

TOPz: Photometric redshifts using template fitting applied to the GAMA survey

E. Tempel^{1,2,*}, J. Laur¹, Z. R. Jones¹, R. Kipper¹, L. J. Liivamägi¹, D. Pandey¹, G. Sakteos¹, A. Tamm¹,
A. N. Triantafyllaki¹, and T. Tuvikene¹

¹ Tartu Observatory, University of Tartu, Observatooriumi 1, 61602 Tõravere, Estonia

² Estonian Academy of Sciences, Kohtu 6, 10130 Tallinn, Estonia

Received 6 January 2025 / Accepted 5 October 2025

ABSTRACT

Context. Accurate photometric redshift (photo-z) estimation is crucial for cosmological and galaxy evolution studies, especially with the advent of large-scale photometric surveys.

Aims. We developed a photo-z estimation code called TOPz (Tartu Observatory Photo-z) and applied it to the GAMA photometric catalogue. Using nine-band photometric data from the GAMA project, we assessed the accuracy of TOPz by comparing its photo-z estimates to available spectroscopic redshifts from GAMA and DESI DR1. The latter extends to $z < 2$ and $m_z < 24$, allowing the photo-z accuracy to be validated beyond the GAMA limits.

Methods. TOPz employs a Bayesian template-fitting approach to estimate photo-z from marginalised redshift posteriors. We generated synthetic galaxy spectra using the CIGALE software and ran template set optimisation. We improved the photometry by applying flux and flux uncertainty corrections. An analytical prior was then imposed on the resulting posteriors to refine the redshift estimates.

Results. The photo-z estimates produced by TOPz show good agreement with the spectroscopic redshifts in the low-redshift regime ($z < 0.5$) where the majority (95%) of the GAMA spectroscopic redshifts are. We demonstrate the redshift accuracy across various magnitude bins and tested how the flux corrections and posteriors reflect the actual uncertainty of the estimates. For the GAMA sample, the $\sigma_{\text{NMAD}} = 0.012$ for $m_z < 18$ and increases to $\sigma_{\text{NMAD}} = 0.021$ for $m_z > 19$. The outlier fraction ($|\Delta z|/(1+z) > 0.1$) in the same magnitude bins increases from 1 to 5%. We show that the TOPz results are consistent with those obtained from other photo-z codes (EAZY and SFM) applied to the same data set. Additionally, TOPz estimates stellar masses as a by-product, comparable to those calculated by other methods. We have made the full GAMA photo-z catalogue and all the codes and scripts used for the analysis and figures publicly available.

Conclusions. TOPz is an advanced photo-z estimation code that integrates flux corrections, physical priors, and template set optimisation to provide state-of-the-art photo-z among competing template-based redshift estimators. Future work will focus on incorporating additional photometric data and applying the TOPz algorithm to the J-PAS narrow-band survey, further validating and enhancing its capabilities.

Key words. methods: observational – methods: statistical – techniques: photometric – catalogs – galaxies: distances and redshifts – galaxies: general

1. Introduction

Photometric redshifts (photo-z) are essential in cosmology and galaxy evolution studies, allowing the expansion of galaxy samples limited by spectroscopic surveys. New surveys such as the Rubin Observatory, Euclid, and Roman Space Telescope will provide deep photometry over large sky areas that allow the inclusion of faint galaxies in the analysis. Estimating accurate photo-z for them is crucial since the number of objects is too numerous for spectroscopic follow-up. Salvato et al. (2019) and Newman & Gruen (2022) provide an excellent overview of how current photo-z codes have to be improved to take advantage of the next-generation photometric surveys.

A wide variety of methods have been developed for estimating photo-z. They are divided into two primary categories based on machine learning or template fitting. Machine-learning methods include neural networks (Sadeh et al. 2016; Bilicki et al. 2018; D’Isanto & Polsterer 2018; Pasquet et al. 2019; Schuld et al. 2021; Jones et al. 2024; Treyer et al. 2024; Pathi et al. 2025),

random forests or decision trees (Gerdes et al. 2010; Carrasco Kind & Brunner 2013), Gaussian processes (Almosallam et al. 2016a,b; Soo et al. 2021), a genetic algorithm (Hogan et al. 2015), or other machine-learning approaches (Hatfield et al. 2020; Ansari et al. 2021; Razim et al. 2021; Duncan 2022).

In the machine-learning approach, there is a need for a spectroscopic training sample, which is also used in methods using direct neighbourhood fitting (De Vicente et al. 2016) or scaled flux matching (Baldry et al. 2021). While these methods provide excellent results, they are limited to the set of galaxy types and redshift ranges covered by the spectroscopic training set (e.g. Hildebrandt et al. 2008; Beck et al. 2017; Toribio San Cipriano et al. 2024). Solutions to overcome this are proposed by augmenting the spectroscopic samples with simulated data (Moskowitz et al. 2024) or aiding the photo-z estimation with a classification algorithm (Fotopoulou & Paltani 2018).

Template-fitting-based photo-z estimation overcomes many problems related to methods that use spectroscopic samples as a training set. Template fitting in its general form allows a Bayesian inference (Laplace 1774) to estimate photo-z posteriors. However, not all template-fitting algorithms are Bayesian in nature.

* Corresponding author: elmo.tempel@ut.ee

Additionally, when input templates are physical, template-fitting methods can provide a rough estimate for the physical properties of galaxies. Examples of template-fitting application include the following: EAZY (Brammer et al. 2008), LePhare (Arnouts et al. 1999; Ilbert et al. 2006), BPZ (Benítez 2000), ZEBRA (Feldmann et al. 2006), HyperZ (Bolzonella et al. 2000), and TOPz (Laur et al. 2022). Lately, hybrid methods that combine machine learning and template fitting are emerging (Hatfield et al. 2022; Tanigawa et al. 2024) or hybrid techniques that combine spectroscopic training sets with template fitting (Beck et al. 2016) that improve the photo-z accuracy (Cavuoti et al. 2017).

Many factors influence the accuracy and performance of different template-fitting photo-z codes and methods. The dependence of the used set of templates has been studied in Bolzonella et al. (2000) or the combination of different sets of templates (Duncan et al. 2018), both aiming to improve the accuracy of the photo-z estimation. Brinchmann et al. (2017) demonstrate that template-fitting photo-z codes are reliable down to the 30th magnitude utilising the MUSE observations (Bacon et al. 2017). The need to constantly improve photo-z estimation methods has been a result of applying it to numerous photometric surveys in the past, such as the PAU survey (Navarro-Girónés et al. 2024), KiDS survey (Bilicki et al. 2018), ALHAMBRA survey (Molino et al. 2014), DESI Legacy Imaging Survey (Zou et al. 2019), miniJPAS survey (Hernán-Caballero et al. 2021), or covering the whole extra-Galactic sky (Bilicki et al. 2016). Wittman et al. (2016) point out that photo-z methods tend to be overconfident and underestimate the actual uncertainty when photo-z is compared with spectroscopic measurements, resulting in many occasions of overly optimistic estimates (Brescia et al. 2019).

Comparisons between different photo-z codes have been made in various contexts. They all have allowed the conclusion to be drawn that there is no particular method or set of spectral templates that significantly outperforms others and further work is required to identify how to best select between machine-learning and template-fitting approaches for each individual galaxy (Abdalla et al. 2011; Tanaka et al. 2018; Euclid Collaboration: Desprez et al. 2020; Schmidt et al. 2020). Dahlen et al. (2013) showed that methods that adopted emission lines in spectral templates and added a zero-point offset to observed fluxes performed better. The best method or combination of techniques is goal-dependent and requires science-specific performance metrics. Further improvements in photo-z can be achieved by constraining the photo-z posterior with the underlying cosmic web from spectroscopic surveys (Shuntov et al. 2020; Tosone et al. 2023). Validation techniques, such as cross-correlation (Gatti et al. 2018; Naidoo et al. 2023) and methods based on normalising flows (Crenshaw et al. 2024), have also been developed. The full photo-z posterior is beneficial for extracting complete information from photo-z codes (López-Sanjuan et al. 2017; Teixeira et al. 2024). Currently, there is no preferred code for photo-z estimation; this gives space to improve the current codes available or even create new ones.

In our previous paper Laur et al. (2022), we used a template-fitting approach to photo-z estimation where the physically motivated templates were generated using the CIGALE software (Boquien et al. 2019; Yang et al. 2022). With the upcoming photometric information from surveys utilising narrow-band filters, template fitting is expected to achieve an even higher accuracy than what is described in Laur et al. (2022), especially with the upcoming J-PAS survey (Bonoli et al. 2021), which provides photometry in more than 50 filters. The currently available narrow-band photometric surveys with fewer filters, such as COMBO-17 (Wolf et al. 2004), COSMOS (Salvato et al. 2009),

or ALHAMBRA (Molino et al. 2014), have shown how photo-z estimation benefits from the narrow-band filters. Benítez (2000) demonstrated the impact of the width of the band, the number of bands, and the spacing between bands while computing photo-z with template fitting.

In this paper, we present the improved version of the TOPz (Tartu Observatory Photo-z) code, where we specifically improved the observed flux corrections (zero-point offsets) and template set selection and applied physically motivated priors. We extended its application to the GAMA (Galaxy And Mass Assembly) sample of galaxies. We aim to demonstrate that this approach will be valuable for future surveys, such as 4MOST WAVES (Driver et al. 2019), which is based on the same photometry as the GAMA survey.

The outline of the paper is the following: in Sect. 2, we describe the Bayesian photo-z estimation method. The GAMA sample is described in Sect. 3. In Sect. 4, we describe the scheme for correcting observational flux and flux uncertainties for photo-z estimation. In Sect. 5, we describe the construction and selection of the templates for the TOPz photo-z code and in Sect. 6 we describe how the TOPz can be used for star-galaxy identification. In Sect. 7, we describe the construction of physical priors based on the survey volume and luminosity function of galaxies. In Sect. 8, we present the TOPz photo-z results applied to the GAMA data, and in Sect. 9 we conclude the paper and discuss future directions. The TOPz code has been made publicly available and the computed photo-z for the GAMA data are available in the CDS (see Appendix A). In Appendix B we validate our photo-z estimates using DESI DR1 spectroscopic data beyond the GAMA spectroscopic limits.

2. TOPz: Bayesian photometric redshifts using template fitting

Template-fitting methods involve matching the observed photometric data of galaxies to a set of pre-defined spectral templates, which represent different types of galaxies at various redshifts. The goal is to compute the likelihood that a given template at a particular redshift matches the observed data, allowing us to derive the probability distribution function (PDF) of the galaxy's photo-z.

Within the framework of Bayesian probability, the task of finding redshift $\zeta = \ln(1+z)$ ¹ given the galaxy observed fluxes F , flux uncertainties² and total magnitude m_0 is to find the conditional probability $p(\zeta|F, m_0)$. To simplify the notations, the observed galaxy properties F and m_0 for a single galaxy are denoted as G in some equations below, hence $p(\zeta|F, m_0) \equiv p(\zeta|G)$. The photo-z probability $p(\zeta|F, m_0)$ for a single galaxy using template fitting can be expressed as (Benítez 2000; Laur et al. 2022)

$$p(\zeta|F, m_0) = \sum_{T \in \mathcal{T}} p(\zeta, T|F, m_0) \propto \sum_{T \in \mathcal{T}} p(\zeta, T|m_0)p(F|\zeta, T), \quad (1)$$

where $p(F|\zeta, T)$ gives the likelihood that observed galaxy fluxes F correspond to the template T at redshift ζ (see Sect. 2.1) and $p(\zeta, T|m_0)$ gives the prior information that galaxy with a reference magnitude m_0 has a specific spectral template T and redshift ζ (see Sect. 7).

¹ Following Baldry (2018) we use $\zeta = \ln(1+z)$ instead of the redshift z .

² In all equations, the notation F includes both, fluxes and flux uncertainties. When necessary, flux and flux uncertainty in a given passband α are designated as F_α and σ_{F_α} .

In the photo- z probability estimation, we use fluxes F to estimate the redshift likelihood $p(F|\zeta, T)$ and we use total magnitude m_0 for the prior $p(\zeta, T|m_0)$. Magnitude m_0 in a reference passband represents the total magnitude of a galaxy, which may differ from the fluxes F utilised in template fitting. For template fitting, galaxy colours (flux ratios) must be accurate. For this reason, the galaxy fluxes for template fitting are usually measured using aperture-corrected fluxes and the same sky mask in all observed bands. Hence, the observed fluxes that provide accurate colours are underestimating the total flux of a galaxy.

For the template fitting, we have to define a set of templates $\mathbf{T}$ that covers all different types of galaxies. The generation of galaxy templates and subsequent template selection for photo- z estimation are described in Sect. 5.

2.1. Redshift likelihood estimation

The redshift likelihood for galaxies can be expressed as (see Benítez 2000)

$$p(F|\zeta, T) \propto \frac{1}{\sqrt{F_{\text{TT}}}} \exp \left[-\frac{\chi^2(\zeta, T, a_m)}{2} \right], \quad (2)$$

where $F_{\text{TT}}^{-1/2}$ is a normalisation parameter that comes from integrating over a nuisance parameter a (see Eqs. (3) and (5)). As is common in template fitting, the χ^2 is defined as

$$\chi^2(\zeta, T, a) = \sum_{\alpha} \frac{(F_{\alpha} - aF_{T\alpha})^2}{\sigma_{F_{\alpha}}^2}, \quad (3)$$

where summation is over all observed fluxes, F_{α} and $\sigma_{F_{\alpha}}$ are observed flux and flux uncertainty in passband α , and $F_{T\alpha}$ is model flux in passband α for template T . The parameter a is a nuisance parameter that minimises the $\chi^2(\zeta, T, a)$ for a given template T and redshift ζ pair. Eq. (3) is minimised when $a = a_m = F_{\text{OT}}/F_{\text{TT}}$. The χ^2 when $a = a_m$ is given as

$$\chi^2(\zeta, T, a_m) = F_{\text{OO}} - \frac{F_{\text{OT}}^2}{F_{\text{TT}}}. \quad (4)$$

In the last equation, the quantities F_{XX} are defined as

$$F_{\text{OO}} = \sum_{\alpha} \frac{F_{\alpha}^2}{\sigma_{F_{\alpha}}^2}, \quad F_{\text{TT}} = \sum_{\alpha} \frac{F_{T\alpha}^2}{\sigma_{F_{\alpha}}^2}, \quad F_{\text{OT}} = \sum_{\alpha} \frac{F_{\alpha} F_{T\alpha}}{\sigma_{F_{\alpha}}^2}. \quad (5)$$

The redshift likelihood normalisation parameter $F_{\text{TT}}^{-1/2}$ in Eq. (2) depends on the internal normalisation of the template. For a single template T , the template normalisation is irrelevant. However, when combining redshift likelihoods from multiple templates, the internal normalisation of each template affects the redshift likelihood $p(F|\zeta, T)$. To overcome this problem, we renormalise each template T to achieve the minimum χ^2 when $a_m = 1$. Under this assumption, for arbitrarily normalised templates, the redshift likelihood can be written as

$$p(F|\zeta, T) \propto \frac{\sqrt{F_{\text{TT}}}}{F_{\text{OT}}} \exp \left[-\frac{\chi^2(\zeta, T, a_m)}{2} \right]. \quad (6)$$

In practical application, Eq. (6) is used to compute the redshift likelihood for a given template T at redshift ζ .

Model flux $F_{T\alpha}$ in passband α is computed as

$$F_{T\alpha} = \frac{\int f_{T,\zeta}(\lambda) W_{\alpha}(\lambda) d\lambda}{\int W_{\alpha}(\lambda) d\lambda}, \quad (7)$$

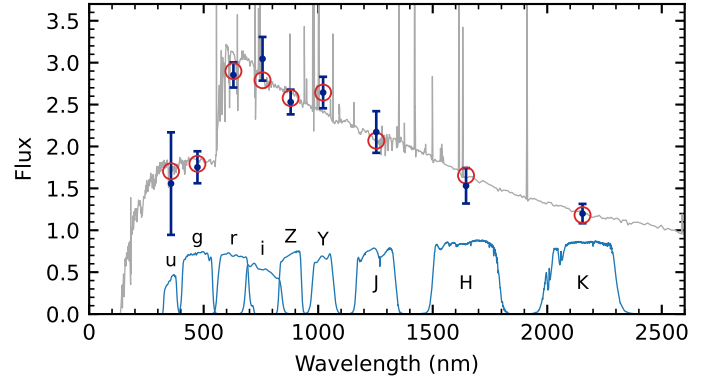


Fig. 1. Illustration of the template-fitting approach. The model spectrum is shown with a grey line, and red circles denote the model fluxes in different passbands. The passband transmission curves used for the GAMA data are shown as blue lines. The blue points with error bars are observed fluxes with 2σ errors for a given galaxy. The template spectrum is normalised to best match the observed fluxes.

where $f_{T,\zeta}(\lambda)$ is the specific flux of the template spectra T at redshift ζ and W_{α} is window function or transmission curve for passband α . Given a model template spectra $f_T(\lambda)$ at redshift zero, the redshifted spectrum $f_{T,\zeta}(\lambda)$ is computed as

$$f_{T,\zeta}(\lambda) \propto f_T(\lambda e^{-\zeta}). \quad (8)$$

Figure 1 illustrates the basics of the template-fitting approach. We compute model fluxes from the model template spectrum in different passbands using Eq. (7). The model fluxes are compared against observed fluxes while including the observed flux uncertainties, and the likelihood that a given model template is good is estimated using Eq. (6).

2.2. TOPz: Photometric redshift code

The Bayesian photo- z estimation based on template fitting is a general approach implemented in many codes. Although all the codes share the basics of the template-fitting approach, the codes do not provide identical photo- z estimates. The reason is the practical implementation and technical details of the template-fitting approach. To have complete control over implemented features and to fully understand the technical details, we developed the TOPz code (Tartu Observatory Photo- z). The TOPz code is a Bayesian photo- z code based on template fitting, incorporating many aspects of the previously developed template-fitting codes. The precursor of the current implementation of the code is given in Laur et al. (2022), but many of the details in the current TOPz implementation are significantly improved upon.

To achieve the best possible photo- z using the template-fitting approach, we have to pay attention to the following aspects: galaxy template spectra generation and subsequent template set optimisation (see Sect. 5), observed flux and flux uncertainty corrections (see Sect. 4), physically motivated priors for photo- z (see Sect. 7). All these aspects are equally important and affect the accuracy and precision of the estimated photo- z . In Fig. 2, we show the schematic overview of the TOPz code, highlighting the key inputs and outputs.

The TOPz code is written in Fortran programming language and is publicly available. Together with the Fortran code, we make available all Python scripts used to run the TOPz code and generate all the figures presented in this paper.

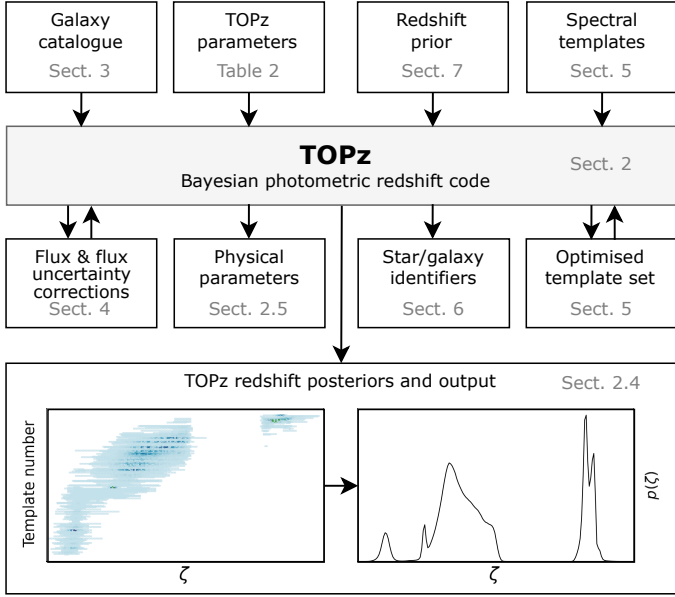


Fig. 2. Schematic overview of the TOPz code, illustrating key inputs and outputs. The two-way arrows indicate that an iterative approach can be applied for further optimisation. Refer to the sections indicated in the corresponding boxes for additional information. See Fig. 3 for the detailed description of the TOPz posteriors and output.

2.3. TOPz: Marginalised redshift posterior

The main output of the TOPz code is a two-dimensional probability distribution $p(\zeta, T | F, m_0)$ for each input source given the observed fluxes, flux uncertainties and total magnitude. This probability distribution is a combination of redshift likelihood and prior information. An example of a two-dimensional probability distribution for a rare source with many maxima is shown in Fig. 3. Most of the sources have only one dominant maximum (see Sect. 8.1 for the statistics in the GAMA sample).

For redshift estimation, the two-dimensional probability $p(\zeta, T | F, m_0)$ is marginalised over all input templates (see Eq. (1)), which provides us a redshift posterior distribution $p(\zeta | F, m_0)$. An example of a redshift posterior for a single galaxy is shown in the lower panel in Fig. 3. For the majority of galaxies, there is only one dominant redshift peak. However, for many fainter galaxies the redshift posterior has more than one peak. To simplify the notations below, the one-dimensional redshift posterior distribution for galaxy G (with parameters F and m_0) given an input template set $\mathbf{T}$ is designated as $p_{\mathbf{T}}^G(\zeta)$.

It is known that the width of the peaks in the photo-z posterior usually underestimates the true uncertainty of the redshift estimate. To mitigate this problem, the marginalised redshift distribution can be convolved with a kernel that smooths the redshift posterior distribution and increases the width of the individual redshift peaks. In TOPz code, we have the option to convolve the redshift posterior $p_{\mathbf{T}}^G(\zeta)$ with a Gaussian kernel $K(\zeta; m_0)$, where the width of the kernel depends on the galaxy total magnitude m_0 . The kernel width dependence on galaxy magnitude should be explicitly estimated for a given data set. In all equations given in Sect. 2.4, we assume that the redshift posterior distribution is a convolved distribution.

2.4. TOPz: Output

In TOPz code, we identify three dominant peaks in the redshift posterior $p_{\mathbf{T}}^G(\zeta)$, designated as ‘best’, ‘alt1’, and ‘alt2’. These

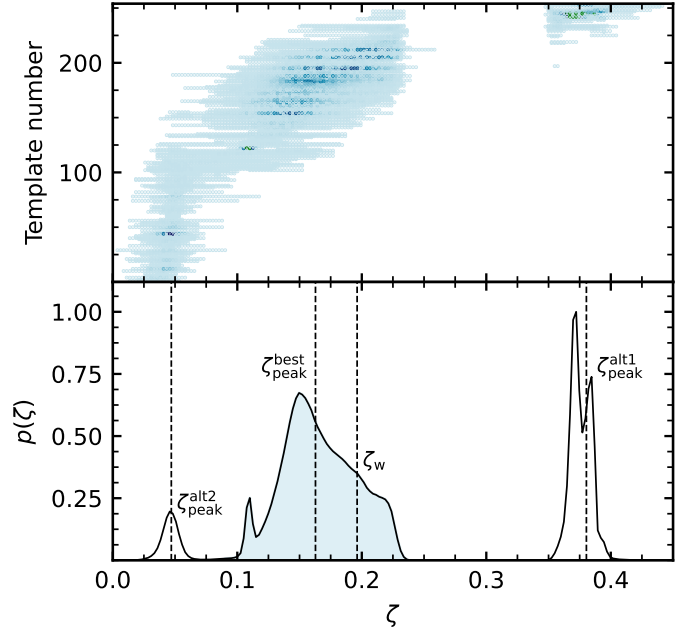


Fig. 3. TOPz redshift posterior for a source with many maxima and a spectroscopic redshift of $z \approx 0.21$ ($\zeta \approx 0.19$). The upper panel displays a 2D posterior distribution, with all templates arranged in order of increasing weighted ζ values. The colour of each point represents the photo- z likelihood of the corresponding template. The lower panel illustrates the marginalised posterior distribution for the same source. Dashed vertical lines indicate the ζ values for all peaks along with ζ_w . For details on the definitions of the three peaks and the probability-weighted ζ_w , refer to Sect. 2.4. The shaded area in the lower panel represents the range of the best peak, bounded by the minimum ($\zeta_{\min}^{\text{best}}$) and maximum ($\zeta_{\max}^{\text{best}}$) values.

peaks in the posterior distribution are defined so they do not overlap in redshift. The peaks are ordered so that

$$p_{\mathbf{T}}^G(\zeta_{\text{peak}}^{\text{best}}) > p_{\mathbf{T}}^G(\zeta_{\text{peak}}^{\text{alt1}}) > p_{\mathbf{T}}^G(\zeta_{\text{peak}}^{\text{alt2}}). \quad (9)$$

Alternatively, the peaks can be ordered based on the area below the peaks (see Eq. (14)). As follows, we define parameters for the dominant ‘best’ peak. The parameters for ‘alt1’ and ‘alt2’ peaks are defined similarly.

For the best peak $\zeta_{\text{peak}}^{\text{best}}$ we define the peak boundaries $\zeta_{\min}^{\text{best}}$ and $\zeta_{\max}^{\text{best}}$ as follows

$$\zeta_{\min}^{\text{best}} = \arg \min_{\zeta < \zeta_{\text{peak}}^{\text{best}}} \left\{ \zeta_{\text{peak}}^{\text{best}} - \zeta \left| \frac{dp_{\mathbf{T}}^G(\zeta)}{d\zeta} = 0 \text{ \& \; } \frac{p_{\mathbf{T}}^G(\zeta)}{p_{\mathbf{T}}^G(\zeta_{\text{peak}}^{\text{best}})} < f_{\text{peak}} \right. \right\}, \quad (10)$$

$$\zeta_{\max}^{\text{best}} = \arg \min_{\zeta > \zeta_{\text{peak}}^{\text{best}}} \left\{ \zeta - \zeta_{\text{peak}}^{\text{best}} \left| \frac{dp_{\mathbf{T}}^G(\zeta)}{d\zeta} = 0 \text{ \& \; } \frac{p_{\mathbf{T}}^G(\zeta)}{p_{\mathbf{T}}^G(\zeta_{\text{peak}}^{\text{best}})} < f_{\text{peak}} \right. \right\}. \quad (11)$$

The peak boundaries are defined as the first minima before and after the peak maximum that are lower than $f_{\text{peak}} \cdot p_{\mathbf{T}}^G(\zeta_{\text{peak}}^{\text{best}})$, where $f_{\text{peak}} < 1.0$ is a free parameter in the TOPz code. It is recommended to set the f_{peak} to a rather low value (e.g. 0.005) to capture the majority of the peak and to avoid spurious small minima in the posterior distribution. The main aim of f_{peak} is to identify peaks that are clearly separated from each other as shown in Fig. 3. Setting the $f_{\text{peak}} = 0.0$ identifies the full

posterior as one dominant peak. The redshift value, where the posterior peak has a maximum, is designated as $\zeta_{\text{peak}}^{\text{best}}$ and defined as

$$\zeta_{\text{peak}}^{\text{best}} = \arg \max_{\zeta_{\text{min}}^{\text{best}} < \zeta < \zeta_{\text{max}}^{\text{best}}} p_{\mathbf{T}}^G(\zeta). \quad (12)$$

To simplify the notations, we do not specifically indicate that $\zeta_{\text{peak}}^{\text{best}}$ for galaxy G depends on the input template set $\mathbf{T}$. Where needed, the dependence is explicitly noted as $\zeta_{\text{peak}}^{\text{best}}(G, \mathbf{T}) \equiv \zeta_{\text{peak}}^{\text{best}}$. The same simplification applies to other peak properties defined in this section.

Additionally, we define the probability-weighted redshift value for each peak

$$\zeta_w^{\text{best}} = \frac{\int_{\zeta_{\text{min}}^{\text{best}}}^{\zeta_{\text{max}}^{\text{best}}} \zeta p_{\mathbf{T}}^G(\zeta) d\zeta}{\int_{\zeta_{\text{min}}^{\text{best}}}^{\zeta_{\text{max}}^{\text{best}}} p_{\mathbf{T}}^G(\zeta) d\zeta}. \quad (13)$$

The summed probability in the redshift posterior that is covered by the best peak is defined as

$$p_{\text{peak}}^{\text{best}} = \frac{\int_{\zeta_{\text{min}}^{\text{best}}}^{\zeta_{\text{max}}^{\text{best}}} p_{\mathbf{T}}^G(\zeta) d\zeta}{\int_{-\text{inf}}^{\text{inf}} p_{\mathbf{T}}^G(\zeta) d\zeta}. \quad (14)$$

We also provide the posterior weighted redshift value over the full redshift range. It is defined as

$$\zeta_w = \frac{\int_{-\text{inf}}^{\text{inf}} \zeta p_{\mathbf{T}}^G(\zeta) d\zeta}{\int_{-\text{inf}}^{\text{inf}} p_{\mathbf{T}}^G(\zeta) d\zeta}. \quad (15)$$

The difference between ζ_w and ζ_w^{best} together with the $p_{\text{peak}}^{\text{best}}$ value can be used to evaluate the reliability of the estimated photo-z. If the difference between ζ_w and ζ_w^{best} is large or the $p_{\text{peak}}^{\text{best}}$ value is low, then it indicates that the posterior has more than one peak and the photo-z solution is not unique.

The marginalised redshift posterior $p_{\mathbf{T}}^G(\zeta)$ distribution provides the uncertainty of the TOPz photo-z estimate. For approximate calculations, we compute the standard deviation of the redshift posterior distribution. The standard deviation is a good measure of the redshift uncertainty when the posterior distribution has only one peak. The computed standard deviation estimates should be treated with care for multi-peaked posteriors. The standard deviation, the photo-z uncertainty estimate for ζ_w , is computed as

$$\sigma_{\zeta_w}^2 = \frac{\int_{-\text{inf}}^{\text{inf}} p_{\mathbf{T}}^G(\zeta) (\zeta - \zeta_w)^2 d\zeta}{\int_{-\text{inf}}^{\text{inf}} p_{\mathbf{T}}^G(\zeta) d\zeta}. \quad (16)$$

Photo-z uncertainty estimates for ζ_w^{best} , ζ_w^{alt1} , and ζ_w^{alt2} are computed similarly while limiting the integration over a given peak.

2.5. TOPz: Physical parameters of galaxies

For template fitting, we use physical model templates (see Sect. 5). Using the physical model templates allows us to estimate a galaxy's physical parameters, such as total stellar mass, star formation rate, etc., to a good approximation, as these integrated galaxy parameters corresponding to each model template are known a priori.

Below, we describe how we estimate the total stellar mass. All other physical parameters can be estimated similarly and are not explicitly defined here. The total stellar mass M_{star} of a galaxy is estimated while marginalising over the two-dimensional photo-z posterior

$$M_{\text{star}} = \frac{\sum_{T \in \mathbf{T}} \sum_{\zeta} p(F | \zeta, T) a_m M_{\text{star}}^T \left(\frac{D_L(\zeta)}{D_L(\zeta_T)} \right)^2}{\sum_{T \in \mathbf{T}} \sum_{\zeta} p(F | \zeta, T)}, \quad (17)$$

where M_{star}^T and ζ_T are model template stellar mass and redshift, respectively. The $D_L(\zeta)$ is a luminosity distance at redshift ζ , and a_m is the model template scaling factor from template fitting. In general, the stellar mass can also be estimated at a fixed redshift (e.g. from spectroscopic observations) while marginalising only over templates $\mathbf{T}$.

In Sect. 8, we show how the TOPz estimated stellar masses compare with the stellar masses provided in the GAMA data (see Sect. 3). We emphasise that the TOPz does not perform any physical modelling of galaxies, it relies on the physical spectral templates. A dedicated software (e.g. CIGALE or ProSpect) should be used to model galaxy spectral energy distribution in detail.

3. GAMA galaxy sample

Galaxy And Mass Assembly (GAMA) survey is a spectroscopic survey carried out on the Anglo-Australian Telescope with the AAOmega multi-object spectrograph (Driver et al. 2011). GAMA results are released in ‘Data Management Units’ and value-added catalogues are released using other publicly available independent observations to complement its data. In this paper, we take the photometric input data from the GAMA project DR4 (Driver et al. 2022) gkvScienceCatv02 catalogue, which is a ‘science-ready’ subset of the entire GAMA input catalogue (Bellstedt et al. 2020). We exclude duplicated and masked objects from our analysis.

Photometry over a wide range of wavelengths is incorporated into GAMA from various imaging sources such as GALEX, WISE, Herschel observations, and most notably for this paper in optical and near-infrared bands from the European Southern Observatory (ESO) VST KiDS (de Jong et al. 2013) and ESO VISTA VIKING surveys (Edge et al. 2013). For GAMA DR4, the image data from all the included surveys were uniformly regridded and retiled onto rescaled KiDS tiles. Source detection and determination of fluxes (as well as corresponding errors) in all filters were subsequently done using the ProFound package (see Robotham et al. 2018; Bellstedt et al. 2020, for details). The photometry is corrected for foreground Galactic extinction employing Planck maps and potentially contaminating bright stars are masked out using Gaia DR2 data. We use the total fluxes with global sky subtraction (*flux_t* columns in the GAMA database), which is expected to represent the total galaxy flux.

In the current paper, we test TOPz on deep nine-band photometry available in the GAMA survey footprint: u, g, r, i from

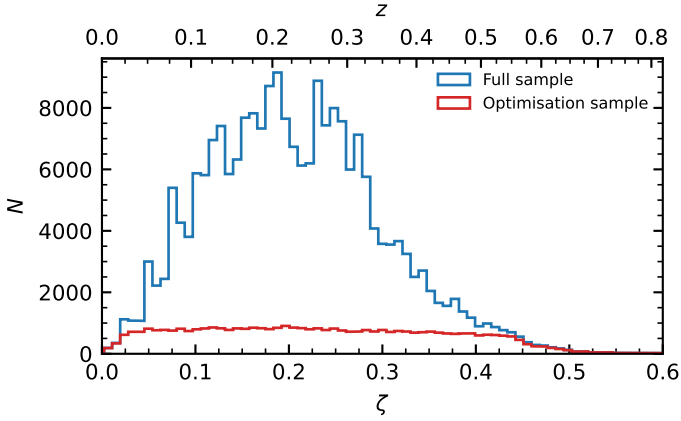


Fig. 4. Spectroscopic redshift distribution in the GAMA catalogue (blue line) and a subset of the spectroscopic sample (red line) used for template set optimisation (see Sect. 5).

the KiDS and Z, Y, J, H, Ks from the VIKING imaging surveys. Despite more filters being available, we restrict ourselves to the optical passbands as [Graham et al. \(2020\)](#) analysed the impact of near-infrared and near-ultraviolet filters on the photo-z estimation and concluded that deeper optical photometry is more beneficial than adding additional filters. It requires a dedicated study to understand the impact of ultraviolet and infrared filters on the GAMA data and how these filters affect the photo-z performance for different types of galaxies with varying magnitudes and redshifts.

In terms of photometry, the GAMA survey is based on the same photometry as the upcoming WAVES ([Driver et al. 2019](#)). The target selection for the WAVES is utilising the photo-z estimates based on the same nine-band photometry. For the WAVES target selection, TOPz is one of the four photo-z methods that are optimally combined (in prep.).

GAMA provides a catalogue of photometrically derived stellar mass estimates for galaxies determined through stellar population synthesis, which is taken from the StellarMasses-GKv24 catalogue in the GAMA database for the equatorial and G23 survey regions subsets. The procedure is described in detail in [Taylor et al. \(2011\)](#) while any updates are listed in [Bellstedt et al. \(2020\)](#). We will use these as a test to compare the stellar masses with those corresponding to the templates selected by our method (see Sect. 2.5).

Since GAMA does not strictly include only galaxies but rather gives sources a type designation such as ‘galaxy’, ‘star’ or ‘ambiguous’ (additional option being ‘artefact’), we will also test if we can recover this classification with appropriate templates (see Sect. 6). For classification analysis, we use only objects marked as ‘galaxy’ or ‘star’ in the GAMA database.

To construct our spectroscopic input sample, we applied the following conditions. We include galaxies within the spectroscopic redshift range between $0.002 < z < 1.0$ and require the normalised redshift quality NQ to be greater than 3. We are further considering only sources within the GAMA spectroscopic survey regions, removing duplicates and objects affected by nearby bright stars. The final input catalogue consists of 221 108 galaxies. In Figs. 4 and 5, we show the redshift distribution for the GAMA spectroscopic catalogue used in the current paper. Note that the spectroscopic sample also includes redshifts from external sources (DESI and DEVILS), which go beyond the GAMA spectroscopic redshift limit. However, the spectroscopic sample for the fainter sources is incomplete.

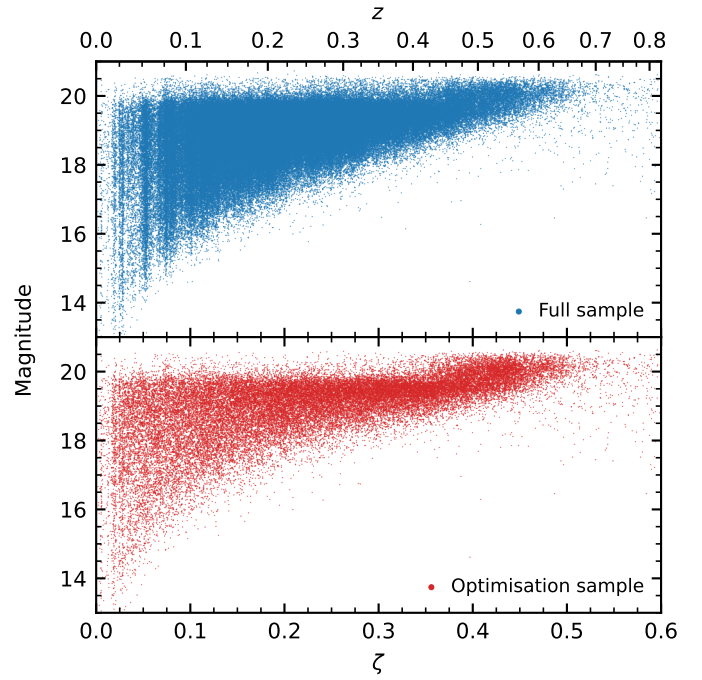


Fig. 5. Galaxy redshift z as a function of observed magnitude m_r for the full sample (blue dots) and for the galaxy sample (red dots) used in template set optimisation.

The full photometric GAMA catalogue contains 1 146 985 objects and includes all types of sources: ‘galaxy’, ‘star’ or ‘ambiguous’ (specified by the `uberclass` variable). Only objects classified as galaxies with spectroscopic redshifts are used for the photo-z analysis in this paper. Additionally, we will run the TOPz code using the full photometric catalogue and make available the photo-z for all photometric sources in the GAMA database (See Appendix A).

4. Observed flux and flux uncertainty corrections

When we fit model templates with observed fluxes, we assume that observed fluxes are accurate (bias-free) and flux errors take into account all statistical and systematic uncertainties. In practice, this assumption is not always valid, and some corrections to fluxes and flux uncertainties might be necessary. In the TOPz code, we have implemented zero-point corrections. First, we test and correct any systematic bias (e.g. photometric zero-point offsets) in observed fluxes. Second, we adjust observed flux uncertainties while adding a potentially missing systematic error component.

For flux and flux uncertainty corrections, we use the full GAMA spectroscopic redshift sample, where we eliminate any systematic outliers. The spectroscopic sample that we use to correct observed fluxes and flux uncertainties satisfies the following criteria:

- $|\zeta_w^{\text{best}} - \zeta_{\text{spec}}| < \text{delta_}\zeta$, where $\text{delta_}\zeta = 0.1$;
- the reduced χ^2 is smaller than 2.0 for the best-fitting template at spectroscopic redshift.

This limits the sample to galaxies, where clear photo-z outliers and bad SED fits are excluded. For individual fluxes in different passbands, we use the additional criteria:

- both, flux and flux uncertainty, must exist and observed flux should be flagged as good (e.g. not affected by neighbouring source or imaging artefacts);

- computed model fluxes are limited to the best-fitting template at spectroscopic redshift;
- clear outliers (bad observed flux estimates) are excluded while requiring $|F_\alpha - a_m F_{T\alpha}| / \sigma_{F_\alpha} < \text{delta_flux}^3$, where $\text{delta_flux} = 3.0$;
- observed flux should have sufficient signal-to-noise ratio, $F_\alpha / \sigma_{F_\alpha} > \text{SN_limit}$, where $\text{SN_limit} = 5.0$.

For each galaxy in the sample, the number of used passbands can be different, and for each passband, we can have a different number of galaxies used for flux and flux uncertainty corrections. We designate $\mathbf{S}_\alpha$ as the sample of galaxies for a given passband α that satisfies all the abovementioned criteria. In the GAMA sample, we removed about 10 per cent of objects in all passbands, except the u -band, where we removed about one third of the objects.

Observed fluxes are corrected assuming a linear bias correction f_{bias}^α and the corrected flux F'_α in passband α is given as

$$F'_\alpha = f_{\text{bias}}^\alpha F_\alpha. \quad (18)$$

Assuming that the flux bias correction originates from photometric zero-point calibration, we apply the same bias correction to observed flux uncertainties, where the corrected flux error σ'_{F_α} is given as

$$\sigma'_{F_\alpha} = f_{\text{bias}}^\alpha \sigma_{F_\alpha}. \quad (19)$$

The flux bias f_{bias}^α is estimated as

$$f_{\text{bias}}^\alpha = \text{med} \{ f_i^\alpha, i \in \mathbf{S}_\alpha \}, \quad (20)$$

where f_i^α is the ratio of best-fitting model flux over observed flux in a given passband α

$$f_i^\alpha = \frac{a_m^i F_{T\alpha}^i}{F_\alpha^i}. \quad (21)$$

In Eq. (20), the median is taken over all galaxies and fluxes that satisfy the abovementioned criteria.

Observed flux uncertainties are corrected using the following equation:

$$(\sigma_{F_\alpha}^{\text{cor}})^2 = (f_{\sigma, \text{scaling}}^\alpha)^2 \left[(\sigma'_{F_\alpha})^2 + (f_{\sigma, \text{rel}}^\alpha F'_\alpha)^2 \right] + \sigma_{\text{model}}^2(\alpha, G). \quad (22)$$

In the last equation, $f_{\sigma, \text{rel}}^\alpha$ defines the additional relative flux uncertainty added to the observed flux uncertainty in quadrature. The factor $f_{\sigma, \text{scaling}}^\alpha \geq 1.0$ scales the observed flux uncertainties so that the mean reduced χ^2 value is close to 1.0. Both factors, $f_{\sigma, \text{rel}}^\alpha$ and $f_{\sigma, \text{scaling}}^\alpha$, are computed using the full sample for photometric corrections and are the same for all galaxies. These factors depend only on the passband. The last factor $\sigma_{\text{model}}^2(\alpha, G)$ in Eq. (22) adds additional model flux uncertainty that is computed individually for each galaxy. Below, we describe how these factors are computed. For template fitting in Eq. (3) we use $F_\alpha^{\text{cor}} \equiv F'_\alpha$ and $\sigma_{F_\alpha}^{\text{cor}}$ instead of F_α and σ_{F_α} .

The additional relative error $f_{\sigma, \text{rel}}^\alpha$ is computed assuming that the scaled difference between observed and model flux is independent of the observed flux. In mathematical terms,

the parameter $f_{\sigma, \text{rel}}^\alpha$ is determined so that the linear regression between $\log(F'_\alpha)$ and $\log(\Delta F_\sigma)$ has zero slope, where

$$\Delta F_\sigma^2 = \frac{(F'_\alpha - a_m F_{T\alpha})^2}{(\sigma'_{F_\alpha})^2 + (f_{\sigma, \text{rel}}^\alpha F'_\alpha)^2}. \quad (23)$$

When observed fluxes are bias-free and flux uncertainties include all relevant statistical and systematic error components, then the required additional relative error $f_{\sigma, \text{rel}}^\alpha \approx 0$.

The uncertainty scaling factor $f_{\sigma, \text{scaling}}^\alpha$ is estimated as

$$(f_{\sigma, \text{scaling}}^\alpha)^2 = \max \left\{ 1.0, \frac{1}{|\mathbf{S}_\alpha|} \sum_{\mathbf{S}_\alpha} \Delta F_\sigma^2 \right\}, \quad (24)$$

where $|\mathbf{S}_\alpha|$ is the number of galaxies in a sample $\mathbf{S}_\alpha$. We only scale the observed fluxes upwards when the flux uncertainties underestimate the statistical difference between sigma-scaled observed and model fluxes.

To account for the variance in template fitting and potentially insufficient coverage of model templates, we defined an additional model template dependent systematic uncertainty $\sigma_{\text{model}}^2(\alpha, G)$ as

$$\sigma_{\text{model}}^2(\alpha, G) = \min \left\{ f_{\text{mod}}^{\text{lim}} \sigma_{F_\alpha}^{\text{mod}2}, \sigma_{F_{T\alpha}}^{\text{mod}2} \right\} + (f_{\text{mod}}^{\text{diff}} |F_{T\alpha}^{\text{mod}} - F_\alpha|)^2, \quad (25)$$

where $f_{\text{mod}}^{\text{lim}}$ and $f_{\text{mod}}^{\text{diff}}$ are user defined scaling factors. Model flux $F_{T\alpha}^{\text{mod}}$ and model flux uncertainty $\sigma_{F_{T\alpha}}^{\text{mod}2}$ are estimated as weighted average over all templates

$$F_{T\alpha}^{\text{mod}} = \frac{\sum_{T \in \mathbf{T}} \sum_{\zeta} p(F|\zeta, T) a_m F_{T\alpha}}{\sum_{T \in \mathbf{T}} \sum_{\zeta} p(F|\zeta, T)}, \quad (26)$$

$$\sigma_{F_{T\alpha}}^{\text{mod}2} = \frac{\sum_{T \in \mathbf{T}} \sum_{\zeta} p(F|\zeta, T) (a_m F_{T\alpha} - F_{T\alpha}^{\text{mod}})^2}{\sum_{T \in \mathbf{T}} \sum_{\zeta} p(F|\zeta, T)}, \quad (27)$$

where the summation is over all model templates $\mathbf{T}$ and redshifts. Weights p are proportional to the model likelihoods given with Eq. (6).

As explained above, the corrections are computed one passband at a time. In principle, a correction in one passband affects the corrections in another since the best-fitting template might change after corrections. To account for this, we compute the flux and flux-uncertainty corrections iteratively, including the corrections from previous iterations. The corrections converge very quickly, and we used five iterations for the GAMA sample. The same approach was used in Laur et al. (2022). This iterative process takes into account the flux scaling over the entire SED. To compute the flux and flux uncertainty corrections, we need a set of model templates $\mathbf{T}$, which are generated as described in Sect. 5. The first set of templates is generated using observed fluxes and uncertainties. When we computed the flux and flux uncertainty corrections, we generated new templates using corrected values. In practice, the corrections change very little, and there was no need to run the template generation process iteratively.

³ Notations are the same as in Eq. (3), where α indicates a given passband and $a_m F_{T\alpha}$ is the best-fitting model flux.

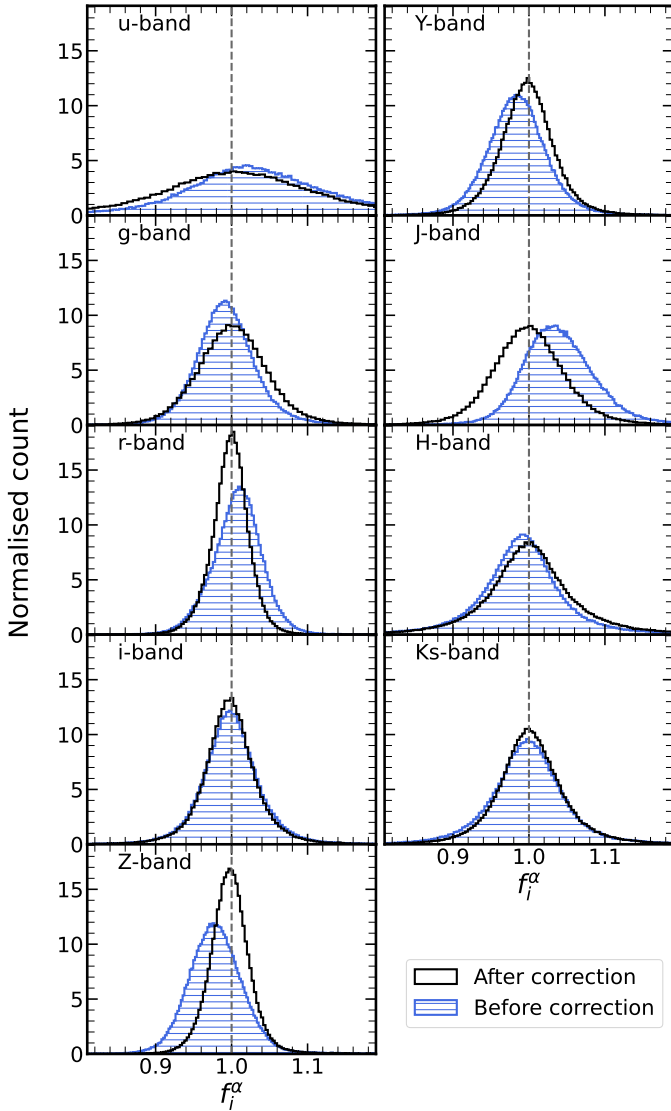


Fig. 6. Flux bias corrections for the GAMA spectroscopic sample. The distribution of flux ratios f_i^α (model flux divided by observed flux) for the best-fitting spectral template at spectroscopic redshift. Each panel shows the f_i^α distribution for different passbands. The area under the histograms is normalised to unity. The blue-shaded region shows the flux bias before any corrections. The black line shows the flux bias f_i^α distribution after corrections. The flux bias correction is defined with Eqs. (18)–(20). See Sect. 4 for a detailed description of the flux and flux uncertainty corrections. As expected, the flux bias correction shifts the distribution median closer to unity (dashed vertical line). Table 1 shows the flux bias correction values for each passband.

Figure 6 shows the distribution of the model flux over the observed flux (see Eq. (21)) for all nine GAMA filters. It is visible that there is a bias between observed and model fluxes. After adding the flux corrections (see Table 1), the distributions are centred around unity as expected for bias-free flux estimates. We emphasise that the flux corrections do not indicate that the observed fluxes themselves are biased. The bias is measured between the observed and model fluxes computed from the optimised set of templates (see Sect. 5). The bias might also indicate that the used spectral templates are systematically biased or do not cover the entire variety of galaxy templates. The flux corrections in Table 1 are derived for the photo-z estimation in the TOPz code. For most filters, the flux bias correction is less than

Table 1. Flux and flux uncertainty corrections applied to the GAMA data for nine passbands.

Passband	f_{bias}^α	$f_{\sigma,\text{rel}}^\alpha$	$f_{\sigma,\text{scaling}}^\alpha$
u	1.0585	0.0500	1.3221
g	0.9935	0.0000	1.3246
r	1.0098	0.0000	1.0000
i	0.9890	0.0000	1.0599
Z	0.9566	0.0000	1.1819
Y	0.9702	0.0000	1.2295
J	1.0493	0.0000	1.2112
H	1.0083	0.0000	1.4387
Ks	1.0198	0.0000	1.6254

Notes. f_{bias}^α is linear flux bias correction, $f_{\sigma,\text{rel}}^\alpha$ is relative uncertainty added to the observed flux uncertainty and $f_{\sigma,\text{scaling}}^\alpha$ defines the scaling parameter for flux uncertainty. See Sect. 4 for more details.

a few per cent, which is less than the intrinsic scatter of the difference between the model and observed fluxes.

Figure 7 shows the distribution between the model and observed fluxes normalised by the flux uncertainty. The distribution is shown before and after flux uncertainty scaling (see Eq. (24)). The flux uncertainty scaling parameters are given in Table 1. The theoretical expectation is that the normalised differences follow a Gaussian distribution with unit variance. Comparing the distributions before and after flux uncertainty scaling, we see that the scaling moves the distributions closer to the theoretical expectation. Altogether, the flux and flux uncertainty corrections presented in this section are derived following theoretical expectations. For the GAMA sample, the corrections given in Table 1 improve the photo-z estimates as we will demonstrate in Sect. 8.3.

5. Template set generation for photometric redshift estimation

To estimate the photo-z of galaxies as described in Sect. 2, we take a similar approach as Laur et al. (2022). To model the redshifts of all galaxies, we need their description in the form of a spectral template set $\mathbf{T}$ that covers adequately all possible galaxy types. To generate the spectral templates of galaxies, we use the CIGALE⁴ (Code Investigating GALaxy Emission) software version 2022.1 (Boquien et al. 2019; Yang et al. 2022), which is a versatile tool to generate a large variety of galaxy spectra. To construct galaxy spectra, we use CIGALE by including star formation history with *sfh2exp* module (modelled as a combination of two exponential profiles), initial mass function, stellar evolution tracks, dust attenuation with *dustatt_modified_CF00* module, nebular emission lines, and redshifting. Csörnyei et al. (2021) show that emission lines with varying strength in the template spectra are improving the photo-z estimation in template-fitting methods (see also Ilbert et al. 2006; Dahlen et al. 2013; Hsu et al. 2014). Hence, we allowed CIGALE to fit a variety of emission lines in the spectra. We did not use dust emission and AGN contribution since we use dominantly optical wavelengths (the smallest dust particles can provide thermal contribution up to $\lambda \sim 3 \mu\text{m}$ (Ryden & Pogge 2021) in rare occasions, but mostly confined to $\lambda \sim 120 \mu\text{m}$).

⁴ <https://cigale.lam.fr>

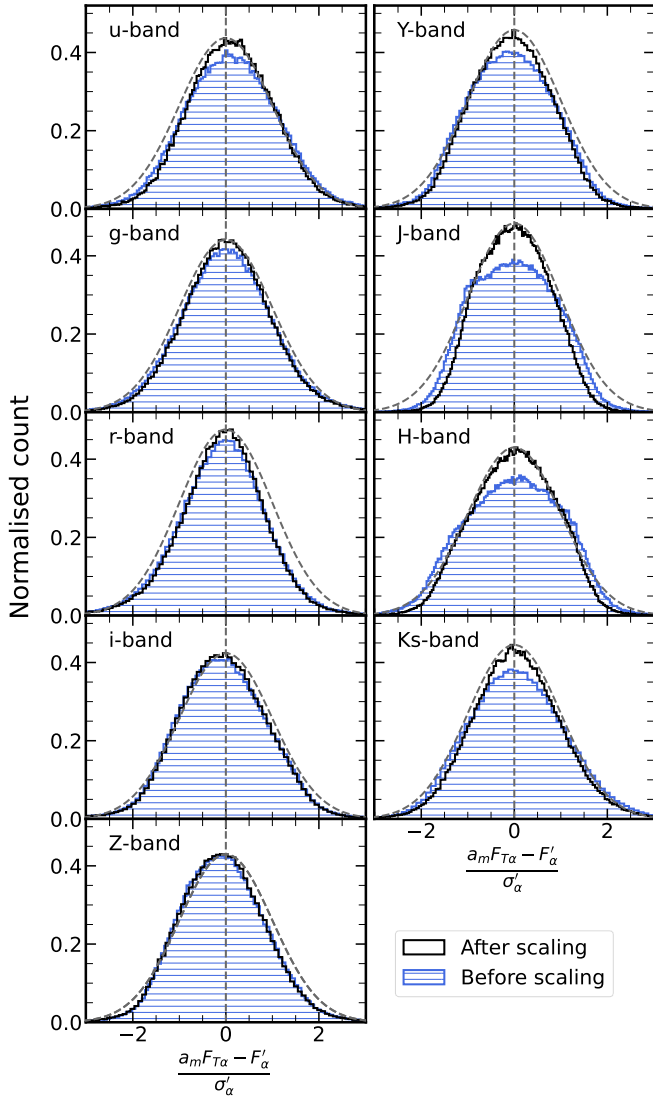


Fig. 7. Difference between the model and observed flux divided by the observed flux uncertainty for the best-fitting spectral template at spectroscopic redshift. Each panel shows the distribution for different passbands. The blue-shaded histogram shows the distribution before flux uncertainty scaling. The solid black histogram shows the distribution after flux uncertainty scaling. The dashed grey line shows the Gaussian distribution with unit variance. See Sect. 4 for a detailed description of the flux and flux uncertainty corrections. As expected, the flux uncertainty scaling is shifting the normalised distribution closer to the theoretical Gaussian distribution. Table 1 gives the flux uncertainty scaling parameters for different passbands.

To generate templates with CIGALE, we used the following parameters:

- Chabrier initial mass function, stellar evolution tracks:
 - Metallicity $Z = 0.0001, 0.0004, 0.004, 0.008, 0.02, 0.05$
 - separation_age = 10 Myr
- Double exponential star formation history:
 - tau_main = 0.5, 1, 2, 3, 4, 5, 6.5, 8 Gyr
 - tau_burst = -200, -100, 50.0, 100, 200, 600 Myr
 - f_burst = 0.002, 0.01, 0.03, 0.05, 0.1, 0.25, 0.5
 - age = 8000, 5000, 3000, 2000 Myr
 - burst_age = 10.0, 30.0, 100.0, 500.0, 1000.0 Myr
- Dust attenuation:
 - A_V _ISM = 0.0, 0.1, 0.25, 0.4, 0.7, 1.0, 2.0

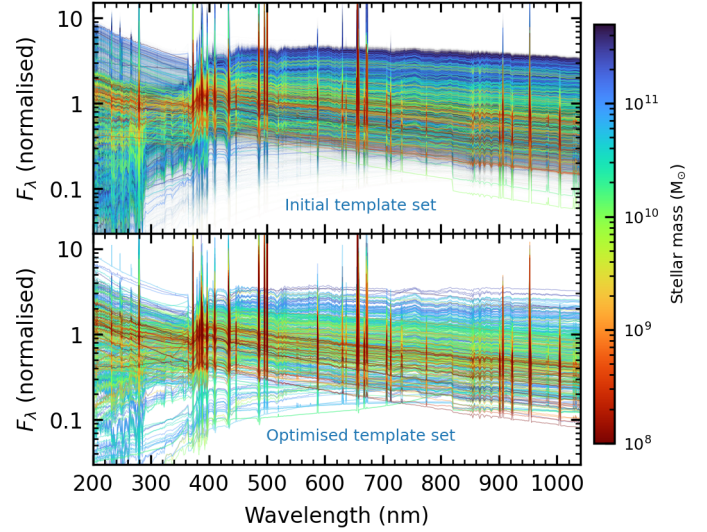


Fig. 8. Spectral templates generated using the CIGALE software (see Sect. 5). The upper panel shows all $\sim 40\,000$ spectral templates, while the lower panel shows the selected 555 templates after template set optimisation. Template fluxes are normalised to unit stellar mass. The colour coding shows the stellar masses from the CIGALE software.

- Nebular emission lines:
 - $\log U = -4.0, -3.5, -3.0, -2.0, -1.5, -1.0$
 - $z_{\text{gas}} = 0.0004, 0.001, 0.004, 0.011, 0.022, 0.041$
 - $n_e = 100$
 - $f_{\text{esc}} = 0.0$
 - $f_{\text{dust}} = 0.0$
 - lines_width = 300.0 km s^{-1} .

We refer to the CIGALE papers and manual for a detailed description of these parameters. The output spectral templates from the CIGALE software cover the wide wavelength range from UV to IR (10^1 – 10^9 nm), which is sufficiently wide to cover all GAMA nine-band photometry over a large redshift range. In the current analysis, we use CIGALE to generate templates for relatively low-redshift ($z < 1$) galaxies. For high-redshift galaxies, a different set of physically motivated templates is necessary (Luberto et al. 2024).

Galaxy template generation for TOPz is a two-step process. In the first step, we use a spectroscopic redshift sample and CIGALE to generate a large set of templates. We fix the redshift to the spectroscopic redshift and generate a best-fitting template for each input galaxy. To keep the number of generated templates reasonable, we downsampled the entire spectroscopic GAMA catalogue to roughly 40 000 galaxies. To avoid removing galaxies that are underrepresented in the spectroscopic catalogue, we downsampled to have uniform coverage in ζ space. Fig. 5 shows the distribution of sources in the full and downsampled GAMA spectroscopic catalogue. The downsampled catalogue covers fairly uniformly the redshift and magnitude ranges in the spectroscopic sample, except for the regions where we are limited by observational data. The upper panel in Fig. 8 shows the spectral templates generated using the CIGALE software. The generated spectral templates cover a wide variety of galaxy types. In TOPz code, we assume that a single template can represent a galaxy. In general, galaxy photometry can be blended, which can be included in template-fitting methods (Jones & Heavens 2019) but is computationally too demanding for large surveys.

For photo- z estimation, we need a large enough template set to cover a wide range of galaxy templates. Still, at the same time,

it should be sufficiently small to be computationally feasible. Frontera-Pons et al. (2023) highlight that template set optimisation improves the accuracy of the photo-z codes. To reduce the template set, we use the template set optimisation that selects only templates that maximise the number of correct photo-z. The template set optimisation is performed using the spectroscopic galaxy sample. The metric that we are optimising is the following:

$$S(\mathbf{T}^*) = \frac{\sum_{G \in \mathbf{G}_{\text{spec}}} S_G^{\mathbf{T}^*} w_G}{\sum_{G \in \mathbf{G}_{\text{spec}}} w_G}, \quad (28)$$

where w_G is a spectroscopic galaxy weight for optimisation and S_G is an optimisation metric per galaxy G , which is defined as

$$S_G^{\mathbf{T}^*} = p_{\text{peak}}^{\text{best}}(G, \mathbf{T}^*) \mathbb{1}_{\{\zeta_{\text{min}}^{\text{best}} < \zeta_{\text{spec}} < \zeta_{\text{max}}^{\text{best}}\}} + w_{\text{spec}} p_{\mathbf{T}^*}^G(\zeta_{\text{spec}}), \quad (29)$$

where $\mathbb{1}\{\cdot\}$ is an indicator function that is one if the spectroscopic redshift ζ_{spec} value is inside the best peak in the photo-z posterior and w_{spec} is a user-defined weight parameter. The optimisation metric S_G maximises two aspects: the probability that the spectroscopic redshift is located inside the best peak in the photo-z posterior, and the photo-z posterior value at ζ_{spec} . The parameter w_{spec} sets the weight between these two criteria.

There is no unique way to define the weights w_G for galaxies. The simplest option is to set $w_G = 1.0$ for each galaxy. In TOPz code, we have also implemented an option to set the weights so that the weighted redshift distribution is uniform in ζ space. This will suppress the inhomogeneities in the spectroscopic redshift sample so that the optimised template set $\mathbf{T}_{\text{opt}}$ is not only optimised for overrepresented redshift ranges.

In Eq. (29), the $\mathbf{T}^*$ is an arbitrary set of templates drawn from the full template set $\mathbf{T}$ that was generated using the CIGALE software. The optimised set of templates $\mathbf{T}_{\text{opt}}$ is defined as

$$\mathbf{T}_{\text{opt}} = \arg \max_{\mathbf{T} \subseteq \mathbf{T}} \{S(\mathbf{T}^*)\}. \quad (30)$$

The optimisation is carried out using Metropolis-Hastings algorithms with birth and death moves. The Metropolis-Hastings algorithm will simultaneously optimise the best number of templates and the actual set of templates.

For maximal photo-z performance, the template set optimisation can be carried out iteratively. The first round of optimisation defines a set of templates that finds a good photo-z for the majority of the galaxies. Using the optimised template set $\mathbf{T}_{\text{opt}}$, we can identify the set of galaxies where photo-z deviates significantly from the spectroscopic redshift, i.e. the catastrophic outlier sample. This set of catastrophic outlier galaxies can be used to generate additional templates with the CIGALE software. Combining the original optimised set $\mathbf{T}_{\text{opt}}$ with the newly generated templates for catastrophic outliers, we have a new full set of templates $\mathbf{T}$ that can be used to find the new optimised set $\mathbf{T}_{\text{opt}}$. There is no guarantee that this process will converge. In practical applications, every iteration of the template set optimisation reduces the number of catastrophic outliers and increases the general performance of the photo-z estimation. In the current paper, for the GAMA sample, we used three iterations. In every iteration, we selected approximately 10 000 galaxies with bad photo-z estimates and generated new templates with CIGALE that were added to the template set $\mathbf{T}$.

The lower panel of Fig. 8 shows the final set of 555 spectral templates used for the photo-z estimation. Compared with

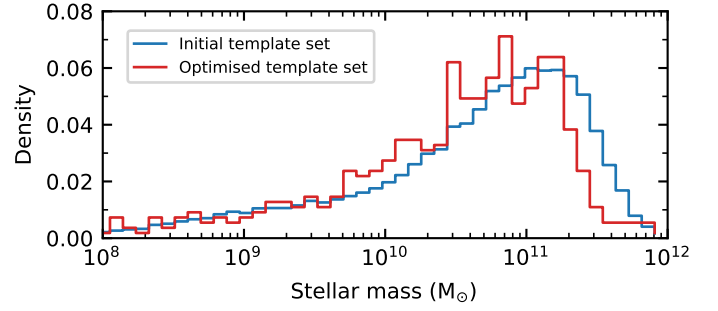


Fig. 9. Normalised histogram of stellar masses for the initial set of templates (blue) and an optimised set of templates (red). Templates are generated using the CIGALE software (see Sect. 5), and stellar masses are total stellar masses provided by the CIGALE.

the initial template set (upper panel), the optimised set covers all possible galaxy types uniformly. Fig. 9 shows the distribution of stellar masses for the initial and optimised set of templates. The distributions cover the entire stellar mass range, confirming that the optimised template set covers all possible galaxy types. There are slightly fewer templates in the optimised set for the most massive templates (see Figs. 8 and 9). This is expected since most massive galaxies are ellipticals whose spectra are more similar to each other than the spectra of spiral galaxies are to one another. This similarity arises from their homogeneous stellar populations, lack of star formation, and minimal dust and gas content. Hence, fewer such templates are sufficient for photo-z estimation in the TOPz code.

6. Star-galaxy identification with template fitting

Different types of astronomical objects have different spectral energy distributions (SED). The SED of galaxies is generally different from the SED of stellar objects. Comparing the model spectral templates with the SED of observed objects, it is possible to find the template that best matches the observed fluxes of an object. Hence, including the stellar templates in the set of model templates, it is possible to perform a simple star-galaxy identification using template fitting.

The stellar templates included in the TOPz code are taken from the TRDS Pickles Atlas⁵ (Pickles 1998). It provides comprehensive spectral coverage of 131 flux-calibrated stellar spectra, which encompass all typical spectral types and luminosity classes at solar abundance, as well as metal-weak and metal-rich F-K dwarf and G-K giant stars. The variety of stellar templates in the Pickles library is shown in Fig. 10.

We treat the Pickles stellar templates differently from the galaxy templates for the template fitting. Stellar templates are never redshifted and are all located at redshift zero. The total number of stellar templates is significantly smaller than that of redshifted galaxy templates. Hence, including stellar templates has only a small impact on the total computation time.

To perform the star-galaxy identification, we estimate the probability that a given object is a star given the observed fluxes F . The probability p_{star} is defined as

$$p_{\text{star}} = \frac{p_{\text{star}}^{\text{max}}}{p_{\text{star}}^{\text{max}} + p_{\text{gal}}^{\text{max}}}, \quad (31)$$

⁵ <https://www.stsci.edu/hst/instrumentation/reference-data-for-calibration-and-tools/astronomical-catalogs/pickles-atlas>

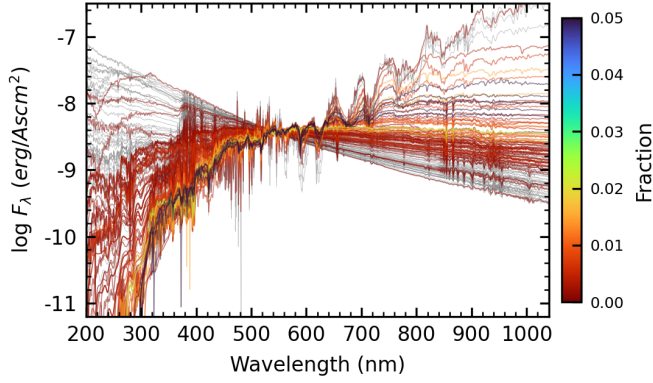


Fig. 10. All 131 stellar templates taken from the TRDS Pickles Atlas. The colour coding shows the fraction of objects that have been assigned a $p_{\text{star}} > 0.5$ for that specific stellar template, despite all of the objects being designated as galaxies in the GAMA catalogue. Dark blue lines show the stellar templates that are most often assigned to galaxies, while templates shown in grey are stellar templates that are not confused with galaxies in the GAMA data set.

where

$$p_{\text{star}}^{\text{max}} = \max_{T \in \mathbf{T}_{\text{star}}} p(F|T), \quad (32)$$

$$p_{\text{gal}}^{\text{max}} = \max_{\zeta, T \in \mathbf{T}_{\text{gal}}} p(F|\zeta, T). \quad (33)$$

The probability p_{star} does not depend directly on the number of used templates. However, the p_{star} depends on the coverage of different stellar types in the input library. Below, we will also use an extended stellar library to analyse this aspect specifically. In computing the p_{star} , we compare the best-fitting galaxy template against the best-fitting stellar template. Additionally, we ignore the prior while computing the p_{star} . Consequently, the p_{star} only depends on the observed object SED and gives the probabilistic estimate that the object has a stellar SED. In the last equations, $\mathbf{T}_{\text{star}}$ and $\mathbf{T}_{\text{gal}}$ refer to stellar and galaxy template libraries, respectively.

In Fig. 10 we show in colour the Pickles stellar templates that are most often confused with galaxies. The colour coding shows the fraction of how many times each stellar template best matches an object compared with all stellar templates. We see that redder stellar templates are most often confused with galaxies. In Fig. 11 we show the p_{star} probability as a function of object magnitude for stars and galaxies as classified by the GAMA survey (see Sect. 3). For stars in the GAMA database, many objects are classified as galaxies in the TOPz. However, there are only a few galaxies where the p_{star} probability is larger than 0.5. Hence, very few galaxies (less than 0.2 per cent) are misclassified by TOPz. When visually investigating some of these misclassified galaxies, we find that those with p_{star} probabilities close to 1 are very small, very red objects that are most often assigned as M-type stars by TOPz. These galaxies are roughly split equally between being visually isolated and surrounded by other objects in the sky. Some of those objects are genuine Milky Way stars projected next to a background galaxy. In Fig. 12, we provide examples of GAMA galaxies classified as stars in the TOPz code.

To test the star-galaxy classification dependence on the choice of the used stellar library, we extended the original Pickles library with additional stellar templates. We added Kurucz ODFNEW/NOVER theoretical stellar spectra (Castelli et al. 1997; Castelli & Kurucz 2003) and added white dwarf models from Koester (2010). In total, we have 9287 spectra in our

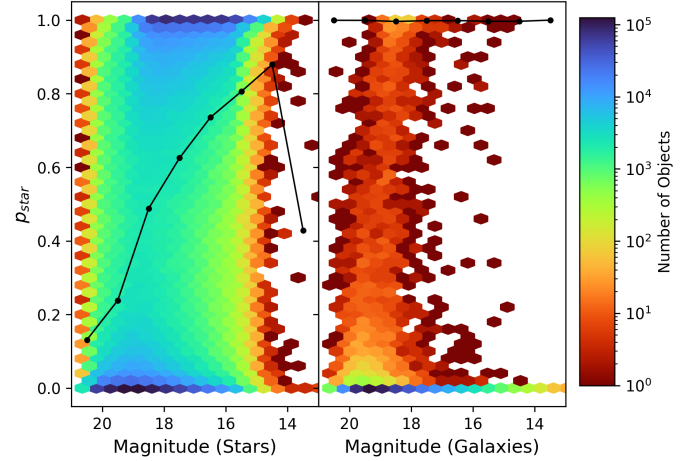


Fig. 11. Objects from the GAMA catalogue plotted on the magnitude- p_{star} plane and separated depending on whether they are designated as stars (left panel) or galaxies (right panel) in the GAMA catalogue. There are 620 131 objects in the left panel and 481 130 objects in the right panel. The colour coding shows the number of objects in each section of the panels. The black lines in each panel show the fraction of objects that are correctly identified by TOPz ($p_{\text{star}} \geq 0.5$ for stars, and $p_{\text{star}} < 0.5$ for galaxies) per magnitude bin. For galaxies, the fraction of correctly classified objects over all magnitude bins is 99.8 per cent.

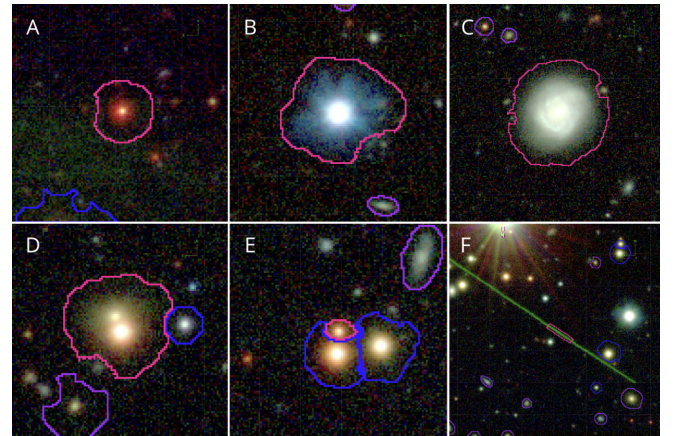


Fig. 12. Example of objects classified as galaxies in the GAMA database but classified as stars by TOPz, with $p_{\text{star}} \approx 1$. Pink segments mark the misclassified objects. Purple segments indicate galaxies and blue segments indicate stars (according to the GAMA classification). In certain cases, the GAMA catalogue assigns galaxy classifications to sources that are evidently stars or image artefacts: (A) a compact red galaxy classified as an M-type star in TOPz; (B) a bright blue star; (C) a luminous spiral galaxy; (D) two blended sources treated as a single object; (E) a small, faint object near several larger, brighter sources; (F) a satellite trail.

extended stellar library. Since the GAMA field is located in the Milky Way halo, we will ignore the interstellar reddening, which is mostly a concern in the Milky Way disc region. Figure 13 shows the cumulative classification probability for objects classified as stars or galaxies in the GAMA database. For galaxies, the extended stellar library makes a small difference, while for stars, the extended stellar library significantly reduces the fraction of wrongly classified objects.

We emphasise that the test was carried out using relatively bright sources in the GAMA photometric catalogue. The confusion between stellar and galaxy templates is significantly increased for fainter targets with larger uncertainties for observed

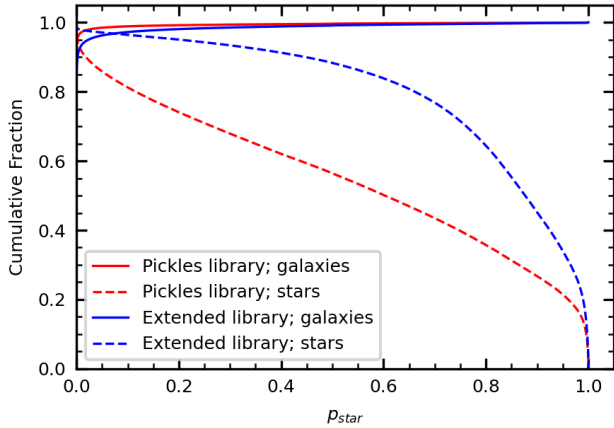


Fig. 13. Cumulative fractions of object classifications using the Pickles stellar library (red curves) and the extended stellar library (blue curves). The solid lines show the forward cumulative distributions of galaxies and the dashed lines show the reversed cumulative distributions of stars. Classification of stars and galaxies is taken from the GAMA database.

fluxes. Hence, for fainter sources, the simple SED classification is only a rough indication of the type and cannot be used as a reliable star-galaxy classification. For accurate star-galaxy classification, the object morphology (e.g. an indication of a point-like source) should be included as well (e.g. Cook et al. 2024).

7. Luminosity function prior for photometric redshifts

Photo- z prior $p(\zeta, T | m_0)$ gives a likelihood that a galaxy with a total magnitude m_0 is located at redshift ζ and has a spectral type T . The prior $p(\zeta, T | m_0)$ can be further expressed as

$$p(\zeta, T | m_0) = p(T | m_0) p(\zeta | m_0, T), \quad (34)$$

where $p(\zeta | m_0, T)$ is the redshift distribution of galaxies with total magnitude m_0 and spectral template T , and $p(T | m_0)$ defines the prior for different spectral templates that in general depends on magnitude m_0 . The $p(T | m_0)$ depends on the template set $\mathbf{T}$ and should be independently estimated for a given template set. For simplicity, we use a non-informative prior in a given paper and set $p(T | m_0) = \text{const}$. This is justified since our optimised set of templates is selected based on the observed data (see Sect. 5).

For template-fitting photo- z codes, Tanaka (2015) emphasises the importance of physically motivated priors. Hence, in the current paper, the redshift distribution prior $p(\zeta | m_0, T)$ is defined using a luminosity function of galaxies in a fixed passband. The prior is defined as

$$p(\zeta | m_0, T) \propto \Phi(L | \zeta) V(\zeta), \quad (35)$$

where $\Phi(L | \zeta)$ is the galaxy luminosity function at redshift ζ , where L is the absolute luminosity of a galaxy. We assume that the number count of galaxies (luminosity function normalisation) does not depend on the redshift. The term $V(\zeta)$ defines the redshift-dependent volume term in the prior.

We employ the Schechter luminosity function that allows us to write

$$\Phi(L | \zeta) = \Phi(L | \zeta, L^*, \beta) \propto \left(\frac{L}{L^*} \right)^\beta \exp\left(-\frac{L}{L^*} \right), \quad (36)$$

where the luminosity function depends on characteristic luminosity L^* and power law index β . To account for the evolution correction of galaxies, the characteristic luminosity L^* is

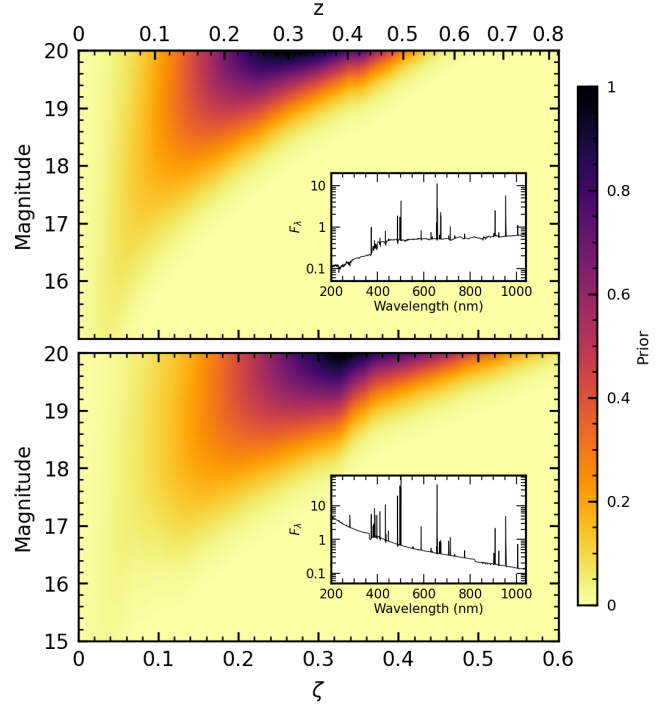


Fig. 14. Redshift prior dependence on redshift ζ and apparent galaxy magnitude for two different templates (upper and lower panel). The inset panels show the template spectra. The two templates were selected arbitrarily from the optimised template set (see Fig. 8) and serve as examples. The priors are arbitrarily normalised. The decrease of prior towards redshift zero is because of the volume prior, while the decrease at larger redshifts is due to the luminosity function prior. The differences in spectral templates affect the higher redshift end due to different k -corrections. The slight deviation from a smooth prior is caused by the features (e.g. emission lines) in the spectral templates.

defined as

$$L^* \equiv L^*(\zeta) = L_0^* \cdot 10^{-0.4(zQ)}, \quad (37)$$

where $Q < 0$ means that galaxies were brighter in the past (see e.g. Blanton et al. 2003; Tempel et al. 2014). While specifying the luminosity function parameters L_0^* , β , and Q , we can analytically calculate the redshift-dependent prior for a galaxy with absolute luminosity L . The luminosity function is defined for a fixed passband α_{LF} and the corresponding absolute luminosity L should be measured in the same passband at the rest frame. To convert the observed total magnitude m_0 measured in passband α_{tot} into absolute luminosity L measured in passband α_{LF} we do the following

$$L_{\alpha_{\text{LF}}}(\zeta, T) = f_{v,0} 10^{-0.4m_0} \frac{D_L(\zeta)}{10 \text{ pc}} \frac{F_{T\alpha_{\text{LF}}}(\zeta_{\text{LF}})}{F_{T\alpha_{\text{tot}}}(\zeta)}, \quad (38)$$

where $f_{v,0}$ is zero point of the magnitude system. If m_0 is given in AB magnitude system then $f_{v,0} = 3630.78 \text{ Jy}$. The $D_L(\zeta)$ is a luminosity distance at redshift ζ computed assuming a specified cosmology. The last term accounts for the k -correction and the difference between the passbands of total magnitude and luminosity function. It is estimated for each template T individually using model fluxes computed with Eq. (7).

Figure 14 shows the used prior for two different templates. The prior dependence on observed magnitude is defined by the luminosity function prior. The prior dependence on redshift is a

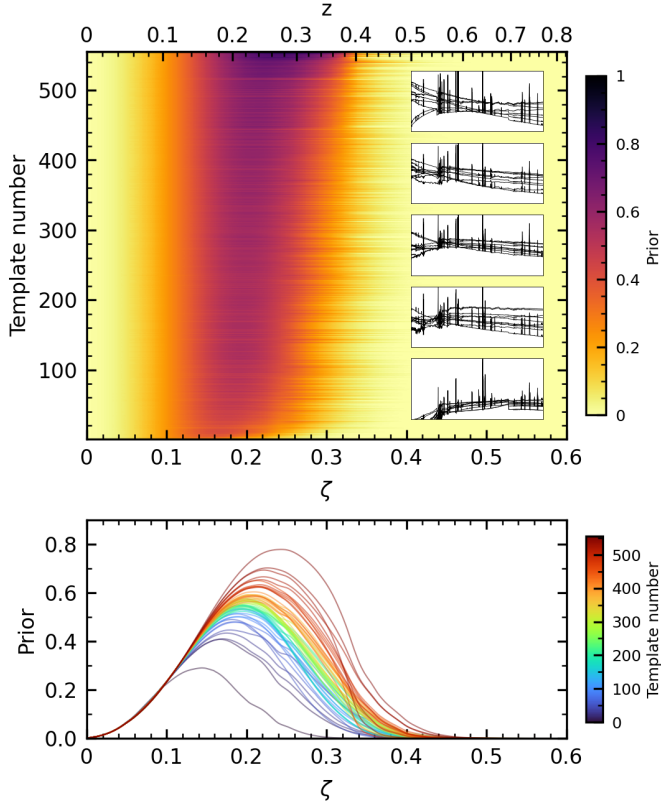


Fig. 15. Upper panel: redshift prior for a fixed apparent magnitude ($m_z = 19$ mag) for all templates in the optimised template set (see lower panel in Fig. 8). The templates are ordered based on the prior weighted mean redshift. The inset panels show how the template spectra change from top to bottom. Lower panel: redshift prior for different templates drawn from the upper panel.

combination of a volume prior and a luminosity function prior. The decrease of prior towards redshift zero comes from the volume prior, while the decrease at larger redshift comes from the luminosity function prior. The general shape of the chosen prior is similar to the previously proposed general analytical prior (Benítez 2000).

Figure 15 shows the redshift prior for the complete set of optimised templates at a fixed apparent magnitude. At lower redshifts, the prior is roughly the same for all the templates. At higher redshifts, the prior for different templates varies slightly due to the distinct k-corrections associated with each template. In general, the redshift range, where the prior is significant, is roughly the same for all templates. Hence, the shape of the luminosity function is more important than the spectral differences between the templates.

8. TOPz performance using GAMA sample

To test the performance and accuracy of the TOPz code, we have applied the TOPz algorithm described in the previous sections to the GAMA spectroscopic sample (see Sect. 3). Table 2 lists all the parameters defined in previous sections and used for the GAMA data. Applied photometric flux corrections are given in Table 1 and described in Sect. 4. Sect. 7 presents the prior used in the current analysis. In this section, we will present how well the TOPz algorithm is able to estimate the photo-z and redshift uncertainties of galaxies compared to the spectroscopic redshift sample.

Table 2. TOPz parameters for the GAMA data set.

Parameter	Value	Section
f_{peak}	0.005	Sect. 2.4
w_{spek}	10.0	Sect. 5
$\text{delta_}\zeta$	0.1	Sect. 4
delta_flux	3.0	Sect. 4
SN_limit	5.0	Sect. 4
$f_{\text{lim}}^{\text{lim}}$	0.5	Sect. 4
$f_{\text{mod}}^{\text{diff}}$	0.01	Sect. 4
L_0^*	-21.4	Sect. 7
β	-0.7	Sect. 7
Q	-0.6	Sect. 7

Table 3. TOPz photo-z parameters in three magnitude bins from Fig. 16.

	$m_z < 18$	$18 < m_z < 19$	$m_z > 19$
BIAS			
median(dz)	-0.001094	-0.001811	-0.002050
median($d\zeta$)	-0.001095	-0.001813	-0.002052
OUTLIERS			
$ dz > 0.1$	1.04%	2.44%	4.99%
$ dz > 0.15$	0.48%	1.00%	2.05%
$ d\zeta > 0.1$	1.04%	2.38%	4.97%
$ d\zeta > 0.15$	0.45%	0.97%	2.05%
$ d\zeta /\sigma_{\text{phot}} > 3$	1.84%	2.37%	3.35%
PRECISION			
$\sigma(dz)$	0.0300	0.0420	0.0530
$\sigma_{\text{MAD}}(dz)$	0.0195	0.0266	0.0329
$\sigma_{\text{NMAD}}(dz)$	0.0121	0.0168	0.0211
$\sigma(d\zeta)$	0.0300	0.0420	0.0530
$\sigma_{\text{MAD}}(d\zeta)$	0.0195	0.0266	0.0329
$\sigma_{\text{NMAD}}(d\zeta)$	0.0121	0.0168	0.0211

Notes. $dz = (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$ and $d\zeta = \zeta_{\text{phot}} - \zeta_{\text{spec}}$.

8.1. Photometric redshifts for the GAMA sample

Figure 16 shows the comparison between the spectroscopic redshifts and photo-z for three different magnitude bins. The galaxy redshifts are recovered visually well without a clear bias in photo-z estimates. All magnitude bins behave similarly, while the scatter is larger for fainter galaxies. This is expected since photo-z accuracy depends on the accuracy of the input photometry. Table 3 quantifies numerically the TOPz photo-z estimates. Fig. 17 shows the photo-z accuracy as a function of galaxy total magnitude. The median of the estimated photo-z is close to zero, indicating that the TOPz estimated photo-z is nearly bias-free. The scatter around the median strongly depends on the galaxy's total magnitude due to the increased uncertainty of photometric flux measurements. The lower panel in Fig. 17 shows the normalised difference between spectroscopic redshift and photo-z estimates; the redshift difference is normalised with the TOPz photo-z uncertainty. The normalised difference is independent of apparent magnitude, indicating that the photo-z uncertainties adequately take into account the observed flux uncertainties. In Fig. 18 we show how the photo-z uncertainty depends on

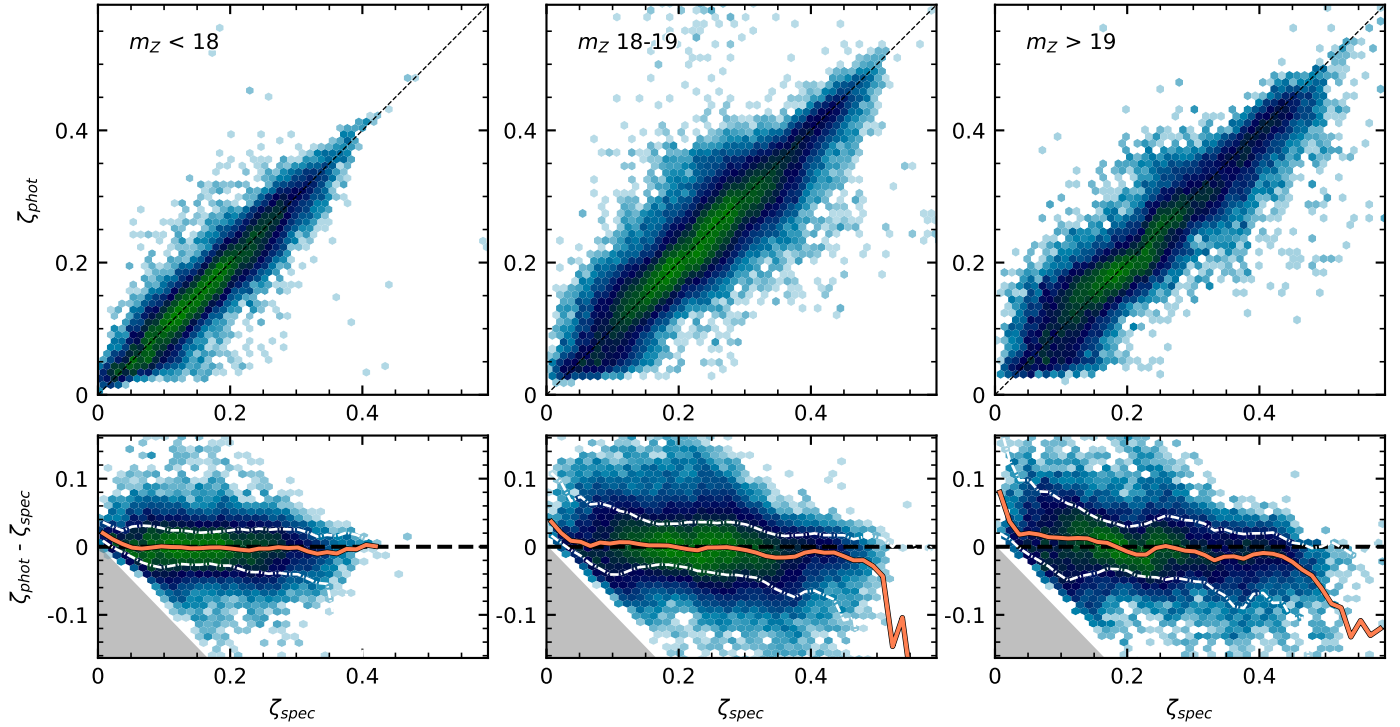


Fig. 16. Correspondence between spectroscopic redshifts and TOPz photo-z estimations. Upper panel: the colour of the hexagons denotes the number of galaxies on a logarithmic scale. The one-to-one relation is drawn with a diagonal dashed black line. See Fig. 19 for the distribution around the one-to-one line. From left to right, the three panels include galaxies in different Z-band magnitude ranges: $m_Z < 18$, $18 < m_Z < 19$ and $m_Z > 19$, respectively. Lower panel: photo-z accuracy as a function of spectroscopic redshift. The three panels are comprised of galaxies at the same magnitude ranges as on the top panels, and hexagon colours are similar. The coral line shows the median accuracy in different spectroscopic redshift bins, and the lower and upper white dash-dot lines indicate the 10 and 90% data quantiles, respectively. The grey area at the bottom-left of each panel indicates an exclusion zone where ζ_{phot} would be negative, causing the increase in the median (coral line) when approaching redshift zero.

magnitude and spectroscopic redshift. As expected, the uncertainty dominantly depends on the magnitude (flux uncertainty). In Sect. 8.4, we specifically analyse the accuracy of the photo-z probability distribution functions (posteriors).

Figure 19 shows the distribution of differences between the photometric and spectroscopic redshifts. It is visible that the peak of the distribution is around zero for all magnitude bins. However, the width of the distribution depends on the magnitude and, as expected, is narrower for brighter objects. The lower panel in Fig. 19 shows the distribution differences between spectroscopic and photo-z normalised by the photo-z uncertainty. The dashed grey line shows the normalised Gaussian distribution with a standard deviation of one. In general, the width of photo-z differences is close to the normal distribution, indicating that the photo-z uncertainties are reasonable and that they are not over- or under-estimated. This is valid for all magnitude bins, showing that the photo-z uncertainties are statistically correct for all galaxies, independent of their observed magnitude. Additionally, we see that the distribution of observed differences peaks slightly more than the expected normal distribution. This might indicate that the photo-z errors are slightly overestimated or that photo-z uncertainties do not follow exactly the theoretical normal distribution. The latter is possible since the template fitting is based on physical galaxy templates, which might improve the photo-z accuracy. More detailed analysis of this should be based on the synthetic galaxies, which require extensive analysis and are beyond the scope of the current paper.

The TOPz code provides the photo-z estimation for up to three best peaks in the posterior distribution. In the GAMA

photometric sample, 99 per cent of galaxies have only one dominant peak where $p_{\text{peak}}^{\text{best}} > 0.5$. The second and third peaks are usually very weak in the posterior distribution. About 5 per cent of objects have $p_{\text{peak}}^{\text{alt1}} > 0.05$ and only 0.5 per cent of objects have $p_{\text{peak}}^{\text{alt2}} > 0.05$. In some cases, the best peak dominates in the photo-z posterior, but the weaker peaks might be closer to the true spectroscopic redshift. In the GAMA spectroscopic sample, for 2 per cent of objects, the second best peak is closer to the true redshift than the best peak. The third peak is closest to the true redshift only for 0.3 per cent of objects. Hence, for general analysis, considering only the dominant peak is sufficient.

In Fig. 20, we divided the GAMA sample into two roughly equal subsamples based on the difference between the $|\zeta_{\text{peak}} - \zeta_{\text{w}}|$. The figure illustrates that the photo-z estimates are statistically more accurate when the difference is small. However, the difference alone is not a good measure to quantify the accuracy of individual photo-z estimates.

8.2. Photo-z outliers, bias, and σ_{NMAD}

The photo-z performance compared to the spectroscopic sample is usually estimated using three quantities: the median bias, the outlier fraction, and the scatter around the spectroscopic redshift. Photo-z accuracy is generally estimated using the normalised median absolute deviation between the photometric and spectroscopic redshift samples. Mathematically, it is expressed as

$$\sigma_{\text{NMAD}} = 1.48 \times \text{median}(|\Delta\zeta - \text{median}(\Delta\zeta)|) \quad (39)$$

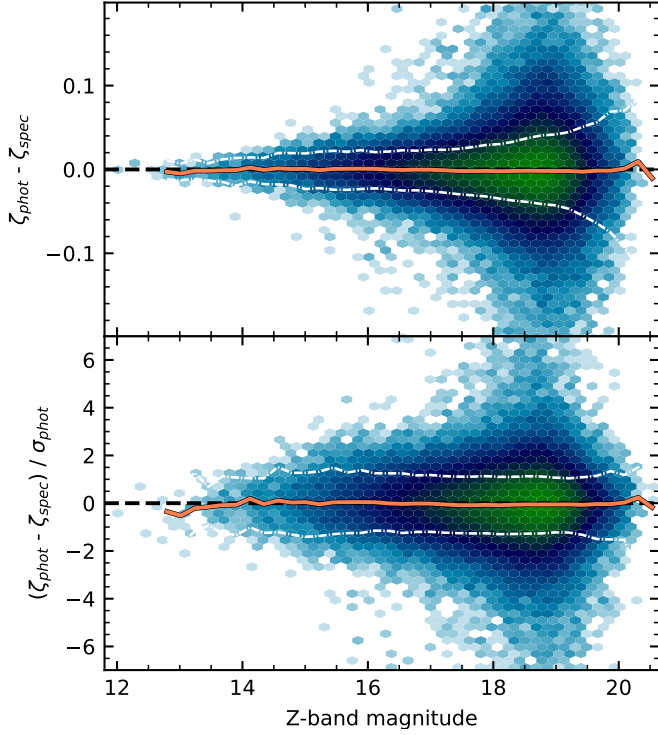


Fig. 17. Photo-z accuracy dependence on total observed magnitude m_0 . The hexagon colours indicate the number of galaxies on a logarithmic scale. The coral line shows the median difference, and the white dash-dot lines indicate the 10 and 90% quantiles. Upper panel: Photo-z accuracy for single galaxies. Lower panel: Photo-z accuracy normalised by the TOPz photo-z uncertainty σ_{phot} value (see Eq. (16)).

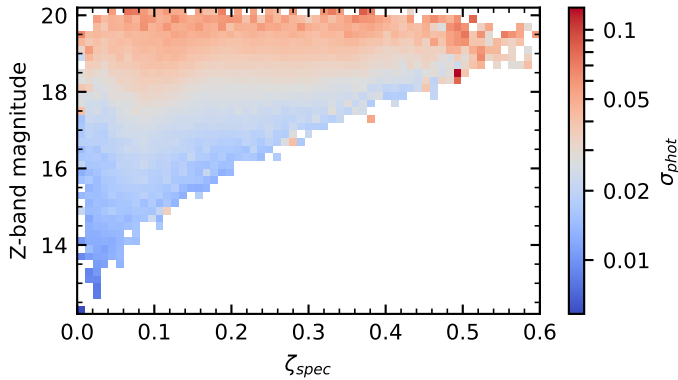


Fig. 18. Photo-z peak width depending on the galaxy's spectroscopic redshift and Z-band magnitude. For clarity, each magnitude and redshift bin shows the average peak width of the galaxies in that bin.

where $\Delta\zeta$ is the difference between the photometric and spectroscopic redshifts expressed in ζ ,

$$\Delta\zeta = \zeta_{phot} - \zeta_{spec}. \quad (40)$$

The bias in the photo-z estimate is defined as median($\Delta\zeta$), and outliers are defined where $|\Delta\zeta|$ exceeds some fixed value. In this paper, we defined the outliers using two thresholds, $|\Delta\zeta| > 0.05$ and $|\Delta\zeta| > 0.1$. The distribution of differences in the GAMA sample is shown in Fig. 19. In Table 3 we quantify the photo-z bias, precision, and outlier fraction for three different magnitude bins. The values are given both for redshift z and $\zeta = \ln(1+z)$.

Figure 21 shows the outlier fraction and σ_{NMAD} for different magnitude bins. As expected (see Sect. 8.1), the σ_{NMAD}

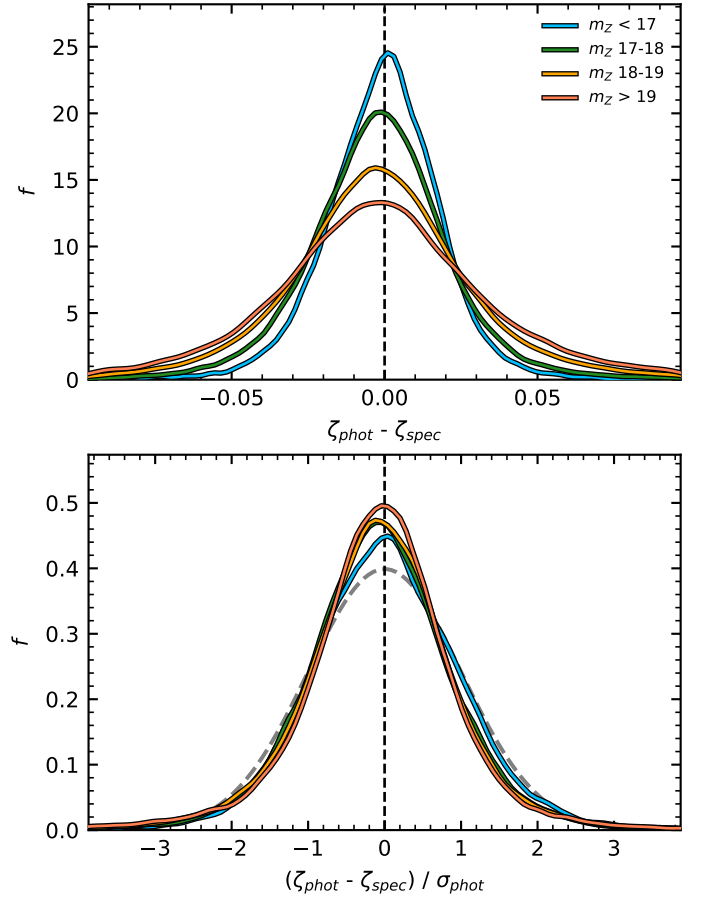


Fig. 19. Photo-z accuracy in multiple magnitude ranges. Upper panel: the colours show the distribution of photo-z accuracy, which is given as the difference between the spectroscopic and photo-z estimates. The dashed black line is drawn at zero to identify any biases in the distributions. Lower panel: the distribution of redshift accuracies that are normalised by the width of their photo-z peak computed using Eq. (16). The dashed grey line shows a Gaussian distribution with a standard deviation of one.

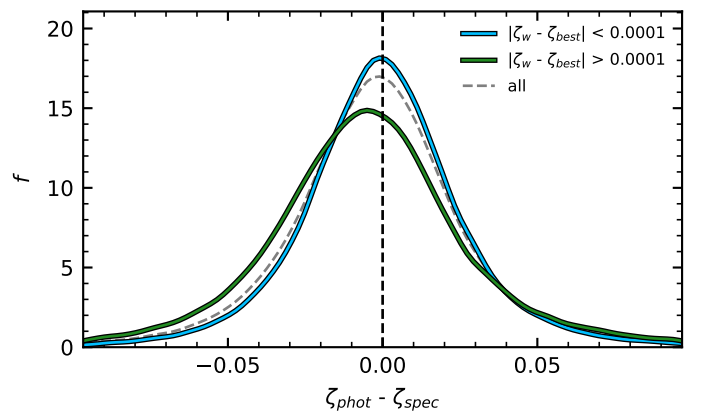


Fig. 20. Photo-z accuracy distribution for two TOPz subsets. The blue line shows the subset where ζ_{best} and ζ_w are closer than 0.0001, and the green line shows the subset with the rest of the galaxies. The dashed grey line is the distribution for the whole sample. The value 0.0001 was chosen so the subsets would contain roughly comparable numbers of galaxies.

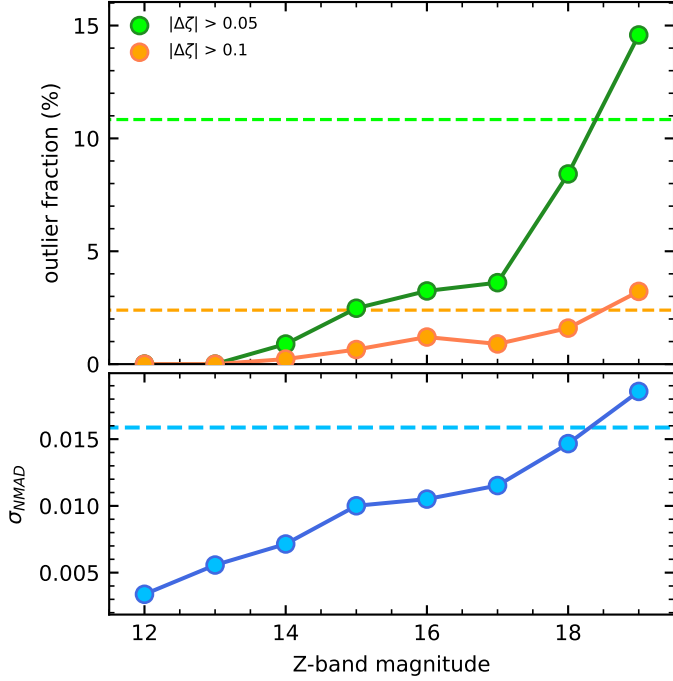


Fig. 21. TOPz photo-z statistics as a function of the apparent Z-band magnitude. Photo-z accuracy is defined as σ_{NMAD} (see Sect. 8.2 and Eq. (39)). Outlier fraction is defined using redshift ζ absolute deviation from spectroscopic redshift for two different $\Delta\zeta$ thresholds: 0.05 and 0.1. Upper panel: outlier fraction as a function of apparent magnitude. The outlier fractions for two different thresholds are marked in green and orange colours. The coloured dashed lines show the outlier fractions over the whole galaxy sample. Lower panel: photo-z accuracy distribution widths as described by the σ_{NMAD} value. The dashed line shows the σ_{NMAD} value calculated over the whole galaxy set.

depends on the observed magnitude, while the photo-z accuracy decreases towards fainter magnitudes. The upper panel in Fig. 21 shows the outlier fraction as a function of galaxy magnitude. For the outlier threshold $|\Delta\zeta| > 0.1$, the fraction of outliers is below two per cent; the only exception is the faintest magnitude bin, where it is around three per cent. However, when we look at the outlier fraction defined with $|\Delta\zeta| > 0.05$, the number of outliers significantly increases for galaxies fainter than $m_Z > 17$ mag. For brighter galaxies, the outlier fraction ($|\Delta\zeta| > 0.05$) remains below four per cent.

Figure 22 shows the σ_{NMAD} , median bias (median($\Delta\zeta$)) and outlier fraction ($|\Delta\zeta| > 0.1$) as a function of spectroscopic redshift and apparent magnitude. The lowest redshift bin is affected by the photo-z prior $\zeta_{\text{phot}} > 0.0$ (see Sect. 8.1 and Fig. 16). The faintest magnitude bin and the largest redshift bins contain relatively few sources. We consider these aspects while interpreting the Fig. 22.

In Fig. 22, the σ_{NMAD} and outlier fraction behave as expected, and both of them mainly depend on the apparent magnitude (see also Fig. 21). The largest redshift bins have a higher outlier fraction for every magnitude range. However, since the number of sources in those bins is relatively low, we cannot draw any firm conclusions. Looking at how the median bias (median($\Delta\zeta$)) depends on redshift and magnitude, we see that there is a clear dependence on redshift. Additionally, the bias depends on the magnitude for redshifts $\zeta < 0.2$. Comparing the bias values with σ_{NMAD} , we see that the bias is always many times smaller than the σ_{NMAD} . Hence, the slight systematic bias does not dominate, and the width of the photo-z posterior

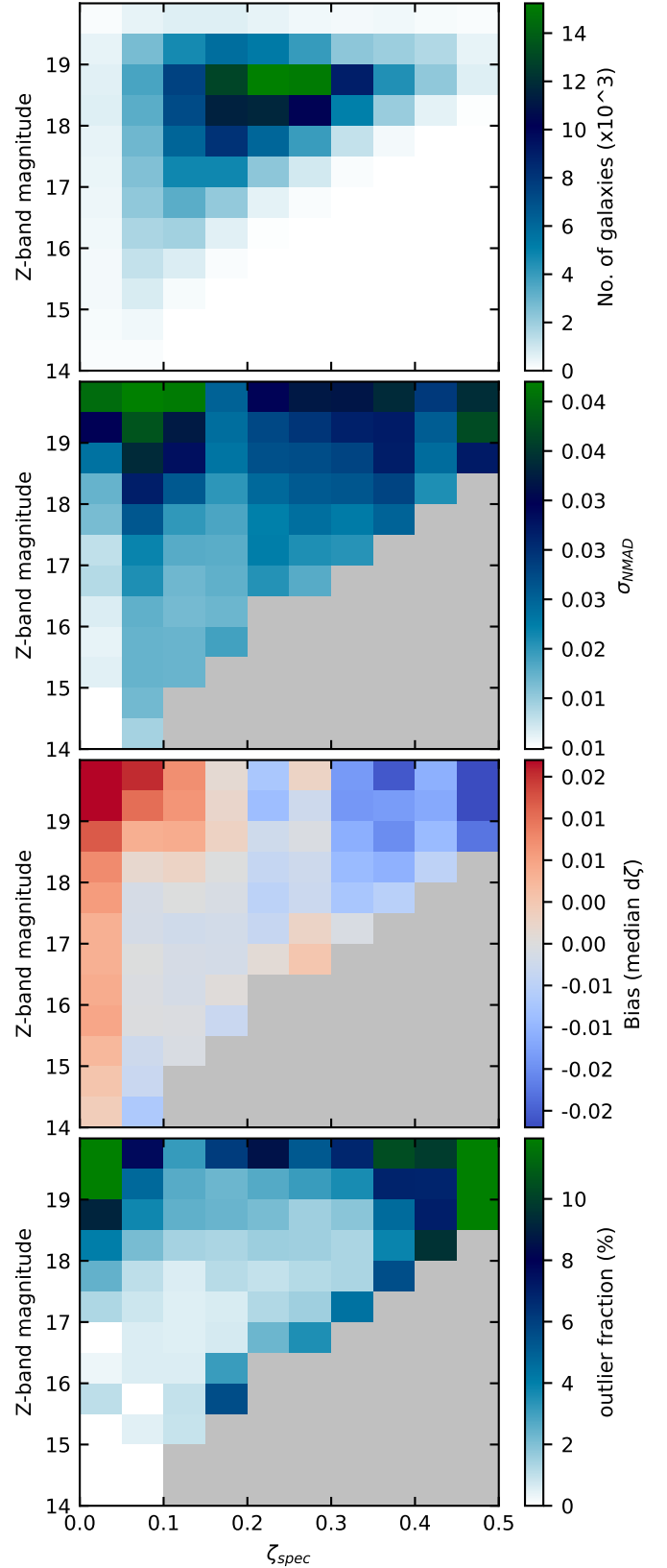


Fig. 22. Photo-z estimation parameter map on a magnitude - ζ_{spec} grid. The top panel shows the number of galaxies in each bin. The three parameters from the second to the bottom panel are photo-z accuracy distribution width (σ_{NMAD} value), the bias of that distribution as the median value of the photometric accuracy, and the outlier rate as the number of galaxies with photometric accuracy worse than 0.1. Bins with fewer than 100 galaxies are drawn as grey areas.

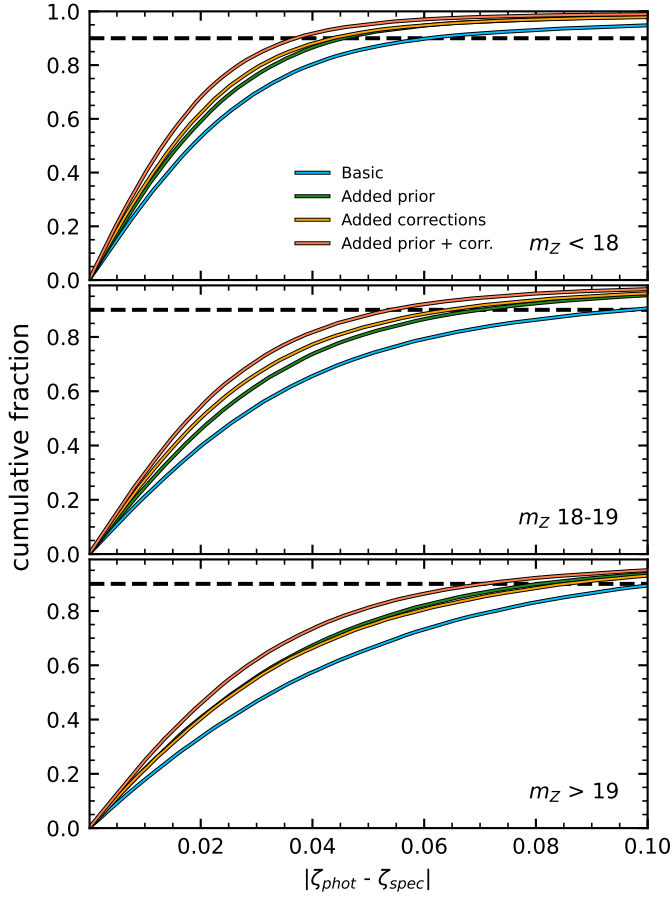


Fig. 23. TOPz feature comparison in three magnitude ranges (different panels). The coloured lines show the cumulative fraction of objects with a certain photo-z accuracy. Each coloured line represents a different TOPz run with certain features turned on or off, as described by the legend on the top panel. The three panels have different sets of galaxies based on their Z-band magnitude (as shown in the bottom-right corner of each panel). The dashed black line indicates the 90% fraction to help better distinguish the accuracies on each panel. See Fig. 19 for the differential distribution of differences between spectroscopic redshifts and photo-z.

predominantly determines the photo-z accuracy. Whether the systematic bias comes from the template-fitting algorithm or the GAMA observational data requires detailed analysis and a large sample of synthetic galaxy spectra.

8.3. Importance of flux corrections and physical priors

In Sect. 4, we described how we correct the input galaxy fluxes and flux uncertainties. In Sect. 7, we introduced a physically motivated prior for the TOPz algorithm. In this section, we will show how these two aspects improve the accuracy of photo-z in general. In our previous paper, Laur et al. (2022) showed that both are improving the photo-z estimation. In Laur et al. (2022), the method was applied to the miniJPAS data, and the flux corrections and priors were defined differently than in the current paper. Below, we will repeat the same analysis using the GAMA data and use the updated flux corrections and priors as described in this paper.

Figure 23 shows the cumulative fraction of objects, where $|\Delta z|$ is smaller than indicated in the x-axis. The different lines show the cumulative fraction for four cases where the flux

corrections or priors were turned on or off. It is clear that when both of them are turned off (titled ‘basic’ on the plot), the photo-z accuracy is significantly lower than when both are used. This is valid for all three magnitude bins. When the flux corrections or priors were turned off individually, the photo-z accuracy was between the two previous cases. We see that both are equally important in improving the photo-z accuracy. For brighter galaxies, the flux corrections have a slightly larger impact.

8.4. Tests on photometric redshift PDFs

In this section, we analyse how well the photo-z probability distribution functions (posteriors) reflect the actual uncertainty of the estimated photo-z. In Sect. 8.1 (see Fig. 19 lower panel), we showed that the TOPz photo-z uncertainty computed with Eq. (16) adequately estimates the actual difference between the photometric and spectroscopic redshift estimates. However, this was estimated for a single object at a time. In this section, we perform integral tests over the entire sample to assess whether, statistically, the TOPz photo-z posteriors are adequate.

In Polsterer et al. (2016), the probability integral transform (PIT) is utilised as an analytical tool to assess the calibration and precision of the photo-z PDFs. The probability value for any single galaxy G is defined as:

$$c_{\text{PIT}}(G) = \sum_{\zeta < \zeta_{\text{spec}}^G} p_{\zeta}^G(\zeta), \quad (41)$$

where we assumed that the photo-z PDF is given in a grid and the PDF is normalised to $\sum p_{\zeta}(\zeta) = 1.0$. The ζ_{spec} is the true redshift, which is the spectroscopic redshift measurement in our case. A PIT test assesses the distribution of the PIT values for the whole data set. When plotted as the cumulative fraction of the single galaxy PIT values $\hat{F}(c_{\text{PIT}})$, a uniform distribution that indicates well-calibrated PDFs will be close to the one-to-one relation. However, any significant deviations from that relation indicate that the PDFs are either over- or underconfident.

Another test to evaluate the PDF calibration involves calculating the fraction of galaxies ($\hat{F}(c_{\text{CI}})$) for which the spectroscopic redshift is within a specified confidence interval (CI). This approach is detailed in Wittman et al. (2016), who define the CI as the area under the PDF that reaches the probability threshold calculated at the spectroscopic redshift. For a single galaxy G , the confidence is defined as:

$$c_{\text{CI}}(G) = \sum_{\zeta \in p_{\zeta}^G(\zeta) \geq p_{\zeta}^G(\zeta_{\text{spec}})} p_{\zeta}^G(\zeta). \quad (42)$$

For well-calibrated PDFs, one would expect 10% of galaxies to fall within a 10% CI, 20% within a 20% CI, and so on. Thus, similarly to the PIT test, if the cumulative fraction $\hat{F}(c_{\text{CI}})$ of the whole data set aligns with the CI values, the PDFs are considered well-calibrated and will be close to the one-to-one relation. Again, any deviations below or above this expectation indicate under- or overconfidence, respectively.

To assess the calibration of our PDFs, we conducted both the PIT and the CI tests, and the results can be seen in Fig. 24. We performed the test for four magnitude bins, where only the faintest magnitude bin in the PIT test showed a slight deviation from the one-to-one line. Generally, the photo-z PDFs from the TOPz code are statistically representative and reflect the estimated photo-z’s actual uncertainty.

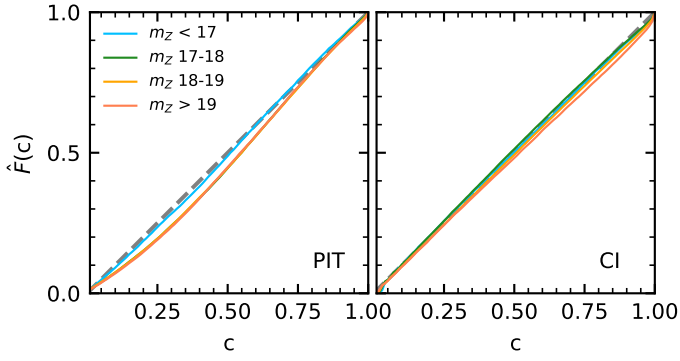


Fig. 24. PIT and CI tests that were conducted on the TOPz output PDFs. The two panels from left to right show the cumulative fraction of the underlying PIT and CI values, respectively. For more info on PIT and CI tests, see Sect. 8.4. The one-to-one relation that shows the ideal distribution is drawn with a diagonal dashed grey line. The cumulative fraction is given in four magnitude ranges with the values and colours identical to Fig. 19. Only the faintest magnitude bin shows a slight deviation from the theoretical expectation.

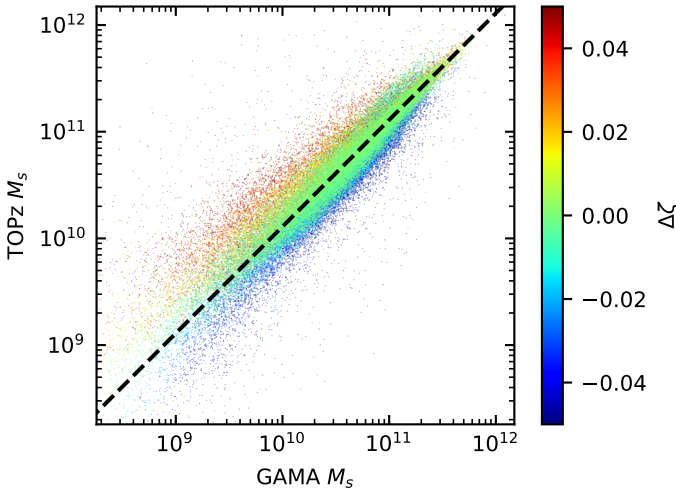


Fig. 25. Comparison between TOPz and GAMA stellar masses for each galaxy. The x -axis shows the GAMA stellar masses acquired using GKV ProFound photometry (see Sect. 3). The y -axis shows the TOPz stellar masses that are computed using the spectral templates from the CIGALE software (see Sect. 2.5). The colour denotes the TOPz photo- z accuracy. The dashed black line is a median stellar mass ratio relation, where $M_{\star, \text{TOPz}} \approx 1.3 \cdot M_{\star, \text{GAMA}}$. Galaxies where TOPz photo- z are close to the spectroscopic redshifts follow the stellar masses from the GAMA database very well.

8.5. Stellar masses from the TOPz code

Template fitting in the TOPz code uses the physical galaxy spectral templates generated by the CIGALE software (see Sect. 5). In Sect. 2.5, we explained how the CIGALE spectral templates can be used to estimate galaxy physical parameters. In this paper, we used the TOPz code to compute the stellar masses for GAMA galaxies.

In Fig. 25, we compare the TOPz stellar masses with the stellar masses available in the GAMA database (see Sect. 3). The GAMA stellar masses are computed for the spectroscopic sample using the spectroscopic redshifts, while TOPz stellar masses are gathered from the best-fitting templates. In general, TOPz stellar masses follow the GAMA stellar masses very nicely. Most of the scatter visible in Fig. 25 is caused by the difference between

the photometric and spectroscopic redshifts. The 1.3 times difference between the stellar masses in the GAMA database and from the TOPz code (CIGALE models) is probably caused by the differences between the assumptions, initial mass functions, and the stellar population synthesis models used by the GAMA team and in the CIGALE software.

8.6. Comparison with EAZY and SFM photometric redshifts

The photo- z for the GAMA data has been previously computed using the EAZY code and the Scaled Flux Matching (SFM) technique. See Driver et al. (2022) for a detailed description of the GAMA data release. We will use these two previous photo- z estimates and compare them with the photometric redshifts derived using the TOPz code. EAZY and TOPz codes are both based on the template-fitting method. The main difference between the two codes is in their details. Both codes use different spectral templates, flux corrections and priors⁶. The second method, the SFM technique, relies on the spectroscopic sample and assigns the photo- z based on the similarity of the observed galaxy SED in different passbands. In general, the SFM technique is closer to the traditional machine-learning photo- z codes, where the main limitation is the coverage of the spectroscopic sample. For the GAMA sample used in the current paper, the spectroscopic coverage is excellent, and it is expected that the SFM code performs better than the template-fitting approach.

To compare the three codes, we use the spectroscopic sample where we have the photo- z estimate from all three codes, which limits our sample to 188 984 galaxies. From EAZY code, we used the z_{peak} values and from SFM code, the z_{exp} values. The performance of the TOPz, SFM, and EAZY codes can be seen in the upper panel of Fig. 26. In general, all three codes are comparable, while the TOPz seems to slightly outperform the EAZY results. Looking at the outliers, it is visible that the SFM is performing the best, which is somewhat expected since the spectroscopic training sample for the SFM technique is sufficiently representative of the GAMA spectroscopic sample. The lower panel in Fig. 26 shows the comparison between the three codes themselves. It is visible that the distribution is narrower than for the comparison with the spectroscopic sample. This indicates that the photometry limits the photo- z accuracy in all codes. Since all codes use the same photometry, they are expected to provide very similar photo- z estimates. A more detailed look reveals that the two codes most identical to each other are the TOPz and SFM. In Table 4 we quantify the bias, precision, and outlier fraction for the three codes.

The photo- z accuracy for the TOPz, SFM, and EAZY codes is shown in Fig. 27. The distribution of differences between photometric and spectroscopic redshifts is shown in the upper panel, where one can see that the EAZY code has the largest bias and the distribution is wider than for the other two codes. The TOPz and SFM distributions are centred around zero and have fairly similar widths, while SFM shows marginally better results. On the lower panel of Fig. 27, we normalised the redshift difference with photo- z uncertainty. For SFM code, we used the zerr values as the σ_{phot} . For EAZY code, we calculated the σ_{phot} as $(u68 - l68)/2$. For both cases, we took care to convert the σ_{phot} values from the redshift units to ζ units. The normalised distributions confirm that the TOPz and SFM codes perform equally well. The normalised distributions are slightly more peaked than

⁶ The EAZY code applied to the GAMA sample includes no flux calibration offsets (E. N. Taylor, private communication).

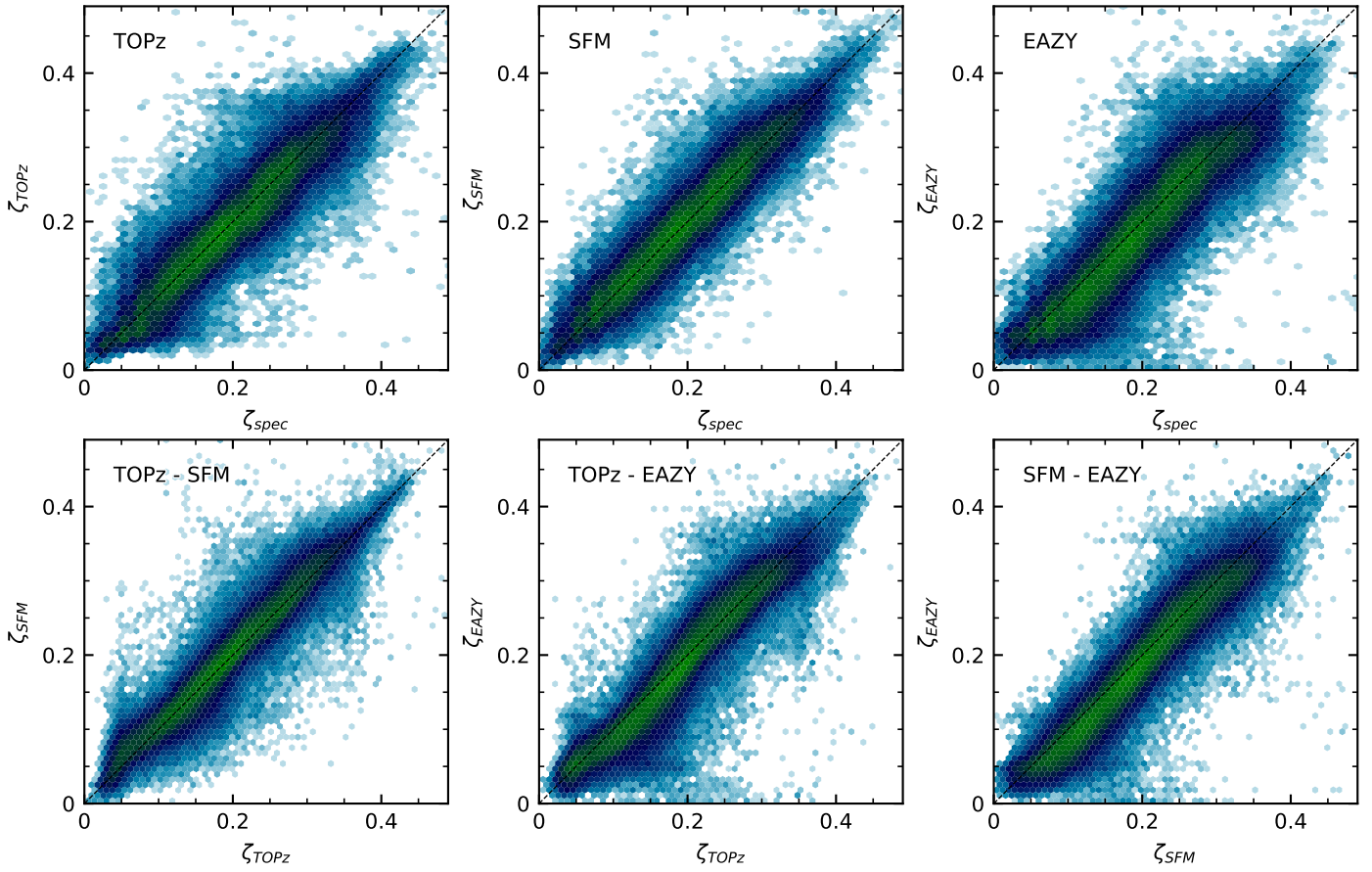


Fig. 26. Comparison between photo-z codes using GAMA data. All the colours and markings are identical to the top panels in Fig. 16. Upper panels: the relation between photo-z estimates and spectroscopic redshifts for the TOPz, SFM, and EAZY codes, respectively. See Sect. 8.6 for a brief description of each method. Lower panels: the relation between the photo-z estimates for each photo-z code pair.

Table 4. Photo-z parameters for the TOPz, EAZY, and SFM codes.

	TOPz	EAZY	SFM
BIAS			
median(dz)	−0.001609	−0.008851	0.001390
median($d\zeta$)	−0.001611	−0.008890	0.001389
OUTLIERS			
$ dz > 0.1$	1.86%	2.69%	1.00%
$ dz > 0.15$	0.71%	1.21%	0.37%
$ d\zeta > 0.1$	1.83%	2.80%	0.97%
$ d\zeta > 0.15$	0.68%	1.24%	0.35%
$ d\zeta /\sigma_{\text{phot}} > 3$	2.06%	5.79%	3.04%
PRECISION			
$\sigma(dz)$	0.0370	0.1440	0.0310
$\sigma_{\text{MAD}}(dz)$	0.0241	0.0305	0.0221
$\sigma_{\text{NMAD}}(dz)$	0.0152	0.0188	0.0139
$\sigma(d\zeta)$	0.0370	0.0810	0.0310
$\sigma_{\text{MAD}}(d\zeta)$	0.0241	0.0306	0.0221
$\sigma_{\text{NMAD}}(d\zeta)$	0.0152	0.0190	0.0139

Notes. $dz = (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$ and $d\zeta = \zeta_{\text{phot}} - \zeta_{\text{spec}}$.

the Gaussian distribution, indicating that the photo-z uncertainty might be slightly underestimated for some sources.

The TOPz performs very well, outperforming slightly the similar template-fitting code EAZY. Compared with SFM, the

two codes are equally good, while SFM has a slightly lower number of outliers. It is somewhat expected since SFM is similar to the machine-learning codes and is trained using the spectroscopic redshift sample. A clear path forward to improve the TOPz photo-z is to combine the template-fitting methods with an approach that uses the spectroscopic redshift sample.

In the analysis carried out in this section, the TOPz used the GAMA spectroscopic sample to optimise the galaxy template library. Similarly, SFM used the same spectroscopic sample for photo-z estimation. When the same spectroscopic dataset is used for comparison, both codes are expected to perform very well. However, the photo-z estimates are not guaranteed to be reliable for fainter galaxies and at higher redshifts. In Appendix B we use the independent DESI spectroscopic redshifts to validate the TOPz photo-z estimates for fainter and slightly higher redshift galaxies. As shown in the Appendix B, the conclusions drawn in this section are valid.

9. Conclusions

In this paper, we have introduced the updated TOPz code, a Bayesian photo-z estimation tool based on template fitting. Our approach integrates several critical features that significantly enhance the accuracy and reliability of photo-z estimates.

We have demonstrated the necessity of both flux corrections and flux uncertainty corrections to account for systematic biases and uncertainties in the observed photometric measurements.

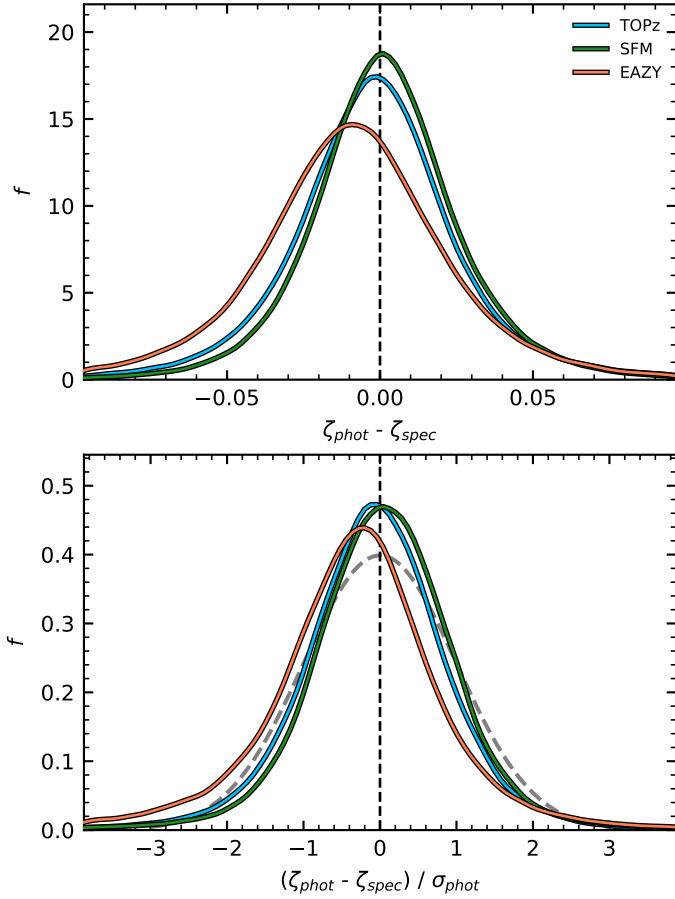


Fig. 27. Photo- z accuracy distributions for different photo- z codes. The panels are identical to Fig. 19 except for the colours that now denote the various codes instead.

These corrections ensure that the input photometry is accurate and unbiased for the SED fitting. Additionally, the physical prior, derived from the luminosity function and survey volume prior, improves the redshift estimates by incorporating expected galaxy distributions. This approach is crucial for achieving reliable photo- z , especially for faint galaxies with more than one peak in their photo- z posterior.

A vital aspect of any template-fitting photo- z codes is the selection of spectral templates. We have demonstrated the importance of template set optimisation in adequately representing the diversity of galaxy types. This optimisation eliminates the need for an independent template prior and ensures that the templates cover all possible galaxy types (see Section 5). The physical templates generated using the CIGALE software provide a good basis for the template set optimisation, allowing us to cover a wide variety of galaxy spectral types needed for template-fitting photo- z estimation.

Our results show that the TOPz photo- z posteriors are statistically reliable and accurately reflect the uncertainty in photo- z estimates. The photo- z accuracy across various magnitude bins shows that the flux corrections and posteriors effectively capture the actual uncertainty of the estimates (see Section 8).

Applying TOPz to the GAMA spectroscopic sample, we found no significant bias in the photo- z estimates (see Table 3). The σ_{NMAD} values are uniform, and the outlier fraction is low, indicating the robustness of our approach. The $\sigma_{\text{NMAD}} = 0.012$ for $m_Z < 18$ and increases to $\sigma_{\text{NMAD}} = 0.021$ for $m_Z > 19$. The

outlier fraction ($|dz| > 0.1$) in the same magnitude bins increases from 1 to 5%. Table 3 lists more quantitative measures. The comparison with other photo- z codes, such as EAZY and SFM, shows that TOPz performs competitively, providing accurate and reliable redshift estimates. In our comparison with the EAZY and SFM photo- z codes, we found that TOPz performs very well, slightly outperforming the EAZY code and showing comparable results to the SFM technique (see Table 4 for a quantitative comparison). The latter, which relies on the spectroscopic sample for training, had a slightly lower number of outliers, which is expected given its similarity to machine-learning approaches. However, the TOPz and SFM codes both show excellent performance, with the main differences attributed to the underlying methodologies and the use of spectroscopic training samples.

Using the physical templates from CIGALE, we estimated the stellar masses for GAMA galaxies. The comparison with the GAMA database shows that TOPz stellar masses are consistent, with most of the scatter being attributed to differences between photometric and spectroscopic redshifts. The ability to estimate physical parameters, such as stellar masses, directly from photometric data highlights the potential of the TOPz approach.

In this paper, we used the GAMA sample as a proof-of-concept. Extending this methodology to other surveys requires additional testing and validation. To further improve photo- z accuracy, we plan to incorporate additional photometric data, address systematics in photometry, and reduce outliers. Including machine-learning methodology in the template-fitting code is also a promising avenue. We aim to apply the TOPz algorithm to the J-PAS narrow-band survey and explore the combination of the J-PAS narrow bands with SDSS photometry. This will validate and enhance the capabilities of the TOPz code, particularly for high-redshift galaxies. In Appendix B we show an independent comparison between TOPz photo- z and DESI DR1 spectroscopic redshifts, which extend to higher redshifts and fainter magnitudes not used in the TOPz photo- z optimisation process.

In summary, the TOPz code significantly advances photo- z estimation based on the template-fitting approach. By integrating flux corrections, physical priors, and optimised template sets, TOPz provides accurate, reliable, and physically motivated redshift estimates. These capabilities are crucial for advancing our understanding of the universe through cosmological and galaxy evolution studies. Combining the spectroscopic redshift samples with photometric redshift data allows us to study the faint dwarf galaxy population in the cosmic web. This is a promising avenue in the upcoming 4MOST WAVES and 4HS surveys (de Jong et al. 2019; Driver et al. 2019; Taylor et al. 2023).

Data availability

Table presented in Appendix A is available at the CDS via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/703/A261>.

The TOPz code is entirely written in modern Fortran and is publicly available on GitHub (<https://github.com/etempel/TOPz>). In addition to the code, we make available all Python scripts that were used to generate the figures presented in this paper. The figure scripts are part of the same GitHub repository. All the internal data sets used during the paper preparation are available upon reasonable request. The filter transmission curves used in this paper are taken from the Spanish Virtual Observatory Filter Profile Service (Rodrigo et al. 2012; Rodrigo & Solano 2020).

Acknowledgements. The authors thank the anonymous referee for their detailed reading of the paper and helpful recommendations, resulting in multiple improvements. We thank Edward (Ned) Taylor for valuable discussions about applying the EAZY code to the GAMA sample. This work was funded by the Estonian Ministry of Education and Research (grant TK202), Estonian Research Council (grants PRG1006, PRG3034, PRG2159, PSG700) and the European Union's Horizon Europe research and innovation programme (EXCOSM, grant No. 101159513). Extensive exploratory analysis in this work was done using the TOPCAT software (Taylor 2005). Most of the figures were generated using the Matplotlib package (Hunter 2007). GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes, including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP, providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is <https://www.gama-survey.org>. This research has made use of the Spanish Virtual Observatory (<https://svo.cab.inta-csic.es>) project funded by MCIN/AEI/10.13039/501100011033 through grant PID2020-112949GB-I00.

References

- Abdalla, F. B., Banerji, M., Lahav, O., & Rashkov, V. 2011, *MNRAS*, **417**, 1891
- Almosallam, I. A., Jarvis, M. J., & Roberts, S. J. 2016a, *MNRAS*, **462**, 726
- Almosallam, I. A., Lindsay, S. N., Jarvis, M. J., & Roberts, S. J. 2016b, *MNRAS*, **455**, 2387
- Ansari, Z., Agnello, A., & Gall, C. 2021, *A&A*, **650**, A90
- Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, *MNRAS*, **310**, 540
- Bacon, R., Conseil, S., Mary, D., et al. 2017, *A&A*, **608**, A1
- Baldry, I. K. 2018, arXiv e-prints [arXiv:1812.05135]
- Baldry, I. K., Sullivan, T., Rani, R., & Turner, S. 2021, *MNRAS*, **500**, 1557
- Beck, R., Dobos, L., Budavári, T., Szalay, A. S., & Csabai, I. 2016, *MNRAS*, **460**, 1371
- Beck, R., Lin, C. A., Ishida, E. E. O., et al. 2017, *MNRAS*, **468**, 4323
- Bellstedt, S., Driver, S. P., Robotham, A. S. G., et al. 2020, *MNRAS*, **496**, 3235
- Benítez, N. 2000, *ApJ*, **536**, 571
- Bilicki, M., Peacock, J. A., Jarrett, T. H., et al. 2016, *ApJS*, **225**, 5
- Bilicki, M., Hoekstra, H., Brown, M. J. I., et al. 2018, *A&A*, **616**, A69
- Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003, *ApJ*, **592**, 819
- Bolzonella, M., Miralles, J. M., & Pelló, R. 2000, *A&A*, **363**, 476
- Bonoli, S., Marín-Franch, A., Varela, J., et al. 2021, *A&A*, **653**, A31
- Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, *A&A*, **622**, A103
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, **686**, 1503
- Brescia, M., Salvato, M., Cavuoti, S., et al. 2019, *MNRAS*, **489**, 663
- Brinchmann, J., Inami, H., Bacon, R., et al. 2017, *A&A*, **608**, A3
- Carrasco Kind, M., & Brunner, R. J. 2013, *MNRAS*, **432**, 1483
- Castelli, F., & Kurucz, R. L. 2003, *IAU Symp.*, **210**, A20
- Castelli, F., Gratton, R. G., & Kurucz, R. L. 1997, *A&A*, **318**, 841
- Cavuoti, S., Tortora, C., Brescia, M., et al. 2017, *MNRAS*, **466**, 2039
- Cook, T. L., Bandi, B., Philipsborn, S., et al. 2024, *MNRAS*, **535**, 2129
- Crenshaw, J. F., Kalmbach, J. B., Gagliano, A., et al. 2024, *AJ*, **168**, 80
- Csörnyei, G., Dobos, L., & Csabai, I. 2021, *MNRAS*, **502**, 5762
- Dahlen, T., Mobasher, B., Faber, S. M., et al. 2013, *ApJ*, **775**, 93
- de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, *Exp. Astron.*, **35**, 25
- de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, *The Messenger*, **175**, 3
- De Vicente, J., Sánchez, E., & Sevilla-Noarbe, I. 2016, *MNRAS*, **459**, 3078
- DESI Collaboration (Abdul-Karim, M., et al.) 2025, arXiv e-prints [arXiv:2503.14745]
- D'Isanto, A., & Polsterer, K. L. 2018, *A&A*, **609**, A111
- Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, *MNRAS*, **413**, 971
- Driver, S. P., Liske, J., Davies, L. J. M., et al. 2019, *The Messenger*, **175**, 46
- Driver, S. P., Bellstedt, S., Robotham, A. S. G., et al. 2022, *MNRAS*, **513**, 439
- Duncan, K. J. 2022, *MNRAS*, **512**, 3662
- Duncan, K. J., Brown, M. J. I., Williams, W. L., et al. 2018, *MNRAS*, **473**, 2655
- Edge, A., Sutherland, W., Kuijken, K., et al. 2013, *The Messenger*, **154**, 32
- Euclid Collaboration (Desprez, G., et al.) 2020, *A&A*, **644**, A31
- Feldmann, R., Carollo, C. M., Porciani, C., et al. 2006, *MNRAS*, **372**, 565
- Fotopoulou, S., & Paltani, S. 2018, *A&A*, **619**, A14
- Frontiera-Pons, J., Sureau, F., Bobin, J., & Kilbinger, M. 2023, *Astron. Comput.*, **44**, 100735
- Gatti, M., Vielzeuf, P., Davis, C., et al. 2018, *MNRAS*, **477**, 1664
- Gerdes, D. W., Sypniewski, A. J., McKay, T. A., et al. 2010, *ApJ*, **715**, 823
- Graham, M. L., Connolly, A. J., Wang, W., et al. 2020, *AJ*, **159**, 258
- Hatfield, P. W., Almosallam, I. A., Jarvis, M. J., et al. 2020, *MNRAS*, **498**, 5498
- Hatfield, P. W., Jarvis, M. J., Adams, N., et al. 2022, *MNRAS*, **513**, 3719
- Hernán-Caballero, A., Varela, J., López-Sanjuan, C., et al. 2021, *A&A*, **654**, A101
- Hildebrandt, H., Wolf, C., & Benítez, N. 2008, *A&A*, **480**, 703
- Hogan, R., Fairbairn, M., & Seeburn, N. 2015, *MNRAS*, **449**, 2040
- Hsu, L.-T., Salvato, M., Nandra, K., et al. 2014, *ApJ*, **796**, 60
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *A&A*, **457**, 841
- Jones, D. M., & Heavens, A. F. 2019, *MNRAS*, **483**, 2487
- Jones, E., Do, T., Boscoe, B., et al. 2024, *ApJ*, **964**, 130
- Koester, D. 2010, *Mem. Soc. Astron. Italiana*, **81**, 921
- Laplace, P.-S. 1774, *Memoir on the Probability of Causes of Events* (USA: University of Chicago)
- Laur, J., Tempel, E., Tamm, A., et al. 2022, *A&A*, **668**, A8
- López-Sanjuan, C., Tempel, E., Benítez, N., et al. 2017, *A&A*, **599**, A62
- Luberto, J., Furlanetto, S., & Mirocha, J. 2024, arXiv e-prints [arXiv:2409.20519]
- Molino, A., Benítez, N., Moles, M., et al. 2014, *MNRAS*, **441**, 2891
- Moskowitz, I., Gawiser, E., Crenshaw, J. F., et al. 2024, *ApJ*, **967**, L6
- Naidoo, K., Johnston, H., Joachimi, B., et al. 2023, *A&A*, **670**, A149
- Navarro-Girónés, D., Gaztañaga, E., Crocce, M., et al. 2024, *MNRAS*, **534**, 1504
- Newman, J. A., & Gruen, D. 2022, *ARA&A*, **60**, 363
- Pasquet, J., Bertin, E., Treyer, M., Arnouts, S., & Fouchez, D. 2019, *A&A*, **621**, A26
- Pathi, I. M., Soo, J. Y. H., Wee, M. J., et al. 2025, *J. Cosmology Astropart. Phys.*, **2025**, 097
- Pickles, A. J. 1998, *PASP*, **110**, 863
- Polsterer, K. L., D'Isanto, A., & Gieseke, F. 2016, arXiv e-prints [arXiv:1608.08016]
- Razim, O., Cavuoti, S., Brescia, M., et al. 2021, *MNRAS*, **507**, 5034
- Robotham, A. S. G., Davies, L. J. M., Driver, S. P., et al. 2018, *MNRAS*, **476**, 3137
- Rodrigo, C., & Solano, E. 2020, in *XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society*, 182
- Rodrigo, C., Solano, E., & Bayo, A. 2012, SVO Filter Profile Service Version 1.0, IVOA Working Draft 15 October 2012
- Ryden, B., & Pogge, R. 2021, *Interstellar and Intergalactic Medium* (Cambridge: Cambridge University Press)
- Sadeh, I., Abdalla, F. B., & Lahav, O. 2016, *PASP*, **128**, 104502
- Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, *ApJ*, **690**, 1250
- Salvato, M., Ilbert, O., & Hoyle, B. 2019, *Nat. Astron.*, **3**, 212
- Schmidt, S. J., Malz, A. I., Soo, J. Y. H., et al. 2020, *MNRAS*, **499**, 1587
- Schuldt, S., Suyu, S. H., Cañameras, R., et al. 2021, *A&A*, **651**, A55
- Shuntov, M., Pasquet, J., Arnouts, S., et al. 2020, *A&A*, **636**, A90
- Soo, J. Y. H., Joachimi, B., Eriksen, M., et al. 2021, *MNRAS*, **503**, 4118
- Tanaka, M. 2015, *ApJ*, **801**, 20
- Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, *PASJ*, **70**, S9
- Tanigawa, S., Glazebrook, K., Jacobs, C., Labbe, I., & Qin, A. K. 2024, *MNRAS*, **530**, 2012
- Taylor, M. B. 2005, *ASP Conf. Ser.*, **347**, 29
- Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, *MNRAS*, **418**, 1587
- Taylor, E. N., Cluver, M., Bell, E., et al. 2023, *The Messenger*, **190**, 46
- Teixeira, G., Bom, C. R., Santana-Silva, L., et al. 2024, *Astronomy and Computing*, **49**, 100886
- Tempel, E., Tamm, A., Gramann, M., et al. 2014, *A&A*, **566**, A1
- Toribio San Cipriano, L., De Vicente, J., Sevilla-Noarbe, I., et al. 2024, *A&A*, **686**, A38
- Tosone, F., Cagliari, M. S., Guzzo, L., Granett, B. R., & Crespi, A. 2023, *A&A*, **672**, A150
- Treyer, M., Ait Ouahmed, R., Pasquet, J., et al. 2024, *MNRAS*, **527**, 651
- Wittman, D., Bhaskar, R., & Tobin, R. 2016, *MNRAS*, **457**, 4005
- Wolf, C., Meisenheimer, K., Kleinheinrich, M., et al. 2004, *A&A*, **421**, 913
- Yang, G., Boquien, M., Brandt, W. N., et al. 2022, *ApJ*, **927**, 192
- Zou, H., Gao, J., Zhou, X., & Kong, X. 2019, *ApJS*, **242**, 8

Appendix A: TOPz photo-z for the GAMA sample

The TOPz photo-z catalogue for the GAMA photometric sample is made available through the Strasbourg Astronomical Data Centre (CDS). The output table contains the following information (column numbers are given in square brackets):

1. [1] `uberID` – unique GAMA III ID of object;
2. [2] `CATAID` – ID of best matching GAMA II object;
3. [3] `zspec` – spectroscopic redshift from the GAMA database;
4. [4–5] `ra`, `dec` – right ascension and declination (deg);
5. [6] `mag` – magnitude in the $r + Z$ detection band;
6. [7] `nflux` – number of used fluxes for photo-z estimation;
7. [8] `zeta_w` – posterior-weighted redshift ($\zeta = \ln(1 + z)$);
8. [9] `mass_w` – posterior-weighted stellar mass estimate;
9. [10] `zeta_best` – posterior-weighted redshift from the best peak;
10. [11] `zeta_best_p` – summed probability in the best peak;
11. [12–13] `zeta_best_min`, `zeta_best_max` – minimum and maximum best-peak redshift values;
12. [14] `zeta_best_sigma` – weighted standard deviation of the best-peak redshift;
13. [15] `zeta_alt1` – posterior-weighted redshift from the first alternative peak;
14. [16] `zeta_alt1_p` – summed probability in the first alternative peak;
15. [17–18] `zeta_alt1_min`, `zeta_alt1_max` – minimum and maximum first alternative peak redshift values;
16. [19] `zeta_alt1_sigma` – weighted standard deviation of the first alternative peak;
17. [20] `zeta_alt2` – posterior-weighted redshift from the second alternative peak;
18. [21] `zeta_alt2_p` – summed probability in the second alternative peak;
19. [22–23] `zeta_alt2_min`, `zeta_alt2_max` – minimum and maximum second alternative peak redshift values;
20. [24] `zeta_alt2_sigma` – weighted standard deviation of the second alternative peak redshift;
21. [25] `npeak` – number of identified peaks in the photo-z posterior;
22. [26] `chi2_mean` – posterior-weighted mean χ^2 value;
23. [27] `is_star` – TOPz probability that this object is a star.

Appendix B: TOPz and EAZY results in the DESI DR1 sample

To validate the TOPz and EAZY photo-z results, we use the DESI DR1 (Dark Energy Spectroscopic Instrument Data Release 1)⁷ (DESI Collaboration 2025). We select DESI sources, where `ZCAT_PRIMARY = True`, `ZWARN = 0` and `OBJTYPE = TGT`. We limited our sample to $z < 2.0$. To have an independent validation set, we eliminated sources that were part of the GAMA spectroscopic sample.

In Figs. B.1 and B.2 we show the comparison between DESI, EAZY, and TOPz redshift estimates. Visually, the scatter between TOPz and EAZY is smaller than between photo-z and DESI spectroscopic redshift. In Fig. B.3, we show the absolute and uncertainty-scaled difference as a function of magnitude. While both EAZY and TOPz perform similarly, there is a slightly

Table B.1. Photo-z parameters for DESI sample.

	TOPz	EAZY
BIAS		
$\text{median}(dz)$	0.000357	-0.018056
$\text{median}(d\zeta)$	0.000357	-0.018221
OUTLIERS		
$ dz > 0.1$	24.97%	27.07%
$ dz > 0.15$	15.71%	17.64%
$ d\zeta > 0.1$	24.94%	27.50%
$ d\zeta > 0.15$	15.77%	18.20%
$ d\zeta /\sigma_{\text{phot}} > 3$	7.62%	19.74%
PRECISION		
$\sigma(dz)$	0.156	0.149
$\sigma_{\text{MAD}}(dz)$	0.0644	0.0712
$\sigma_{\text{NMAD}}(dz)$	0.0467	0.0506
$\sigma(d\zeta)$	0.174	0.180
$\sigma_{\text{MAD}}(d\zeta)$	0.0644	0.0716
$\sigma_{\text{NMAD}}(d\zeta)$	0.0467	0.0511

Notes. $dz = (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$ and $d\zeta = \zeta_{\text{phot}} - \zeta_{\text{spec}}$.

smaller bias in TOPz photo-z estimates. This is also confirmed quantitatively in Table B.1. In Fig. B.3, we show the distribution of uncertainty-scaled differences between EAZY and TOPz photo-z and DESI spectroscopic redshifts. It is visible that TOPz results are less biased and the uncertainties follow more closely the expected Gaussian distribution.

The analysis carried out using independent DESI spectroscopic redshifts confirms that the conclusions drawn in the main body of the paper are valid. It is true even for the fainter and higher redshift galaxies that were not used during the TOPz photo-z optimisation process.

⁷ <https://data.desi.lbl.gov/doc/releases/dr1/>

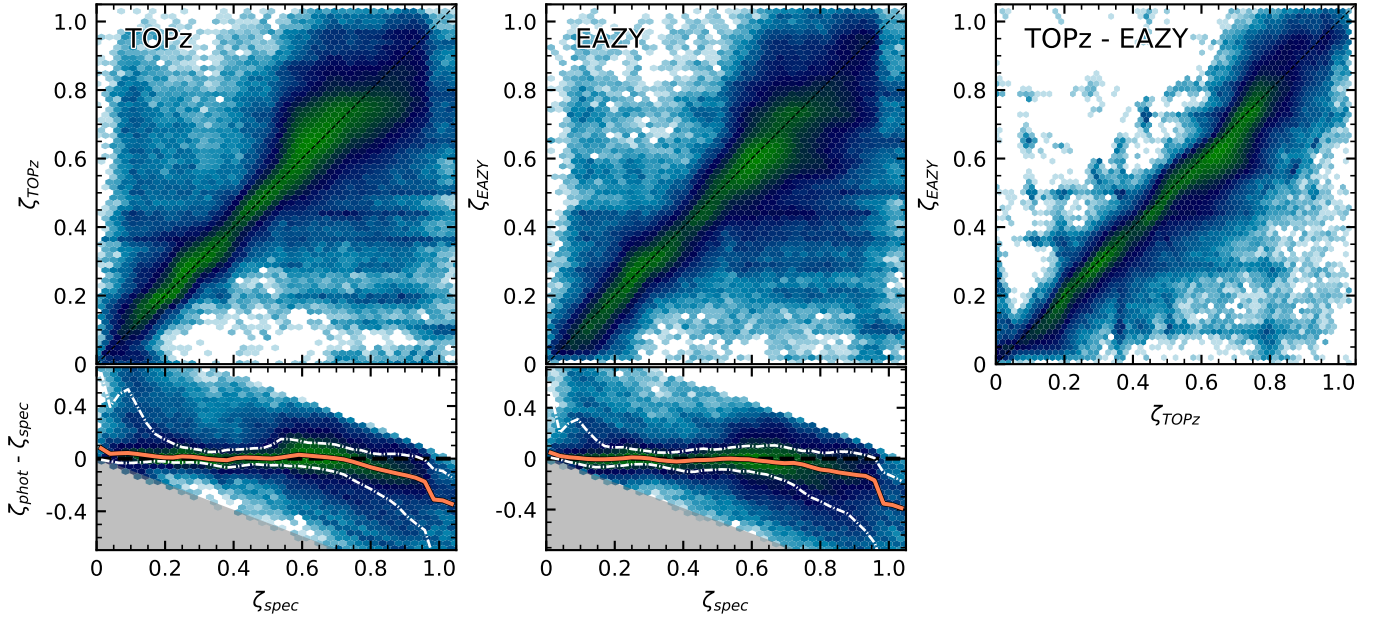


Fig. B.1. Comparison between photo-z results of TOPz and EAZY in the DESI sample. **Upper panels:** The two leftmost panels show the individual photo-z results compared to the spectroscopic redshift ($\zeta = \ln(1+z)$) while the right panel shows the relation between the TOPz and EAZY photo-z. The layout is similar to Fig. 26. **Lower panels:** Photo-z accuracy as a function of spectroscopic redshift for TOPz and EAZY codes, respectively. The layout is similar to the lower panels in Fig. 16.

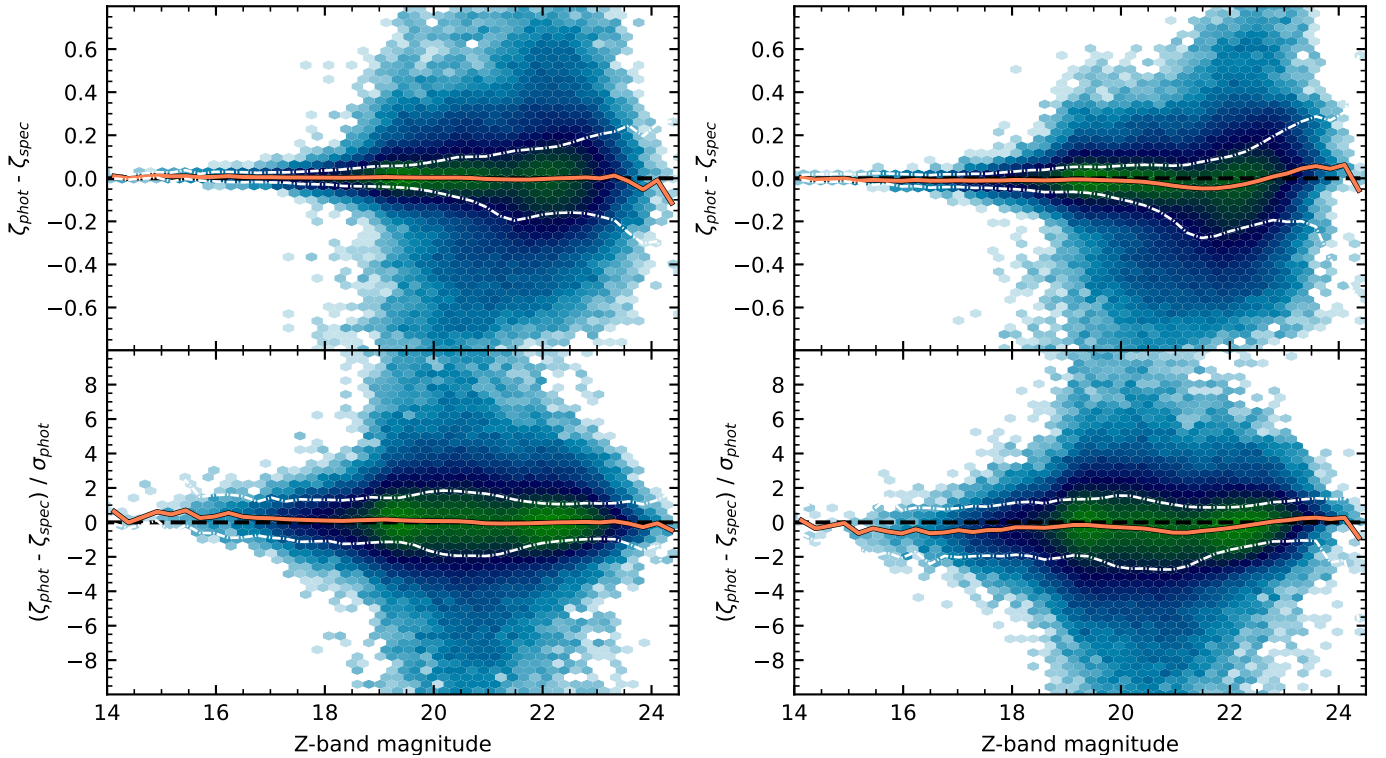


Fig. B.2. Magnitude dependence on the photo-z accuracy for TOPz (left) and EAZY (right) code in the DESI sample. The layout is identical to Fig. 17.

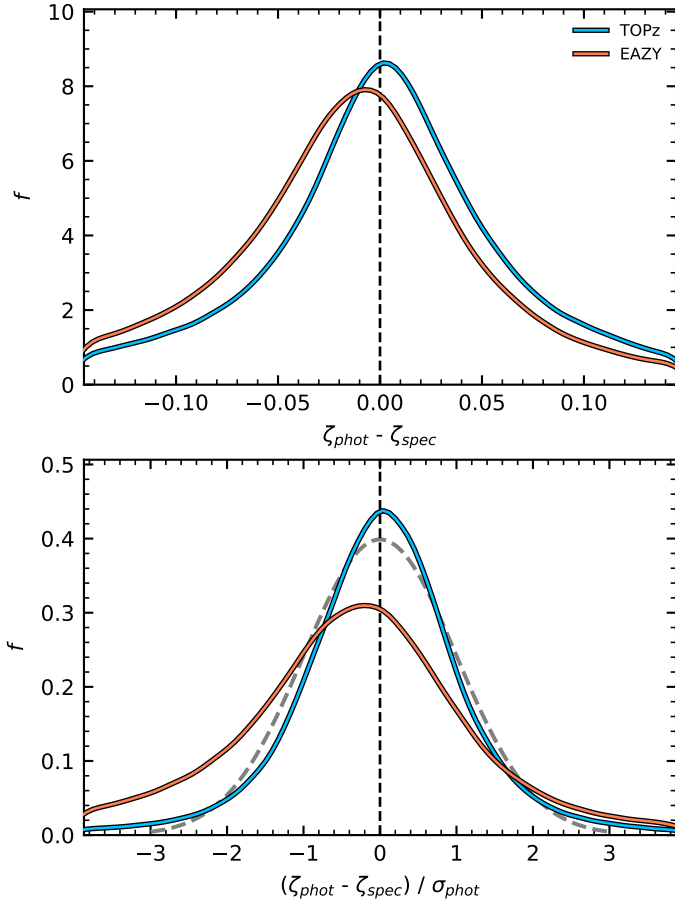


Fig. B.3. Photo-z accuracy for the TOPz and EAZY codes in the DESI sample. The layout is identical to Fig. 27.