

Photometric Redshifts for the Vera C. Rubin Observatory Data Preview

2

ERIC CHARLES ^{1,2} TIANQING ZHANG ³ SAMUEL J. SCHMIDT ⁴
DAN S. TARANU ⁵ JULIA GSCHWEND ⁶ MELISSA L. GRAHAM ^{7,8} AND
ARUN KANNAWADI ^{9,5}

¹*SLAC National Accelerator Laboratory, 2575 Sand Hill Rd., Menlo Park, CA 94025, USA*

²*Kavli Institute for Particle Astrophysics and Cosmology, SLAC National Accelerator Laboratory, 2575 Sand Hill Rd., Menlo Park, CA 94025, USA*

³*Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA*

⁴*Physics Department, University of California, One Shields Avenue, Davis, CA 95616, USA*

⁵*Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*

⁶*Laboratório Interinstitucional de e-Astronomia, Av. Pastor Martin Luther King Jr, 126 - Del Castilho Nova América Offices, Torre 3000 / sala 817. 20765-000 - Rio de Janeiro - RJ, Brasil*

⁷*University of Washington, Dept. of Astronomy, Box 351580, Seattle, WA 98195, USA*

⁸*Institute for Data-intensive Research in Astrophysics and Cosmology, University of Washington, 3910 15th Avenue NE, Seattle, WA 98195, USA*

⁹*Department of Physics, Duke University, Durham, NC 27708, USA*

(Dated: September 9, 2026)

ABSTRACT

We present the photometric redshifts (photo-z) inferred using algorithms implemented in the Redshift Assessment Infrastructure Layers (RAIL) for the NSF-DOE Vera C. Rubin Observatory Data Preview 2 (DP2). We produce a compilation of reference redshift catalog using spectroscopic, grism and many band photometric redshift dataset hosted on the LLineA Photo-z Server. We curate training and testing set for assessing the scientific and technical performance of Rubin photo-z. The algorithm applied to the object catalog are FlexZBoost, BPZ, kNN, GPz, DNF and TPz; with a combination of 6-band and 4-band photo-z depending on availability of u and y photometry. The redshift point estimates and uncertainty estimation in tabular format through the Large Survey DataBase (LSDB).

1. INTRODUCTION

The NSF-DOE Vera C. Rubin Observatory's Data Preview 2 (DP2; [Vera C. Rubin Observatory Team 2026](#); [NSF-DOE Vera C. Rubin Observatory 2026a](#)) represents a significant step forward in preparation for the upcoming Legacy Survey of Space and Time (LSST; [Ž. Ivezić et al. 2019](#)), providing an essential testbed for scientific tools and analytical pipelines using precursor imaging data. A central goal of LSST

is estimating photometric redshifts (photo- z s) for billions of galaxies, supporting extragalactic astrophysical research and cosmological studies that depend on reliable redshift measurements and distributions. In support of this objective, the Rubin project established a strategic plan for delivering high-quality photo- z s to the research community (M. L. Graham et al. 2022).

While photo- z estimates are not part of the official Data Preview 1 (DP1) or DP2 data products, the "Photo- z Science Unit" undertook the task of producing photo- z values for all galaxies in DP2 using available multi-band imaging data on a best-effort basis, establishing a foundation for future large-scale implementations. This science unit delivered redshifts for DP1 (E. Charles et al. 2025; Vera C. Rubin Observatory Team et al. 2026).

For this purpose, we used the RAIL (Redshift Assessment Infrastructure Layers) software framework, a versatile and modular tool developed for photo- z estimation and validation (The RAIL Team et al. 2025). In particular, we leveraged RAIL to train photo- z estimation models using high-quality spectroscopic redshift training samples matched to the DP2 photometric catalog.

This document presents the resulting best-effort photo- z catalogs, which should be considered highly preliminary, along with the methodologies and procedures employed to create them. Furthermore, we discuss key characteristics of the DP2 photometric data, the spectroscopic calibration samples, and the resulting photo- z catalogs. Supplementary information regarding data products and distribution is provided in the appendices.

We anticipate receiving input from scientific users as they examine the data and identify issues we may have overlooked, and we plan to integrate this feedback into subsequent efforts.

2. DATA

2.1. *Rubin DP2*

The Rubin Observatory's Data Preview 2 dataset is the first public release of Rubin observatory imaging data from the full LSST camera processed through the LSST Science Pipelines (Rubin Observatory Science Pipelines Developers et al. 2026), serving as a critical testbed for scientific and technical validation ahead of full LSST operations. DP2 is based on observations from the LSST camera (LSSTCam) and includes multi-band optical imaging in u, g, r, i, z and y filters over > 1000 square degrees of sky. The dataset consists of processed images, source catalogs, and associated metadata, all formatted using the Rubin Data Butler system (T. Jenness et al. 2022). Although somewhat smaller in scale than future LSST datasets, DP2 offers realistic photometric measurements, object detection, and data structures, making it an invaluable resource for developing and testing algorithms for tasks such as photo- z estimation, object classification, and data quality assessment. The DP2 footprint spans

several disjoint regions of sky and exhibits substantial variation in imaging depth and band coverage, which we describe in Sec. 2.1.1.

Critically, the fields selected for commissioning observation included several calibration fields, for which many high-quality redshifts exist from other surveys.

2.1.1. Preparation of photometric object catalogs

The Rubin Data Management (Rubin DM) pipeline measures multiple types of object photometry, the first stage in creating a photo- z catalog was to determine which measured photometry we would use as inputs. In this note, we generally use the 1.0 arc-second Gaussian Aperture fluxes and their associated errors, e. g. `u_gaap1p0Flux` and `u_gaap1p0FluxErr`. These fluxes should provide good measures of consistent galaxy colors within the defined aperture, ideal for photo- z estimation, though they may not necessarily reflect the colors of the overall galaxy if there is a significant color gradient and the galaxy is larger than the 1.0 arc-second aperture. This choice of photometric measurement may not be ideal, and investigation into the optimized set of photometric inputs will continue into the future. This is similar to studies done with DP1 and described in the associated technote and paper (E. Charles et al. 2025; T. Zhang et al. 2026).

The DP2 imaging is highly inhomogeneous in both depth and band coverage, which directly shapes how we prepare the catalogs and train our models. In Fig. 1, we show the median 5σ extended-source depth in the i band, derived from the 1.0 arc-second GAaP fluxes, for each of the 1391 tracts in DP2. The median depth across the footprint is 24.27 mag, but the tract-to-tract variation is large, spanning roughly 23.5 to 25.5 mag. In addition, not all tracts are observed in all six bands: as shown in the bottom panel of Fig. 1, 931 tracts have full six-band ($ugrizy$) coverage while 460 tracts are observed in only four bands ($griz$). Because the available bands set the input features for photo- z estimation, we cannot apply a single model uniformly across the footprint. We therefore train, for each algorithm in this work, one model on the six-band tracts and a separate model on the four-band tracts, and apply each model to the corresponding subset of the footprint. We note that the large depth variation implies that the effective magnitude limit and hence the photo- z quality also varies across the footprint.

Preparing object catalog (NSF-DOE Vera C. Rubin Observatory 2026b) data for photo- z algorithm training and estimation included five steps:

1. Applying quality cuts to the object catalog. We developed a selection for the training and test datasets and two additional selections for full DP2 catalogs: one applicable for fields with observations in only four bands ($dp2_4band$, requiring per-tract median depth in 'griz' $< 23mag$ the other for fields with observations in all six bands ($dp2_6band$, requiring per-tract median depth in 'ugrizy' $< 22mag$).
2. Converting fluxes in nJy to AB magnitudes ($m_{AB} = -2.5 \log_{10}(f_{\nu}/\text{nJy}) + 31.4$)

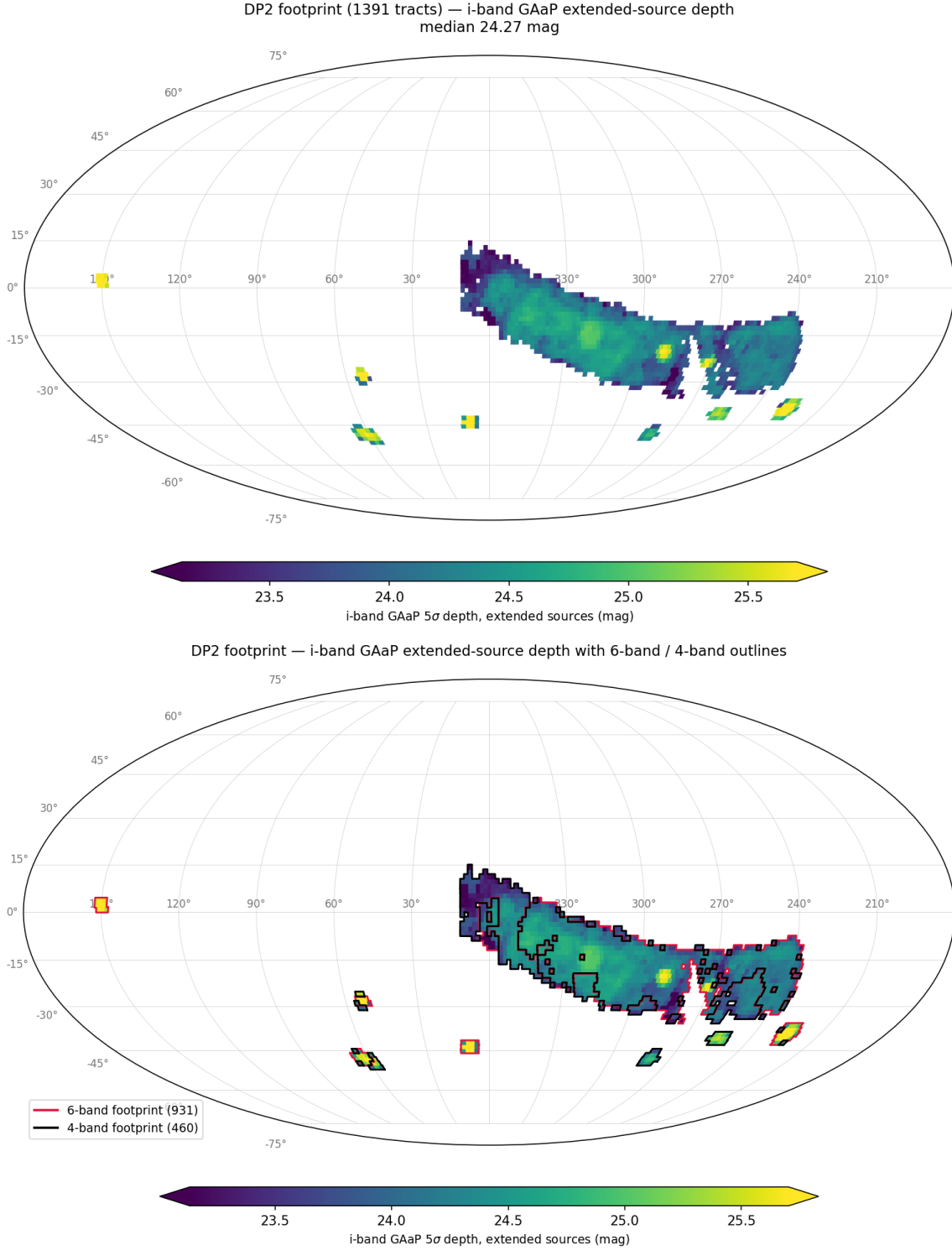


Figure 1. The DP2 survey footprint shown in a Mollweide projection. *Top panel:* the median 5σ extended-source depth in the *i* band, computed from the 1.0 arc-second GAaP fluxes, for each of the 1391 tracts in DP2. The color scale (bottom) runs from ~ 23.5 to ~ 25.5 mag, and the median depth across the footprint is 24.27 mag. The depth varies substantially from tract to tract. *Bottom panel:* the same depth map, with tracts outlined according to their band coverage, the red outline marks the 931 tracts with full six-band (*ugrizy*) photometry, and the black outline marks the 460 tracts with only four-band (*griz*) photometry. For each algorithm in this work we train one model on the six-band tracts and a separate model on the four-band tracts.

3. De-reddening to account for Galactic dust. We use the SFD dust maps (E. F. Schlafly & D. P. Finkbeiner 2011).
4. Cross-matching objects with reference catalogs that include redshift information as described in Sec. 2.2.
5. Shuffling and splitting the resulting catalog into `dp2_v3p1_training` and `dp2_v3p1_test` data sets.

2.1.2. Data properties

We have developed tools to generate diagnostic plots of both the input object catalogs and the photo- z estimates as part of our data analysis. Fig. 2 shows histograms of the AB magnitudes in 1.0 arc-second apertures of objects in the `dp2_v3p1_test` dataset, which includes an explicit magnitude cut, $m_i < 26.0$. Fig. 3 shows the correlation between magnitude and redshift for all objects in the `dp2_v3p1_test` data set. Fig. 4 shows the “adjacent band colors”, i.e., $u - g$, $g - r$, $r - i$, $i - z$, $z - y$ versus redshift for the same, with a series of SED templates (as used by template-fitting algorithms, see Sec. 3.1) overlaid. Finally, Fig. 5 shows the color-color plots for the same adjacent colors. In all cases, the 1.0 arc-second aperture magnitudes are used for plotting.

2.2. Redshift reference sample

The photo- z estimation algorithms in RAIL require highly accurate and precise redshift estimates cross-matched to the DP2 photometric catalog to provide labeled [DATASET] data sets. We created such datasets using data drawn from publicly available catalogs, comprising galaxies with known spectroscopic redshifts, grism redshifts, and high-quality photo- z estimates from deep, multi-band imaging. These training sets enable machine learning models within RAIL to learn the mapping between galaxy colors and redshifts, thereby enabling photo- z estimation for every galaxy detected in DP2. In this subsection, we describe how we assemble the reference redshift catalog and curate it into the training and test sets used in this work.

We first built a spectroscopic reference sample within the DP2 footprint by concatenating a large compilation of ~ 43 million redshifts from 38 reference redshift catalogs using the `Combine Redshift Catalogs` pipeline¹⁰ from the LLineA PZ Server¹¹. The pipeline uses LSDB to perform an efficient all-to-all cross-matching with the reference catalogs to identify multiple measurements of the same objects and resolve these duplicates according to data quality criteria based on previously homogenized quality flags, prioritizing spectroscopic data over grism measurements and grism measurements over high-quality photo- z estimates from multi-band imaging (Gschwend et al., in preparation). As a result, we produced two redshift reference catalogs, one as a

¹⁰ https://docs.linea.org.br/en/sci-platforms/pz_server_crc.html

¹¹ <https://pzserver.linea.org.br>

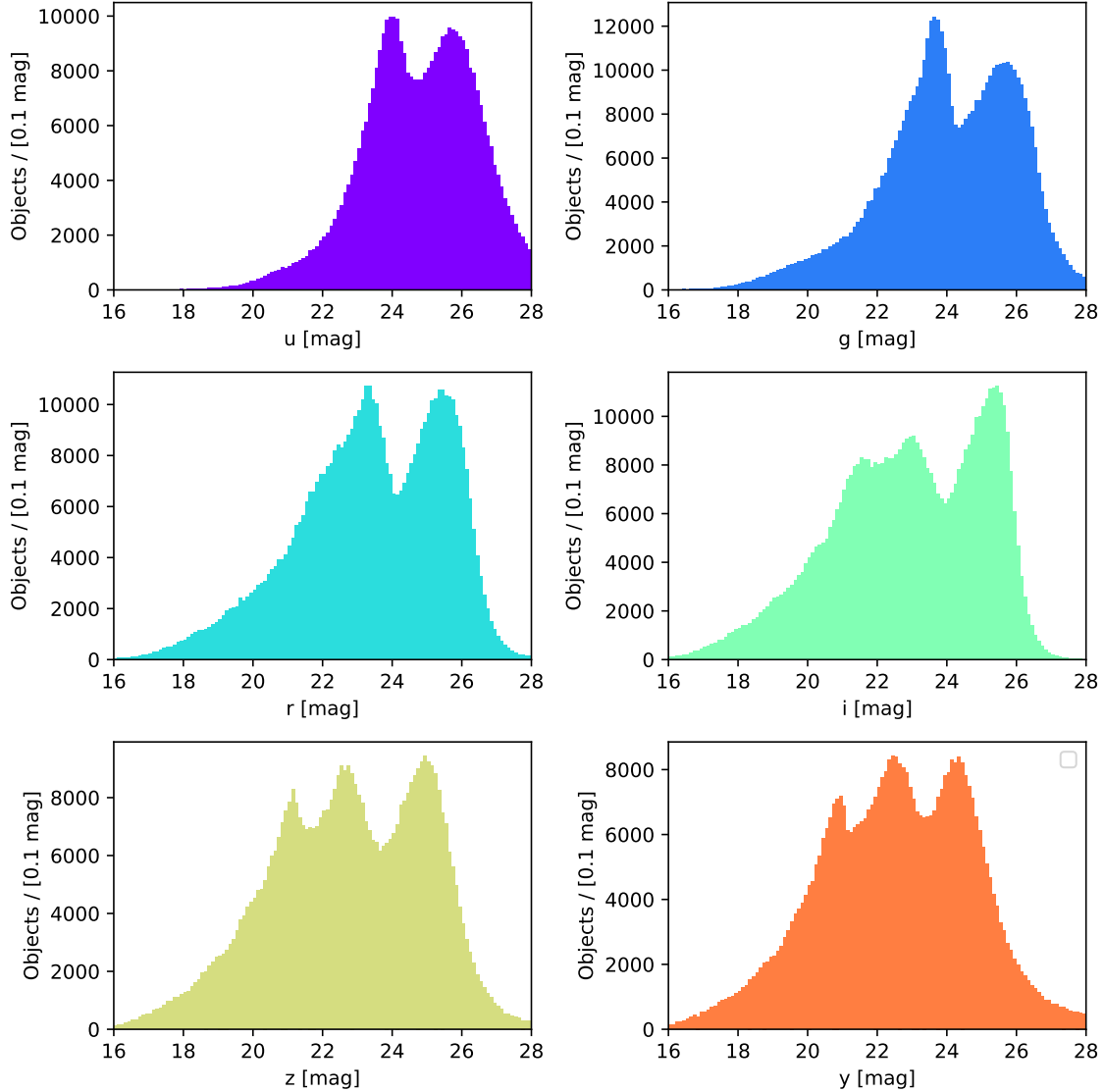


Figure 2. 1.0 arc-second Gaussian Aperture (Gaap) magnitudes (in AB system) of objects in each of the six Rubin filter bands in the `dp2_v3p1_test` dataset. These aperture magnitudes were used as the inputs for the photo- z algorithms.

compilation with redshifts of ~ 6.6 million unique galaxies from 37 reference surveys, and DESI DR1, separately, with homogenized flags.

We then use the **Training Set Maker** pipeline¹², which also uses LSDB as its backend, to cross-match the compilation and DESI DR1 against the DP2 object catalog, separately, retaining the nearest photometric counterpart within a fixed matching radius of 1.0 arcsec. The resulting reference sets span both the deep drilling fields, including COSMOS, ELAIS, ECDFS, and EDFs, for which many high-quality redshifts exist, and several wide-field spectroscopic surveys, most notably DESI and, for the compilation, SDSS. In Table 1, we report the number of cross-matched galaxies

¹² https://docs.linea.org.br/en/sci-platforms/pz_server_tsm.html

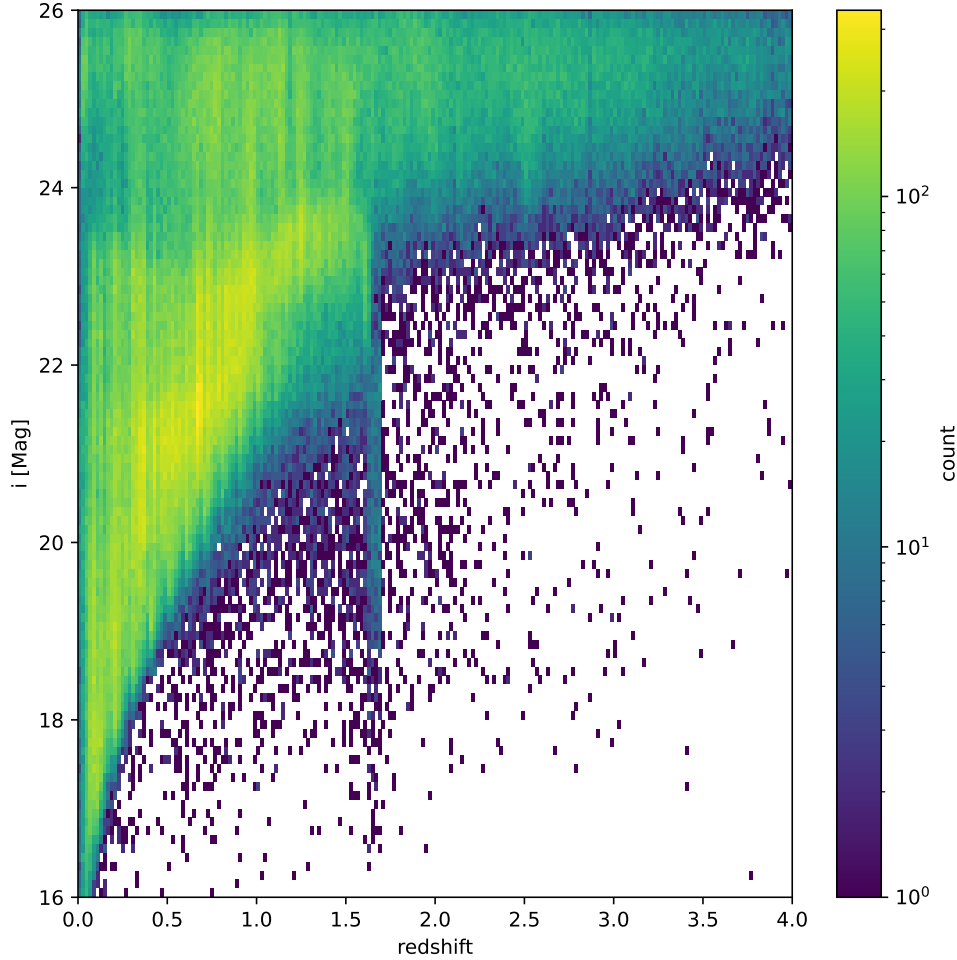


Figure 3. 1.0 arc-second Gaussian Aperture (Gaap) i-band magnitude versus redshift for all objects in the `dp2_v3p1_test` dataset.

contributed by each survey, listing all surveys with more than 100 objects and grouping the remainder into a single “Other” row. The cross-matched data are strongly dominated by the wide-field DESI, SDSS spectroscopic samples and the COSMOS deep-field compilations (counted as one reference source). In Fig. 6, we show the right-ascension–declination distribution of the training galaxies, color-coded by their originating survey.

To construct the training and test sets, we take the concatenation of the two samples described above including the reference compilation and the DESI cross-match, and split them 80/20 into a pre-training and a pre-testing set. We exclude quasars, which we identify as point-like sources in the second-moment trace versus magnitude plane (removing objects whose resolved size, computed from the difference between

Table 1. Composition of the reference redshift catalog after cross-matching against the DP2 object catalog. N is the number of cross-matched objects and $\bar{z}$ is the median redshift. Type codes: s = spectroscopic, p = many-band photo- z s, g = grism. Surveys with fewer than 100 objects are grouped in the “Other” row.

Survey	Type	N	$\bar{z}$
DESI DR1	s	1,544,298	0.66
SDSS DR17	s	503,965	0.50
COSMOS2020 Classic	p	355,527	1.27
COSMOS SRC	s	176,691	0.56
PRIMUS	s	22,166	0.48
2MRS	s	21,100	0.06
WiggleZ	s	15,590	0.58
VIPERS PDR2	s	15,497	0.67
OzDES	s	15,027	0.51
DEEP2	s	7,876	0.98
2dFGRS	s	7,684	0.11
JADES DR5 (photo- z)	p	7,383	1.70
HELP DMU23	p	7,005	0.45
ASTRODEEP-JWST	p	6,382	1.57
COSMOS-Web	p	5,799	1.00
VVDS	s	5,606	0.52
SAGA	s	3,924	0.25
ACES	s	2,853	0.62
SDSS DR19	s	2,383	0.69
3D-HST	g	2,171	1.14
6dFGS	s	1,578	0.06
ASTRODEEP	p	1,498	1.03
VIMOS	s	1,049	0.69
CANDELS	p	947	1.05
ATLAS	s	748	0.31
NED	s	601	1.18
VANDELS	s	564	3.35
MUSE	s	522	0.73
JADES	s	404	2.39
MOSDEF	s	314	2.28
SpARCS	s	291	0.87
CLASH-VLT	s	204	0.52
SPT-GMOS	s	161	0.43
VUDS	s	156	1.91
CDB	s	149	0.60
Other (<100 each)	—	273	—
Total		2,738,386	

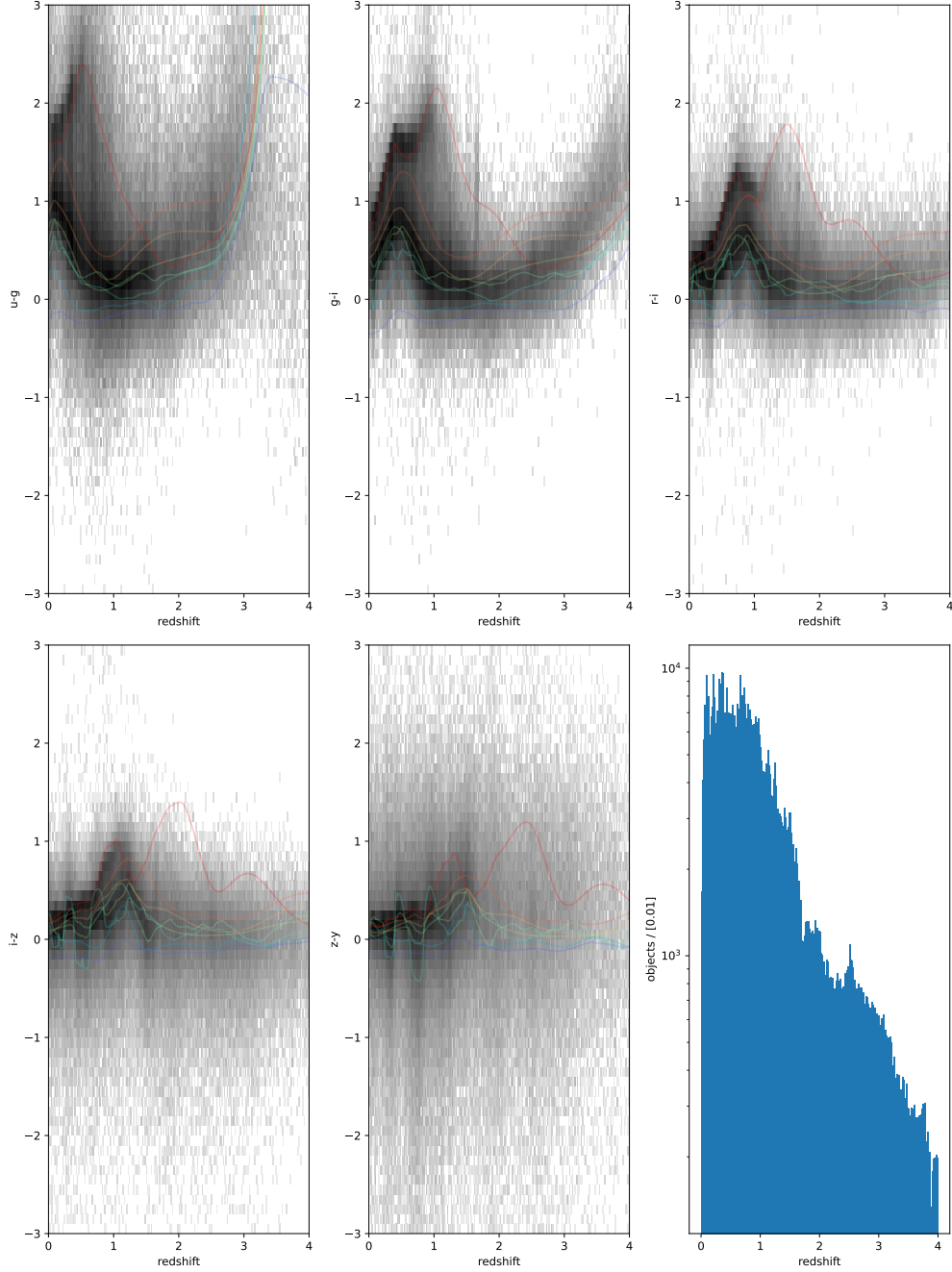


Figure 4. “Adjacent band colors”, i.e., $u - g$, $g - r$, $r - i$, $i - z$, $z - y$, in 1.0 arc-second apertures versus redshift for all objects in the `dp2_v3p1_test` dataset. Colored lines represent the expected colors for the eight “CWWSB” SEDs described in Sec. 3.1, and should very roughly show the predicted range of color evolution expected for our galaxy sample. The bottom right plot shows a histogram of the redshifts in the test dataset.

the i -band moment trace and the PSF trace, is below 0.01 arcsec) and, where available, using the spectroscopic type flags from DESI and SDSS. We also remove all reference objects with redshift $z < 0.001$. The full procedure is implemented in `make_pretrain_pretest.py`.

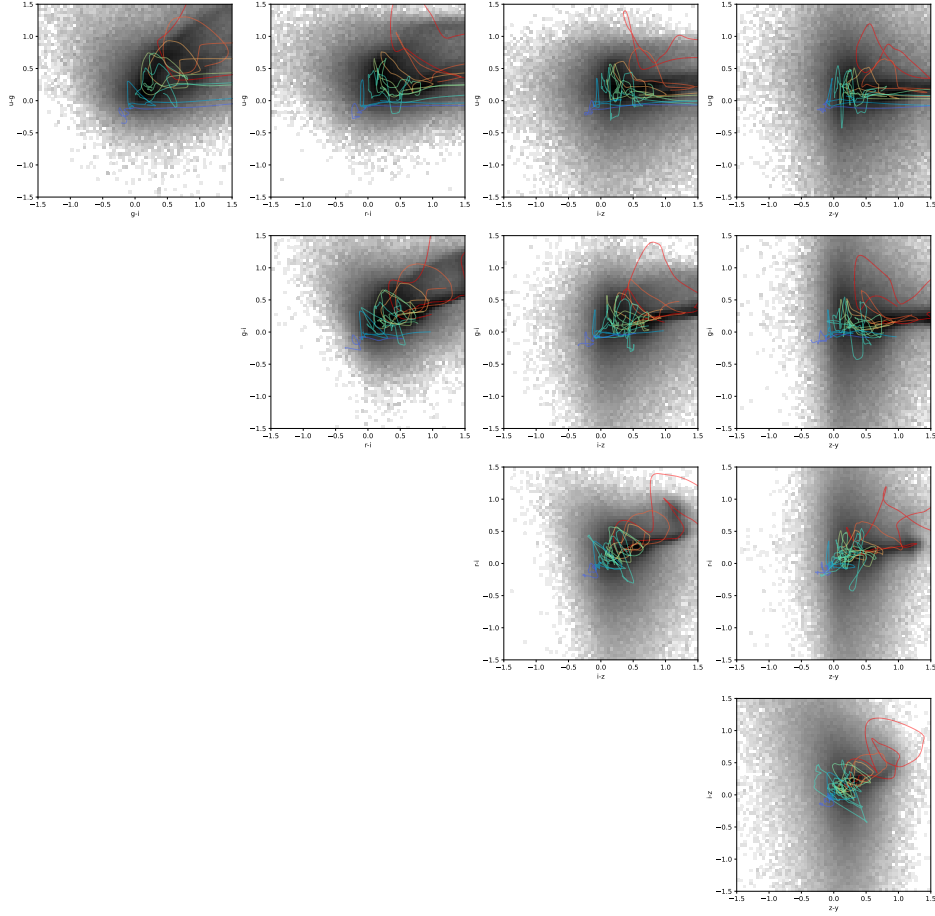


Figure 5. Color-color plots for all objects in the `dp2_v3p1_test` dataset. Colored lines represent the expected colors for the eight “CWSB” SEDs described in Sec. 3.1, and should very roughly show the predicted color evolution expected for our galaxy sample.

For the training set, the joint redshift and magnitude distribution of the 80% pre-training catalog is highly uneven, as is apparent from the survey composition in Table 1. This is a problem for photo- z training, because a machine-learning model will imprint the label distribution of the training set as an effective prior on its output, forcing the redshift solutions preferred by the abundant bright, low-redshift galaxies onto fainter galaxies whose true color-redshift mapping may differ and whose magnitude and redshift distribution is completely different (T. Zhang et al. 2026). To mitigate this, we adopt the two-dimensional down-sampling methodology of R. Zhou et al. (2021): we bin the objects of the dominant surveys on a two-dimensional grid of i -band magnitude and redshift, using a bin width of 0.1 mag and 0.1 in redshift,

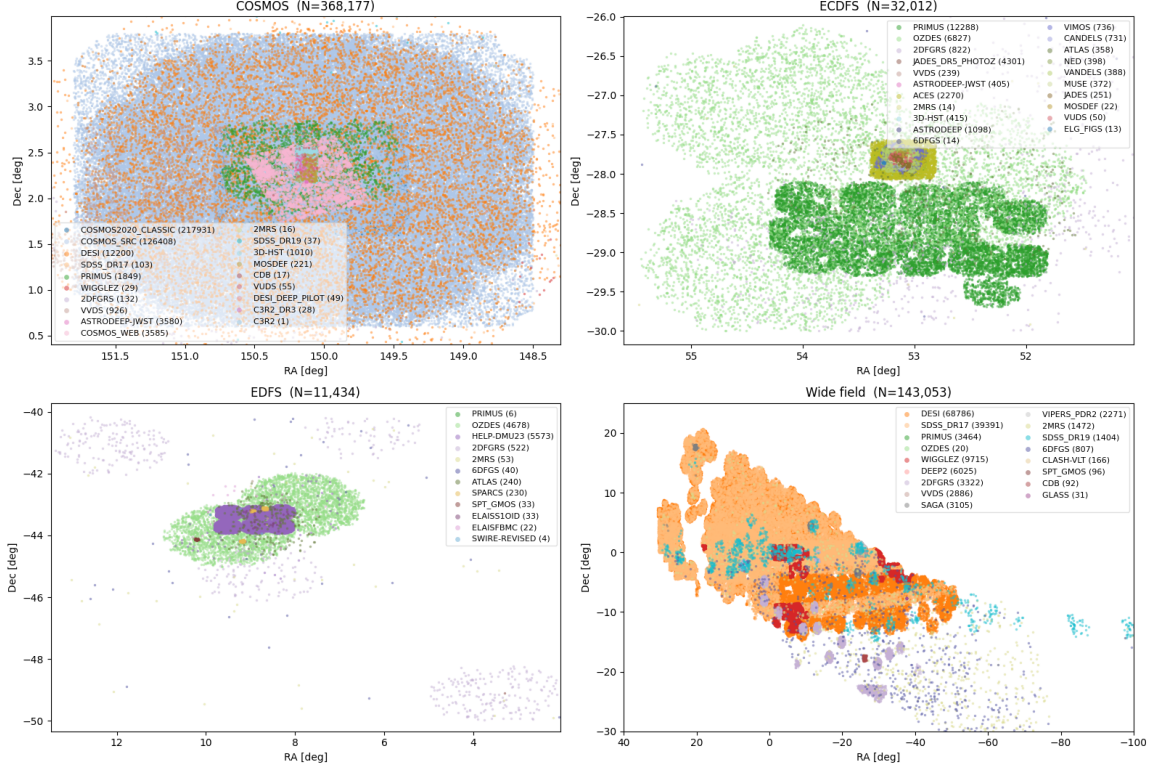


Figure 6. Right-ascension–declination distribution of the training galaxies used for photo- z estimation, with each survey shown in a distinct color (see legend). The sample combines the wide-field DESI and SDSS spectroscopic surveys with the COSMOS, ELAIS, ECDFS, and EDFs deep drilling fields.

and randomly down-sample the cells of specific surveys to a per-cell cap. Galaxies from all other surveys are kept without modification. The caps we apply are listed in Table 2.

Table 2. Per-cell caps applied in the two-dimensional magnitude–redshift down-sampling of the training set (Table 1). Each cell is $0.1 \text{ mag} \times 0.1$ in redshift; within each cell, at most the listed number of objects from the given survey are retained. Surveys not listed are kept in full.

Survey	Cap per bin
DESI DR1	100
SDSS DR17	100
2MRS	20
6dFGS	20

In Fig. 7, we show the i -band magnitude (left) and redshift (right) distributions of the resulting training set, color-coded by survey.

For the test set, we instead want the sample to correctly reflect the color–magnitude distribution of the DP2 galaxy sample, so that the reported performance metrics are representative of the objects to which the models are ultimately applied. We therefore construct the test set using a Self-Organizing Map (SOM; T. Kohonen 1982)

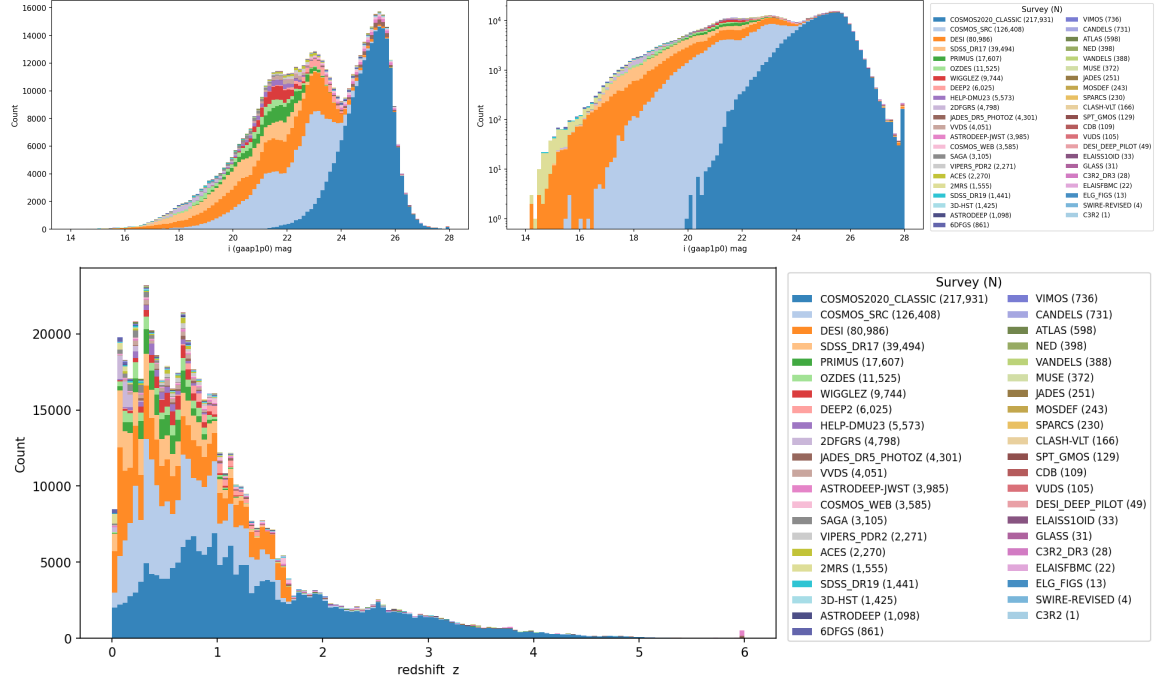


Figure 7. Magnitude (top) and redshift (bottom) distributions of the down-sampled training set, shown as stacked histograms with each survey contribution in a distinct color (see legend). The left and right panels of the magnitude figure show the same distribution on linear and logarithmic count axes, respectively.

resampling approach. In brief, we compute i -band GAAP magnitudes and the $g-r$, $r-i$, and $i-z$ colors for both the DP2 photometric sample and the pre-test spectroscopic pool, and train a SOM on a large subsample of the DP2 photometry so that its cells tessellate the color-magnitude space occupied by DP2. We then assign every DP2 object and every pre-test object to its best-matching cell, and draw from the pre-test pool a number of objects per cell proportional to the DP2 occupation fraction of that cell. The resulting test set reproduces the color-magnitude distribution of the DP2 sample, up to the coverage of the spectroscopic pool. In Fig. 8, we show the i -band magnitude (top) and redshift (bottom) distributions of the SOM-resampled test set, color-coded by survey. We note that the SOM resampling matches only the observable photometric properties, so the redshift distribution of the test set still reflects the redshift coverage of the underlying spectroscopic surveys.

In Fig. 9, we compare the joint distributions of redshift, i -band magnitude, and the $g-r$, $r-i$, and $i-z$ colors for the training set, the SOM-resampled test set, and a subsample of the DP2 photometric population. We find that the test set matches the DP2 subsample very well in color-magnitude space, confirming that the SOM resampling produces a test sample that is representative of the objects to which the models are applied. The training set, in contrast, covers a wider range in color-magnitude space, which benefits the model training by exposing the algorithms to a broader region of feature space.

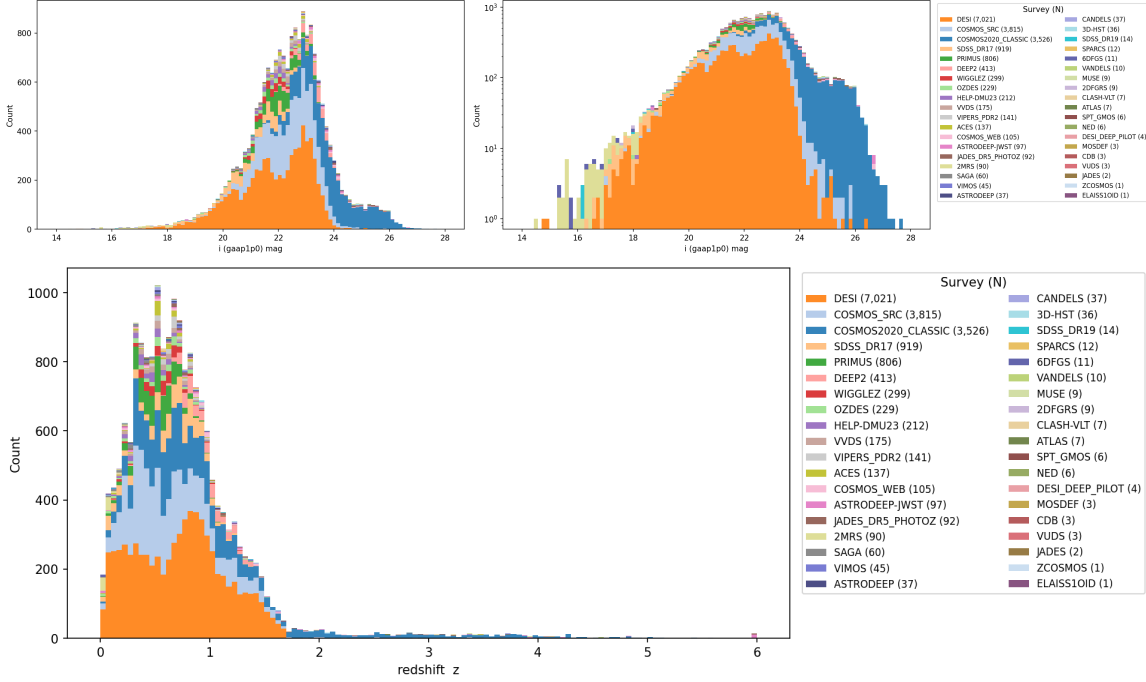


Figure 8. Magnitude (top) and redshift (bottom) distributions of the SOM-resampled test set, shown as stacked histograms with each survey contribution in a distinct color (see legend). The left and right panels of the magnitude figure show the same distribution on linear and logarithmic count axes, respectively.

3. METHODOLOGY

We used RAIL as the core tool for training and applying photo- z estimation models. We also used RAIL to evaluate the model performance through a suite of diagnostic metrics using photo- z point estimates, including redshift bias, scatter (e.g., normalized median absolute deviation), and catastrophic outlier rate and to make diagnostic plots of the algorithm’s performance. Specifically, we used the methodology described below to evaluate all the algorithms listed in Tab. 3.

As part of the Rubin photo- z roadmap process, KNN and BPZ were chosen as “testing algorithms” from the shortlist of algorithms considered in (M. L. Graham et al. 2022), the expectation is that these will be supported by project data management team for Data Preview 2, while the others will be supported by the wider community, the RAIL development team and DESC collaboration.

3.1. Template fitting based estimators

RAIL’s template-based fitting algorithms (Lephare and BPZ, see references in Tab. 3 for more details) estimate photometric redshifts by comparing observed galaxy photometry to a set of predefined theoretical or empirical galaxy templates. These templates represent a range of galaxy spectral energy distributions (SEDs) that are meant to span the expected range of galaxy types observed in the Universe. Synthetic model fluxes are computed by red-shifting each SED to a set grid of values and convolving with the Rubin filter curves, which characterizes the expected variation in observed

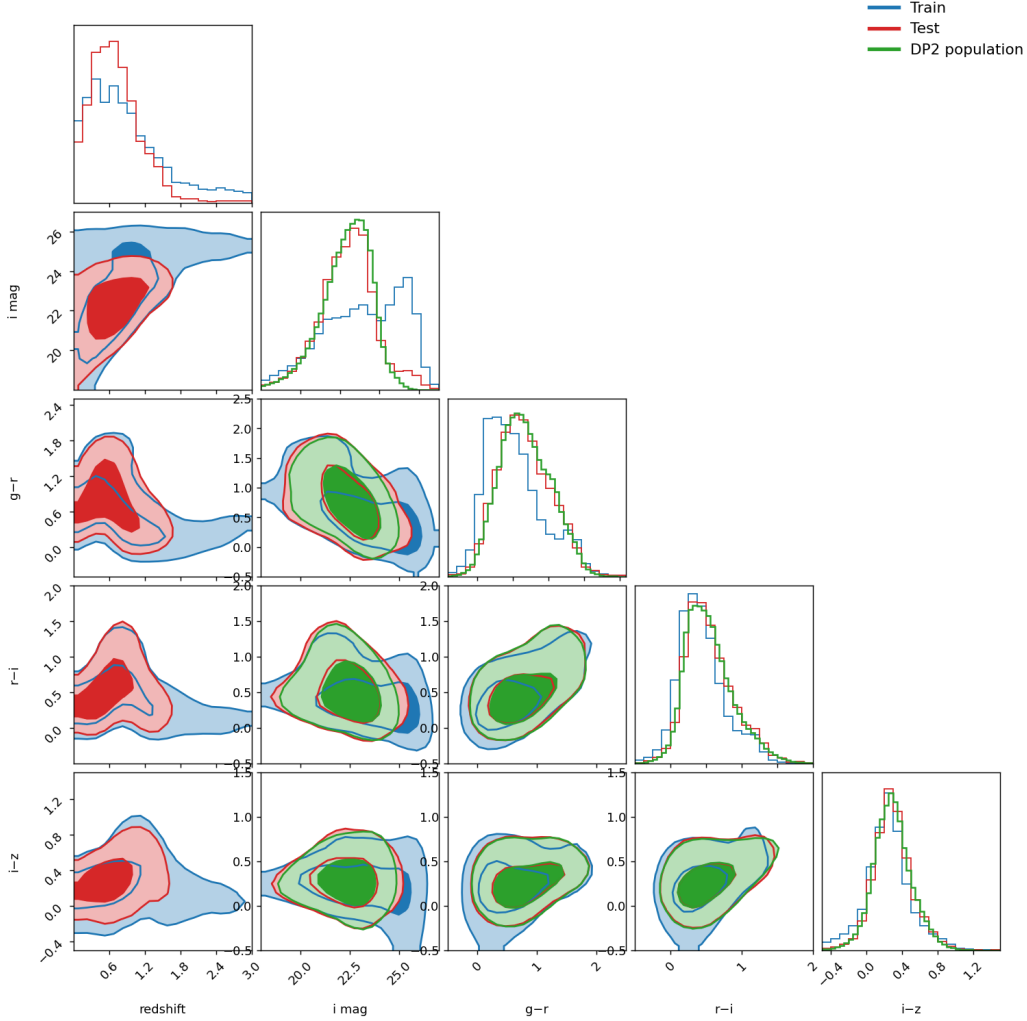


Figure 9. Corner plot comparing the training set (blue), the SOM-resampled test set (red), and a subsample of the DP2 photometric population (green) in redshift, i -band magnitude, and the $g-r$, $r-i$, and $i-z$ colors. The diagonal panels show the one-dimensional marginal distributions, and the off-diagonal panels show the two-dimensional distributions, with the filled and outlined contours enclosing the densest regions of each sample. The DP2 population has no reference redshift and therefore does not appear in the redshift panels. The test set (red) closely follows the DP2 population (green) in color-magnitude space, while the training set (blue) spans a wider range.

colors with redshift. Our algorithms calculate the chi-square/likelihood for each SED at each grid point by comparing the model fluxes in our photometric bands to the observed fluxes and uncertainties. This process enables the algorithm to determine the relative likelihood for each redshift for a given galaxy by identifying the template that best matches its observed color signature. An optional empirical Bayesian prior can also be applied to produce a posterior probability rather than a straight likelihood if information on the expected distributions are available. The training phase for these algorithms consist mostly of preparing pre-computed tables for the expected fluxes in each filter band for each SED template.

Algorithm name	Home package	Reference
Project Supported Testing		
BPZ	rail-bpz	N. Benítez (2000)
KNN	rail-sklearn	The RAIL Team et al. (2025)
Community Supported		
DNF	rail-dnf	J. De Vicente et al. (2016)
FlexZBoost	rail-flexzboost	R. Izbicki & A. B. Lee (2017)
GPz	rail-gpz-v1	I. A. Almosallam et al. (2016)
LePHARE	rail-lephare	S. Arnouts et al. (1999)
TPZ	rail-tpz	M. Carrasco Kind & R. J. Brunner (2013)

Table 3. Summary of the pre-wrapped estimators/summarizers/classifiers used in this paper and described detail in ([The RAIL Team et al. 2025](#)).

Two sets of templates are included in the `rail_base` software package and used in this note: the “CWWSB” templates that are the default templates used by `rail_bpz` and described in [D. Coe et al. \(2006\)](#), consisting of eight total SEDs, one Elliptical template, two Spiral templates, and five Irregular/Starburst template. In this note, these SEDs are used to compute synthetic colors that we compare to the observational data in Figures 4 and 5, but are otherwise not used. The second template set consists of those described in [O. Ilbert et al. \(2009\)](#), consisting of 31 synthetic SEDs (including some that are “interpolated” between adjacent templates by taking the mean of the two SEDs). These SEDs are included with `rail_lephare` as “COSMOS_mod.list” in the `lephare-data` package. These SEDs are used by both template codes `rail_bpz` and `rail_lephare`. The 31 base SEDs contain no internal dust extinction, `rail_lephare` is configured to include dust automatically, while `rail_bpz` required dust to be added explicitly, and additional SED templates that include internal extinction added to the input SED list. `LePhare` also includes a set of stellar and AGN SEDs, which are available in the `lephare-data` package.

3.2. Machine learning based estimators

RAIL’s machine learning algorithms (CMNN, DNF, FlexZBoost, GPZ, KNN, TPZ) are described in detail in the publications listed in Tab. 3. In short these algorithms attempt to perform a regression analysis to estimate the redshift from the input photometric data. It is a well-known limitation of machine learning algorithms that they exhibit biases when presented with non-representative data, or are asked to extrapolate results outside region of their training data.

3.3. Analysis framework and bookkeeping software

The `rail_projects` software package within the RAIL ecosystem provides an essential framework for managing and organizing large-scale photometric redshift estimation workflows. It acts as a project management and book-keeping tool that helps users streamline their research, especially when working with complex or large datasets, like those associated with the Rubin DP2 data. The primary focus of

`rail_projects` is to offer a systematic way to track different stages of data processing, model training, evaluation, and results across multiple experiments, ensuring that all tasks are well-documented and reproducible.

Additionally, `rail_projects` facilitates the management of large datasets by organizing data into structured directories and providing interfaces for batch processing. It allows users to scale up their work to handle not only large catalogs of objects but also multiple datasets and redshift estimation tasks. The package ensures that datasets and models are kept in sync across various stages, from raw input data to intermediate results to final outputs, and makes it easier to systematically export data coherent products.

3.4. *Data exploration and algorithm optimization*

We explored dozens of different configuration settings, covering analysis choices such as which flux measurements to use, the machine learning hyper-parameters, which sets of SED templates to use, among others. The `dp2/dp2.yaml` configuration file (see Sec. A.1) captures the set of configurations that we tested and labels each set into an analysis “flavor” to facilitate bookkeeping.

After examining the performance of the algorithms under different configurations, we settled on a set of optimized parameters for each algorithm and on the `gaap_1p0` fluxes to produce the best-effort catalogs for DP2. These optimized configurations parameters are captured in the `dp2/dp2_v3p1.yaml` configuration file.

3.5. *Production of redshifts catalogs*

We also used the `meas_pz` and `meas_pz_extensions` software packages to create photo- z catalogs in the Rubin data management framework.

Specifically, we imported trained models produced in the `rail_project` into the DP2 data Butler and produced photo- z estimates for every object in DP2 for each algorithm.

We did this for the `baseline` analysis flavor, (consisting of “out-of-the-box” algorithms, with all settings at default values) and the optimized 6 band (`dp2_optimize`) flavor defined in `dp2/dp2_v3p1.yaml`. The resulting data products are internally available via data Butlers at USDF and NERSC as described in App. B.1. As with the `rail_projects` created catalogs, it is expected some of the products will be distributed via other methods in the near future (see App. B.2 and App. B.3).

4. PERFORMANCE

We evaluated all the algorithms listed in Tab. 3 for both scientific and technical performance.

4.1. *Redshift estimator performance*

4.1.1. *Point-estimate performance*

In Fig. 10 and Fig. 11, we show the z_{best} point estimate versus the reference redshift z_{ref} for the six algorithms that go into the DP2 production: `FlexZBoost`, `KNN`,

DP2 test set | 6band

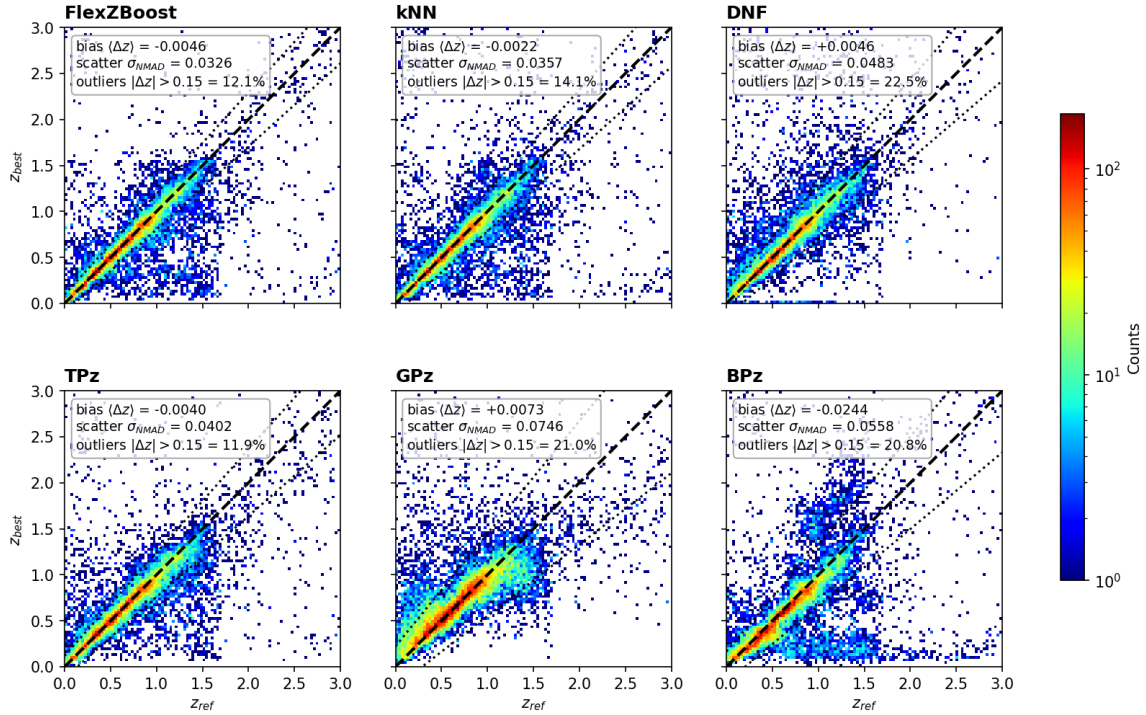


Figure 10. Two-dimensional histograms of the z_{best} point estimate versus the reference redshift z_{ref} for the six production algorithms (FlexZBoost, KNN, DNF, TPZ, GPz, BPz), evaluated on the six-band DP2 test set. The color scale shows the object counts per pixel on a logarithmic scale. The dashed line marks the one-to-one relation, and the dotted lines mark the $|\Delta z| = 0.15$ catastrophic outlier boundaries. The inset in each panel lists the bias $\langle \Delta z \rangle$, the scatter σ_{NMAD} , and the outlier fraction.

DNF, TPZ, GPz, and BPZ, evaluated on the six-band and four-band DP2 test sets, respectively. For each algorithm we quote the bias $\langle \Delta z \rangle$, the scatter σ_{NMAD} , and the catastrophic outlier fraction ($|\Delta z| > 0.15$), where $\Delta z = (z_{\text{best}} - z_{\text{ref}})/(1 + z_{\text{ref}})$. We find that all six algorithms perform normally, with the bulk of the objects concentrated along the one-to-one relation and small biases $|\langle \Delta z \rangle| \lesssim 0.025$ in both the six-band and four-band configurations. Among the six, FlexZBoost achieves the best scatter ($\sigma_{\text{NMAD}} = 0.033$ for six-band and 0.038 for four-band), while TPZ achieves the best catastrophic outlier fraction (11.9% for six-band and 14.4% for four-band). As expected, the performance of every algorithm is somewhat degraded in the four-band configuration relative to the six-band one, reflecting the reduced color information.

4.1.2. Redshift estimation quality flags

Applying a simple cut on the RMS of the $p(z)$ distribution can dramatically reduce the catastrophic outlier rate. In Fig. 12 we show how the efficiency and purity obtained on the test sample varies with a single additional cut, x in $\sigma_{p(z)} < x$. In Fig. 13 we show how applying a cut at $\sigma_{p(z)} < 0.2$ improves the scatter of the photometric redshift point-estimate versus spectroscopic redshift distribution. As fainter objects

DP2 test set | 4band

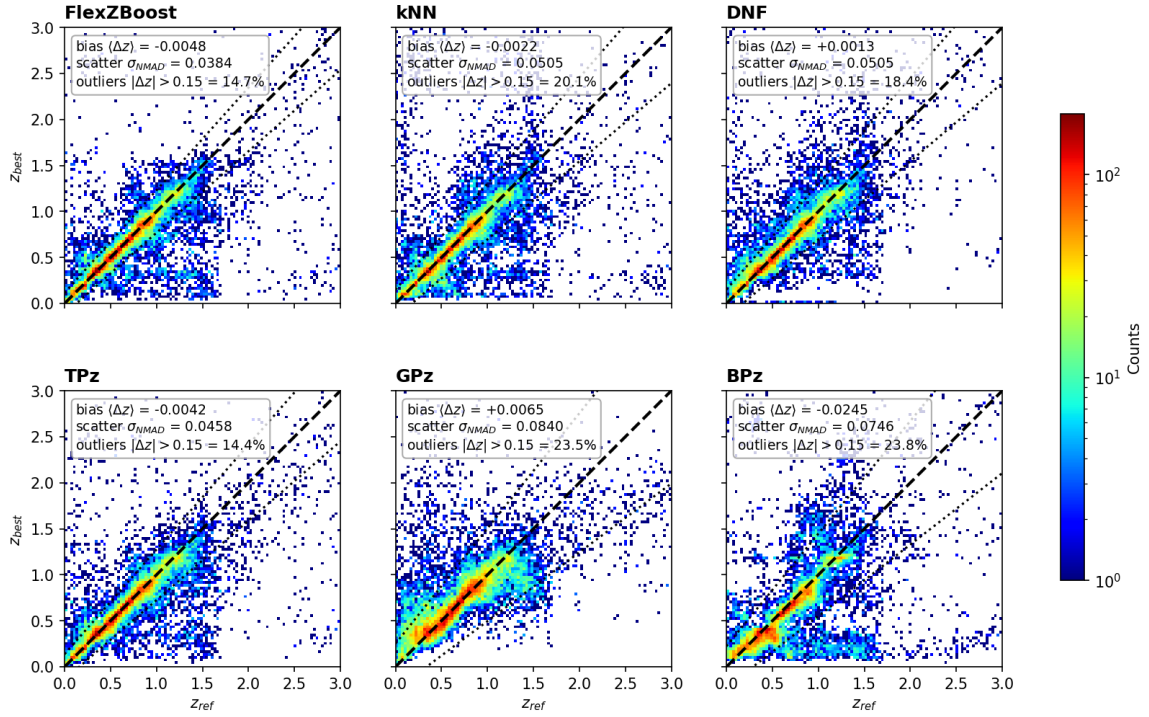


Figure 11. As Fig. 10, but for the four-band (*griz*) DP2 test set.

and objects at higher redshift often have larger $p(z)$ RMS values, cuts on RMS will likely result in relative shifts in the magnitude and redshift distribution depending on the cut value.

4.2. Redshift estimates across the footprint

In Fig. 15, we show the mean **FlexZBoost** z_{best} point estimate in each DP2 tract, projected onto a Mollweide map of the survey footprint. The map aggregates the point estimates of 651,640,921 objects across the 1391 recommended tracts. We observe that the mean estimated redshift varies substantially across the footprint, from ~ 0.4 in the shallowest tracts to ~ 1.4 in the deepest. This variation closely tracks the tract-to-tract depth variation shown in Fig. 1: deeper tracts detect fainter and, on average, higher-redshift galaxies, and therefore have a higher mean estimated redshift, whereas shallower tracts are limited to brighter, lower-redshift objects. We note that this spatial pattern is therefore primarily a selection effect driven by the varying imaging depth (and band coverage) across the footprint, rather than a physical redshift gradient, and it should be accounted for when using the DP2 photo- z catalogs for any spatially resolved analysis.

4.3. Technical performance

Tables 4 and 5 show per-object estimation times and the total training time for each algorithm. Table 6 shows the total size in DP2 of all the photo- z estimates for

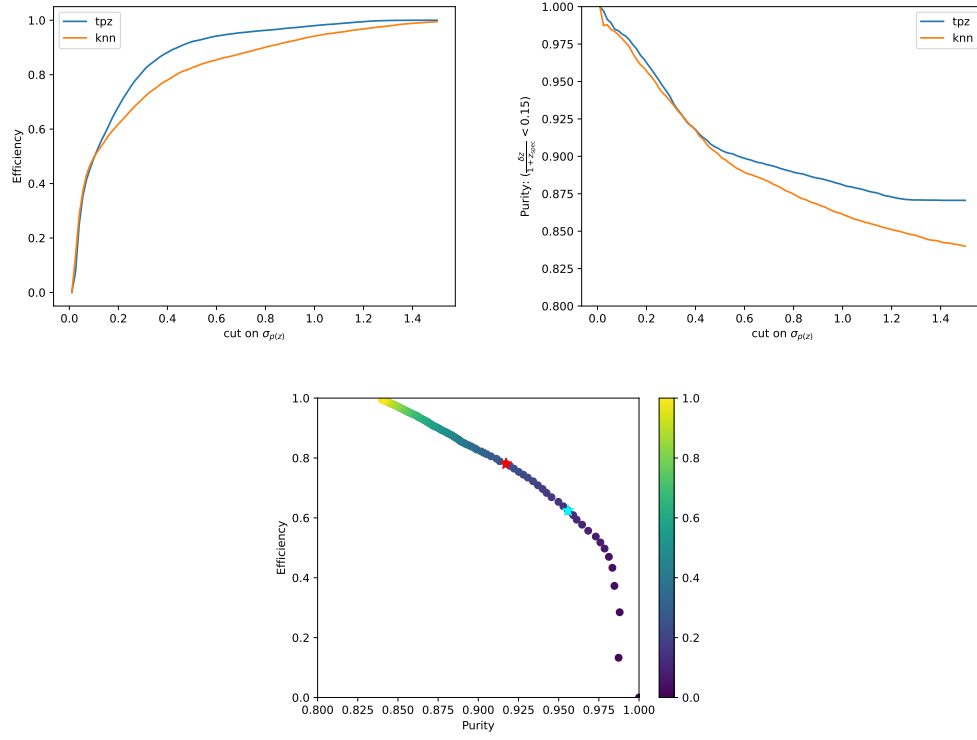


Figure 12. Efficiency (top left) and “purity”, i.e., fraction of objects with $\frac{\delta z}{1+z_{\text{spec}}} < 0.20$ (top right) versus quality cut, x in $\sigma_{p(z)} < x$. Bottom, purity version efficiency with cut value shown by the color scale. The red star shows the point for a cut value $\sigma_{p(z)} < 0.4$.

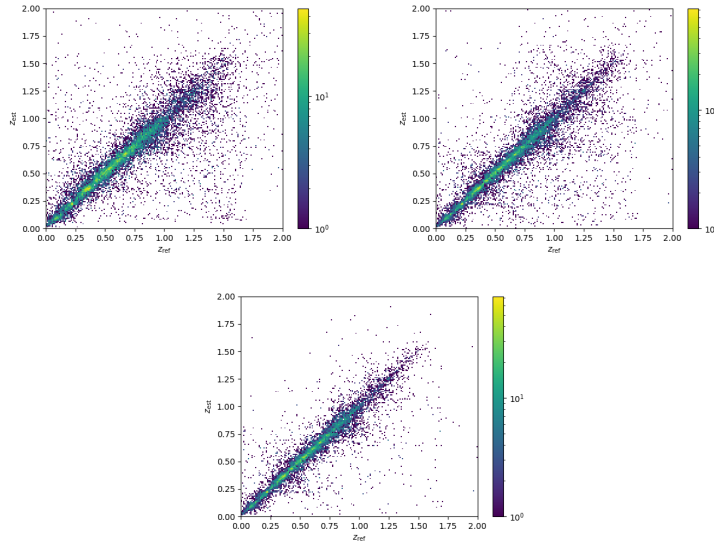


Figure 13. Estimator performance for TPZ with no cut(left) and quality cuts of $\sigma_{p(z)} < 0.4$ (center) and $\sigma_{p(z)} < 0.2$ (right).

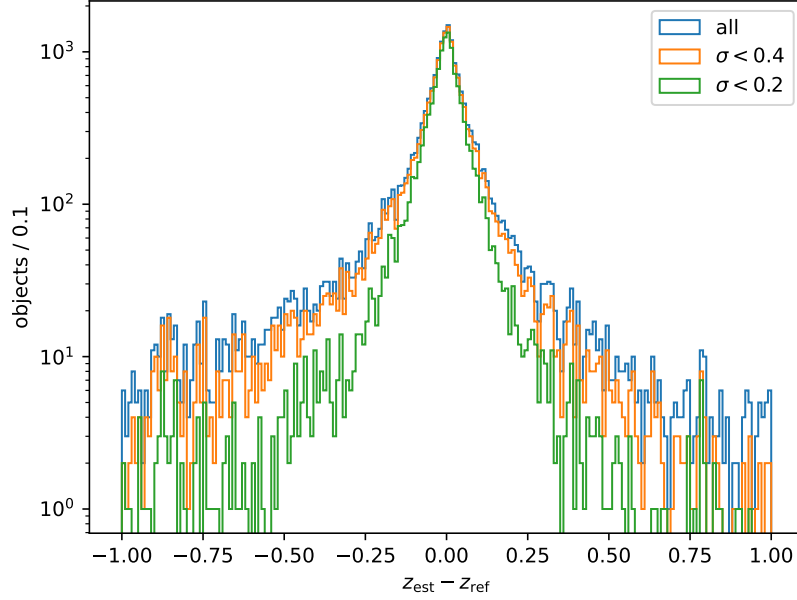


Figure 14. Estimator $\delta z = z_{\text{est}} - z_{\text{ref}}$ for TPZ for difference quality cuts.

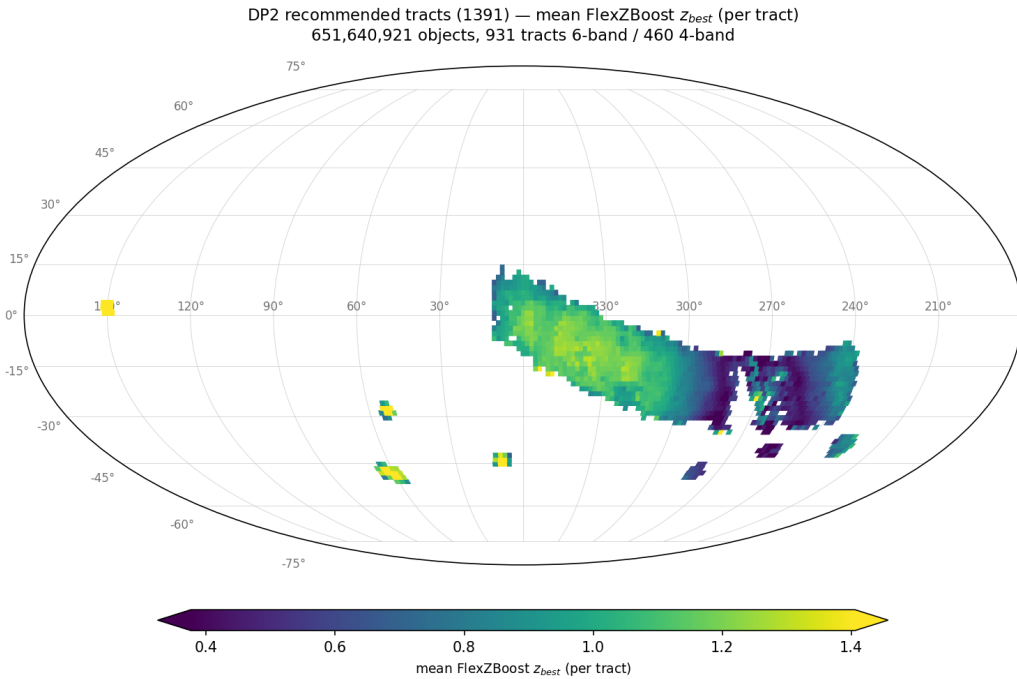


Figure 15. Mean FlexZBoost z_{best} point estimate in each DP2 tract, shown on a Mollweide projection of the footprint (astronomical convention, with right ascension increasing to the left). The color scale runs from ~ 0.4 (dark) to ~ 1.4 (bright). The map aggregates 651,640,921 objects across the 1391 recommended tracts (931 with six-band and 460 with four-band coverage). The mean estimated redshift varies across the footprint following the imaging-depth variation of Fig. 1.

Algorithm	4 Band (ms/obj)	6 Band (ms/obj)
BPZ	15.95	18.29
DNF	1.86	2.25
FlexZBoost	1.27	1.34
GPz	0.03	0.03
KNN	1.04	0.46
LePHARE	229.43	253.14
TPZ	4.17	3.88

Table 4. Per-object estimation times for each of the algorithms for both 4-band and 6-band data in ms.

Algorithm	4 Band (min)	6 Band (min)
BPZ	0.01	0.01
DNF	0.01	0.02
FlexZBoost	2.74	2.81
GPz	76.27	107.04
KNN	45.10	48.46
LePHARE	25.20	27.44
TPZ	2.68	2.71

Table 5. Times to inform (i.e., train) each of the algorithms for both 4-band and 6-band data in minutes. There were 554676 objects in the training data set.

Algorithm	4 Band (TB)	6 Band (TB)
BPZ	1.1	1.5
DNF	1.1	1.5
FlexZBoost	1.1	1.5
GPz	0.05	0.06
KNN	0.15	0.19
TPZ	2.2	2.9

Table 6. Total sizes in TB of the DP2 photo- z estimates for each algorithm. The different in sizes is almost entirely attributable to the different representations used to store the $p(z)$ distributions.

each algorithm. We note that no particular efforts to reduce the file sizes have been made, the $p(z)$ estimates are stored using 64-bit floats in uncompressed hdf5 files. We expect be able to reduce the sizes of the larger files by factors of 8 or more with minimal loss of information.

5. LIMITATIONS AND CAVEATS

The photo- z estimates described in this note were done on a best-effort basis by members of the “Photo- z Science Unit” and are not official Rubin data products. As such, the release of the various DP1-related data products comes with a number of caveats.

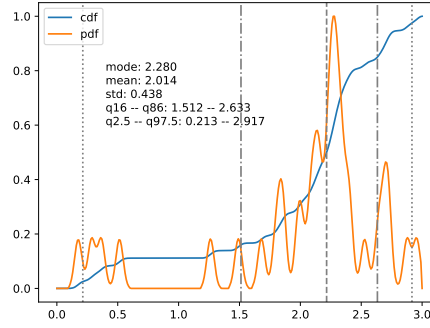


Figure 16. Example of a $p(z)$ distribution for a faint object is not detected in all bands. (In this case the object was only significantly detected in 'g' and 'r' bands).

- The size and depth of the training set seriously limits the robustness of the training past a redshift of $z \simeq 1.5$. See in particular Fig. 3.
- In general, photo- z algorithms do not perform well with objects with marginal detections or with non-detections in several bands. Simply put, the photometric uncertainties in such objects often overwhelms the information that is present and the photometric estimates are very uncertain. See Fig. 16 for an example of the PDF of such an object. Similarly, see, Fig 17 as an illustration of how the accuracy of a photo- z point-estimate degrades significantly for faint objects.
- Taking points (1) and (2) together, we do not advise to use any of the provided DP1 photo- z estimates without applying cuts on the detection significance or the width of the $p(z)$ distribution, or both.
- While we did put some work into optimizing the performance of the various algorithms, this was by no means comprehensive, and we urge caution in drawing any conclusions about the relative merits of the algorithms.
- The training and test sets were drawn from the same set of spectroscopically matched objects. As such, the training set is fairly representative of the test set. On the other hand, the full DP1 catalog does not have any spectroscopic selections applied, so the training and test sets will not be representative of the full DP1 data set.

6. SUMMARY AND CONCLUSION

This note describes an initial calculation of photometric redshifts on the Rubin Data Preview 2 dataset using the tools and workflows developed specifically for this purpose by the Photo- z Science Unit. These early results are very promising, we show that we can successfully match Rubin data with spectroscopic samples with deep six-band Rubin coverage to create a reference sample and run our software pipelines to generate a catalog of individual object redshifts consisting of both point 1D redshift PDFs and

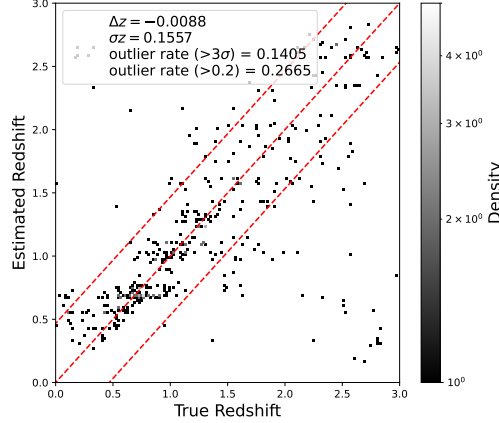


Figure 17. Point-estimates versus spectroscopic redshifts for all galaxies in the reserved test set with $m_i > 24.0$, showing the degradation in estimation performance for faint objects.

point estimate redshifts. We show results for six photo- z algorithms using a reserve set of redshifts, and see that photo- z performance is in-line with expectations both in terms of our redshift predictions and for compute times. We describe available data products and anticipated methods to access them. Overall, this is a very successful initial test of photo- z pipelines.

However, work will continue on several fronts to further optimize the performance of our redshift predictions. As mentioned earlier in the note, the definition of the true redshift reference sample used to train our algorithms has a large impact on results, being the “truth” used to define the flux/color-to-redshift mapping that is inherent to photo- z estimation. We will work to refine our reference sample definition, which objects to include/exclude based on quality flags, explore the tradeoffs of including grism and many-band photo- z estimates with larger redshift uncertainties in our training sample, and other such effects. As more and more Rubin data is taken, areal coverage is increasing, including coverage of additional deep calibration fields with existing spec- z datasets. This will naturally improve photo- z performance. Expansion of reference redshifts will also enable the development of improved object flagging for identifying which redshifts are trustworthy and those which are likely to be incorrect.

We will also continue to examine the Rubin photometry and evaluate the performance of multiple measurement algorithms. In this note we used Gaussian Aperture (Gaap) fluxes and magnitudes, as they are expected to produce consistent colors. We will undertake a thorough exploration of multiple flux measurements (e. g. cModel, Sersic, additional Gaap apertures) to test which combinations lead to the best photo- z estimates, as well as combinations like using multiple aperture magnitudes that may contain information on the galaxy light profile that could help to constrain redshift. As we are still in the engineering and testing phase of the project, there is ongoing work to characterize system performance, and improved understanding of the system will likely lead to better photo- z performance. For example, there may be some hints that u-band fluxes are overestimated and u-band uncertainties under-

estimated at faint magnitudes (see the high redshift u-g colors in Fig. 4 and figures in [Vera C. Rubin Observatory Team et al. \(2026\)](#)). DP2 data was obtained using LSSTCam during commissioning while future data releases will use LSSTCam data from survey operations, and thus performance for DR1 and beyond may be shaped by improvements in operations, data reduction and photometric calibrations.

In this analysis, while there was some optimization of code-specific parameters for the individual estimators, e. g. the number of neighbors used for KNN, the number of trees used for TPZ, it was not comprehensive. In addition, the optimal values will likely change slightly as we refine both our spectroscopic reference sample and our photometric inputs. A thorough exploration of the code parameters will happen as we converge on our final photometric and spectroscopic setup.

As stated in the introduction, we expect feedback from science users as they explore the data and discover issues that we have not anticipated, and that we will incorporate that feedback in future work.]]] Key takeaway: while very promising, these results are far from final, there is ongoing work to optimize photo- z performance. We have plans in place and will continue to refine as new data arrives, and these plans should lead to improved redshift performance.

APPENDIX

A. DATA PRODUCTS

In the course of this work, we generated several data products, including redshift estimates and a variety of others that can be used to reproduce those estimates and to facilitate thorough evaluation of the algorithm performance. These products include: configuration files (Sec. A.1), ancillary inputs (Sec. A.2), trained models (Sec. A.3), redshift estimates (Sec. A.4.1), summary statistics (Sec. A.4.2), and performance monitoring plots (Sec. A.5), all of which are essential for understanding the quality of the photometric redshift estimates and for refining the algorithms. Together, these data products provide a comprehensive framework for conducting, managing, and evaluating photometric redshift estimation workflows in Rubin DP2. They ensure that the process is not only scientifically rigorous but also organized and reproducible, enabling effective collaboration and ongoing refinement of photometric redshift techniques.

A.1. *Configuration files*

The configuration files, collected in the GitHub [rail_project_config](#) repository, provide the necessary parameters and settings to control the various stages of the redshift estimation process, are a key element of `rail_projects` workflow. These files typically include specifications for the photometric bands used (e.g., g, r, i, z), the algorithm choices (e.g., template fitting or machine learning methods), details about data pre-processing, such as feature normalization and handling of missing data. These configuration files also specify the training and validation dataset splits, hyper-parameters for machine learning models, and paths for input/output data. These files are essential for ensuring reproducibility and for sharing the exact settings used in different redshift estimation runs, enabling other researchers to replicate or extend the analysis.

A.2. *Ancillary input files*

In addition to the configuration files, we require ancillary inputs such as the galaxy spectral energy distribution (SED) templates and the filter throughputs. Galaxy SED templates are collections of theoretical or observed spectra for galaxies at different redshifts and with different properties (e.g., galaxy type, age, star formation history). These templates are used by template-fitting algorithms to model the expected galaxy colors as a function of redshift, allowing for the estimation of photometric redshifts by comparing observed colors to those predicted by the templates. Two SED sets are employed in this note, they are each described in Section 3.1.

The filter throughputs specify the characteristics of the fiducial observational filter transmission curves, and are used in Rubin DP2, including their corresponding central wavelengths. These filter curves are employed in calculating the synthetic fluxes or magnitudes expected from each of the SED templates used in template-fitting algo-

rithms, and for matching observational data to predicted theoretical values, e. g. the predicted colors shown in Figure 4.

A.3. *Estimator data models*

After training, the trained models for each photometric redshift estimation algorithm are stored as serialized files, either in Pickle (for Python-based models) or YAML (for model configurations) format. These models encapsulate the learned relationships between photometric features (such as magnitudes and colors) and redshift values, allowing them to be applied to new data for redshift estimation. These files store the final state of the model, including the weights, biases, and other learned parameters for machine learning models. In the case of template-fitting methods, the corresponding model files may include the template sets and the fitting parameters. The Pickle or YAML format ensures that the models can be easily loaded, applied to new datasets, and evaluated in future studies.

A.4. *Redshift estimates*

A.4.1. *Probability distributions*

The per-object redshift estimates generated by the photometric redshift algorithms are stored in **qp** files. These files serve as containers for storing the redshift predictions for each object in the dataset before any detailed statistical analysis or final reporting. In the **qp** files, each galaxy’s redshift estimate is stored in a format that is compatible with the algorithm providing the estimate. Specifically, for each object we store distribution $p(z)$, which, depending on the algorithm, maybe represent a posterior probability, a likelihood, a conditional likelihood, or just a hunch as to redshift of the object in question. These **qp** files are designed to be lightweight and easy to query, allowing users to quickly retrieve the redshift estimates for individual objects. The structure of **qp** files is optimized for efficient access and data retrieval, making it easier for users to process and analyze large numbers of photometric redshift estimates across large datasets like Rubin DP2. Additional information, including usage examples, about **qp** and **qp** files is available at <https://qp.readthedocs.io/en/main/>.

A.4.2. *Point estimates and summary statistics*

In addition to the raw redshift estimates, the **qp** files also store per-object point estimates in the form of an ancillary table. These estimates include crucial information about the quality and reliability of each photometric redshift estimate, such as the uncertainty in the redshift prediction (e.g., the confidence interval or standard deviation or, the likelihood score (in the case of probabilistic models). This table will eventually includes flags for identifying objects with low-confidence estimates or those that may be outliers. These summary statistics are important for evaluating the overall performance of the redshift estimation algorithms on an object-by-object basis and are often used to filter out low-quality or problematic redshift estimates before conducting larger statistical analyses.

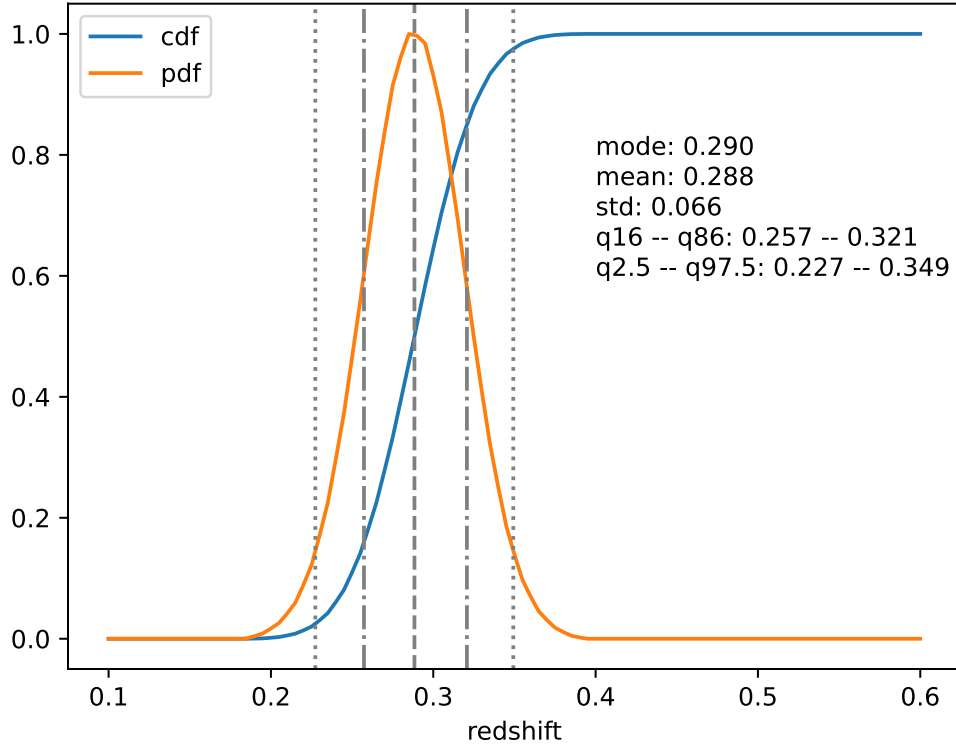


Figure 18. Single object $p(z)$ estimate, showing both the PDF, and the CDF, as well as the summary statistics described in the text.

We also generate meta catalog file that collects the Object ID, magnitude information, and point estimates. The point estimates are named by “{algorithm}_z_{point estimate name}”. Here is a list of the point estimates and summary statistics that are stored for each algorithm and the corresponding column names:

1. **mean**: is the per-object expectation of redshift given by the PDF.
2. **median**: is the 50% percentile of the PDF.
3. **mode**: is the redshift that yields maximum PDF evaluated on 301 grid points between $z = 0$ -3.
4. **err68_lower(upper)**: is the 16th (84th) percentile of the per-galaxy PDF, which correspond to the 1σ confidence interval.
5. **err95_lower(upper)**: is the 2.5th (97.5th) percentile of the per-galaxy PDF, which correspond to the 2σ confidence interval.

Fig. 18 shows an example of the $p(z)$ distribution, point-estimates and summary statistics for a single object.

Yes.

A.5. Performance Monitoring Plots

Finally, we produced standardized performance monitoring plots as part of the photometric redshift workflow. These plots provide visual representations of the model’s performance, allowing users to assess how well the redshift estimates align with the true redshifts from the spectroscopic training data. Common plots include:

- Input dataset characterization plots, as shown in Sec. 2.1.2.
- Scatter plots to visualize the relationship between the predicted and true redshifts, highlighting any systematic biases or non-linearities.
- Redshift comparison plots, e.g., the bias or width of the redshift residuals, $\frac{z_{\text{phot}} - z_{\text{spec}}}{1 + z_{\text{spec}}}$, as a function of redshift or object magnitude.

To robustly summarize the distribution of residuals while minimizing the influence of outliers, we compute biweighted statistics following a two-step procedure. First, we apply an iterative 3σ clipping to the input array using `scipy.stats.sigmaclip`, repeating the clipping process `nclip` times (default is 3). This removes extreme outliers before computing robust estimators. We then calculate the *biweight location* and *biweight scale* of the clipped subset using the `astropy.stats` functions, which yield robust analogs to the mean and standard deviation, respectively.

In addition, we compute two outlier rates: the *relative outlier rate*, defined as the fraction of values in the original sample that deviate from zero by more than $3\times$ the biweight scale, and the *absolute outlier rate*, defined as the fraction of values exceeding a fixed threshold set by `self.config.abs_out_thresh`. This framework provides both robust central moments and a diagnostic of extreme deviations in the full sample.

Examples of the two latter types of plots for the BPZ and KNN algorithms are shown in Sec. 4

B. DATA DISTRIBUTION

To support both the Rubin community and the DESC, we will distribute these data products in several different ways:

1. Via the Rubin Data Butler (see Sec. B.1).
2. Via the Photo-z Server (see Sec. B.2).
3. Via LSDB (see Sec. B.3)

In each of these distribution mechanisms there are a number of metadata that are used as fields in defining specific data products. In particular:

- **algo**: specifies a particular photo- z estimation algorithm (see Tab. 3),
- **selection**: specifies a particular data selection (see Sec. 2.1.1),

Dataset type	Description
Collection: <code>pretrained_models/pz/DP2/{selection}/{flavor}</code>	
<code>pzModel_{algo}</code>	Estimator models (see A.3)
Collection: <code>LSSTComCam/runs/DRP/DP2/pz/DM-51523/{selection}/{flavor}/{version}</code>	
<code>pz_estimate_{algo}</code>	QP ensembles (see A.4.1)
<code>pz_{algo}_config</code>	Configuration parameters (see A.1)
<code>pz_{algo}_log</code>	Log files
<code>pz_{algo}_metadata</code>	Processing metadata

Table 7. Photo- z related objects stored in the Rubin Data Butler. Note that `{selection}` describes the selection applied to the trained data set.

- **flavor:** specifies a particular set of configuration parameters,
- **dataset:** specifies a particular dataset (e.g. `dp2`).
- **version:** specifies a processing version (e.g., `v1`, `v2`).

B.1. *Distribution via the Rubin Data Butler*

Creation of photometric redshift estimates using RAIL in the Rubin DM framework and distribution via the Rubin Data Butler is supported the `meas_pz` software package for DM supported algorithms and by the `meas_pz_extensions` software package for the community-supported algorithms described in this note.

For DP1, the following objects stored in the data butler at USDF and NERSC are listed in Tab. 7.

B.2. *Distribution via the Photo- z Server*

B.2.1. *Introducing the PZ Server*

The [PZ Server](#) is an online tool to support photo- z related science in LSST. During the survey operations (DR1 to DR11), the PZ Server will provide access to curated official photo- z -related data products vetted by Rubin DM and data products created by the PZ Server team as part of the Brazilian in-kind contribution program.

For DP1 and DP2, preliminary data products generated by the previous [E. Charles et al. \(2025\)](#) and this work are available under the "User-generated Data Products" section with suffixes "SITCOMTN-154" and "RTN-124", respectively, in the product names. To facilitate users to find data products of interest, there is also a [documentation page](#) with all the links and useful information organized by DP and product type.

In addition, similar to the RSP, the PZ Server provides an API interface that enables users to access data through Python scripts from any location, including RSP notebook aspect, provided they know the product name as registered on the server. If it is the first time using the library, it must be installed via `pip` in the terminal or in a notebook cell: `pip install --no-cache-dir pzserver`

Reference Redshift Catalogs			
Product Name	pzserver library identifier	Source Catalogs	Redshifts
REF_Z_CLEAN_NO_DESI_LSS	318_ref_z_clean_no_desi_lss	37	6,616,173
DESI DR1 LITE	314_desi_dr1_lite	1	28,425,963

Table 8. Redshift reference catalogs described in Section 2.2, removed duplicates and objects with bad quality (`homogenized_flag < 3`). The links to all the original reference catalogs are available at the PZ Server Data documentation page.

Training and Test Sets		
Product Name	pzserver library identifier	Redshifts
MATCHED_CLEAN RTN-124	417_matched_clean_rtn124	1,194,088
MATCHED_DESI RTN-124	418_matched_desi_rtn124	1,544,298
CLIPPED_TRAIN RTN-124	415_clipped_train_rtn124_parquet	554,676
SOM_TEST RTN-124	412_som_test_rtn124_parquet	18,400

Table 9. Raw matched redshift catalogs, clipped train and test sets described in Section 2.2

Then, the `PzServer` class opens the remote connection to the PZ Server database. An access token is required for authentication. The token can be generated by users on the PZ Server website (top right corner menu on the home page).

```
from pzserver import PzServer
pz_server = PzServer(token="<paste your access token here>")
```

To display the product metadata and download it to the local working directory (if not in a Jupyter notebook, replace `display` for `get`).

```
pz_server.display_product_metadata("<product_id>")
pz_server.download_product("<product_id>", save_in=".")
```

Alternatively, it is possible to load a table directly into memory as Astropy Table. For instance, to get the Combined Redshift Reference Catalog described above, converted into a Pandas DataFrame:

```
z_catalog = pz_server.get_product("318_ref_z_clean_no_desi_lss").to_pandas()
```

A [tutorial notebook](#) with examples for all `pzserver` methods is available on the `pzserver` library’s repository on GitHub.

B.2.2. RTN-124 Data Products

Links to the data products pages and `pzserver` library identifiers are provided in Tables 8 to 12.

B.3. Distribution via LSDB

The DP2 data is published in Hierarchical Adaptive Tiling Scheme (HATS) format on the Rubin Science Platform (RSP). HATS data can be read with any Parquet

Gold baseline	
Product Name	pzserver library identifier
BPZ model DP2 gold baseline RTN-124	419_bpz_model_dp2_gold_baseline_rtn124
DNF model DP2 gold baseline RTN-124	420_dnf_model_dp2_gold_baseline_rtn124
FlexZBoost model DP2 gold baseline RTN-124	421_flexzboost_model_dp2_gold_baseline_rtn124
GPz model DP2 gold baseline RTN-124	422_gpz_model_dp2_gold_baseline_rtn124
kNN model DP2 gold baseline RTN-124	423_knn_model_dp2_gold_baseline_rtn124
LePhare model DP2 gold baseline RTN-124	424_lephare_model_dp2_gold_baseline_rtn124
TPZ model DP2 gold baseline RTN-124	425_tpz_model_dp2_gold_baseline_rtn124

Table 10. Gold baseline training results.

Gold 4 bands	
Product Name	pzserver library identifier
BPZ model DP2 gold 4 bands RTN-124	426_bpz_model_dp2_gold_4_bands_rtn124
DNF model DP2 gold 4 bands RTN-124	427_dnf_model_dp2_gold_4_bands_rtn124
FlexZBoost model DP2 gold 4 bands RTN-124	428_flexzboost_model_dp2_gold_4_bands_rtn124
GPz model DP2 gold 4 bands RTN-124	429_gpz_model_dp2_gold_4_bands_rtn124
kNN model DP2 gold 4 bands RTN-124	430_knn_model_dp2_gold_4_bands_rtn124
LePhare model DP2 gold 4 bands RTN-124	431_lephare_model_dp2_gold_4_bands_rtn124
TPZ model DP2 gold 4 bands RTN-124	432_tpz_model_dp2_gold_4_bands_rtn124

Table 11. Gold 4 bands training results.

Gold 6 bands	
Product Name	pzserver library identifier
BPZ model DP2 gold 6 bands RTN-124	433_bpz_model_dp2_gold_6_bands_rtn124
DNF model DP2 gold 6 bands RTN-124	434_dnf_model_dp2_gold_6_bands_rtn124
FlexZBoost model DP2 gold 6 bands RTN-124	435_flexzboost_model_dp2_gold_6_bands_rtn124
GPz model DP2 gold 6 bands RTN-124	436_gpz_model_dp2_gold_6_bands_rtn124
kNN model DP2 gold 6 bands RTN-124	437_knn_model_dp2_gold_6_bands_rtn124
LePhare model DP2 gold 6 bands RTN-124	438_lephare_model_dp2_gold_6_bands_rtn124
TPZ model DP2 gold 6 bands RTN-124	440_tpz_model_dp2_gold_6_bands_rtn124

Table 12. Gold 6 bands training results.

reader, but the LSDB package (N. Caplar et al. 2025) enables loading large datasets with limited memory, performing efficient all-sky analysis, and fast cross matching to other surveys like DESI DR1 and Gaia DR3.

The photo- z point estimates (mean, mode, median and “best”), 1 and 2 sigma redshift intervals for DP2 object catalog are served as tabular catalogs in the HATS format. The location on the RSP for this catalog is `/rubin/lsdb_data/dp2/object_photoz`.

ACKNOWLEDGMENTS

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.

Facilities: Rubin:Simonyi,Rubin:USDAC

Software: LSST Science Pipelines ([Rubin Observatory Science Pipelines Developers et al. 2026](#))

REFERENCES

- Almosallam, I. A., Jarvis, M. J., & Roberts, S. J. 2016, MNRAS, 462, 726, doi: [10.1093/mnras/stw1618](#)
- Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540, doi: [10.1046/j.1365-8711.1999.02978.x](#)
- Benítez, N. 2000, ApJ, 536, 571, doi: [10.1086/308947](#)
- Caplar, N., Beebe, W., Branton, D., et al. 2025, arXiv e-prints, arXiv:2501.02103, doi: [10.48550/arXiv.2501.02103](#)
- Carrasco Kind, M., & Brunner, R. J. 2013, MNRAS, 432, 1483, doi: [10.1093/mnras/stt574](#)
- Charles, E., Crenshaw, J. F., Zhang, T., et al. 2025, Initial studies of photometric redshifts with LSSTComCam from DP1, Commissioning Technical Note SITCOMTN-154, NSF-DOE Vera C. Rubin Observatory, doi: [10.71929/rubin/2571480](#)
- Coe, D., Benítez, N., Sánchez, S. F., et al. 2006, AJ, 132, 926, doi: [10.1086/505530](#)
- De Vicente, J., Sánchez, E., & Sevilla-Noarbe, I. 2016, MNRAS, 459, 3078, doi: [10.1093/mnras/stw857](#)
- Graham, M. L., Bosch, J., Guy, L. P., & The DM System Science Team. 2022, A Roadmap to Photometric Redshifts for the LSST Object Catalog, Data Management Technical Note DMTN-049, NSF-DOE Vera C. Rubin Observatory. <https://dmtn-049.lsst.io/>
- Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236, doi: [10.1088/0004-637X/690/2/1236](#)
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: [10.3847/1538-4357/ab042c](#)
- Izbicki, R., & Lee, A. B. 2017, Electronic Journal of Statistics, 11, 2800, doi: [10.1214/17-EJS1302](#)
- Jenness, T., Bosch, J. F., Salnikov, A., et al. 2022, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 12189, Software and Cyberinfrastructure for Astronomy VII, 1218911, doi: [10.1117/12.2629569](#)
- Kohonen, T. 1982, Biological Cybernetics, 43, 59, doi: [10.1007/BF00337288](#)
- NSF-DOE Vera C. Rubin Observatory. 2026a, Legacy Survey of Space and Time Data Preview 2 [Data set], NSF-DOE Vera C. Rubin Observatory, doi: [10.71929/rubin/3382528](#)

- NSF-DOE Vera C. Rubin Observatory.
2026b, Legacy Survey of Space and
Time Data Preview 2: object dataset
type [Data set], NSF-DOE Vera C.
Rubin Observatory,
doi: [10.71929/rubin/3382539](https://doi.org/10.71929/rubin/3382539)
- Rubin Observatory Science Pipelines
Developers, AlSayyad, Y., Axelrod,
T. S., et al. 2026, The LSST Science
Pipelines Software: Optical Survey
Pipeline Reduction and Analysis
Environment, Project Science Technical
Note PSTN-019, NSF-DOE Vera C.
Rubin Observatory,
doi: [10.71929/rubin/2570545](https://doi.org/10.71929/rubin/2570545)
- Schlafly, E. F., & Finkbeiner, D. P. 2011,
ApJ, 737, 103,
doi: [10.1088/0004-637X/737/2/103](https://doi.org/10.1088/0004-637X/737/2/103)
- The RAIL Team, van den Busch, J. L.,
Charles, E., et al. 2025, arXiv e-prints,
arXiv:2505.02928,
doi: [10.48550/arXiv.2505.02928](https://doi.org/10.48550/arXiv.2505.02928)
- Vera C. Rubin Observatory Team. 2026,
The Vera C. Rubin Observatory Data
Preview 2, Technical Note RTN-115,
NSF-DOE Vera C. Rubin Observatory,
doi: [10.71929/rubin/3377440](https://doi.org/10.71929/rubin/3377440)
- Vera C. Rubin Observatory Team,
Acero-Cuellar, T., Acosta, E., et al.
2026, AJ, 171, 360,
doi: [10.3847/1538-3881/ae521f](https://doi.org/10.3847/1538-3881/ae521f)
- Zhang, T., Charles, E., Crenshaw, J. F.,
et al. 2026, MNRAS, 550, stag1012,
doi: [10.1093/mnras/stag1012](https://doi.org/10.1093/mnras/stag1012)
- Zhou, R., Newman, J. A., Mao, Y.-Y.,
et al. 2021, MNRAS, 501, 3309,
doi: [10.1093/mnras/staa3764](https://doi.org/10.1093/mnras/staa3764)