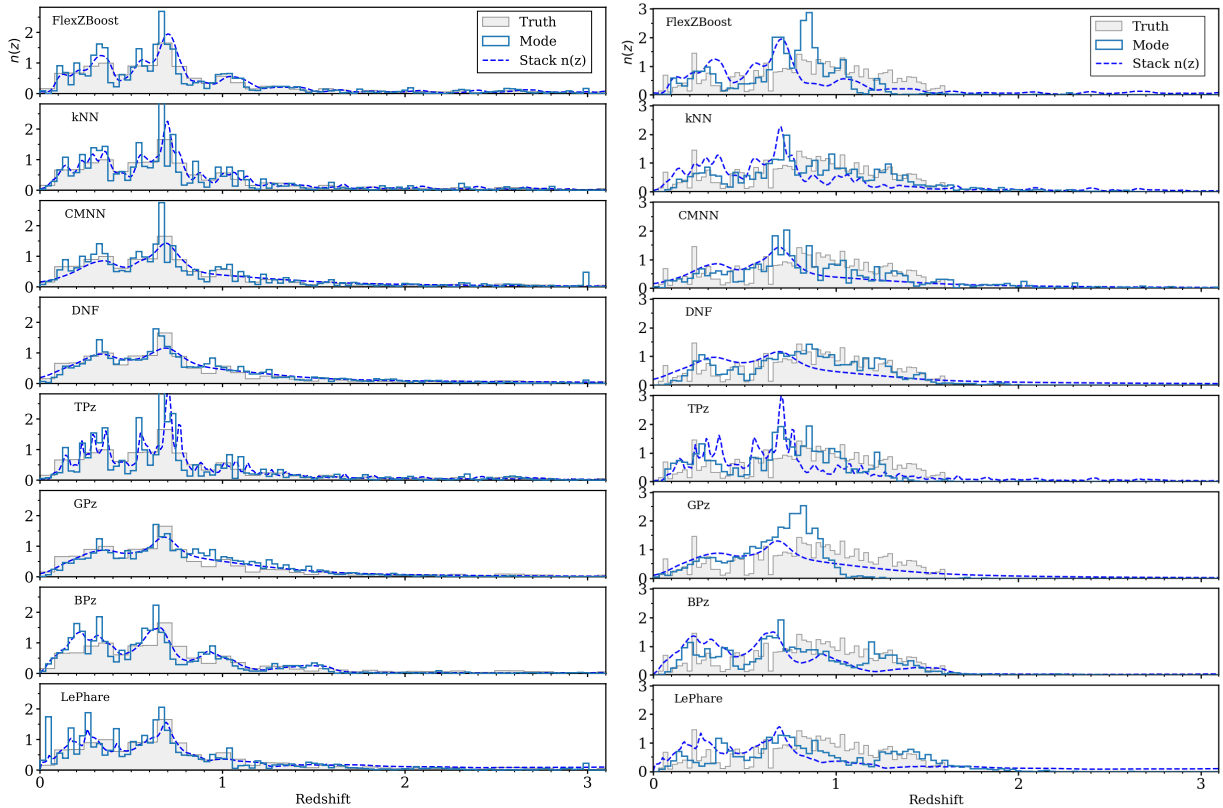


Figure 8. Comparisons of stacked redshift distributions estimated by each photo- z algorithm with the true redshift distributions for the ECDFS test set (left panels) and the independent DESI cross-match in the SV_38_7 field (right panels). Each panel shows, for one algorithm, the stacked redshift distribution $n(z)$ (blue dashed lines) estimated from the photometric redshift PDFs, the histogram of point-estimate photo- z values (blue step lines), and the histogram of the true redshift distribution (grey distribution). These plots demonstrate the ability of each algorithm to recover the underlying redshift distribution. All profiles in this figure are normalized.

vs. redshift, e.g., the template-fitting methods tend to exhibit more significant bias. We also noticed that the bias for BPz and DNF are particularly significant at low redshift ($z_{\text{ref}} < 0.5$), while DNF provides the lower bias, scatter, and outlier rate at the highest redshift. The bias of the photo- z algorithms across the redshift exceeds the

LSST Y1 requirement on the mean redshift (The LSST Dark Energy Science Collaboration et al. 2018), highlighting the importance of redshift calibration via techniques like clustering redshift (Newman 2008). The scatter of the photo- z are within the LSST Y1 requirement for $z_{\text{ref}} < 1.2$.

Table 2. Performance metrics for photo- z algorithms using 6-band DP1 data in the ECDFS field. Columns show the redshift bias $\mathbb{E}[\Delta z]$, the scatter measured by σ_{NMAD} , the outlier rate $\eta_{0.15}$, and the tomographic binning accuracy.

Algorithm	$\mathbb{E}[\Delta z]$	σ_{NMAD}	$\eta_{0.15}$	bin. acc.
FlexZBoost	-0.0005	0.0280	0.117	0.781
kNN	-0.0035	0.0379	0.149	0.689
CMNN	0.0001	0.0416	0.145	0.723
DNF	-0.0001	0.0361	0.152	0.698
TPZ	0.0010	0.0547	0.138	0.689
GPz	0.0025	0.0551	0.204	0.634
BPZ	-0.0179	0.0459	0.192	0.587
LePhare	-0.0143	0.0400	0.171	0.628

The variety of performance at different redshift and magnitude ranges means the optimal photo- z algorithm depends on the science cases one wants to explore. This is a particular advantage of RAIL, as it is relatively easy to compute photo- z by multiple methods.

In Fig. 6, we show the PIT-QQ curves of all methods we are evaluating. The general trend of the empirical CDF follows the identity lines. However, we note that some methods are giving overconfident error bars, e.g., CMNN. We also notice the template fitting methods Bpz and LePhare have PDF skewed toward high redshift, thus a PIT-QQ curve lower than the diagonal line.

In addition, we also train models using four-band photometry (*griz*), and estimate the photo- z on the SV_38_7 field. We also evaluate performance for galaxy ensembles binned by redshift and magnitude. We find that each of our photo- z algorithms performs well in the range of $z = 0.2$ to 1.2. We refer readers to Charles et al. (2025) for discussion of the performance with magnitude and redshift, and details about the four-band photo- z .

4.3 photo- z Including Euclid NIR Photometry

We study the impact on the photo- z performance of using galaxies with DP1 + Euclid NIR photometry, described in Section 2.2.3. For simplicity, we only used the FlexZBoost algorithm, which was trained and tested on this dataset both with and without the Euclid NIR photometry.

In Fig. 7, we show the comparisons of the bias, scatter, and outlier rate as a function of redshift for FlexZBoost with/without Euclid NIR photometry. Since our train/test galaxies concentrate at redshift $z < 1.2$, the infrared information does not significantly improve the photo- z bias and scatter overall. We found considerable improvement in the bias, scatter and outlier rate for test galaxies at $z_{\text{ref}} > 1.2$. To study the full improvement of including near-infrared photometry in LSST photometry, future studies need to obtain more high-redshift reference samples to reduce the uncertainties of the metrics.

4.4 Redshift Distribution Estimation

In this section, we compare the redshift distributions predicted by each photo- z algorithm against the true redshift distributions from the reference sample. We evaluate both stacked $n(z)$ estimates, obtained by summing per-galaxy photometric redshift PDFs, and the distributions of point estimates. These comparisons enable us to evaluate the ability of each algorithm to recover the true redshift distribution, which is essential for robust weak lensing and clustering measurements for DESC.

Fig. 8 shows the stacked $n(z)$ and the mode histogram compared to the true $n(z)$ distribution for each algorithm. For the ECDFS test

set with six-band photometry, most photo- z algorithms' stacked $n(z)$ successfully recover the true redshift distribution. This suggests that, given deep multiband data and a representative training set, both machine-learning and template-fitting methods can provide reliable ensemble redshift estimates. Again, we notice that the template-fitting algorithms have a peak at redshift $z_{\text{ref}} \sim 0.2$, which may be caused by misidentifying Lyman break galaxies as low redshift Balmer break galaxies.

In the four-band SV_38_7 validation field, some machine-learning methods exhibit imprints of the training set redshift distribution, producing tails at high redshift that the true redshift distribution does not have, as well as a spike at redshift $z \sim 0.7$. Fundamentally, this is caused because the ECDFS training galaxies and SV_38_7 galaxies have significantly different color-magnitude-redshift distribution. Of all the methods, we find that the mode histogram of DNF best matches the true distribution. Several of the stacked $n(z)$ distribution exhibit significant systematics that warrant further investigation, e.g., the stacked $n(z)$ of CMNN, GPz, and LePhare. Similar spikes can also be seen in the stacked PDF of these methods, shown in Appendix A.

We note that all training and testing galaxies are drawn from a relatively small field. Therefore, our photo- z are subject to training set sample variance (Cunha et al. 2012) when applying the photo- z algorithm to a different field, e.g., SV_38_7.

5 CONCLUSION

In this work, we use the Rubin Observatory's DP1 dataset to demonstrate the capability of RAIL to produce a photometric redshift for LSST and DESC. These products include per-galaxy PDF, their point estimates, and ensemble $n(z)$, using various machine-learning and template-fitting algorithms.

We construct a reference set in the ECDFS and cross-match to DP1 galaxies, and cross-match the DP1 galaxies to Euclid and DESI to obtain infrared photometry and external validation redshifts. The cross-matched galaxies are used as training and testing galaxies for the machine-learning models, and calibration galaxies for the template-fitting methods.

We have demonstrated that the algorithms in RAIL can produce photo- z results with promising accuracy and precision for real LSST data. With a representative training set, the biases of the machine-learning algorithms are less than 0.005, which is the LSST Y1 requirement (The LSST Dark Energy Science Collaboration et al. 2018) for overall redshift bias. The photo- z scatter can reach ~ 0.03 level, below the LSST Y1 requirement of 0.1, and the outlier rate can reach $\sim 10\%$ for a high SNR sample. We study the improvement of photo- z by including the Euclid NIR photometry, and find minor improvement for photo- z bias and outlier rate at high redshift. We demonstrate that the point estimate of the photo- z can be used for tomographic binning, with FlexZBoost reaching 0.78 binning accuracy. The stacked $n(z)$ shows good agreement with the true distribution in the ECDFS test set. For the SV_38_7 field, we find that the DNF point estimates result in the least biased $n(z)$, while some methods display a significant level of systematic error.

The results we present have several notable caveats that are worth noting: The size and depth of the training set limit the robustness of the photo- z past a redshift of $z \approx 1.2$. Some photo- z algorithms perform poorly with non-detections in some bands. We undertook limited effort in optimizing the hyperparameters of the algorithms, so additional performance gains likely could be found in future work. The red leak, i.e., non-zero infrared transmission, of the LSSTCom-

Cam g-band might impact the g-band photometry and the photo-z. We do not expect this issue to arise in LSSTCam.

We identify several potential future improvements. To name a few: (a) non-detection/negative fluxes in a subset of bands need to be better handled by a few algorithms; (b) algorithm-specific quality flags should be calculated and recorded; (c) combining different flux measurements may provide additional flux information; this will be explored in the future; (d) training set choices can be further examined to maximize performance on the galaxies in Rubin’s object catalog.

In summary, this work lays a solid foundation for applying RAIL to real LSST datasets. The DP1 photo-z results show promising results for future analysis. The work demonstrates the readiness of the pipeline, but it also highlights some areas that require future progress.

CONTRIBUTION STATEMENTS

T. Zhang: analysis and coordination, reference catalog construction, algorithm optimization, developer of RAIL, writing the manuscript.
E. Charles: core developer of RAIL and `rail_projects`, photo-z data production, project oversight.

J.F. Crenshaw: construction of the reference sample and Euclid cross-match.

S. Schmidt: core developer of RAIL, algorithm optimization, and paper editing, and project coordination.

P. Adari: construction of the DESI cross-match and analysis.

J. Gschwend: developer of the PZ server.

S. Mau: developer of `rail_projects`.

ACKNOWLEDGEMENTS

TZ, AM, DO, OL are supported by Schmidt Sciences. TZ thanks SLAC National Accelerator Laboratory for providing hospitality and an excellent research environment during the course of this study. JFC acknowledges support from the U.S. Department of Energy, Office of Science, Office of High Energy Physics Cosmic Frontier Research program under Award Number DE-SC0011665. This research is supported by the U.S. Department of Energy under awards DE-SC0025309 and DE-SC0023387 received by PA. This research received support from the National Science Foundation (NSF) under grant No. NSF DGE-1656518 through the NSF Graduate Research Fellowship received by SM. MJJ acknowledges support for the current research from the National Research Foundation (NRF) of Korea under the programs 2022R1A2C1003130 and RS-2023-00219959. SL is supported by the U.S. Department of Energy under grant number DE-1161130-116-SDDTA and under Contract No. DE-AC02-76SF00515 with the SLAC National Accelerator Laboratory. The work of AAPM was supported by the U.S. Department of Energy under contract number DEAC02-76SF00515. AAPM thanks the Department of Physics and the Laboratory of Particle Astrophysics and Cosmology at Harvard University, the Cosmology Group at Boston University, and the Department of Physics at Washington University in St. Louis. EU is supported by the U.S. Department of Energy and the Cosmic Frontier grant DE-SC0007881. AvdL is supported by the US Department of Energy under award under award DE-SC0025309. CWW was supported by Department of Energy, grant DE-SC0010007.

The photo-z Server uses computational resources of IDAC-Brazil

at the Laboratório Interinstitucional de e-Astronomia (LIneA) with financial support from INCT do e-Universo (Process no. 465376/2014-2).

ISN, LTSC and JdV are partially supported by Spanish Ministerio de Ciencia, Innovacion y Universidades under grant PID2021-123012. ISN, LTSC and JdV gratefully acknowledges funding from the MAD4SPACE-CM TEC-2024/TEC-182 project funded by Comunidad de Madrid.

This material is based upon work supported in part by the National Science Foundation through Cooperative Agreements AST-1258333 and AST-2241526 and Cooperative Support Agreements AST-1202910 and AST-2211468 managed by the Association of Universities for Research in Astronomy (AURA), and the Department of Energy under Contract No. DE-AC02-76SF00515 with the SLAC National Accelerator Laboratory managed by Stanford University. Additional Rubin Observatory funding comes from private donations, grants to universities, and in-kind support from LSST-DA Institutional Members.

This paper has undergone internal review in the LSST Dark Energy Science Collaboration. The internal reviewers were Qianjun Hang and Brett Andrews. We thank them for their valuable contribution. We thank Yihao Zhou and Lei Hu for their helpful input on the paper.

The DESC acknowledges ongoing support from the Institut National de Physique Nucléaire et de Physique des Particules in France; the Science & Technology Facilities Council in the United Kingdom; and the Department of Energy and the LSST Discovery Alliance in the United States. DESC uses resources of the IN2P3 Computing Center (CC-IN2P3-Lyon/Villeurbanne - France) funded by the Centre National de la Recherche Scientifique; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; STFC DiRAC HPC Facilities, funded by UK BEIS National E-infrastructure capital grants; and the UK particle physics grid, supported by the GridPP Collaboration. This work was performed in part under DOE Contract DE-AC02-76SF00515.

DATA AVAILABILITY

The DP1 photo-z data products can be accessed in several ways.

Rubin Science Platform

DP1 photo-z data can be accessed at the Rubin Science Platform¹¹. During the DP1 commissioning, we have demonstrated functionality to a) ingest model files produced by training the various algorithms in the Rubin Data Butler (Butler), b) use the Rubin Data Management software stack along with photo-z specific plugins¹² to produce pre-object PDF estimates in the data management framework, with the estimates automatically being written to the Butler, c) retrieve those $p(z)$ estimates as we would retrieve any other Rubin data product. Work is ongoing to fully support this functionality in Rubin Data Preview 2 (DP2), with the estimates for the kNN and BPZ algorithms being supported and distributed by the Rubin Observatory for DP2, and other algorithms being supported and distributed by the Rubin science community.

¹¹ data.lsst.cloud

¹² https://github.com/lsst-dm/meas_pz

Photo-z Server

The Photo-z Server¹³ is a web-based service available for the LSST community to create and host lightweight PZ-related data products. All data products described in this document will be hosted on the Photo-z Server, along with their respective metadata and documentation. A list with access instructions and links to product pages will be available on the data product documentation page¹⁴.

LSDB

The Large Scale DataBase (LSDB) (Caplar et al. 2025) hosts DP1 data as well as the photo-z point estimates from this work at https://data.lsdbs.io/#Rubin_DP1/object_photoz.

REFERENCES

- Abbott T. M. C., et al., 2019, *ApJ*, 872, L30
- Abbott T. M. C., et al., 2022, *Phys. Rev. D*, 105, 023520
- Almosallam I. A., Jarvis M. J., Roberts S. J., 2016, *MNRAS*, 462, 726
- Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, *MNRAS*, 310, 540
- Balestra I., et al., 2010a, *A&A*, 512, A12
- Balestra I., et al., 2010b, *A&A*, 512, A12
- Benítez N., 2000, *ApJ*, 536, 571
- Blake C., et al., 2016, *MNRAS*, 462, 4240
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Caplar N., et al., 2025, *arXiv e-prints*, p. [arXiv:2501.02103](https://arxiv.org/abs/2501.02103)
- Carrasco Kind M., Brunner R. J., 2013, *MNRAS*, 432, 1483
- Charles E., et al., 2025, Commissioning Technical Note SITCOMTN-154, Initial studies of photometric redshifts with LSSTComCam from DP1, <https://sitcomtn-154.lsst.io/>. NSF-DOE Vera C. Rubin Observatory, doi:10.71929/rubin/2571480, <https://sitcomtn-154.lsst.io/>
- Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, *AJ*, 132, 926
- Collaboration D., et al., 2025, Data Release 1 of the Dark Energy Spectroscopic Instrument ([arXiv:2503.14745](https://arxiv.org/abs/2503.14745)), <https://arxiv.org/abs/2503.14745>
- Colless M., et al., 2001, *MNRAS*, 328, 1039
- Cool R. J., et al., 2013, *ApJ*, 767, 118
- Cooper M. C., et al., 2012, *MNRAS*, 425, 2116
- Croom S. M., Smith R. J., Boyle B. J., Shanks T., Loaring N. S., Miller L., Lewis I. J., 2001, *MNRAS*, 322, L29
- Cunha C. E., Huterer D., Busha M. T., Wechsler R. H., 2012, *MNRAS*, 423, 909
- DESI Collaboration et al., 2025, *arXiv e-prints*, p. [arXiv:2503.14745](https://arxiv.org/abs/2503.14745)
- D'Eugenio F., et al., 2025, *ApJS*, 277, 4
- Daddi E., et al., 2004, *ApJ*, 600, L127
- De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, *MNRAS*, 459, 3078
- Doherty M., Bunker A. J., Ellis R. S., McCarthy P. J., 2005, *MNRAS*, 361, 525
- Euclid Collaboration et al., 2025, *arXiv e-prints*, p. [arXiv:2503.15302](https://arxiv.org/abs/2503.15302)
- Garilli B., et al., 2021, *A&A*, 647, A150
- Giavalisco M., 2002, *ARA&A*, 40, 579
- Graham M. L., Connolly A. J., Ivezić Ž., Schmidt S. J., Jones R. L., Jurić M., Daniel S. F., Yoachim P., 2018, *AJ*, 155, 1
- Graham M. L., Bosch J., Guy L. P., The DM System Science Team. 2022, Data Management Technical Note DMTN-049, A Roadmap to Photometric Redshifts for the LSST Object Catalog, <https://dmtn-049.lsst.io/>. NSF-DOE Vera C. Rubin Observatory, <https://dmtn-049.lsst.io/>
- Grazian A., et al., 2006, *A&A*, 449, 951
- Hathi N. P., Malhotra S., Rhoads J. E., 2008, *ApJ*, 673, 686
- Helou G., Madore B. F., Schmitz M., Bica M. D., Wu X., Bennett J., 1991, in Albrecht M. A., Egret D., eds, *Astrophysics and Space Science Library* Vol. 171, Databases and On-line Data in Astronomy. pp 89–106, doi:10.1007/978-94-011-3250-3_10
- Huchra J. P., et al., 2012, *ApJS*, 199, 26
- Ilbert O., et al., 2009, *ApJ*, 690, 1236
- Ilbert O., Cohen-Tanugi J., Shirley R., Arnouts S., Salvato M., 2024, LePHARE : PHotometric Analysis for Redshift Estimation, doi:10.5281/zenodo.14162574, <https://doi.org/10.5281/zenodo.14162574>
- Ivezić Ž., et al., 2019, *ApJ*, 873, 111
- Izbicki R., Lee A. B., 2017, *Electronic Journal of Statistics*, 11, 2800
- Jones D. H., et al., 2009, *MNRAS*, 399, 683
- Kilbinger M., 2015, *Rept. Prog. Phys.*, 78, 086901
- Kodra D., et al., 2023, *ApJ*, 942, 36
- Kriek M., et al., 2007, *ApJ*, 669, 776
- Kriek M., et al., 2015a, *ApJS*, 218, 15
- Kriek M., et al., 2015b, *ApJS*, 218, 15
- Kron R. G., 1980, *ApJS*, 43, 305
- Le Fèvre O., et al., 2004, *A&A*, 428, 1043
- Le Fèvre O., et al., 2013a, *arXiv e-prints*, p. [arXiv:1307.0545](https://arxiv.org/abs/1307.0545)
- Le Fèvre O., et al., 2013b, *A&A*, 559, A14
- Le Fèvre O., et al., 2015, *A&A*, 576, A79
- Lidman C., et al., 2020, *MNRAS*, 496, 19
- Malhotra S., et al., 2005, *ApJ*, 626, 666
- McLure R. J., et al., 2018, *MNRAS*, 479, 25
- Merlin E., et al., 2021, *A&A*, 649, A22
- Merlin E., et al., 2024, *A&A*, 691, A240
- Mignoli M., et al., 2005, *A&A*, 437, 883
- Momcheva I. G., et al., 2016a, *ApJS*, 225, 27
- Momcheva I. G., et al., 2016b, *ApJS*, 225, 27
- Morris A. M., et al., 2015, *AJ*, 149, 178
- NSF-DOE Vera C. Rubin Observatory 2025a, Legacy Survey of Space and Time Data Preview 1 [Data set], doi:10.71929/RUBIN/2570308, <https://www.osti.gov/servlets/purl/2570308>
- NSF-DOE Vera C. Rubin Observatory 2025b, Legacy Survey of Space and Time Data Preview 1: deep_coadd dataset type [Data set], doi:10.71929/RUBIN/2570313, <https://www.osti.gov/servlets/purl/2570313>
- NSF-DOE Vera C. Rubin Observatory 2025c, Technical Note RTN-095, The Vera C. Rubin Observatory Data Preview 1, <https://rtn-095.lsst.io/>. NSF-DOE Vera C. Rubin Observatory, doi:10.71929/rubin/2570536, <https://rtn-095.lsst.io/>
- Newman J. A., 2008, *ApJ*, 684, 88
- Newman J. A., Gruen D., 2022, *Annual Review of Astronomy and Astrophysics*
- Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, *A&A*, 507, 1793
- Norman C., et al., 2002, *ApJ*, 571, 218
- Pasquali A., et al., 2006, *ApJ*, 636, 115
- Pentericci L., et al., 2018a, *A&A*, 616, A174
- Pentericci L., et al., 2018b, *A&A*, 619, A147
- Polletta M., et al., 2007, *ApJ*, 663, 81
- Popesso P., et al., 2009, *A&A*, 494, 443
- Ravikumar C. D., et al., 2007, *A&A*, 465, 1099
- Rhoads J. E., et al., 2009, *ApJ*, 697, 942
- SLAC National Accelerator Laboratory NSF-DOE Vera C. Rubin Observatory 2024, LSST Commissioning Camera, doi:10.71929/RUBIN/2561361, <https://www.osti.gov/servlets/purl/2561361>
- Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, 737, 103
- Scolnic D. M., et al., 2018, *ApJ*, 859, 101
- Silverman J. D., et al., 2010, *ApJS*, 191, 124
- Straughn A. N., et al., 2008, *AJ*, 135, 1624

¹³ <https://pzserver.linea.org.br>

¹⁴ https://docs.linea.org.br/en/data/pz_server_data.html#data-preview-1

- Straughn A. N., et al., 2009, *AJ*, 138, 1022
- Szokoly G. P., et al., 2004, *ApJS*, 155, 271
- Tasca L. A. M., et al., 2017, *A&A*, 600, A110
- The LSST Dark Energy Science Collaboration et al., 2018, *arXiv e-prints*, p. [arXiv:1809.01669](https://arxiv.org/abs/1809.01669)
- The RAIL Team et al., 2025, *arXiv e-prints*, p. [arXiv:2505.02928](https://arxiv.org/abs/2505.02928)
- Tomida K., Tomisaka K., Matsumoto T., Hori Y., Okuzumi S., Machida M. N., Saigo K., 2013, *ApJ*, 763, 6
- Trump J. R., et al., 2015, *ApJ*, 811, 26
- Van Waerbeke L., et al., 2000, *A&A*, 358, 30
- Vanzella E., et al., 2008, *A&A*, 478, 83
- Vanzella E., et al., 2009, *ApJ*, 695, 1163
- Wolf C., et al., 2004, *A&A*, 421, 913
- Wuyts S., Labbé I., Förster Schreiber N. M., Franx M., Rudnick G., Brammer G. B., van Dokkum P. G., 2008, *ApJ*, 682, 985
- Wuyts S., van Dokkum P. G., Franx M., Förster Schreiber N. M., Illingworth G. D., Labbé I., Rudnick G., 2009, *ApJ*, 706, 885
- van der Wel A., Franx M., van Dokkum P. G., Rix H. W., 2004, *ApJ*, 601, L5
- ¹Department of Physics and Astronomy and PITT PACC, University of Pittsburgh, Pittsburgh, PA 15260, USA
- ²SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA
- ³Kavli Institute for Particle Astrophysics and Cosmology (KIPAC), Stanford University, Stanford, CA 94305, USA
- ⁴Department of Physics, University of Washington, Seattle, WA 98195, USA
- ⁵DIRAC Institute, University of Washington, Seattle, WA 98195, USA
- ⁶Department of Physics and Astronomy, University of California, One Shields Avenue, Davis, CA 95616, USA
- ⁷Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794-3800, USA
- ⁸Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
- ⁹Université Paris Cité, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France
- ¹⁰Department of Physics and Astronomy, San José State University, San José, CA 95192-0106, USA
- ¹¹Department of Physics, University of Wisconsin-Madison, Madison, WI 53706, USA
- ¹²Laboratoire d'Annecy de Physique des Particules (LAPP), 9 Chem. de Bellevue, B.P. 110, 74941 Annecy, France
- ¹³Laboratoire de Physique des 2 Infinis Irène Joliot-Curie (IJCLab), Bât. 100, 15 rue Georges Clémenceau, 91405 Orsay, France
- ¹⁴Ohio University, Athens, OH 45701, USA
- ¹⁵Université Clermont-Auvergne, CNRS, LPCA, 63000 Clermont-Ferrand, France
- ¹⁶LPSC-IN2P3, Laboratoire de Physique Subatomique et Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, 53, rue des Martyrs, 38026 Grenoble Cedex, France
- ¹⁷Steward Observatory, The University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721, USA
- ¹⁸NSF-DOE Vera C. Rubin Observatory, 950 N. Cherry Ave., Tucson, AZ 85719, USA
- ¹⁹Laboratoire de Physique Nucléaire et des Hautes Énergies (LPNHE), Sorbonne Université, Université Paris Cité, CNRS/IN2P3, F-75005 Paris, France
- ²⁰Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
- ²¹Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA
- ²²Vera C. Rubin Observatory, Avenida Juan Cisternas #1500, La Serena, Chile
- ²³Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- ²⁴Laboratoire d'Astrophysique, Marseille Aix-Marseille University, 38 Rue Frédéric Joliot Curie, 13013 Marseille, France
- ²⁵University of Pennsylvania, Philadelphia, PA 19104, USA
- ²⁶Department of Physics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, South Korea
- ²⁷Departments of Physics and Astronomy, University of California, Berkeley, CA, USA
- ²⁸Department of Physics, Duke University, Durham, NC 27708, USA
- ²⁹Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
- ³⁰McWilliams Center for Cosmology and Astrophysics, Department of Physics, Carnegie Mellon University, Pittsburgh, PA, USA
- ³¹Sidrat Research, 124 Merton Street, Suite 507, Toronto, ON M4S 2Z2, Canada
- ³²Infrared Processing and Analysis Center (IPAC), California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA
- ³³Max-Planck-Institut für extraterrestrische Physik, Giessenbachstr. 1, 85748 Garching, Germany
- ³⁴Centre de Physique des Particules de Marseille, 163, avenue de Luminy Case 902 13288 Marseille cedex 09, France
- ³⁵Department of Physics, Harvard University, Cambridge, MA, USA
- ³⁶Institute of Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, United Kingdom

APPENDIX A: STACKED PDF OF ALL GALAXIES

In Figure A1, we show the stacked PDF of all galaxies in all four DPI fields, binned by their *i*-band magnitude. The faint galaxies have overall higher redshifts, which follows the empirical expectation of a magnitude-limited sample. We can see signs of overfitting for FlexZBoost and kNN at high redshift, as those peaks match across the fields and are likely not real. The double peak feature that appeared in our training set is imprinted onto most of the stacked PDFs. This highlights the importance of smoothing out the training set by reweighting and resampling in the future.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.

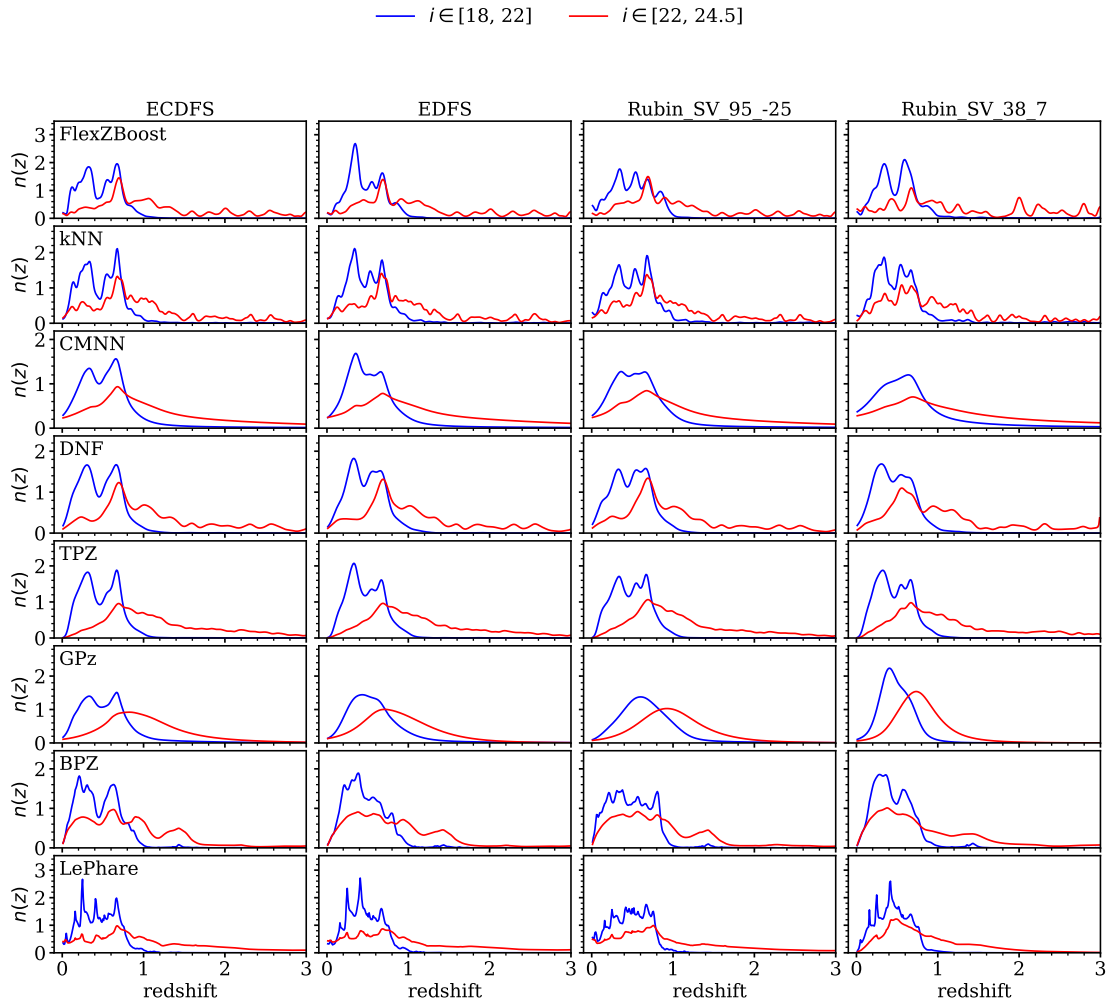


Figure A1. Stacked PDF of all DP1 gold samples in the four fields with multi-band observation, from all eight of our algorithms. The galaxies are binned into a bright bin ($i \in [18, 22]$, blue lines) and a faint ($i \in [22, 24.5]$, red lines). The stacked PDFs of machine learning methods show the imprint of the training sample.