

Full length article

Photometric redshifts for the S-PLUS Survey: Is machine learning up to the task?

E.V.R. Lima^{a,*}, L. Sodré Jr.^a, C.R. Bom^{b,c}, G.S.M. Teixeira^b, L. Nakazono^a, M.L. Buzzo^a, C. Queiroz^{d,e}, F.R. Herpich^a, J.L. Nilo Castellon^{f,g}, M.L.L. Dantas^h, O.L. Dors Jr.ⁱ, R.C. Thom de Souza^{j,k,l}, S. Akas^m, Y. Jiménez-Tejaⁿ, A. Kanaan^o, T. Ribeiro^p, W. Schoennell^q

^a Universidade de São Paulo, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Rua do Matão 1226, CEP 05508-090, São Paulo, SP, Brazil

^b Centro Brasileiro de Pesquisas Físicas, Rua Dr. Xavier Sigaud 150, CEP 22290-180, Rio de Janeiro, RJ, Brazil

^c Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rodovia Mário Covas, lote J2, quadra J Distrito Industrial de Itaguaí, CEP 23810-000, Itaguaí, RJ, Brazil

^d Instituto de Física, Universidade de São Paulo, Rua do Matão 1371, CEP 05508-090, São Paulo, SP, Brazil

^e Departamento de Astronomia, Instituto de Física, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, CEP 91501-970, Porto Alegre, RS, Brazil

^f Dirección de Investigación y Desarrollo, Universidad de La Serena, Avenida Juan Cisternas 1200, La Serena, Chile

^g SIGMA Space Science and Technologies, CL-1700000, La Serena, Chile

^h Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences, ul. Bartycka 18, 00-716, Warsaw, Poland

ⁱ Universidade do Vale do Paraíba, Av. Shishima Hifumi 2911, CEP 12244-000, São José dos Campos, SP, Brazil

^j Universidade Estadual de Maringá, Programa de Pós-graduação em Ciência da Computação, Av. Colombo 5790, CEP 87020-900, Maringá, PR, Brazil

^k Universidade Estadual de Maringá, Programa de Pós-graduação em Engenharia de Produção, Av. Colombo 5790, CEP 87020-900, Maringá, PR, Brazil

^l Universidade Federal do Paraná, Campus Jandaia do Sul, Rua Doutor João Maximiano 426, CEP 86900-900, Jandaia do Sul, PR, Brazil

^m Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, GR 15236 Penteli, Greece

ⁿ Instituto de Astrofísica de Andalucía - CSIC, Glorieta de la Astronomía s/n, E-18008 Granada, Spain

^o Departamento de Física, Universidade Federal de Santa Catarina, CEP 88040-900, Florianópolis, SC, Brazil

^p NOAO, 950 North Cherry Ave. Tucson, AZ 85719, United States

^q GMT Corporation, N. Halstead Street 465, Suite 250, Pasadena, CA 91107, United States

ARTICLE INFO

Article history:

Received 12 March 2021

Accepted 21 October 2021

Available online 9 November 2021

Keywords:

Galaxies: distances and redshifts

Methods: data analysis

Software: development

Techniques: photometric

Surveys

ABSTRACT

The Southern Photometric Local Universe Survey (S-PLUS) is a novel project that aims to map the Southern Hemisphere using a twelve filter system, comprising five broad-band SDSS-like filters and seven narrow-band filters optimized for important stellar features in the local universe.

In this paper we use the photometry and morphological information from the first S-PLUS data release (S-PLUS DR1) cross-matched to unWISE data and spectroscopic redshifts from Sloan Digital Sky Survey DR15. We explore three different machine learning methods (Gaussian Processes with GPz and two Deep Learning models made with TensorFlow) and compare them with the currently used template-fitting method in the S-PLUS DR1 to address whether machine learning methods can take advantage of the twelve filter system for photometric redshift prediction. Using tests for accuracy for both single-point estimates such as the calculation of the scatter, bias, and outlier fraction, and probability distribution functions (PDFs) such as the Probability Integral Transform (PIT), the Continuous Ranked Probability Score (CRPS) and the Odds distribution, we conclude that a deep-learning method using a combination of a Bayesian Neural Network and a Mixture Density Network offers the most accurate photometric redshifts for the current test sample. It achieves single-point photometric redshifts with scatter (σ_{NMAD}) of 0.023, normalized bias of -0.001, and outlier fraction of 0.64% for galaxies with r_{auto} magnitudes between 16 and 21.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Recent developments in both computer science and engineering allowed astronomers to create projects with sky area coverage and photometric depths never seen before. Current surveys, such

* Corresponding author.

E-mail address: erik.vini@usp.br (E.V.R. Lima).

as the Dark Energy Survey (DES, [S nchez, 2010](#)) and the Kilo-Degree Survey (KiDS, [Hildebrandt et al., 2016](#)), and future wide field surveys, such as the Vera C. Rubin Observatory (LSST, [Ivezic et al., 2008](#)), Nancy Grace Roman Space Telescope ([Gehrels and Spergel, 2015](#)) and Euclid survey ([Laureijs et al., 2011](#)), among others, are gathering and/or will be able to gather information on hundreds of millions of objects, and more efficient methods to determine redshifts need to be developed. Due to time and cost constraints, and limited spectroscopic capabilities for faint objects, the determination of spectroscopic redshifts (z_{spec}) in the current generation of spectroscopic surveys for most objects in the aforementioned surveys is unfeasible. Consequently, a new way to estimate distances to celestial objects has been developed in the last decades which rely only on much faster and cheaper photometric observations. These estimates are named photometric redshifts (z_{phot}).

Several ongoing projects, such as the Advanced Large Homogeneous Area Medium Band Redshift Astronomical Survey (ALHAMBRA, [Molino et al., 2014](#)), the Cosmic Evolution Survey (COSMOS, [Laigle et al., 2016](#)), Survey for High- z Absorption Red and Dead Sources (SHARDS, [Barro et al., 2019](#)), Physics of the Accelerating Universe Survey (PAUS, [Alarcon et al., 2021](#)), Javalambre Physics of the Accelerating Universe Survey (J-PAS, [Benitez et al. 2014](#), [Bonoli et al. 2020](#)), the Javalambre-Photometric Local Universe Survey (J-PLUS, [Cenarro et al., 2019](#)), and the Southern Photometric Local Universe Survey (S-PLUS, [Mendes de Oliveira et al., 2019](#)), are able to obtain accurate photometric redshifts due to the use of narrow-band filters. The filter systems of these surveys provide an increase in z_{phot} accuracy by a factor of 4 when compared to the use of broad-band filters only ([Molino et al., 2020](#)).

Redshifts are a measurement of fundamental importance for astronomy, allowing researchers to have insight on different fields of cosmology and astrophysics. Although z_{phot} s are not as precise as the spectroscopic counterpart, they still allow several studies, and previous works have shown what accuracy is necessary for different science cases. For instance, an error of less than $0.004(1 + z_{\text{spec}})$ is necessary to constrain the Baryonic Acoustic Oscillations (BAO) ([Ben tez et al. 2009](#), [Weinberg et al. 2013](#)), while photometric redshifts with $< 1\%$ error can also be used on weak-lensing studies ([Hildebrandt et al. 2012](#), [Kilbinger et al. 2013](#), [Rozo et al. 2009](#)). This approach can also be used to find overdensities of faint galaxies ([Pharo et al., 2018](#)), or to constrain cosmological parameters using clusters ([Battye and Weller, 2003](#)).

Photometric redshifts can be estimated using three different approaches: template fitting, machine learning, and hybrid methods. Template fitting (TF) methods make use of the spectral energy distribution of an object and fit it to a template library, which can be synthetic or empirical, containing spectra of different types of galaxies at different redshifts. Methods that apply this approach are BPZ ([Benitez, 2000](#)), Le Phare ([Arnouts et al. 1999](#), [Ilbert et al. 2006](#)), and EAZY ([Brammer et al., 2008](#)), among others.

Machine learning (ML) methods are based on the assumption that colours and/or magnitudes of galaxies are dependent of their redshift. Using some sort of data (tabular or images) split into training, validation and testing samples, it is possible to train a model to map a relation between these quantities. After a model is trained, it can be applied to a sample of galaxies to estimate their redshifts. Methods that use this approach are GPz ([Almosallam et al., 2016](#)), TPZ ([Kind and Brunner, 2013](#)), ANNz ([Firth et al., 2003](#)), ANNz2 ([Sadeh et al., 2016](#)) and deep learning models created with TensorFlow ([Abadi et al., 2016](#)), PyTorch ([Paszke et al., 2017](#)), or Theano ([Al-Rfou et al., 2016](#)).

Both TF and ML methods have their advantages and disadvantages. The former needs, for example, well-calibrated, unbiased

templates, as well as explicit assumptions about the dust extinction, and are more computationally expensive. The latter requires a training sample with a reasonable number of objects that is both complete and representative of the sample of interest and it is reliable in the range of input parameters of the training sample only. To explore the advantages of both approaches, a few hybrid methods have been developed, combining features from both procedures, such as data augmentation using model templates for machine learning codes ([Hoyle et al., 2015](#)), improving template fitting codes by the use of unsupervised self-organizing maps ([Speagle and Eisenstein, 2017](#)), and the ZEBRA photometric redshift code, which uses a training set of spectroscopic redshifts to improve the robustness of template fitting z_{phot} estimates ([Feldmann et al., 2006](#)).

In this work, we explore the performance of the machine learning methods for the determination of photometric redshifts for galaxies using the information of the twelve bands of the S-PLUS filter system, the ellipticity, the maximum surface brightness, and the full width at half maximum from its catalogue. We have chosen to compare Gaussian Processes with GPz and two Deep Learning models made with TensorFlow. We also compare our results with the template fitting method currently used for the determination of z_{phot} s in S-PLUS, the BPZ2 code ([Molino et al., 2020](#)). With this work, we examine how these machine learning methods perform with the inclusion of narrow-band filters and non-photometric information.

This paper is organized as follows: in Section 2 we describe the S-PLUS survey and the data that will be used in this study. In Section 3 we describe briefly the methods and the metrics applied to evaluate the quality of the estimates. In Section 4 we present our results and discussions and give our final remarks in Section 5.

2. Data

In this work we used the photometric data from the S-PLUS Data Release 1 ([Mendes de Oliveira et al., 2019](#)) and the unWISE project ([Lang, 2014](#)), cross-matched to the spectroscopic information from SDSS Data Release 15 ([Blanton et al., 2017](#)). Further details about their cross-matching are given in Section 2.1. The Southern Photometric Local Universe Survey ([Mendes de Oliveira et al., 2019](#)) is an ongoing project which aims to cover over 9000 square degrees in the southern sky, which will overlap with existing and future surveys (like DES, KiDS and LSST), using the same twelve bands adopted by the J-PLUS survey ([Cenarro et al., 2019](#)), covering the wavelength range between 3700   and 9000  . The S-PLUS survey started in August 2016 and is expected to finish in 2025 using a robotic 0.8 metres diameter telescope located at Cerro Tololo, Chile, named T80-S. The camera used in this project has a 2 square degrees field-of-view and is equipped with a 9k   9k CCD with a 0.55 arcsec/pixel scale. The photometric system has 5 SDSS-like broadband filters (u_{JVA} , g_{SDSS} , r_{SDSS} , i_{SDSS} and z_{SDSS}) and 7 narrow-band filters (J0378, J0395, J0410, J0430, J0515, J0660 and J0861) designed specifically to explore important spectral features in the local universe, such as [OII] 3727,3729  , Ca H+K, H , G-band, Mgb triplet, H , and the Ca triplet, respectively ([Gruel et al., 2012](#); [Mar n-Franch et al., 2012](#)). This number of filters offers a great opportunity to estimate high-quality photometric redshifts in the local universe. The filter transmission curves¹ are presented in Fig. 1.

The S-PLUS DR1 covers 336 square degrees in Stripe-82. This region overlaps with other surveys, presenting a great opportunity to train supervised machine learning models due to the availability of spectroscopic redshifts. The typical depth of the

¹ Obtained from <http://www.splus.iag.usp.br/instrumentation/>.

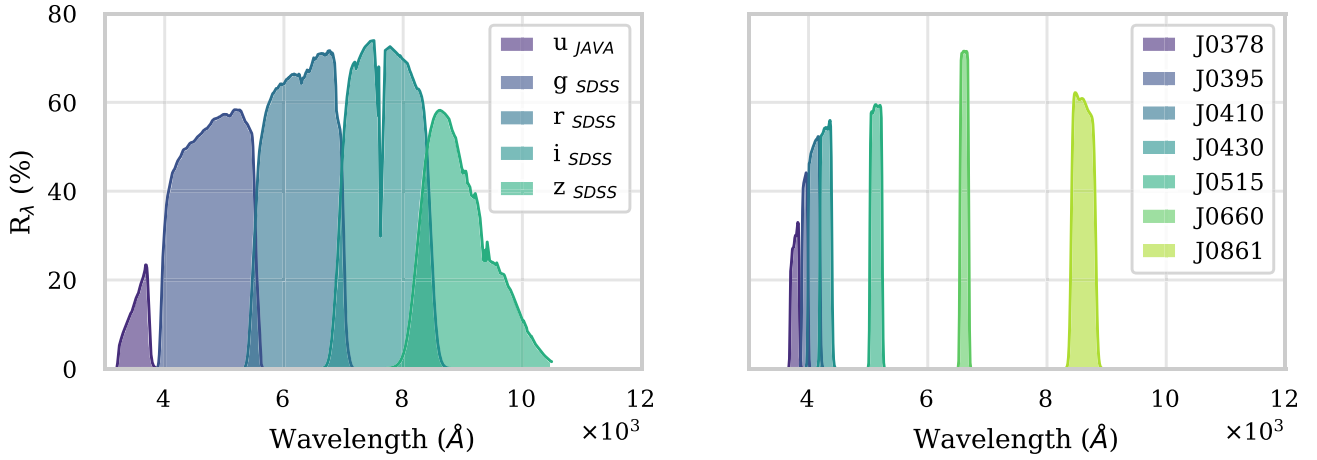


Fig. 1. Transmission curves (R_λ) for the S-PLUS filter system. *Left:* broad-band filters. *Right:* narrow-band filters.

survey is $r \sim 21$ AB magnitudes for objects with signal-to-noise greater than 3.

The survey calibration pipeline provides 3 different types of magnitudes for each filter, following the output of SExtractor (Bertin and Arnouts, 1996): the `auto_restricted` magnitudes, which have the advantage of high signal-to-noise while integrating most of the light from galaxies; the petrosian (petro) magnitudes, with an aperture size defined by the distance from the centre of the galaxy to a region where the signal-to-noise drops to zero in the detection images, reducing possible contamination from neighbouring galaxies; and the `aper` magnitudes, computed in a fixed 3 arcsecond diameter aperture to match the SDSS fibre size. As discussed in Molino et al. (2017), the `auto_restricted`² magnitudes provide the best properties for photometric redshift determination due to the high signal-to-noise while containing most of the galaxy light, and so it was chosen to be used in this work. For further details on the calibration pipeline, please refer to Mendes de Oliveira et al. (2019).

We also use infrared photometry information from unWISE (Lang, 2014). This project improves upon the products of AllWISE by taking advantage of deeper imaging and improved source detection in crowded regions. This allows unWISE photometry to be 0.7 mag deeper in the W1 (3.4 μ m) band and to detect three times more sources than AllWISE.

2.1. Sample construction and pre-processing

The supervised machine learning methods used in this work require a dataset containing the photometric and morphological information (the inputs) and the spectroscopic redshift (our target variable) to train the models. Since S-PLUS does not provide spectroscopic data, a cross-match with other catalogues is necessary to obtain this information.

We construct the samples for training, validation and testing in two steps:

1. A cross-match between unWISE photometric data and SDSS DR15 data in the Stripe-82 is done. Since both catalogues share the same identification number for objects, a simple search between corresponding IDs is made with the TOPCAT (Taylor, 2005) software. This creates a sample containing the spectroscopic redshift from SDSS and the W1 (3.4 μ m) and W2 (4.6 μ m) magnitudes from unWISE;

2. The sample created in the previous step is cross-matched with S-PLUS. This match is done using equatorial coordinates (RA and Dec) with a maximum tolerance radius of 1 arcsecond. The resulting sample contains the information from S-PLUS, the W1 and W2 magnitudes of unWISE and the spectroscopic redshift of SDSS. Furthermore, since a comparison with the currently used method is required, this sample also contains the photometric redshifts obtained with BPZ2 (Molino et al., 2020), available in the S-PLUS DR1.

After these steps we have a sample containing all input features needed for training. In addition to S-PLUS auto magnitudes, we included in our input set the two unWISE magnitudes W1 and W2, the colours calculated with respect to the `r_auto` magnitude, and the SExtractor parameters full width at half maximum (FWHM_n), maximum surface brightness (MUMAX), and the ellipticity of the objects. In total we use 30 different input features but do not take into account the uncertainties of the observations at this moment. We test which input configuration offers the most accurate photometric redshifts in Section 4.2.

The quality of the results of a given machine learning method is dependent on the quality of the data itself. With this in mind, a pre-processing step is done to assure that only high quality data is used during the training process of the models. The constraints on different parameters are listed in Table 1. The parameter `nDet_auto` refers to the number of filters in which a given object is detected and is used to study the effect of missing bands in z_{phot} determination. `PhotoFlag` is related to the quality flag given by SExtractor and is important to avoid poor photometry, such as due to blended objects. `PROB_GAL` is a value related to a star-galaxy separation described in Costa-Duarte et al. (2019) and we use it for the selection of galaxies. z_{spec} is the spectroscopic redshift and it is also used to remove remaining stars misclassified as galaxies after `PROB_GAL` constraints. z_{err} is the error in z_{spec} and is used to assure redshift estimates of good quality. `class_SDSS` is the spectroscopic class determined by SDSS, used to remove quasars. After applying these constraints, the total number of galaxies in our sample is 37 559 when `nDet_auto` = 12 and increases to 57 158 when `nDet_auto` $\geq$ 8 (see Fig. 4).

The sample is split randomly into training and testing samples, containing 70% and 30% of the total number of objects, respectively. These fractions are chosen to allow reasonable sample sizes to map the inputs to the output of the model and to allow a robust estimate of statistics as shown in Section 4. Afterwards, during the training process, the initial training sample is further split into training and validation samples, for cross-validation.

² Throughout this paper we will denote the `auto_restricted` magnitudes as `auto` for simplicity.

Table 1
Parameters, range of values, and source surveys.

Parameter	Value	Survey
Magnitude (r_auto)	[16, 21]	S-PLUS
nDet_auto	[8, 12]	S-PLUS
PhotoFlag	= 0	S-PLUS
PROB_GAL	≥ 0.5	S-PLUS
z _{spec}	≥ 1 × 10 ⁻⁴	SDSS
z _{err}	≤ 0.4	SDSS
class_SDSS	≠ QSO	SDSS

A K -fold cross-validation, with $K = 4$, consists on splitting a sample randomly into K parts. During the training process, the sample containing 70% of all objects is split into 75% for training and 25% for validation. With this approach we have a more robust estimate of the performance of our models while making a compromise between training time and the validation dataset size. It should be noticed that, for all methods, the same training, validation and testing samples are used so we can have a fair comparison.

3. Methods

In this section, we describe the methods and the choice of parameters used in this work, as well as the metrics adopted to evaluate the results.

3.1. Gaussian processes for photometric redshifts (GPz)

GPz (Almosallam et al., 2016) is a photometric redshift estimation code based on Sparse Gaussian Processes, developed to overcome the limitations related to the application of Gaussian Processes to large samples and improve upon simple regression methods.

One of the disadvantages of simpler methods, such as a polynomial regression, is the poor adaptiveness to non-linear data. One way of overcoming this problem is using Gaussian Processes (GPs, Williams and Rasmussen, 1995). GPs are a Bayesian, non-parametric and non-linear regression method that can adapt to the data.

It works by finding a distribution of functions that can map the input space (e.g. magnitudes) to the output space (e.g. redshifts) based on the observations present in the training sample. This method relies on the concept of locality, in which two nearby objects in the input space will also be nearby in the output space (e.g. two galaxies with similar colours will likely have similar redshifts). This increases the prediction power of the method but also increases its complexity and computational demand (see Gomes et al. 2017).

One way to maintain the advantages of Gaussian Processes and reduce its computational cost was introduced in Almosallam et al. (2016), through Sparse Gaussian Processes (SGPs). SGPs assumes that one can map the input-output relation using a reduced number of Gaussian kernel functions without losing accuracy. In the GPz algorithm, each kernel function can have its own hyperparameters in order to describe different densities and patterns of the training sample. The position of each kernel is defined in a way to maintain the distribution of the input data (for a more detailed discussion, see Gomes et al. 2017).

The GPz code can be adapted to a number of scientific problems by changing its configuration parameters. Almosallam et al. (2015) present a detailed description of each parameter and their effects on the model. After extensive testing, we choose the following set of parameters:

- 2000 iterations;

- 200 basis functions;
- Joint: True;
- Heteroscedastic: True;
- Cost-sensitive learning method: Normalized;
- Decorrelate: False.

Both the iterations and basis functions are self-explanatory. “Joint” indicates whether the method should learn a simple linear regression prior from the data. The “Heteroscedastic” parameter is set to True so that the method will model the uncertainty in the z_{phot} estimates as a function of the input (input-dependent noise). “Cost-sensitive learning” is related to the weight of each sample in the dataset, here defined as $1/(1 + z_{\text{sample}})$. The final parameter, “Decorrelate”, is set to False, so the method does not do a PCA pre-processing step. Further details about these and other parameters are presented in Almosallam (2017).

The optimization method is also changed from “BFGS” to “L-BFGS-B”, for lower memory consumption, less time per iteration, and no loss in accuracy.

3.2. Deep learning

In this work we compared two different deep-learning architectures, both developed using Tensorflow (Abadi et al., 2016) and the Keras (Chollet et al., 2015) Application Programming Interface.

Both deep-learning models used in this work are based on Multi-Layer Perceptrons (MLPs, Rosenblatt, 1958). In this approach, called Dense, all units in a given layer are connected to all units in the previous and following layer. The training process occurs via changes of the weights connecting those neurons based on feedback related to the errors in the predictions at each step of feedforward and backpropagation.

3.2.1. Dense Network (DN)

Our Dense Network (DN) architecture consists of an input layer with 30 neurons, one for each input feature (as described in Section 2), followed by 12 hidden layers with 35 to 15 neurons (see Fig. 2), with Batch Normalizations (Ioffe and Szegedy, 2015) in between, and an output layer containing 200 neurons. All activations are ReLU (Glorot et al., 2011) and the Categorical Cross-entropy loss is optimized with the Nadam (Dozat, 2016) algorithm. This architecture has a total of 10 081 parameters and was trained for 200 epochs.

The output layer has a Softmax (Goodfellow et al., 2016) function and contains 200 neurons to represent the photometric redshift PDFs, i.e., the redshift space is split into 200 bins of width 0.0067 between 0 and 1.34 (the maximum spectroscopic redshift in our sample). This approach converts the regression task into a classification task that gives the probability that a galaxy redshift is in a given bin.

3.2.2. Mixture Density Network (MDN)

For this specific model, we have used Tensorflow Probability (Dillon et al., 2017), since it allows the use of probabilistic reasoning in the Tensorflow framework, making the creation of Bayesian Neural Networks easier. They differ from traditional Neural Networks by the description of the network weights. While the latter uses trainable single values for the weights, the former uses trainable distributions. This change allows the estimation of the epistemic and aleatoric uncertainties (H llermeier and Waegeman, 2019). The epistemic error is related to model uncertainty and its predictions, and can be reduced to zero as the sample size increases. The aleatoric uncertainty is due to measurement errors and cannot be reduced to zero as more samples are observed.

Mixture Density Networks (MDN) were introduced by Bishop (1994) and are defined as a combination of a conventional neural

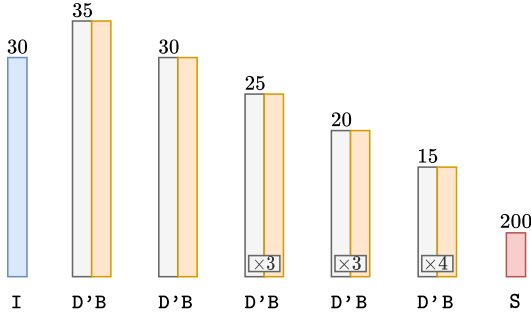


Fig. 2. Architecture of the Dense Network. The input layer (I, in blue) is followed by blocks of Dense (D', white) and BatchNorm (B, orange) layers, and the last layer (S, red) has a Softmax activation. The numbers represent the quantity of neurons in that layer. The numbers in the 3 last hidden blocks represent how many times these blocks are repeated.

network and a mixture density model. They can represent arbitrary probability distributions in the same way that a neural network can represent arbitrary functions. This is interesting for photometric redshift prediction because this problem may have many solutions due to the colour-morphology-redshift degeneracy (objects with different magnitudes having the same redshift) which is more apparent when using only broad-band photometry to estimate photometric redshifts, since they cannot provide a detailed sampling of the spectral features, and the input-output mapping is better expressed with distributions than with single-point estimates. The use of narrow-band filters, such as those in S-PLUS, allow for a better sampling of the spectra and thus lessens the effect of this degeneracy, improving the z_{phot} estimates.

Our MDN outputs 20 Gaussian functions, each described by three values: the mean, standard deviation and weight. The PDFs are the combination of these Gaussian functions (see Section 4.4). The number of components was chosen as the one providing the best single-point and PDF metrics, described in Section 3.3, after several tests.

The network is composed of 3 blocks containing DenseVariational layers (provided by Tensorflow Probability) with 196 units, a Batch Normalization, and a Dropout layer with 10% rate. The output layer is a MixtureNormal, which is responsible for returning the mean (μ), standard deviation (σ) and weight (w) for 20 Gaussian distributions. The architecture is illustrated in Fig. 3. The Negative-Log Likelihood loss was minimized using the Nadam optimizer, with a learning rate of 0.001. All hidden layers have a Leaky ReLU activation and this model was trained for 500 epochs.

A DenseVariational layer is similar to the Dense layer, but replaces the description of the neuron weights with distributions. The MixtureNormal layer generates $N = 20$ independent distributions (the user can choose the number and type of distributions) to generate the PDF output. Further details can be found in the Tensorflow Probability documentation.³

We also tested a shallower architecture for the DN model and a deeper architecture for the MDN model but we did not obtain satisfactory results.

3.3. Evaluation metrics

3.3.1. Single-point estimates

To assess the single-point estimate accuracy of each method, we need to analyse three main quantities: the scatter of the z_{phot} predictions, their bias, and the number of catastrophic outliers

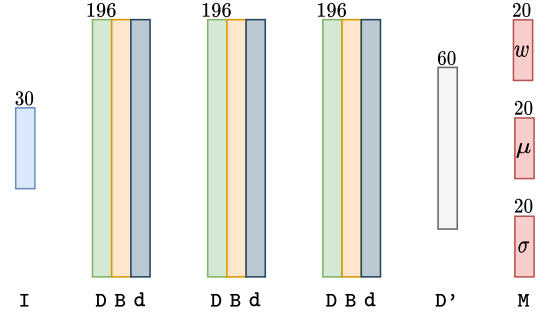


Fig. 3. Architecture of the Mixture Density Network. The input layer (I) is represented in blue and is followed by blocks of DenseVariational (D, white), BatchNorm (B, orange), and Dropout (d, dark grey) layers, a Dense (D', grey) layer, and the output layer (M, a MixtureNormal) is in red. The numbers represent the quantity of neurons in that layer.

in our results. We present these metrics per bin of r-band magnitudes and spectroscopic redshift, as well as for the complete sample (without binning).

The scatter is computed as in Brammer et al. (2008). It is robust, with low sensitivity to outliers:

$$\sigma_{\text{NMAD}} = 1.48 \times \text{median} \left(\frac{\delta z - \text{median}(\delta z)}{1 + z_{\text{spec}}} \right), \quad (1)$$

where $\delta z = (z_{\text{phot}} - z_{\text{spec}})$, z_{phot} is the photometric redshift calculated by the code, and z_{spec} is the true redshift of the object.

The systematic bias of the results, representing whether our predictions have a tendency to over- or underestimate the redshifts, is defined as:

$$\bar{\mu} = \left\langle \frac{\delta z}{1 + z_{\text{spec}}} \right\rangle, \quad (2)$$

and the outliers are defined as the objects for which the condition

$$\frac{|\delta z|}{1 + z_{\text{spec}}} > 0.15, \quad (3)$$

is satisfied (Ilbert et al., 2006). We will represent the outlier fraction by η .

Since we analyse these metrics as a function of bins of magnitude and/or redshift, it is worth mentioning that these quantities are calculated for all objects inside that bin and we take the median of the values. The magnitude bins have a width of 0.5 and the redshift bins have a width of 0.05. These values are chosen in a way to accommodate a reasonable number of objects inside each bin, so the statistics calculated in Section 4 are reliable.

3.3.2. Probability distribution functions

In order to verify the quality of the generated PDFs, we use three different metrics.

The Odds value, as introduced by Benítez (2000), can be used as an estimation of the accuracy of a PDF:

$$\text{Odds} = \int_{z_{\text{peak}} - \Delta z}^{z_{\text{peak}} + \Delta z} p(z) dz, \quad (4)$$

where z_{peak} is the photometric redshift corresponding to the peak of the PDF, Δz is the typical z_{phot} accuracy in the local universe (see below), and $p(z)$ is the PDF itself. Since the Odds value represents the fraction of the PDF contained in the interval $z_{\text{peak}} \pm \Delta z$, a lower value will represent a broad distribution, while a higher value translates to a narrow PDF.

For this work, the value of Δz is chosen as 0.02, which is the typical z_{phot} standard deviation for methods such as BPZ2 (Molino

³ https://www.tensorflow.org/probability/api_docs/python/tfp.

et al., 2020). This particular value was adopted in order to provide a fair comparison between our models and BPZ2.

The second tool used to check the quality of the PDFs is the Probability Integral Transform histogram (PIT, Polsterer et al., 2016). Each PDF has a PIT value calculated as:

$$\text{PIT} = \int_0^{z_{\text{spec}}} p(z) dz, \quad (5)$$

where the shape of its distribution reveals details about the PDFs. The ideal case would be a uniform distribution. If the PDFs are too narrow, the PIT histogram will be concave (u-shaped). In the opposite situation, the histogram will be convex. Also, any bias in the PDFs appears as a slope in the distribution, positive in the case of underestimation and negative in the case of overestimation of the redshift.

We also calculate the Continuous Ranked Probability Score (CRPS, Hersbach, 2000; Polsterer et al., 2016). The CRPS is a metric that measures the distance between the Cumulative Distribution Function (CDF) of a given PDF and its spectroscopic redshift, described as a step function. It is calculated as:

$$\text{CRPS}_i = \int_{-\infty}^{+\infty} [\text{CDF}_i(z) - \text{CDF}_{\text{spec}}(z)]^2 dz,$$

where the subscript i denotes a specific galaxy, CDF_i is the cumulative distribution function of the PDF of i , and CDF_{spec} is the cumulative distribution function of spectroscopic redshift of i and is defined based on the Heaviside step-function:

$$\text{CDF}_{\text{spec}}(z) = H(z - z_{\text{spec}}) = \begin{cases} 0 & \text{for } z < z_{\text{spec}}, \\ 1 & \text{for } z \geq z_{\text{spec}}. \end{cases}$$

4. Results

Now we present and discuss the results for z_{phot} point estimates and PDFs for GPz, DN, and MDN, comparing them with BPZ2. First we verify the impact of missing-band values, due to non-detections or non-observations, on the training process. Then we analyse how the choice of input features affects our results. After selecting the optimal input features, we compare the single-point estimate results from all methods and check the quality of their PDFs. When available, the shaded regions in the plots shown in this section represent the standard deviation from the 4-Fold Cross Validation, except for the BPZ2 method, for which the uncertainties were calculated using bootstrapping.

4.1. Missing feature analysis

In multiband surveys, such as S-PLUS, not all objects are always detected in all filters, and the impact of objects with missing photometry in the z_{phot} determination should be addressed. Fig. 4 (bottom) shows the fraction of galaxies in our sample with missing photometry for the 12 bands of S-PLUS photometric system. Due to galaxy spectra and filter sensitivity, most of the missing bands are in the blue side of the spectrum.

For this analysis, we trained the models with constrained values of $n\text{Det_auto}$, i.e., considering samples with $n\text{Det_auto} \geq 8, \dots, 12$. Following Chollet (2017), we replace missing values by zeros.

Fig. 5 presents the performance of the machine learning methods as a function of r_{auto} for different values of $n\text{Det_auto}$. We can notice, somewhat surprisingly, that the performance increases as $n\text{Det_auto}$ decreases. The improvement is more noticeable in the faint-end, where most of the objects with missing information are located. This can be explained by two major factors: first, as we ease the restrictions on $n\text{Det_auto}$ more objects are included in the training sample, as described in Section 2.1.

Table 3

Results for both Deep Learning methods and GPz using the samples described in Table 2 for $n\text{Det_auto} \geq 8$.

Configuration	$\sigma_{\text{NMAD}} (\times 10^{-2})$	$\bar{\mu} (\times 10^{-2})$	$\eta (\%)$
GPz			
A	3.159	−0.178	0.717
M	3.239	−0.182	0.857
M+E	3.139	−0.180	0.717
C	3.114	−0.159	0.904
C+E	3.017	−0.160	0.682
DN			
A	2.487	−0.182	0.636
M	2.550	−0.629	0.723
M+E	2.494	−0.376	0.548
C	2.509	−0.017	0.717
C+E	2.426	−0.220	0.571
MDN			
A	2.343	−0.104	0.636
M	2.457	−0.141	0.904
M+E	2.354	−0.125	0.665
C	2.437	−0.223	0.934
C+E	2.338	−0.078	0.583

Second, even though these objects were not detected in some filters (mostly in the blue; Fig. 4), the information contained in the redder bands allow for a better input–output mapping during the training process, increasing the generalization capacity of the models (Chollet, 2017). Notice that, in this section, the training, validation, and testing samples change due to different ($n\text{Det_auto}$) values, but all models shared the same datasets in each case. Due to these results, the analyses in the following sections are done using all objects that have $n\text{Det_auto} \geq 8$.

4.2. Input features analysis

We now analyse which combination of input features provide the most accurate photometric redshifts.

The configuration for each input combination is described in Table 2. The results of this section are shown in Fig. 6 as a function of r_{auto} and in Table 3 for the entire sample without binning. The DN, MDN, and GPz models perform better when trained using all input features available. For this reason, and to maintain consistency between the methods, we present only the results for the models trained using the “A” configuration in the following sections.

We also evaluated the training and validation loss of the models and verified that there is no under- or overfitting, indicating that the results are reliable and the models have a good generalization capacity.

4.3. Single-point estimates

Now we analyse the single-point estimates (SPE) of each model in comparison with BPZ2, the method currently used in S-PLUS for z_{phot} estimation. Each method obtains SPE in different ways. GPz provides the mean (used as SPE) and uncertainty, the SPE of the DN is a weighted average of the possible z_{phot} s and their probabilities (the expected value), the z_{phot} of the MDN is chosen as the value corresponding to the peak of the PDF (maximum-a-posteriori), and BPZ2 single-point estimate is given by $z_b = \sum_T \int p(z, T) z dz$, where $p(z, T)$ is the PDF and T is a template spectral type at redshift z (Benítez, 2000).

The results obtained for all methods are present in Fig. 7. The deep learning-based model MDN provides the most accurate z_{phot} s both as a function of magnitude and spectroscopic redshift. The currently used standard for z_{phot} determination in S-PLUS, the

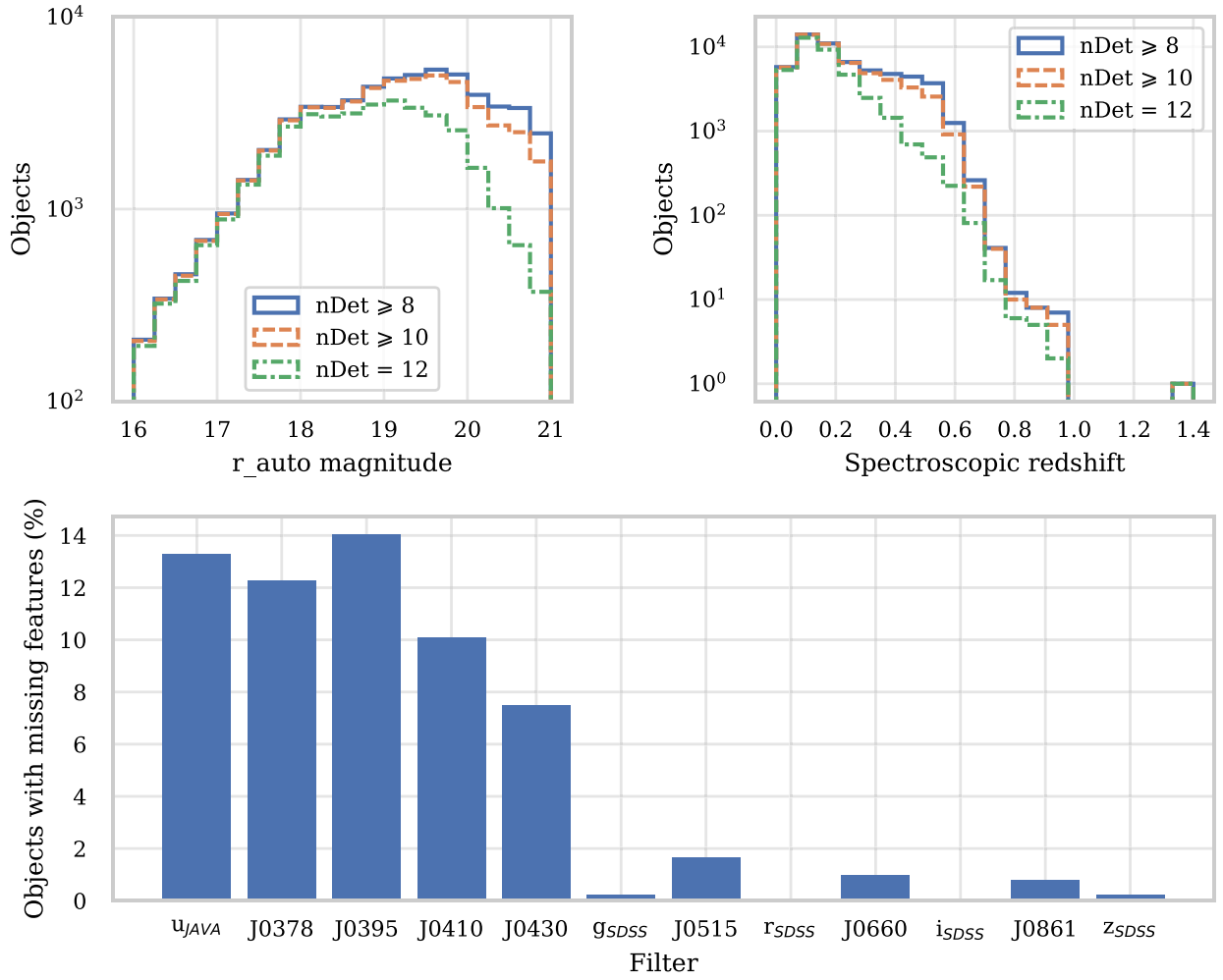


Fig. 4. Top left: r_{auto} distribution for three values of $n\text{Det}_{\text{auto}}$. Top right: z_{spec} distribution for three values of $n\text{Det}_{\text{auto}}$. Bottom: Fraction of objects with missing measurements for each filter in our training sample.

Table 2
Name of samples and the input features considered for the tests described in Section 4.2.

Sample	Notation	Input features ^a
All	A	u , J0378, J0395, J0410, J0430, g , J0515, r , J0660, i , J0861, z , W1, W2 $u-r$, 378- r , 395- r , 410- r , 430- r , $g-r$, 515- r , $r-660$, $r-i$, $r-861$, $r-z$, $r-W1$, $r-W2$ FWHM_n, MUMAX, Ellipticity
Magnitudes	M	u , J0378, J0395, J0410, J0430, g , J0515, r , J0660, i , J0861, z , W1, W2
Magnitudes+Extras	M+E	u , J0378, J0395, J0410, J0430, g , J0515, r , J0660, i , J0861, z , W1, W2 FWHM_n, MUMAX, Ellipticity
Colours	C	r $u-r$, 378- r , 395- r , 410- r , 430- r , $g-r$, 515- r , $r-660$, $r-i$, $r-861$, $r-z$, $r-W1$, $r-W2$
Colours+Extras	C+E	r $u-r$, 378- r , 395- r , 410- r , 430- r , $g-r$, 515- r , $r-660$, $r-i$, $r-861$, $r-z$, $r-W1$, $r-W2$ FWHM_n, MUMAX, Ellipticity

^aSome filter names were abbreviated for simplicity but are the same as described in Section 2.

BPZ2 method, underperformed in this sample in comparison to the other methods, mainly due to higher bias and outlier fraction. The results for the entire sample (without binning in magnitude or redshift) are shown in Table 4.

It should be pointed out that the deep learning-based methods were the only ones sensitive to the “redshift window of opportunity” (Molino et al., 2020). This feature is characterized by two emission lines entering two narrow band filters of S-PLUS (namely the [OIII] and H α lines in the J0660 and J0861 filters, respectively) in the redshift interval of $0.26 \leq z_{\text{spec}} \leq 0.32$ and

also greatly improves the accuracy of template-fitting methods.

In summary, we verified that the deep learning-based MDN model provided the most accurate single-point estimate z_{photS} . This model also had the best performance for faint objects, while having low bias and a negligible outlier fraction. When it is compared with the current standard in S-PLUS, BPZ2, the MDN model achieved 24% lower scatter, 82% lower bias, and 77% lower outlier fraction in the same testing sample.

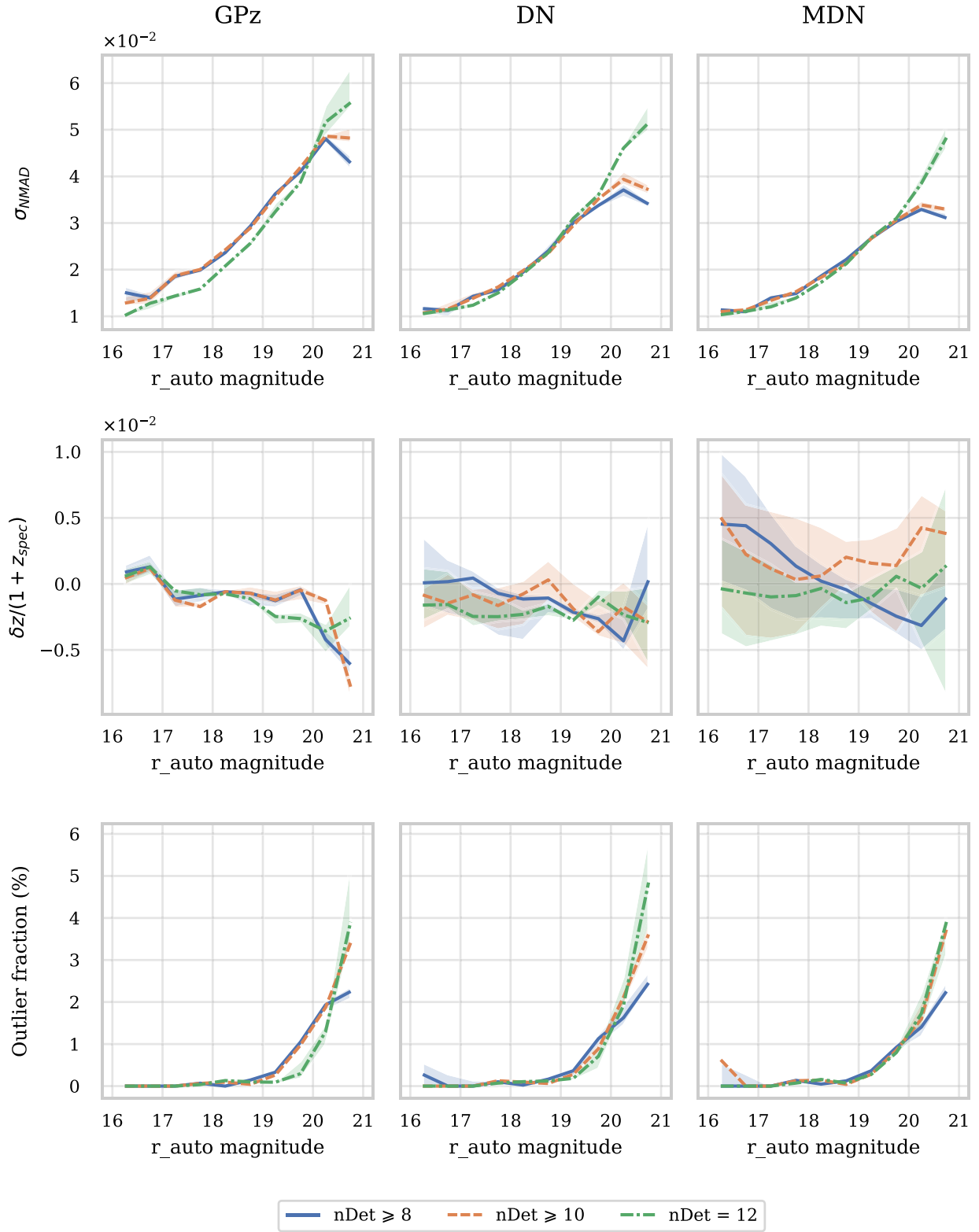


Fig. 5. Results for the scatter, bias, and outlier fraction for each method (columns) as a function of the magnitude in the r -band for different $n_{\text{Det_auto}}$ values. The shaded areas represent the uncertainties obtained from the 4-fold cross validation.

4.4. PDF analysis

Besides providing a probabilistic description of results, probability distribution functions (PDFs) are useful to understand and assess the uncertainties of a model. The PDFs here represent

all possible values of an object photometric redshift, and include errors in both the data measurement process (the aleatoric uncertainty) and in the model itself (the epistemic uncertainty).

Each method generates PDFs in different ways:

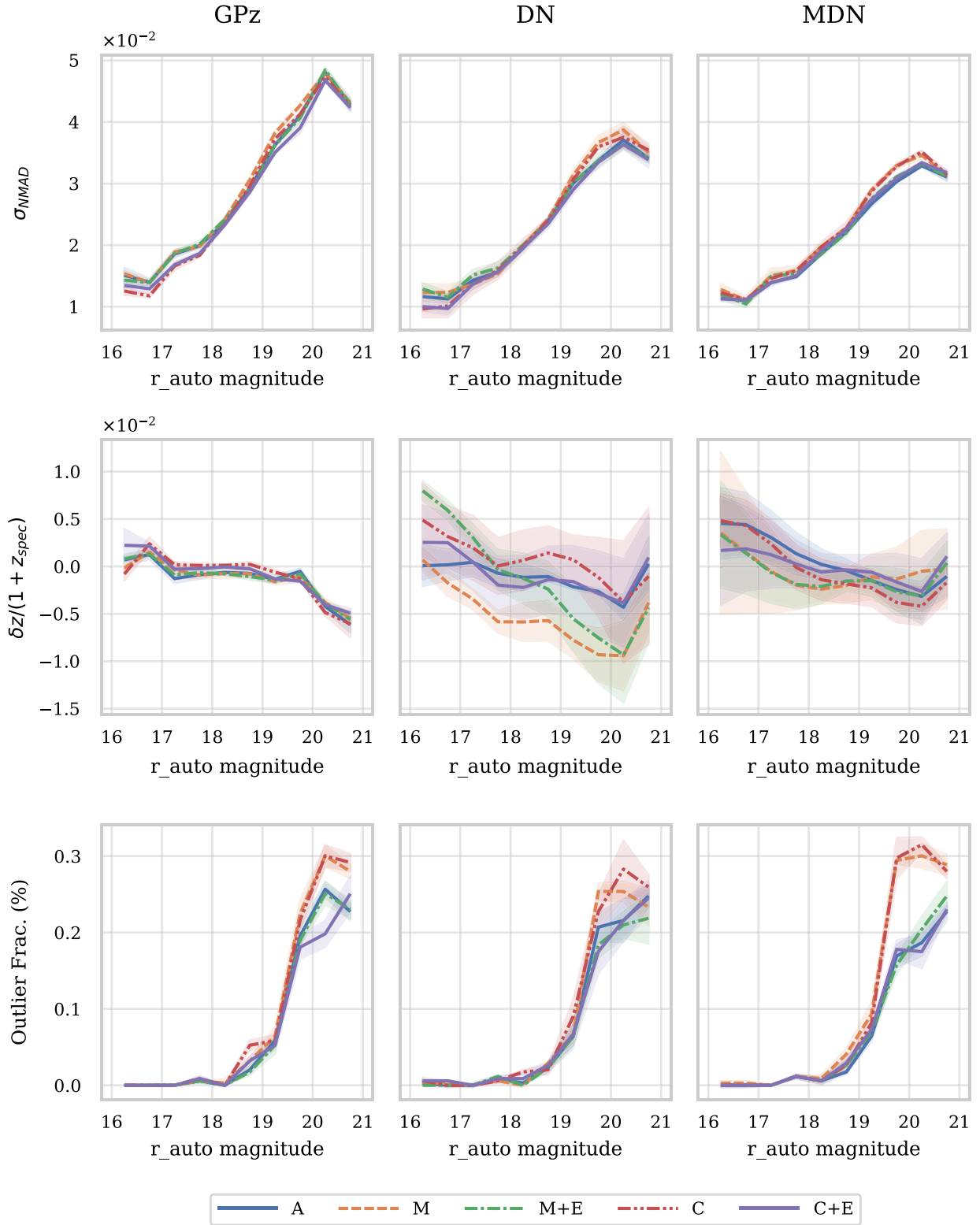


Fig. 6. Comparison of the results as a function of the r_{auto} magnitude for the scatter, bias, and outlier fraction for the three machine learning models, using different input feature configurations as described in Table 2.

- The DN produces PDFs due to the Softmax activation used in the output layer. The PDF is constructed by the probability of the galaxy belonging to each of the 200 bins of redshift between $0 \leq z_{\text{phot}} \leq 1.34$.
- For the MDN, the PDF is obtained as a combination of 20 Gaussian distributions, each with a different mean, standard

deviation, and weight. A combination of these distributions allows the generation of a multi-peak PDF.

- In the case of GPz, each z_{phot} prediction also generates a predictive variance, modelled as a function of input (heteroscedastic noise), so that the PDF is a Gaussian function.

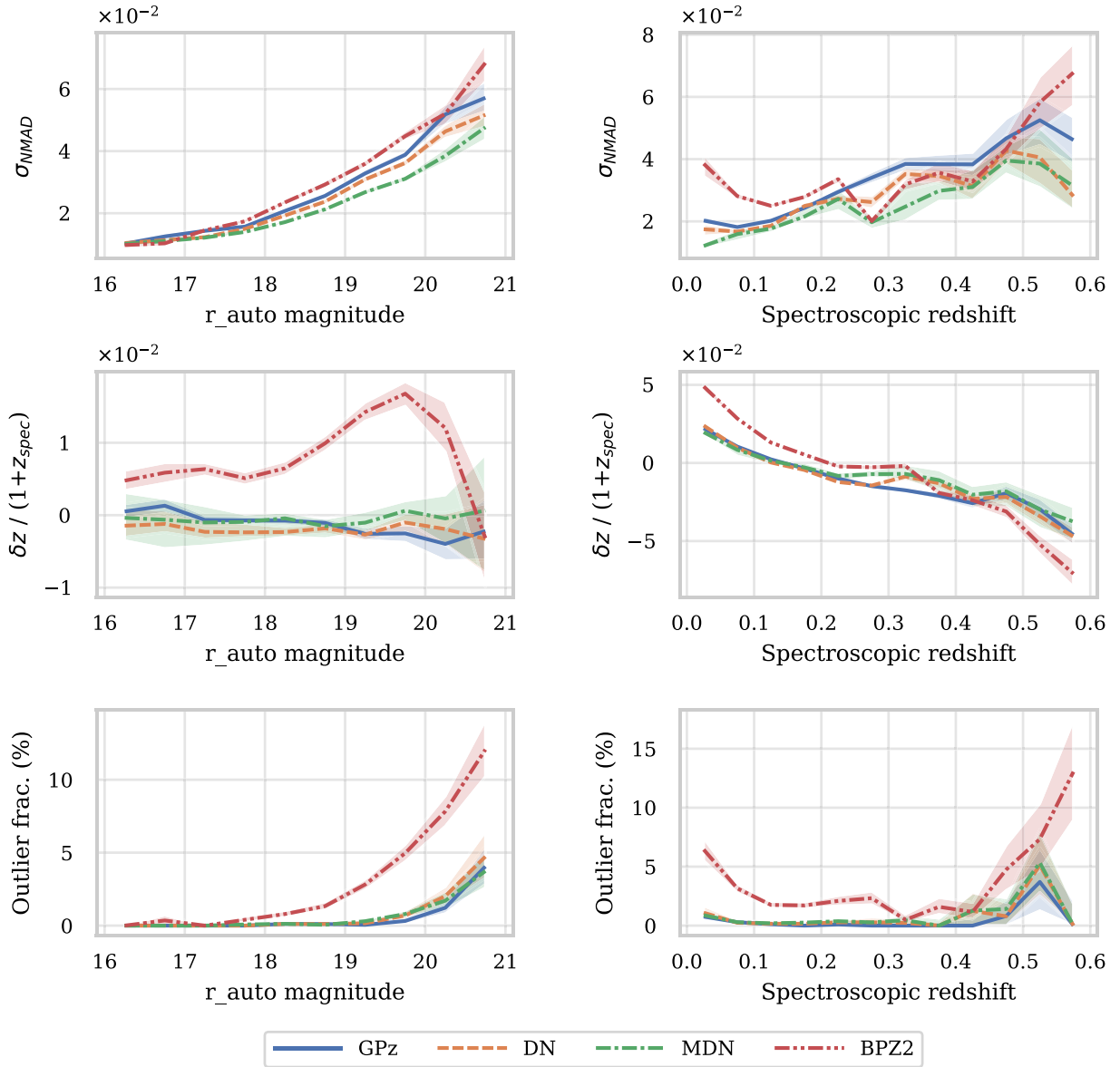


Fig. 7. Comparison between the metrics obtained for all methods. Left column, from top to bottom: scatter, bias and outlier fraction as a function of the r_{auto} magnitude. Right column, from top to bottom: same metrics as a function of the spectroscopic redshift. The GPz results are in blue, DN are in orange, MDN are in green, and BPZ2 results are in red. The shaded areas represent the uncertainties obtained from the 4-fold cross validation (or bootstrapping in the case of BPZ2).

Table 4
Summary of model results for $n_{\text{Det_auto}} \geq 8$.

Method	$\sigma_{\text{NMAD}} (\times 10^{-2})$	$\mu (\times 10^{-2})$	η (%)
GPz	3.159	-0.178	0.717
DN	2.487	-0.182	0.636
MDN	2.343	-0.104	0.636
BPZ2	3.080	0.580	2.758

We use three benchmarks to assess the quality of the PDFs: the Odds values, the Probability Integral Transform (PIT) histogram and the Continuous Ranked Probability Score (CRPS), all described in Section 3.3.

As shown in Fig. 8, the machine learning methods have broader PDFs than BPZ2, with peaks in the distribution of Odds between 0.20 and 0.25 for GPz, and between 0.25 and 0.30 for both the DN and MDN. BPZ2 provides PDFs which are narrower in general, as can be verified by the number of objects with Odds ≥ 0.6 .

The results obtained for the PIT test are presented in Fig. 9. BPZ2 is not included in this comparison because the PDFs of this method were not available. From this figure, we see that GPz PDFs have a near ideal distribution but have a higher quantity of outliers with PIT close to 0 or 1, which can be related to the way this model describes the PDFs (as unimodal Gaussians). The deep learning methods PDFs are better calibrated in relation to the width but present a negative slope in the case of the MDN, indicating that this method overestimates z_{phot} s (this can also be seen in Fig. 7), whereas PDFs are slightly broader for the DN and does not show any signs of bias.

The CRPS histograms are very similar between the different models, but the median values (presented inside the plots in Fig. 9) indicate that the DN and MDN models provide better PDFs in relation to the true redshift values, but with no significant difference between both.

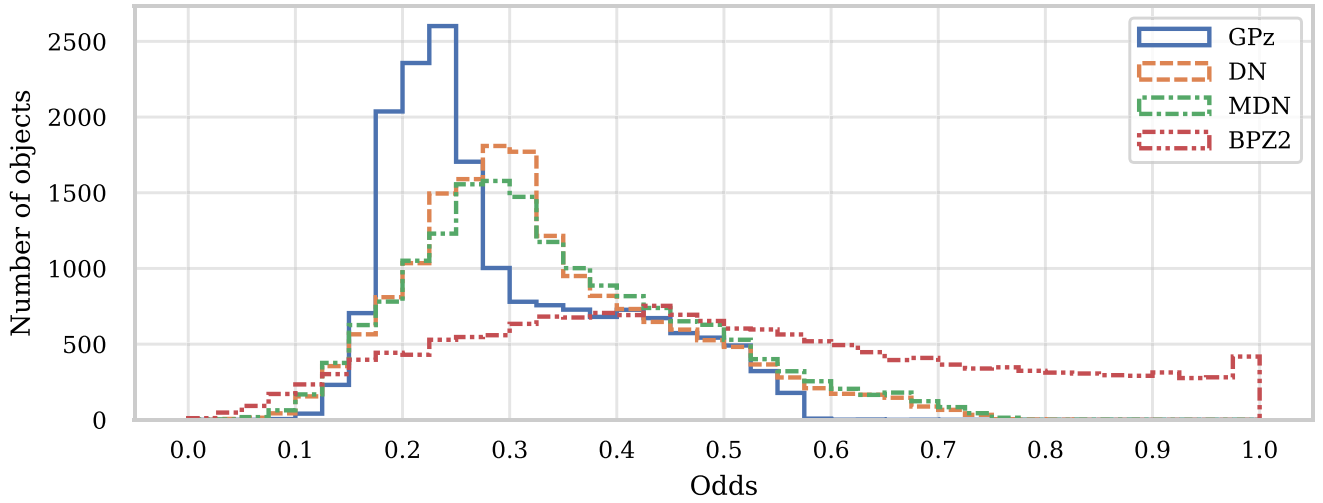


Fig. 8. Odds value distributions for all methods. Values closer to 1 indicate narrower PDFs.

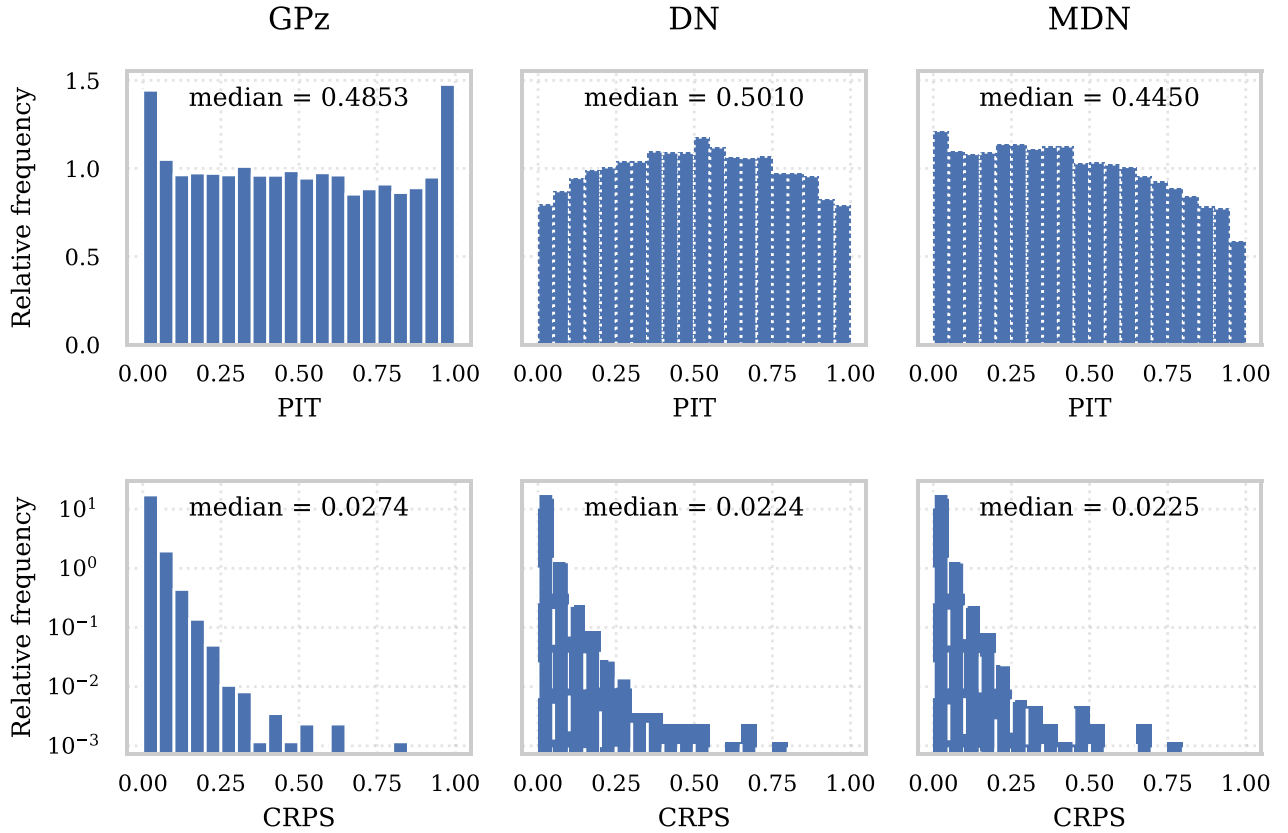


Fig. 9. PIT and CRPS distributions for GPz, DN, and the MDN models.

5. Conclusions

With the advent of all-sky photometric surveys, the need for photometric redshifts to study the large scale universe has increased. Recent advances in computational resources and software are making possible to extract accurate distance estimates from the immense volume of data produced by surveys.

In this work, we have explored three different machine learning methods (GPz, Dense Network and Mixture Density Network), evaluating their performances on the S-PLUS twelve-filter system to generate accurate photometric redshifts. We also compared our results with the method currently used in the S-PLUS DR1, the template-fitting code BPZ2.

We started our analysis by testing how the inclusion of objects with different numbers of missing bands impact photometric redshift estimates. Our findings, as shown in Section 4.1, indicate that even objects with four missing bands provide enough information to improve the results. This can be explained by two main factors: they are generally faint, and adding them to the training sample improves model performance mostly at high magnitudes; and by forcing the model to learn with incomplete information, missing-band objects help increasing the model generalization capacity.

In Section 4.2 we verified how the choice of input features impacts the photometric redshift accuracy. According to our tests for all models, the photo-z prediction scatter (σ_{NMAD}) changed

less than 5% between the different input choices described in Table 2. Also, as can be seen in Fig. 6, although the scatter does not change significantly, the bias and outlier fractions are, in general, lower when we use all 30 input features for training.

Fig. 7 shows that, for the dataset used in this work, all machine learning methods outperform the template-fitting approach currently used in S-PLUS DR1. A special highlight should be given to the deep learning-based methods, which are able to reach a higher accuracy than BPZ2 for faint objects (σ_{NMAD} 30% lower for $20 \leq r_{\text{auto}} \leq 21$). For all objects in our sample, the MDN model achieved 24% lower scatter, 82% lower bias, and 77% lower outlier fraction when compared to BPZ2 results on the same sample. Also, when analysing the results as a function of spectroscopic redshift, we see an increase in photometric redshift accuracy in the interval $0.26 \leq z_{\text{spec}} \leq 0.32$. This only happens due to the [OIII] and H α emission lines entering two narrow-band filters of the survey at this redshift range. This is an evidence of how important the twelve-filter system of S-PLUS is.

The photometric redshifts have a systematic bias as a function of redshift, as shown in Fig. 7. A similar trend is also present in other works, such as Kitching et al. (2016, p.11) and Banerji et al. (2008, p.11). A possible explanation is the lower amount of training samples in the low- and high-redshift regions. A possible solution is the implementation of sample weights, which will be done in future work.

When comparing the PDFs, we verified that none of the methods provided well-calibrated distributions, even for the deep learning-based methods with the best single-point estimate results. A second version of the MDN is under development aiming to solve those issues.

In summary, we concluded that, between the methods compared in this work, the best-performing for this sample is the Mixture Density Network, providing accurate photometric redshifts both as a function of r_{auto} and $\text{spec-}z$. It also provides PDFs that are able to take into account the redshift-colour degeneracy due to its multi-peaked nature. Our conclusions can be extended to the remainder of the S-PLUS survey, so that we are able to provide accurate photometric redshifts for several science cases in the Collaboration.

This method can be easily applied to future S-PLUS data releases offering improved photometric redshift performance when compared to the current template-fitting based standard (Molino et al., 2020). As more regions are observed and the overlaps with spectroscopic data increase, we also expect further improvements of our model.

CRedit authorship contribution statement

E.V.R. Lima: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing. **L. Sodré Jr.:** Conceptualization, Methodology, Supervision, Writing – review & editing. **C.R. Bom:** Methodology, Software, Writing – review & editing. **G.S.M. Teixeira:** Methodology, Software, Writing – review & editing. **L. Nakazono:** Writing – review & editing. **M.L. Buzzo:** Writing – review & editing. **C. Queiroz:** Writing – review & editing. **F.R. Herpich:** Writing – review & editing. **J.L. Nilo Castellon:** Writing – review & editing. **M.L.L. Dantas:** Writing – review & editing. **O.L. Dors Jr.:** Writing – review & editing. **R.C. Thom de Souza:** Writing – review & editing. **S. Akras:** Writing – review & editing. **Y. Jiménez-Teja:** Writing – review & editing. **A. Kanaan:** Investigation, Resources. **T. Ribeiro:** Investigation, Resources. **W. Schoennell:** Investigation, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author acknowledges the financial support given by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES, grant 88887.470064/2019-00) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant 169181/2017-0), during the development of this research. L. S. J. acknowledges support from Brazilian agencies FAPESP (2019/10923-5) and CNPq (304819/2017-4). L. N. acknowledges the financial support provided by the São Paulo Research Foundation (FAPESP), Brazil (grant 2019/01312-2 and 2014/10566-4). M. L. B. acknowledges the financial support provided by the São Paulo Research Foundation (FAPESP), Brazil (grant 2018/09165-6 and 2019/23388-0). C. Q. acknowledges the financial support provided by the São Paulo Research Foundation, Brazil (grant 2015/11442-0 and 2019/06766-1) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil. F. R. H. thanks FAPESP for the financial support through the program 2018/21661-9. J. L. N. C. is grateful for the financial support received from the Southern Office of Aerospace Research and development (SOARD), grants FA9550-18-1-0018 and FA9550-22-1-0037, of the Air Force Office of the Scientific Research International Office of the United States (AFOSR/IO). M. L. L. D. acknowledges the polish NCN, Poland grant number 2019/34/E/ST9/00133. S. A. acknowledges support under the grant 5077 financed by IAASARS/NOA. Y. J-T has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 898633. Y. J-T. also acknowledges financial support from the State Agency for Research of the Spanish MCIU through the "Center of Excellence Severo Ochoa" award to the Instituto de Astrofísica de Andalucía (SEV-2017-0709).

This work has made use of the computing facilities of the Laboratory of Astroinformatics (IAG/USP, NAT/Unicsul), whose purchase was made possible by the Brazilian agency FAPESP (grant 2009/54006-4) and the INCT-A.

The S-PLUS project, including the T80S robotic telescope and the S-PLUS scientific survey, was founded as a partnership between the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), the Observatório Nacional (ON), the Federal University of Sergipe (UFS), and the Federal University of Santa Catarina (UFSC), with important financial and practical contributions from other collaborating institutes in Brazil, Chile (Universidad de La Serena), and Spain (Centro de Estudios de Física del Cosmos de Aragón, CEFA).

The unWISE coadded images and catalogue are based on data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, and NEOWISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology. WISE and NEOWISE are funded by the National Aeronautics and Space Administration.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA

Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

References

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: A system for large-scale machine learning. arXiv e-prints, [arXiv:1605.08695](https://arxiv.org/abs/1605.08695), [arXiv:1605.08695](https://arxiv.org/abs/1605.08695).
- Al-Rfou, R., Alain, G., Almahairi, A., Angermueller, C., Bahdanau, D., Ballas, N., Bastien, F., Bayer, J., Belikov, A., Belopolsky, A., Bengio, Y., Bergeron, A., Bergstra, J., Bisson, V., Blecher Snyder, J., Bouchard, N., Boulanger-Lewandowski, N., Bouthillier, X., de Brébisson, A., Breuleux, O., Carrier, P.-L., Cho, K., Chorowski, J., Christiano, P., Cooijmans, T., Côté, M.-A., Côté, M., Courville, A., Dauphin, Y.N., Delalleau, O., Demouth, J., Desjardins, G., Dieleman, S., Dinh, L., Ducoffe, M., Dumoulin, V., Ebrahimi Kahou, S., Erhan, D., Fan, Z., Firat, O., Germain, M., Glorot, X., Goodfellow, I., Graham, M., Gulcehre, C., Hamel, P., Harlouchet, I., Heng, J.-P., Hidasi, B., Honari, S., Jain, A., Jean, S., Jia, K., Korobov, M., Kulkarni, V., Lamb, A., Lamblin, P., Larsen, E., Laurent, C., Lee, S., Lefrançois, S., Lemieux, S., Léonard, N., Lin, Z., Livezey, J.A., Lorenz, C., Lowin, J., Ma, Q., Manzagol, P.-A., Mastropietro, O., McGibbon, R.T., Memisevic, R., van Merriënboer, B., Michalski, V., Mirza, M., Orlandi, A., Pal, C., Pascanu, R., Pezeshki, M., Raffel, C., Renshaw, D., Rocklin, M., Romero, A., Roth, M., Sadowski, P., Salvatier, J., Savard, F., Schlüter, J., Schulman, J., Schwartz, G., Vlad Serban, I., Serdyuk, D., Shabanian, S., Simon, E., Spieckermann, S., Ramana Subramanyam, S., Synowski, J., Tanguy, J., van Tulder, G., Turian, J., Urban, S., Vincent, P., Visin, F., de Vries, H., Warde-Farley, D., Webb, D.J., Willson, M., Xu, K., Xue, L., Yao, L., Zhang, S., Zhang, Y., Theano Development Team, 2016. Theano: A Python framework for fast computation of mathematical expressions. arXiv e-prints, [arXiv:1605.02688](https://arxiv.org/abs/1605.02688), [arXiv:1605.02688](https://arxiv.org/abs/1605.02688).
- Alarcon, A., Gaztanaga, E., Eriksen, M., Baugh, C.M., Cabayol, L., Casas, R., Carretero, J., Castander, F.J., De Vicente, J., Fernandez, E., García-Bellido, J., Hildebrandt, H., Hoekstra, H., Joachimi, B., Manzoni, G., Miquel, R., Norberg, P., Padilla, C., Renard, P., Sanchez, E., Serrano, S., Sevilla-Noarbe, I., Siudek, M., Tallada-Crespí, P., 2021. The PAU Survey: an improved photometric sample in the COSMOS field. MNRAS 501 (4), 6103–6122. [doi:10.1093/mnras/staa3659](https://doi.org/10.1093/mnras/staa3659), [arXiv:2007.11132](https://arxiv.org/abs/2007.11132).
- Almosallam, I., 2017. *Heteroscedastic Gaussian Processes for Uncertain and Incomplete Data (Ph.D. thesis)*. University of Oxford.
- Almosallam, I.A., Jarvis, M.J., Roberts, S.J., 2016. GPz: non-stationary sparse Gaussian processes for heteroscedastic uncertainty estimation in photometric redshifts. Mon. Not. R. Astron. Soc. 462 (1), 726–739. [doi:10.1093/mnras/stw1618](https://doi.org/10.1093/mnras/stw1618).
- Almosallam, I.A., Lindsay, S.N., Jarvis, M.J., Roberts, S.J., 2015. A sparse Gaussian process framework for photometric redshift estimation. Mon. Not. R. Astron. Soc. 455 (3), 2387–2401. [doi:10.1093/mnras/stv2425](https://doi.org/10.1093/mnras/stv2425).
- Arnouts, S., Cristiani, S., Moscardini, L., Matarrese, S., Lucchin, F., Fontana, A., Giallongo, E., 1999. Measuring and modelling the redshift evolution of clustering: the hubble deep field north. Mon. Not. R. Astron. Soc. 310 (2), 540–556. [doi:10.1046/j.1365-8711.1999.02978.x](https://doi.org/10.1046/j.1365-8711.1999.02978.x).
- Banerji, M., Abdalla, F.B., Lahav, O., Lin, H., 2008. Photometric redshifts for the Dark Energy Survey and VISTA and implications for large-scale structure. MNRAS 386 (3), 1219–1233. [doi:10.1111/j.1365-2966.2008.13095.x](https://doi.org/10.1111/j.1365-2966.2008.13095.x), [arXiv:0711.1059](https://arxiv.org/abs/0711.1059).
- Barro, G., Pérez-González, P.G., Cava, A., Brammer, G., Pandya, V., Eliche Moral, C., Esquej, P., Domínguez-Sánchez, H., Alcalde Pampliega, B., Guo, Y., Koekemoer, A.M., Trump, J.R., Ashby, M.L.N., Cardiel, N., Castellano, M., Conselice, C.J., Dickinson, M.E., Dolch, T., Donley, J.L., Espino Briones, N., Faber, S.M., Fazio, G.G., Ferguson, H., Finkelstein, S., Fontana, A., Galametz, A., Gardner, J.P., Gawiser, E., Giallisco, M., Grazian, A., Grogin, N.A., Hathi, N.P., Hemmati, S., Hernán-Caballero, A., Kocevski, D., Koo, D.C., Kodra, D., Lee, K.-S., Lin, L., Lucas, R.A., Mobasher, B., McGrath, E.J., Nandra, K., Nayyeri, H., Newman, J.A., Pforr, J., Peth, M., Rafelski, M., Rodríguez-Munoz, L., Salvato, M., Stefanon, M., van der Wel, A., Willner, S.P., Wiklind, T., Wuyts, S., 2019. The CANDELS/SHARDS multiwavelength catalog in GOODS-N: Photometry, photometric redshifts, stellar masses, emission-line fluxes, and star formation rates. ApJS 243 (2), 22. [doi:10.3847/1538-4365/ab23f2](https://doi.org/10.3847/1538-4365/ab23f2), [arXiv:1908.00569](https://arxiv.org/abs/1908.00569).
- Battye, R.A., Weller, J., 2003. Constraining cosmological parameters using Sunyaev-Zeldovich cluster surveys. Phys. Rev. D 68 (8), [doi:10.1103/physrevd.68.083506](https://doi.org/10.1103/physrevd.68.083506).
- Benítez, N., 2000. Bayesian photometric redshift estimation. Astrophys. J. 536 (2), 571–583. [doi:10.1086/308947](https://doi.org/10.1086/308947).
- Benítez, N., Dupke, R., Moles, M., Sodré, L., Cenarro, J., Marin-Franch, A., Taylor, K., Cristóbal, D., Fernandez-Soto, A., Mendes de Oliveira, C., Cepa-Nogue, J., Abramo, L.R., Alcaniz, J.S., Overzier, R., Hernández-Monteagudo, C., Alfaro, E.J., Kanaan, A., Carvano, J.M., Reis, R.R.R., Martínez González, E., Ascaso, B., Ballesteros, F., Xavier, H.S., Varela, J., Ederoclite, A., Vázquez Ramio, H., Broadhurst, T., Cypriano, E., Angulo, R., Diego, J.M., Zandivarez, A., Díaz, E., Melchior, P., Umetsu, K., Spinelli, P.F., Zitrin, A., Coe, D., Yepes, G., Vielva, P., Sahni, V., Marcos-Caballero, A., Shu Kitaura, F., Maroto, A.L., Masip, M., Tsujikawa, S., Carneiro, S., Gonzalez Nuevo, J., Carvalho, G.C., Reboucas, M.J., Carvalho, J.C., Abdalla, E., Bernui, A., Pigazzo, C., Ferreira, E.G.M., Chandrachani Devi, N., Bengaly, Jr., C.A.P., Campista, M., Amorim, A., Asari, N.V., Bongiovanni, A., Bonoli, S., Bruzual, G., Cardiel, N., Cava, A., Cid Fernandes, R., Coelho, P., Cortesi, A., Delgado, R.G., Díaz García, L., Espinosa, J.M.R., Galiano, E., González-Serrano, J.L., Falcon-Barroso, J., Fritz, J., Fernandes, C., Gorgas, J., Hoyos, C., Jimenez-Teja, Y., Lopez-Aguerrí, J.A., Lopez-San Juan, C., Mateus, A., Molino, A., Novais, P., Omill, A., Oteo, I., Perez-Gonzalez, P.G., Poggianti, B., Proctor, R., Ricciardelli, E., Sanchez-Blazquez, P., Storch-Bergmann, T., Telles, E., Schoenell, W., Trujillo, N., Vazdekis, A., Viironen, K., Dalfon, S., Aparicio-Villegas, T., Rocha, D., Ribeiro, T., Borges, M., Martins, S.L., Marcolino, W., Martínez-Delgado, D., Pérez-Torres, M.A., Siffert, B.B., Calvao, M.O., Sako, M., Kessler, R., Alvarez-Candal, A., De Pra, M., Roig, F., Lazaro, D., Gorosabel, J., Lopes de Oliveira, R., Lima-Neto, G.B., Irwin, J., Liu, J.F., Alvarez, E., Balmes, I., Chueca, S., Costa-Duarte, M.V., da Costa, A.A., Dantas, M.L.L., Diaz, A.Y., Fabregat, J., Ferrari, F., Gavela, B., Gracia, S.G., Gruel, N., Gutierrez, J.L.L., Guzman, R., Hernandez-Fernandez, J.D., Herranz, D., Hurtado-Gil, L., Jablonsky, F., Laporte, R., Le Tiran, L.L., Licandro, J., Lima, M., Martin, E., Martinez, V., Montero, J.J.C., Penteado, P., Pereira, C.B., Peris, V., Quilis, V., Sanchez-Portal, M., Soja, A.C., Solano, E., Torra, J., Valdivielso, L., 2014. J-PAS: The Javalambre-physics of the accelerated universe astrophysical survey. arXiv e-prints, [arXiv:1403.5237](https://arxiv.org/abs/1403.5237).
- Benítez, N., Gaztañaga, E., Miquel, R., Castander, F., Moles, M., Crocce, M., Fernández-Soto, A., Fosalba, P., Ballesteros, F., Campa, J., Cardiel-Sas, L., Castilla, J., Cristóbal-Hornillos, D., Delfino, M., Fernández, E., Fernández-Sopuerta, C., García-Bellido, J., Lobo, J.A., Martínez, V.J., Ortiz, A., Pacheco, A., Paredes, S., Pons-Bordería, M.J., Sánchez, E., Sánchez, S.F., Varela, J., de Vicente, J.F., 2009. Measuring baryon acoustic oscillations along the line of sight with photometric redshifts: The PAU survey. ApJ 691 (1), 241–260. [doi:10.1088/0004-637X/691/1/241](https://doi.org/10.1088/0004-637X/691/1/241), [arXiv:0807.0535](https://arxiv.org/abs/0807.0535).
- Bertin, E., Arnouts, S., 1996. SExtractor: Software for source extraction. A&AS 117, 393–404. [doi:10.1051/aas:1996164](https://doi.org/10.1051/aas:1996164).
- Bishop, C., 1994. *Mixture Density Networks*. WorkingPaper, Aston University.
- Blanton, M.R., Bershad, M.A., Abolfathi, B., Albaret, F.D., Allende Prieto, C., Almeida, A., Alonso-García, J., Anders, F., Anderson, S.F., Andrews, B., et al., 2017. Sloan digital sky survey IV: Mapping the milky way, nearby galaxies, and the distant universe. Astron. J. 154, 28. [doi:10.3847/1538-3881/aa7567](https://doi.org/10.3847/1538-3881/aa7567), [arXiv:1703.00052](https://arxiv.org/abs/1703.00052).
- Bonoli, S., Marin-Franch, A., Varela, J., Vázquez Ramió, H., Abramo, L.R., Cenarro, A.J., Dupke, R.A., Vilchez, J.M., Cristóbal-Hornillos, D., González Delgado, R.M., Hernández-Monteagudo, C., López-Sanjuan, C., Muniesa, D.J., Civera, T., Ederoclite, A., Hernán-Caballero, A., Marra, V., Baqui, P.O., Cortesi, A., Cypriano, E.S., Dalfon, S., de Amorim, A.L., Díaz-García, L.A., Diego, J.M., Martínez-Solaesche, G., Pérez, E., Placco, V.M., Prada, F., Queiroz, C., Alcaniz, J., Alvarez-Candal, A., Cepa, J., Maroto, A.L., Roig, F., Siffert, B.B., Taylor, K., Benítez, N., Moles, M., Sodré, J., Carneiro, S., Mendes de Oliveira, C., Abdalla, E., Angulo, R.E., Aparicio Resco, M., Balaguera-Antolínez, A., Ballesteros, F.J., Brito-Silva, D., Broadhurst, T., Carrasco, E.R., Castro, T., Cid Fernandes, R., Coelho, P., de Melo, R.B., Doubrawa, L., Fernandez-Soto, A., Ferrari, F., Finoguenov, A., García-Benito, R., Iglesias-Páramo, J., Jiménez-Teja, Y., Kitaura, F.S., Laur, J., Lopes, P.A.A., Lucatelli, G., Martínez, V.J., Maturi, M., Quartín, M., Pigazzo, C., Rodríguez-Martín, J.E., Salzano, V., Tamm, A., Tempel, E., Umetsu, K., Valdivielso, L., von Martens, R., Zitrin, A., Díaz-Martín, M.C., López-Alegre, G., López-Sainz, A., Yanes-Díaz, A., Rueda-Teruel, F., Rueda-Teruel, S., Abril Ibañez, J., Antón Bravo, J.L., Bello Ferrer, R., Bielsa, S., Casino, J.M., Castillo, J., Chueca, S., Cuesta, L., Garzarán Calderaro, J., Iglesias-Marzoa, R., Íñiguez, C., Lamadrid Gutierrez, J.L., Lopez-Martinez, F., Lozano-Pérez, D., Maicas Sacristán, N., Molina-Ibáñez, E.L., Moreno-Signes, A., Rodríguez Llano, S., Royo Navarro, M., Tilve Rua, V., Andrade, U., Alfaro, E.J., Akras, S., Arnalte-Mur, P., Ascaso, B., Barbosa, C.E., Beltrán Jiménez, J., Benetti, M., Bengaly, C.A.P., Bernui, A., Blanco-Pillado, J.J., Borges Fernandes, M., Bregman, J.N., Bruzual, G., Calderone, G., Carvano, J.M., Casarini, L., Chies-Santos, A.L., Coutinho de Carvalho, G., Dimauro, P., Duarte Puertas, S., Figueroa, D., González-Serrano, J.L., Guerrero, M.A., Gurung-López, S., Herranz, D., Huertas-Company, M., Irwin, J.A., Izquierdo-Villalba, D., Kanaan, A., Kehrig, C., Kirkpatrick, C.C., Lim, J., Lopes, A.R., Lopes de Oliveira, R., Marcos-Caballero, A., Martínez-Delgado, D., Martínez-González, E., Martínez-Somonte, G., Oliveira, N., Orsi, A.A., Overzier, R.A., Penna-Lima, M.,

- Reis, R.R.R., Spinoso, D., Tsujikawa, S., Vielva, P., Vitorelli, A.Z., Xia, J.Q., Yuan, H.B., Arroyo-Polonio, A., Dantas, M.L.L., Galarza, C.A., Gonçalves, D.R., Gonçalves, R.S., Gonzalez, J.E., Gonzalez, A.H., Greisel, N., Landim, R.G., Lazzaro, D., Magris, G., Monteiro-Oliveira, R., Pereira, C.B., Rebouças, M.J., Rodríguez-Espinosa, J.M., Santos da Costa, S., Telles, E., 2020. The miniJPAS survey: a preview of the Universe in 56 colours. arXiv e-prints, arXiv:2007.01910. arXiv:2007.01910.
- Brammer, G.B., van Dokkum, P.G., Coppi, P., 2008. EAZY: A fast, public photometric redshift code. *Astrophys. J.* 686 (2), 1503–1513. doi:10.1086/591786.
- Cenarro, A.J., Moles, M., Cristóbal-Hornillos, D., Marín-Franch, A., Ederoclite, A., Varela, J., López-Sanjuan, C., Hernández-Montegudo, C., Angulo, R.E., Ramió, H.V., Viironen, K., Bonoli, S., Orsi, A.A., Hurier, G., Roman, I.S., Greisel, N., Vilella-Rojo, G., Díaz-García, L.A., Logroño-García, R., Gurung-López, S., Spinoso, D., Izquierdo-Villalba, D., Aguerri, J.A.L., Prieto, C.A., Bonatto, C., Carvano, J.M., Chies-Santos, A.L., Daffon, S., Dupke, R.A., Falcón-Barroso, J., Gonçalves, D.R., Jiménez-Teja, Y., Molino, A., Placco, V.M., Solano, E., Whitten, D.D., Abril, J., Antón, J.L., Bello, R., de Toledo, S.B., Castillo-Ramírez, J., Chueca, S., Civera, T., Díaz-Martín, M.C., Domínguez-Martínez, M., Garzán-Calderaro, J., Hernández-Fuertes, J., Iglesias-Marzoa, R., Iñiguez, C., Ruiz, J.M.J., Kruuse, K., Lamadrid, J.L., Lasso-Cabrera, N., López-Alegre, G., López-Sainz, A., Maicas, N., Moreno-Signes, A., Muniesa, D.J., Rodríguez-Llano, S., Rueda-Teruel, F., Rueda-Teruel, S., Soriano-Laguía, I., Tilve, V., Valdivielso, L., Yanes-Díaz, A., Alcaniz, J.S., de Oliveira, C.M., Sodré, L., Coelho, P., de Oliveira, R.L., Tamm, A., Xavier, H.S., Abramo, L.R., Akas, S., Alfaro, E.J., Alvarez-Candal, A., Ascaso, B., Beasley, M.A., Beers, T.C., Fernandes, M.B., Bruzual, G.R., Buzzo, M.L., Carrasco, J.M., Cepa, J., Cortesi, A., Costa-Duarte, M.V., Prá, M.D., Favole, G., Galarza, A., Galbany, L., García, K., Delgado, R.M.G., González-Serrano, J.L., Gutiérrez-Soto, L.A., Hernández-Jiménez, J.A., Kanaan, A., Kuncarayakti, H., Landim, R.C.G., Laur, J., Licandro, J., Neto, G.B.L., Lyman, J.D., Apellániz, J.M., Miralda-Escudé, J., Morate, D., Nogueira-Cavalcante, J.P., Novais, P.M., Oncins, M., Oteo, I., Overzier, R.A., Pereira, C.B., Rebassa-Mansergas, A., Reis, R.R.R., Roig, F., Sako, M., Salvador-Rusiñol, N., Sampedro, L., Sánchez-Blázquez, P., Santos, W.A., Schmidtobrick, L., Sifert, B.B., Telles, E., Vilchez, J.M., 2019. J-PLUS: The javalambre photometric local universe survey. *Astron. Astrophys.* 622, A176. doi:10.1051/0004-6361/201833036.
- Chollet, F., 2017. *Deep Learning with Python*. Manning.
- Chollet, F., et al., 2015. Keras. <https://keras.io>.
- Costa-Duarte, M.V., Sampedro, L., Molino, A., Xavier, H.S., Herpich, F.R., Chies-Santos, A.L., Barbosa, C.E., Cortesi, A., Schoenell, W., Kanaan, A., Ribeiro, T., Mendes de Oliveira, C., Akas, S., Alvarez-Candal, A., Barbosa, C.L., Castellón, J.L.N., Coelho, P., Dantas, M.L.L., Dupke, R., Ederoclite, A., Galarza, A., Gonçalves, T.S., Hernandez-Jimenez, J.A., Jiménez-Teja, Y., Lopes, A., Lopes, P.A.A., Lopes de Oliveira, R., Melo de Azevedo, J.L., Nakazono, L.M., Perottoni, H.D., Queiroz, C., Saha, K., Sodré, J., Telles, E., Thom de Souza, R.C., 2019. The S-PLUS: a star/galaxy classification based on a Machine Learning approach. arXiv e-prints, arXiv:1909.08626. arXiv:1909.08626.
- Dillon, J.V., Langmore, I., Tran, D., Brevdo, E., Vasudevan, S., Moore, D., Patton, B., Alemi, A., Hoffman, M., Saurous, R.A., 2017. TensorFlow distributions. arXiv e-prints, arXiv:1711.10604. arXiv:1711.10604.
- Dozat, T., 2016. *Incorporating Nesterov Momentum into Adam*.
- Feldmann, R., Carollo, C.M., Porciani, C., Lilly, S.J., Capak, P., Taniguchi, Y., Fèvre, O.L., Renzini, A., Scoville, N., Ajiki, M., Aussel, H., Contini, T., McCracken, H., Mobasher, B., Murayama, T., Sanders, D., Sasaki, S., Scarlata, C., Scodreggio, M., Shioya, Y., Silverman, J., Takahashi, M., Thompson, D., Zamorani, G., 2006. The Zurich Extragalactic Bayesian Redshift Analyzer and its first application: COSMOS. *Mon. Not. R. Astron. Soc.* 372 (2), 565–577. doi:10.1111/j.1365-2966.2006.10930.x.
- Firth, A.E., Lahav, O., Somerville, R.S., 2003. Estimating photometric redshifts with artificial neural networks. *Mon. Not. R. Astron. Soc.* 339 (4), 1195–1202. doi:10.1046/j.1365-8711.2003.06271.x.
- Gehrels, N., Spergel, D., 2015. Wide-Field InfraRed Survey Telescope (WFIRST) mission and synergies with LISA and LIGO-virgo. *J. Phys. Conf. Ser.* 610, 012007. doi:10.1088/1742-6596/610/1/012007.
- Glorot, X., Bordes, A., Bengio, Y., 2011. Deep sparse rectifier neural networks. In: Gordon, G., Dunson, D., Dudík, M. (Eds.), *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. In: *Proceedings of Machine Learning Research*, vol. 15, PMLR, Fort Lauderdale, FL, USA, pp. 315–323. URL: <http://proceedings.mlr.press/v15/glorot11a.html>.
- Gomes, Z., Jarvis, M.J., Almosallam, I.A., Roberts, S.J., 2017. Improving photometric redshift estimation using GPz: size information, post processing, and improved photometry. *Mon. Not. R. Astron. Soc.* 475 (1), 331–342. doi:10.1093/mnras/stx3187.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press, <http://www.deeplearningbook.org>.
- Gruel, N., Moles, M., Varela, J., Cristóbal-Hornillos, D., Ederoclite, A., Cenarro, A.J., Marín-Franch, A., Chueca, S., Yanes-Díaz, A., Rueda-Teruel, F., Rueda-Teruel, S., Luis-Simoes, R., López-Sainz, A., Hernández-Fuertes, J., 2012. Calibration plan for J-PAS and J-PLUS surveys. In: Peck, A.B., Seaman, R.L., Comerón, F. (Eds.), *Observatory Operations: Strategies, Processes, and Systems IV*. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 8448, p. 84481V. doi:10.1117/12.925581.
- Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather Forecast.* 15 (5), 559–570. doi:10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2, URL: https://journals.ametsoc.org/view/journals/wefo/15/5/1520-0434_2000_015_0559_dotcrp_2_0_co_2.xml.
- Hildebrandt, H., Erben, T., Kuijken, K., van Waerbeke, L., Heymans, C., Coupon, J., Benjamin, J., Bonnett, C., Fu, L., Hoekstra, H., Kitching, T.D., Mellier, Y., Miller, L., Velander, M., Hudson, M.J., Rowe, B.T.P., Schrabback, T., Semboloni, E., Benítez, N., 2012. CFHTLenS: improving the quality of photometric redshifts with precision photometry. *Mon. Not. R. Astron. Soc.* 421 (3), 2355–2367. doi:10.1111/j.1365-2966.2012.20468.x.
- Hildebrandt, H., Viola, M., Heymans, C., Joudaki, S., Kuijken, K., Blake, C., Erben, T., Joachimi, B., Klaes, D., Miller, L., Morrison, C.B., Nakajima, R., Kleijn, G.V., Amon, A., Choi, A., Covone, G., de Jong, J.T.A., Dvornik, A., Conti, I.F., Grado, A., Harnois-Déraps, J., Herbonnet, R., Hoekstra, H., Köhlinger, F., McFarland, J., Mead, A., Merten, J., Napolitano, N., Peacock, J.A., Radovich, M., Schneider, P., Simon, P., Valentijn, E.A., van den Busch, J.L., van Uitert, E., Waerbeke, L.V., 2016. KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing. *Mon. Not. R. Astron. Soc.* 465 (2), 1454–1498. doi:10.1093/mnras/stw2805.
- Hoyle, B., Rau, M.M., Bonnett, C., Seitz, S., Weller, J., 2015. Data augmentation for machine learning redshifts applied to Sloan Digital Sky Survey galaxies. *Mon. Not. R. Astron. Soc.* 450 (1), 305–316. doi:10.1093/mnras/stv599.
- Hüllermeier, E., Waegeman, W., 2019. Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. arXiv e-prints, arXiv:1910.09457. arXiv:1910.09457.
- Ilbert, O., Arnouts, S., McCracken, H.J., Bolzonella, M., Bertin, E., Fèvre, O.L., Mellier, Y., Zamorani, G., Pellò, R., Iovino, A., Tresse, L., Brun, V.L., Bottini, D., Garilli, B., Maccagni, D., Picat, J.P., Scaramella, R., Scodreggio, M., Vettolani, G., Zanichelli, A., Adami, C., Bardelli, S., Cappi, A., Charlot, S., Cilieg, P., Contini, T., Cucciati, O., Foucaud, S., Franzetti, P., Gagnaud, I., Guzzo, L., Marano, B., Marinoni, C., Mazure, A., Meneux, B., Merighi, R., Paltani, S., Pollo, A., Pozzetti, L., Radovich, M., Zucca, E., Bondi, M., Bongiorno, A., Busarello, G., Torre, S.D.L., Gregorini, L., Lamareille, F., Mathez, G., Merluzzi, P., Ripepi, V., Rizzo, D., Vergani, D., 2006. Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey. *Astron. Astrophys.* 457 (3), 841–856. doi:10.1051/0004-6361:20065138.
- Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. CoRR abs/1502.03167. URL: <http://arxiv.org/abs/1502.03167>. arXiv:1502.03167.
- Ivezic, Z., Axelrod, T., Brandt, W., Burke, D., Claver, C., Connolly, A., Cook, K., Gee, P., Gilmore, D., Jacoby, S., Jones, R., Kahn, S., Kantor, J., Krabbendam, V., Lupton, R., Monet, D., Pinto, P., Saha, A., Schalk, T., Schneider, D., 2008. Large synoptic survey telescope: From science drivers to reference design. *Serbian Astron. J.* (176), 1–13. doi:10.2298/saj0876001i.
- Kilbinger, M., Fu, L., Heymans, C., Simpson, F., Benjamin, J., Erben, T., Harnois-Déraps, J., Hoekstra, H., Hildebrandt, H., Kitching, T.D., Mellier, Y., Miller, L., Waerbeke, L.V., Benabed, K., Bonnett, C., Coupon, J., Hudson, M.J., Kuijken, K., Rowe, B., Schrabback, T., Semboloni, E., Vafaei, S., Velander, M., 2013. CFHTLenS: combined probe cosmological model comparison using 2D weak gravitational lensing. *Mon. Not. R. Astron. Soc.* 430 (3), 2200–2220. doi:10.1093/mnras/stt041.
- Kind, M.C., Brunner, R.J., 2013. TPZ: photometric redshift PDFs and ancillary information by using prediction trees and random forests. *Mon. Not. R. Astron. Soc.* 432 (2), 1483–1501. doi:10.1093/mnras/stt574.
- Kitching, T.D., Verde, L., Heavens, A.F., Jimenez, R., 2016. Discrepancies between cftlens cosmic shear and Planck: new physics or systematic effects? *MNASSA* 459 (1), 971–981. doi:10.1093/mnras/stw707, arXiv:1602.02960.
- Laigle, C., McCracken, H.J., Ilbert, O., Hsieh, B.C., Davidzon, I., Capak, P., Hasinger, G., Silverman, J.D., Pichon, C., Coupon, J., Aussel, H., Le Borgne, D., Caputi, K., Cassata, P., Chang, Y.Y., Civano, F., Dunlop, J., Fynbo, J., Kartaltepe, J.S., Koekemoer, A., Le Fèvre, O., Le Floc'h, E., Leauthaud, A., Lilly, S., Lin, L., Marchesi, S., Milvang-Jensen, B., Salvato, M., Sanders, D.B., Scoville, N., Smolcic, V., Stockmann, M., Taniguchi, Y., Tasca, L., Toft, S., Vaccari, M., Zabl, J., 2016. The COSMOS2015 catalog: Exploring the $1 < z < 6$ universe with half a million galaxies. *ApJS* 224 (2), 24. doi:10.3847/0067-0049/224/2/24, arXiv:1604.02350.
- Lang, D., 2014. UnWISE: Unblurred coadds of the WISE imaging. *AJ* 147 (5), 108. doi:10.1088/0004-6256/147/5/108, arXiv:1405.0308.
- Laureijs, R., Amiaux, J., Arduini, S., Auguères, J.L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvel, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., Pasian, F., Percival, W., Rhodes, J., Saavedra Criado, G., Sauvage, M., Scaramella, R., Valenziano, L., Warren, S., Bender, R., Castander, F., Cimatti, A., Le Fèvre, O., Kurki-Suonio, H., Levi, M., Lilje, P., Meylan, G., Nichol, R., Pedersen, K., Popa, V., Rebolo Lopez, R., Rix, H.W., Rottgering, H., Zeilinger, W., Grupp, F.,

- Hudelot, P., Massey, R., Meneghetti, M., Miller, L., Paltani, S., Paulin-Henriksson, S., Pires, S., Saxton, C., Schrabback, T., Seidel, G., Walsh, J., Aghanim, N., Amendola, L., Bartlett, J., Baccigalupi, C., Beaulieu, J.P., Benabed, K., Cuby, J.G., Elbaz, D., Fosalba, P., Gavazzi, G., Helmi, A., Hook, I., Irwin, M., Kneib, J.P., Kunz, M., Mannucci, F., Moscardini, L., Tao, C., Teyssier, R., Weller, J., Zamorani, G., Zapatero Osorio, M.R., Boulade, O., Foumond, J.J., Di Giorgio, A., Guttridge, P., James, A., Kemp, M., Martignac, J., Spencer, A., Walton, D., Blümchen, T., Bonoli, C., Bortoletto, F., Cerna, C., Corcione, L., Fabron, C., Jahnke, K., Ligorì, S., Madrid, F., Martin, L., Morgante, G., Pampalona, T., Prieto, E., Riva, M., Toledo, R., Trifoglio, M., Zerbi, F., Abdalla, F., Douspis, M., Grenet, C., Borgani, S., Bouwens, R., Courbin, F., Delouis, J.M., Dubath, P., Fontana, A., Frailis, M., Grazian, A., Koppenhöfer, J., Mansutti, O., Melchior, M., Mignoli, M., Mohr, J., Neissner, C., Noddle, K., Portet, M., Scodreggio, M., Serrano, S., Shane, N., Starck, J.L., Surace, C., Taylor, A., Verdoes-Kleijn, G., Vuerli, C., Williams, O.R., Zacchei, A., Altieri, B., Escudero Sanz, I., Kohley, R., Oosterbroek, T., Astier, P., Bacon, D., Bardelli, S., Baugh, C., Bellagamba, F., Benoist, C., Bianchi, D., Biviano, A., Branchini, E., Carbone, C., Cardone, V., Clementi, S., Colombi, S., Conselice, J., Lombardi, M., Deacon, N., Dunlop, J., Fedeli, C., Fontanot, F., Franzetti, P., Giocoli, C., Garcia-Bellido, J., Gow, J., Heavens, A., Hewett, P., Heymans, C., Holland, A., Huang, Z., Ilbert, O., Joachimi, B., Jennings, E., Kerins, E., Kiessling, A., Kirk, D., Kortak, R., Krause, O., Lahav, O., van Leeuwen, F., Lesgourgues, J., Lombardi, M., Magliocchetti, M., Maguire, K., Majerotto, E., Maoli, R., Marulli, F., Maurogordato, S., McCracken, H., McLure, R., Melchiorri, A., Merson, A., Moresco, M., Nonino, M., Norberg, P., Peacock, J., Pello, R., Penny, M., Pettorino, V., Di Porto, C., Pozzetti, L., Quercellini, C., Radovich, M., Rassat, A., Roche, N., Ronayette, S., Rossetti, E., Sartoris, B., Schneider, P., Semboloni, E., Serjeant, S., Simpson, F., Skordis, C., Smadja, G., Smartt, S., Spano, P., Spiro, S., Sullivan, M., Tilquin, A., Trotta, R., Verde, L., Wang, Y., Williger, G., Zhao, G., Zoubian, J., Zucca, E., 2011. Euclid definition study report. arXiv e-prints, [arXiv:1110.3193](https://arxiv.org/abs/1110.3193).
- Marín-Franch, A., Chueca, S., Moles, M., Benítez, N., Taylor, K., Cepa, J., Cenarro, A.J., Cristóbal-Hornillos, D., Ederoclite, A., Gruel, N., Hernández-Fuertes, J., López-Sainz, A., Luis-Simoes, R., Rueda-Teruel, F., Rueda-Teruel, S., Varela, J., Yanes-Díaz, A., Brauneck, U., Danielou, A., Dupke, R., Fernández-Soto, A., Mendes de Oliveira, C., Sodré, L., 2012. Design of the J-PAS and J-PLUS filter systems. In: Navarro, R., Cunningham, C.R., Prieto, E. (Eds.), *Modern Technologies in Space- and Ground-Based Telescopes and Instrumentation II*. In: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 8450, p. 84503S. [doi:10.1117/12.925430](https://doi.org/10.1117/12.925430).
- Mendes de Oliveira, C., Ribeiro, T., Schoenell, W., Kanaan, A., Overzier, R.A., Molino, A., Sampedro, L., Coelho, P., Barbosa, C.E., Cortesi, A., Costa-Duarte, M.V., Herpich, F.R., Hernandez-Jimenez, J.A., Placco, V.M., Xavier, H.S., Abramo, L.R., Saito, R.K., Chies-Santos, A.L., Ederoclite, A., Lopes de Oliveira, R., Gonçalves, D.R., Akra, S., Almeida, L.A., Almeida-Fernandes, F., Beers, T.C., Bonatto, C., Bonoli, S., Cypriano, E.S., Vinicius-Lima, E., de Souza, R.S., Fabiano de Souza, G., Ferrari, F., Gonçalves, T.S., Gonzalez, A.H., Gutiérrez-Soto, L.A., Hartmann, E.A., Jaffe, Y., Kerber, L.O., Lima-Dias, C., Lopes, P.A.A., Menendez-Delmestre, K., Nakazono, L.M.I., Novais, P.M., Ortega-Minakata, R.A., Pereira, E.S., Perottoni, H.D., Queiroz, C., Reis, R.R.R., Santos, W.A., Santos-Silva, T., Santucci, R.M., Barbosa, C.L., Siffert, B.B., Sodré, L., Torres-Flores, S., Westera, P., Whitten, D.D., Alcaniz, J.S., Alonso-García, J., Alencar, S., Alvarez-Cand al, A., Amram, P., Azanha, L., Barbá, R.H., Bernardinelli, P.H., Borges Fernandes, M., Branco, V., Brito-Silva, D., Buzzo, M.L., Caffer, J., Campillay, A., Cano, Z., Carvano, J.M., Castejon, M., Cid Fernandes, R., Dantas, M.L.L., Daflon, S., Damke, G., de la Reza, R., de Melo de Azevedo, L.J., De Paula, D.F., Diem, K.G., Donnerstein, R., Dors, O.L., Dupke, R., Eikenberry, S., Escudero, C.G., Faifer, F.R., Fariás, H., Fernandes, B., Fernandes, C., Fontes, S., Galarza, A., Hirata, N.S.T., Katena, L., Gregorio-Hetem, J., Hernández-Fernández, J.D., Izzo, L., Jaque Arancibia, M., Jatenco-Pereira, V., Jiménez-Teja, Y., Kann, D.A., Krabbe, A.C., Labayru, C., Lazzaro, D., Lima Neto, G.B., Lopes, A.A.R., Magalhães, R., Makler, M., de Menezes, R., Miralda-Escudé, J., Monteiro-Oliveira, R., Montero-Dorta, A.D., Muñoz-Elgueta, N., Nemmen, R.S., Nilo Castellón, J.L., Oliveira, A.S., Ortiz, D., Pattaro, E., Pereira, C.B., Quint, B., Riguccini, L., Rocha Pinto, H.J., Rodrigues, I., Roig, F., Rossi, S., Saha, K., Santos, R., Schnorr Müller, A., Sesto, L.A., Silva, R., Smith Castellí, A.V., Teixeira, R., Telles, E., Thom de Souza, R.C., Thöne, C., Trevisan, M., de Ugarte Postigo, A., Urrutia-Viscarra, F., Veiga, C.H., Vika, M., Vitorelli, A.Z., Werle, A., Werner, S.V., Zaritsky, D., 2019. The Southern Photometric Local Universe Survey (S-PLUS): improved SEDs, morphologies, and redshifts with 12 optical filters. *MNASSA* 489 (1), 241–267. [doi:10.1093/mnras/stz1985](https://doi.org/10.1093/mnras/stz1985), [arXiv:1907.01567](https://arxiv.org/abs/1907.01567).
- Molino, A., Benítez, N., Ascaso, B., Coe, D., Postman, M., Jovel, S., Host, O., Lahav, O., Seitz, S., Medezinski, E., Rosati, P., Schoenell, W., Koekemoer, A., Jimenez-Teja, Y., Broadhurst, T., Melchior, P., Balestra, I., Bartelmann, M., Bouwens, R., Bradley, L., Czikon, N., Donahue, M., Ford, H., Graur, O., Graves, G., Grillo, C., Infante, L., Jha, S.W., Kelson, D., Lazkoz, R., Lemze, D., Maoz, D., Mercurio, A., Meneghetti, M., Merten, J., Moustakas, L., Nonino, M., Orgaz, S., Riess, A., Rodney, S., Sayers, J., Umetsu, K., Zheng, W., Zitron, A., 2017. CLASH: accurate photometric redshifts with 14 HST bands in massive galaxy cluster cores. *Mon. Not. R. Astron. Soc.* 470 (1), 95–113. [doi:10.1093/mnras/stx1243](https://doi.org/10.1093/mnras/stx1243).
- Molino, A., Benítez, N., Moles, M., Fernández-Soto, A., Cristóbal-Hornillos, D., Ascaso, B., Jiménez-Teja, Y., Schoenell, W., Arnalte-Mur, P., Pović, M., Coe, D., López-Sanjuan, C., Díaz-García, L.A., Varela, J., Stefanon, M., Cenarro, J., Matute, I., Masegosa, J., Márquez, I., Perea, J., Del Olmo, A., Husillos, C., Alfaro, E., Aparicio-Villagas, T., Cerviño, M., Huertas-Company, M., Aguerri, J.A.L., Broadhurst, T., Cabrera-Cañó, J., Cepa, J., González, R.M., Infante, L., Martínez, V.J., Prada, F., Quintana, J.M., 2014. The ALHAMBRA Survey: Bayesian photometric redshifts with 23 bands for 3 deg. *MNASSA* 441 (4), 2891–2922. [doi:10.1093/mnras/stu387](https://doi.org/10.1093/mnras/stu387), [arXiv:1306.4968](https://arxiv.org/abs/1306.4968).
- Molino, A., Costa-Duarte, M.V., Sampedro, L., Herpich, F.R., Sodré, J., Mendes de Oliveira, C., Schoenell, W., Barbosa, C.E., Queiroz, C., Lima, E.V.R., Azanha, L., Muñoz-Elgueta, N., Ribeiro, T., Kanaan, A., Hernandez-Jimenez, J.A., Cortesi, A., Akra, S., Lopes de Oliveira, R., Torres-Flores, S., Lima-Dias, C., Castellon, J.L.N., Damke, G., Alvarez-Candal, A., Jiménez-Teja, Y., Coelho, P., Pereira, E., Montero-Dorta, A.D., Benítez, N., Gonçalves, T.S., Santana-Silva, L., Werner, S.V., Almeida, L.A., Lopes, P.A.A., Chies-Santos, A.L., Telles, E., de Souza, T., C. R., Gonçalves, D.R., de Souza, R.S., Makler, M., Buzzo, M.L., Placco, V.M., Nakazono, L.M.I., Saito, R.K., Overzier, R.A., Abramo, L.R., 2020. Assessing the photometric redshift precision of the S-PLUS survey: the Stripe-82 as a test-case. *MNASSA* 499 (3), 3884–3908. [doi:10.1093/mnras/staa1586](https://doi.org/10.1093/mnras/staa1586), [arXiv:1907.06315](https://arxiv.org/abs/1907.06315).
- Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., Lerer, A., 2017. Automatic differentiation in PyTorch. In: *NIPS-W*.
- Pharo, J., Malhotra, S., Rhoads, J., Ryan, R., Tilvi, V., Pirzkal, N., Finkelstein, S., Windhorst, R., Grogan, N., Koekemoer, A., Zheng, Z., Hathi, N., Kim, K., Joshi, B., Yang, H., Christensen, L., Cimatti, A., Gardner, J.P., Zakamska, N., Ferreras, I., Hibon, P., Pasquali, A., 2018. Spectrophotometric redshifts in the faint infrared grism survey: Finding overdensities of faint galaxies. *Astrophys. J.* 856 (2), 116. [doi:10.3847/1538-4357/aaadfc](https://doi.org/10.3847/1538-4357/aaadfc).
- Polsterer, K.L., D'Isanto, A., Gieseke, F., 2016. Uncertain photometric redshifts. arXiv e-prints, [arXiv:1608.08016](https://arxiv.org/abs/1608.08016).
- Rosenblatt, F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* 65–386.
- Rozo, E., Wechsler, R.H., Rykoff, E.S., Annis, J.T., Becker, M.R., Evrard, A.E., Frieman, J.A., Hansen, S.M., Hao, J., Johnston, D.E., Koester, B.P., McKay, T.A., Sheldon, E.S., Weinberg, D.H., 2018. Cosmological constraints from the SDSS maxbcg cluster catalog. *Astrophys. J.* 708 (1), 645–660. [doi:10.1088/0004-637x/708/1/645](https://doi.org/10.1088/0004-637x/708/1/645).
- Sadeh, I., Abdalla, F.B., Lahav, O., 2016. ANNz2: Photometric redshift and probability distribution function estimation using machine learning. *Publ. Astron. Soc. Pac.* 128 (968), 104502. [doi:10.1088/1538-3873/128/968/104502](https://doi.org/10.1088/1538-3873/128/968/104502).
- Sánchez, E., 2010. The dark energy survey. *J. Phys. Conf. Ser.* 259, 012080. [doi:10.1088/1742-6596/259/1/012080](https://doi.org/10.1088/1742-6596/259/1/012080).
- Speagle, J.S., Eisenstein, D.J., 2017. Deriving photometric redshifts using fuzzy archetypes and self-organizing maps – I. Methodology. *Mon. Not. R. Astron. Soc.* 469 (1), 1186–1204. [doi:10.1093/mnras/stw1485](https://doi.org/10.1093/mnras/stw1485).
- Taylor, M.B., 2005. TOPCAT & STIL: Starlink table/VOTable processing software. In: Shopbell, P., Britton, M., Ebert, R. (Eds.), *Astronomical Data Analysis Software and Systems XIV*. In: *Astronomical Society of the Pacific Conference Series*, vol. 347, p. 29.
- Weinberg, D.H., Mortonson, M.J., Eisenstein, D.J., Hirata, C., Riess, A.G., Rozo, E., 2013. Observational probes of cosmic acceleration. *Phys. Rep.* 530 (2), 87–255. [doi:10.1016/j.physrep.2013.05.001](https://doi.org/10.1016/j.physrep.2013.05.001).
- Williams, C.K.I., Rasmussen, C.E., 1995. Gaussian processes for regression. In: *Proceedings of the 8th International Conference on Neural Information Processing Systems*. In: *NIPS'95*, MIT Press, Cambridge, MA, USA, pp. 514–520.