The miniJPAS survey quasar selection – I. Mock catalogues for classification

Carolina Queiroz [iD],[1,2]* L. Raul Abramo,[2] Natália V. N. Rodrigues,[2] Ignasi Pérez-Ràfols [iD],[3,4]
Ginés Martínez-Solaeche [iD],[5] Antonio Hernán-Caballero,[6] Carlos Hernández-Monteagudo,[7,8]
Alejandro Lumbreras-Calle,[6] Matthew M. Pieri,[4] Sean S. Morrison [iD],[4,9] Silvia Bonoli,[10,11]
Jonás Chaves-Montero [iD],[10] Ana L. Chies-Santos [iD],[1,12] L. A. Díaz-García,[5] Alberto Fernandez-Soto,[13,14]
Rosa M. González Delgado,[5] Jailson Alcaniz,[15] Narciso Benítez,[5] A. Javier Cenarro,[16] Tamara Civera,[6]
Renato A. Dupke,[15,17,18] Alessandro Ederoclite,[6] Carlos López-Sanjuan,[16] Antonio Marín-Franch,[16]
Claudia Mendes de Oliveira,[19] Mariano Moles,[5,6] David Muniesa,[16] Laerte Sodré, Jr. [iD],[19] Keith Taylor,[20]
Jesús Varela[16] and Héctor Vázquez Ramió[16]

Affiliations are listed at the end of the paper

ABSTRACT

In this series of papers, we employ several machine learning (ML) methods to classify the point-like sources from the miniJPAS catalogue, and identify quasar candidates. Since no representative sample of spectroscopically confirmed sources exists at present to train these ML algorithms, we rely on mock catalogues. In this first paper, we develop a pipeline to compute synthetic photometry of quasars, galaxies, and stars using spectra of objects targeted as quasars in the *Sloan Digital Sky Survey*. To match the same depths and signal-to-noise ratio distributions in all bands expected for miniJPAS point sources in the range $17.5 \leq r < 24$, we augment our sample of available spectra by shifting the original r-band magnitude distributions towards the faint end, ensure that the relative incidence rates of the different objects are distributed according to their respective luminosity functions, and perform a thorough modelling of the noise distribution in each filter, by sampling the flux variance either from Gaussian realizations with given widths, or from combinations of Gaussian functions. Finally, we also add in the mocks the patterns of non-detections which are present in all real observations. Although the mock catalogues presented in this work are a first step towards simulated data sets that match the properties of the miniJPAS observations, these mocks can be adapted to serve the purposes of other photometric surveys.

Key words: methods: data analysis – techniques: photometric – catalogues – surveys – quasars: general.

1 INTRODUCTION

Ongoing and future photometric surveys will gather large data sets across vast volumes, shedding light on our current understanding about the formation and evolution of galaxies. Examples of such multiband surveys are the Dark Energy Survey (DES; The Dark Energy Survey Collaboration 2005), the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019), *Euclid* (Amendola et al. 2013), and the Javalambre-Physics of the Accelerated Universe Astrophysical Survey (J-PAS; Benitez et al. 2014). In this context, automated classification methods are essential tools to optimally catalogue all the observed sources, and the assembly of a reliable sample of photometrically selected quasars poses as a particularly challenging task.

Quasars are extremely luminous active galactic nuclei (AGNs), powered by accretion of matter on to a central supermassive black hole (Salpeter 1964; Zel'dovich & Novikov 1964). These astronomical objects are not only the brightest and one of the most highly biased tracers of large-scale structure (Porciani, Magliocchetti & Norberg 2004; Croom et al. 2005; Shen et al. 2007; da Ângela et al. 2008; Ross et al. 2009; Leistedt et al. 2013; Leistedt & Peiris 2014; Eftekharzadeh et al. 2015; Laurent et al. 2017), but they also share with their host galaxies mutual mechanisms of self-regulatory feedback processes that impact on the galaxy growth, making them a key ingredient in galaxy evolution models (Kauffmann & Haehnelt 2000; Di Matteo, Springel & Hernquist 2005; Schaye et al. 2015; Sijacki et al. 2015; Harrison 2017). Since they can be seen at large distances, quasars also work as 'lighthouses', serving as background light sources to map the intervening neutral hydrogen gas through the Gunn–Peterson effect (Gunn & Peterson 1965; Lynds 1971; Sargent et al. 1980), resulting in the so-called Lyman α forest.

Their UV–optical spectral energy distributions (SEDs) are characterized by a thermal component from the accretion disc emission in

---

* E-mail: c.queirozabs@gmail.com

the UV/optical, and a non-thermal continuum from the EUV to the X-rays, a series of broad and/or narrow emission lines[1] (e.g. Vanden Berk et al. 2001), blended iron lines (e.g. Vestergaard & Wilkes 2001; Véron-Cetty, Joly & Véron 2004), as well as a degree of dust reddening at times (e.g. Hopkins et al. 2004). In photometric images, quasar candidates typically appear as point-like sources and, thus, they can be easily confused with stars and even unresolved galaxies, especially in the low signal-to-noise ratio (S/N) regime. In nearby galaxies hosting low-luminosity AGNs, the signal from the host can also contaminate the (weaker) AGN emission. The contamination of quasars from intervening populations was first identified with the use of broad-band imaging to pre-select spectroscopic targets for the *Sloan Digital Sky Survey* (SDSS; Richards et al. 2009; Ross et al. 2012; Leistedt et al. 2013; Leistedt & Peiris 2014). Since fibers end up allocated only to the brightest, most clearly distinguished quasars, this pre-selection is unable to avoid contamination by other sources in colour–magnitude and colour–colour diagrams, leading to sub-optimal source classification, and a target success rate that changes across the survey footprint.

Fortunately, medium-to-narrow multiband photometric surveys that continuously cover a large wavelength range, such as ALHAMBRA (Moles et al. 2008), SHARDS (Pérez-González et al. 2013), PAUS (Martí et al. 2014), Subaru COSMOS 20 project (Taniguchi et al. 2015), J-PLUS (Cenarro et al. 2019), and S-PLUS (Mendes de Oliveira et al. 2019), can help to break degeneracies in the quasar identification by resolving some of the broad emission lines of type-I quasars (as well as most broad absorption line objects), and detecting some of the narrow-lines of type-IIs. In particular, Abramo et al. (2012) showed that J-PAS (Benitez et al. 2014) will observe nearly ∼240 quasars per square degree for a limiting magnitude of $g < 23$, which means that a JPAS-like survey of quasars could be the largest and most complete in the redshift range $0.5 \lesssim z \lesssim 4.0$. This creates a significant potential for probing cosmological phenomena, such as baryon acoustic oscillations and redshift space distortions. This in turn is potentially impactful for the study of dark energy models, modified gravity models (e.g. Aparicio Resco et al. 2020), as well as primordial non-Gaussianities and relativistic effects (e.g. Abramo & Bertacca 2017).

Prior to the arrival of the final scientific instrument (JPCam; Taylor et al. 2014; Marín-Franch et al. 2017), the J-PAS telescope (JST/T250) was equipped with a single CCD camera, called JPAS-Pathfinder, which carried out the first observations in nearly 1 deg^2 on the All-wavelength Extended Groth strip International Survey (AEGIS) field, and tested the performance of the J-PAS optical system. This science verification survey, dubbed miniJPAS (Bonoli et al. 2021), is a proof of concept for the forthcoming J-PAS survey, allowing us to test the precision with which J-PAS will be able to classify sources, and extract the photometric redshifts of galaxies and quasars.

In addition to the science that J-PAS will be able to conduct using the quasars it identifies, the collaboration is joining efforts with the WEAVE survey (Dalton 2016), a multi-object spectrograph that will start observing in 2022. Part of the WEAVE strategy is to follow-up high-redshift ($z \geq 2.1$) quasars to conduct a Lyman α forest and metal-line absorption survey (Pieri et al. 2016). The targets for this

WEAVE-QSO survey will be provided mainly by J-PAS, which is currently the only instrument capable of identifying quasars in numbers, and down to the depths needed by WEAVE-QSO to do its science.

Because of their high interest to both cosmology and galaxy evolution, the task of building a complete sample of quasar targets is more pressing than ever. Besides, in this new era of massive data acquisition we need effective statistical methods to classify the millions of sources that will be detected by these multiband photometric surveys, and e.g. identify the quasar candidates. This can be accomplished by using two approaches (or even a combination of both): template-fitting methods (e.g. Chaves-Montero et al. 2017) and machine learning algorithms (ML; e.g. Golob et al. 2021; Martínez-Solaeche et al. 2021; Nakazono et al. 2021).

Recently, ML approaches have become a powerful tool in astronomy, being preferred when dealing with massive data sets. However, this comes at the expense of requiring complete and representative training sets to ensure that the distribution of main features for each class of astronomical object will be reflected in the resulting trained models, and will be recovered with a good predictive performance on the test sets. For some ML applications in the context of source classification (see e.g. Odewahn et al. 1992, 2004; Fadely, Hogg & Willman 2012; Sevilla-Noarbe et al. 2018; Cabayol et al. 2019).

In particular, Baqui et al. (2021) employed several different ML methods to perform a star/galaxy separation in the miniJPAS catalogue using both photometric and morphological information, as well as spectroscopically confirmed miniJPAS sources for the training. To go beyond this morphological classification scheme requires a three-class separation (see e.g. Ball et al. 2006; Brescia, Cavuoti & Longo 2015; Clarke et al. 2020; Nakazono et al. 2021) to have a more thorough assessment of the contaminants in the miniJPAS quasar sample.

However, we lack a representative sample of spectroscopically confirmed sources in the area surveyed by miniJPAS. In particular, only a few hundred spectroscopic quasars were observed in that region. In fact, even if all the sources in miniJPAS had perfect types and redshifts, that still falls far of the numbers needed to train ML methods. Furthermore, it is not clear whether there will be, in the near future, a sufficiently large and sufficiently deep photometric catalogue of objects with secure (spectroscopic) classification. All this motivated us to develop realistic mock catalogues for use until such a time when both our photometric data and spectroscopic follow-up data reach sufficient size.

Generating synthetic fluxes allows us to asses some basic properties of data sets from upcoming surveys, such as selection effects, the uncertainties in derived galaxy properties, and the relative impact of the different sources of errors. They further allow us to make forecasts exploring survey strategies. Hence, our mock catalogues are crucial for assessing the quality of the miniJPAS classification, and for serving our purposes during the initial phases of the J-PAS survey.

In this paper, we describe the methodology adopted for generating simulated fluxes of quasars, galaxies, and stars which are based on the properties of the miniJPAS observations. These mock catalogues will be employed for training and validating the performances of several ML algorithms, classifying the miniJPAS point-like sources, and identifying quasar candidates. The ML codes employed in the classification, as well as their individual and combined performances (on test sets of both simulated fluxes and real observations), will appear in subsequent papers (Rodrigues et al., in preparation; Martínez-Solaeche et al., in preparation; Pérez-Ràfols et al., in preparation). These classifiers provide scores which can be used as probabilities that any given object is a quasar (at low $z < 2.1$

---

[1]In the unified model of AGNs (Antonucci 1993; Urry & Padovani 1995), each object receives a different nomenclature depending on the viewing angle to the centre of the source, and the presence of obscuring material. Throughout this paper, we shall use indistinctly 'quasars' to refer to type-I AGNs – unless otherwise specified.

or high $z \geq 2.1$ redshift), a star or a galaxy. In Pérez-Ràfols et al. (in preparation), we also present a primary catalogue of miniJPAS quasar candidates, which will be more thoroughly investigated with spectroscopic follow-up in the future.

This paper is organized as follows. Section 2 outlines the miniJPAS and *SDSS* data employed in this work, and describes the sample selection criteria. In Section 3, we present our pipeline for generating simulated photospectra of quasars, galaxies, and stars from *SDSS* spectra, and describe the luminosity functions and noise models. In Section 4, we validate the mock catalogues, and compare the main properties of the simulated fluxes with the miniJPAS observations. In Section 5, we suggest some improvements that could further optimize our mock catalogues, and provide additional applications. Finally, we summarize our main findings in Section 6. All magnitudes here are presented in the AB system. More technical information is also available as supplementary material at MNRAS online.

2 DATA PREPARATION

In this section, we describe the data sets used in this work, which consist in photometric observations from the miniJPAS catalogue (Bonoli et al. 2021), and spectra of quasar targets from the *SDSS* Superset catalogue (Pâris et al. 2017).

2.1 miniJPAS

The miniJPAS survey (Bonoli et al. 2021) imaged 0.895 deg^2 of the Extended Groth Strip (EGS) in four overlapping pointings. The observations[2] were conducted with the full J-PAS photometric system (see Fig. 1), which consists of 54 narrow-band (NB) filters ranging from 3780 to 9100 Å plus two medium-band filters centred on 3497 Å (named *uJAVA*) and 9316 Å (named *J*1007), and complemented with three *SDSS*-like broad-bands (*uJPAS*, *gSDSS*, and *rSDSS*). The 54 NB filters present full widths at half maximum (FWHM) of ∼145 Å and are equally spaced every ∼100 Å, whereas the FWHM of the *uJAVA* and *J*1007 bands are 495 and 635 Å, respectively. In addition to these filters, the miniJPAS observations also included the *iSDSS* broad-band (in a total of 60 filters).

The miniJPAS data were calibrated and reduced by the Data Processing and Archiving Unit at CEFCA (Cristóbal-Hornillos et al. 2014). In our analyses, we included only data from the primary catalogue (PDR201912), which contains 64 293 sources with detection in the *r* band. The photometry for all sources in this catalogue was obtained with `sextractor` (Bertin & Arnouts 1996) in the dual-mode configuration, and the different types of apertures were defined using the *r* band as the reference filter.

The dual-mode catalogue provides different types of photometries (in units of magnitudes and fluxes), both total magnitudes and magnitudes in apertures of different sizes, corrected for atmospheric extinction. For the purposes of this work, we employ `APER_3` magnitudes, which are further corrected for Galactic extinction and aperture corrections (as explained in Section 2.3.3). For further details about the observations and data reduction (see Bonoli et al. 2021).
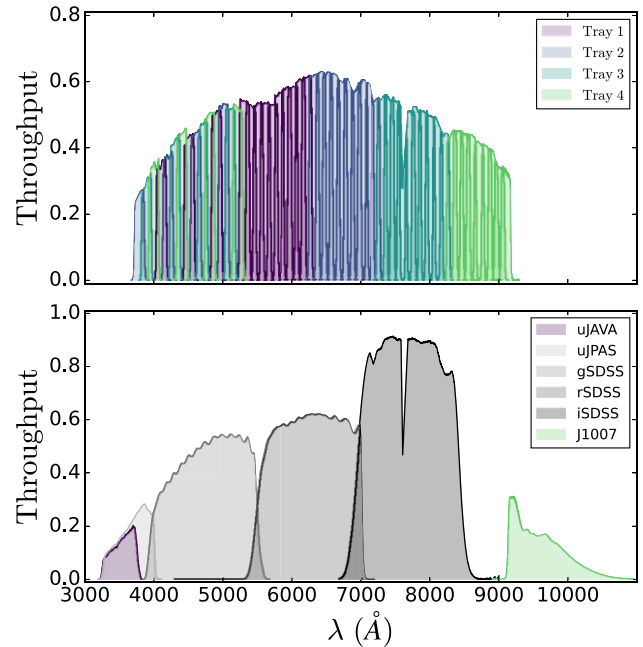

Figure 1. Throughputs of the photometric system employed in the miniJPAS observations. Upper panel: 54 narrow-bands coloured by their distribution in the filter trays. Lower panel: Medium and broad-bands. The throughputs include effects from the CCD quantum efficiency, the entire optical system of the telescope and sky absorption.

2.2 *SDSS* superset

The first step to create the mocks was to assemble a large sample of objects with reliable classification, redshifts, and optical spectra, that covered most of the J-PAS wavelength range. Fortunately, the miniJPAS area was observed by a wealth of multiwavelength facilities, such as AEGIS (Davis et al. 2007), ALHAMBRA (Moles et al. 2008), DEEP2/DEEP3 (Cooper et al. 2011, 2012; Newman et al. 2013), *SDSS* (Dawson et al. 2013), and HSC–SSP (Aihara et al. 2018, 2019). We opted to construct the synthetic photospectra from the publicly available data set *SDSS* DR12Q Superset[3] (Pâris et al. 2017) which contains all quasar targets from the final data release of the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013), as part of the *SDSS*-III Collaboration (Eisenstein et al. 2011).

This superset of visually inspected spectra and redshifts provides a census of not only quasars, but also stars and galaxies whose broad-band colours are consistent with those of quasars. Therefore, the combination of visual inspection plus highly specialized targeting algorithms makes the *SDSS* Superset catalogue ideal to select not only spectroscopically confirmed quasars, but also the main contaminants in the quasar sample. In particular, the Superset catalogue constitutes a reliable starting point for generating the mocks.

Although here we limit ourselves to *SDSS* DR12 (Dawson et al. 2013) targets, because sample veracity is of critical importance for the classification of miniJPAS sources, and *SDSS* DR12Q spectra have all been visually inspected, the spectra and some quality metrics employed in this work come from *SDSS* DR16 (Ahumada et al. 2020; Lyke et al. 2020) with its pipeline refinements such as improved

---

[2]All miniJPAS images and catalogues are publicly available through the CEFCA web portal: http://archive.cefca.es/catalogues/minijpas-pdr201912.

[3]The *SDSS* DR12Q Superset catalogue is available at:
https://data.sdss.org/sas/dr12/boss/qso/DR12Q/Superset_DR12Q.fits

Table 1. Total number of sources in the spectroscopic sample and miniJPAS cross-matched samples after applying some quality cuts.

	Superset spectra	miniJPAS-Superset cross-match	miniJPAS-DEEP3 cross-match
QSO	281 208	117	15
Galaxy	20 021	40	8 779
Star	131 430	115	37

spectrophotometric calibration for BOSS quasars (see e.g. Margala et al. 2016 for further details).

2.3 Sample selection

In this section, we describe the selection of the data sets utilized in this work. In Sections 2.3.1 and 2.3.2, we present the quality cuts applied to the *SDSS* Superset catalogue and the procedure adopted to augment our sample of spectra by generating fainter sources from the original ones, respectively. In Section 2.3.3, we describe the quality cuts and aperture corrections applied to miniJPAS photometry. Finally, in Sections 2.3.4 and 2.3.5 we characterize the miniJPAS spectroscopic and point-like samples, respectively. We summarize the total number of quasars, galaxies, and stars in each sample in Table 1.

2.3.1 Superset original sample

The *SDSS* DR12Q Superset contains 546 856 quasar targets. For the purposes of this work, we selected spectra that satisfied the following criteria: zWARNING = 0 (good-quality spectra); SN_MEDIAN_ALL > 0 (further quality information from the median signal-to-noise ratio per resolution element); Z_CONF < 3 (large confidence rating for the visually inspected redshift); CLASS_PERSON = 1, 3, 30, and 4 (object classification via visual inspection as star, quasar, broad-absorption line quasar, and galaxy, respectively), and apparent magnitudes in the range $17.5 \leq r < 24$. Note, however, that the spectra available for galaxies are limited to the range $18.7 \leq r < 24$. This results in a sample of 281 208 quasars at $z < 4.3$, 20 021 galaxies at $z < 0.9$, and 131 430 stars – including main-sequence stars, white dwarfs (WD), carbon stars (C), and cataclysmic variables (CV).

From Data Release 2 and beyond, the *SDSS* final calibrated spectra are not corrected for Galactic dust reddening. We corrected for Galactic extinction following the same prescription as *SDSS* by using the conversions from $E(B–V)$ to total extinctions as tabulated in table 6 of Schlafly & Finkbeiner (2011). We then adopted the extinction law from Fitzpatrick & Massa (2007) implemented in the extinction PYTHON module.[4]

2.3.2 Superset faint sample

The *SDSS*-III/BOSS survey selected quasar candidates by various selection algorithms – for a detailed description of the BOSS quasar target selection (see Ross et al. 2012). This selection is designed to be sensitive to quasars in the range $2.15 < z < 3.5$, and is limited to $r \lesssim 21.85$. To construct a fair sample of simulated fluxes that resembles the luminosity properties from the miniJPAS observations and reaches $r \sim 24$, we need thus a parent sample of spectra which includes, on average, many more faint sources than the original Superset catalogue does.
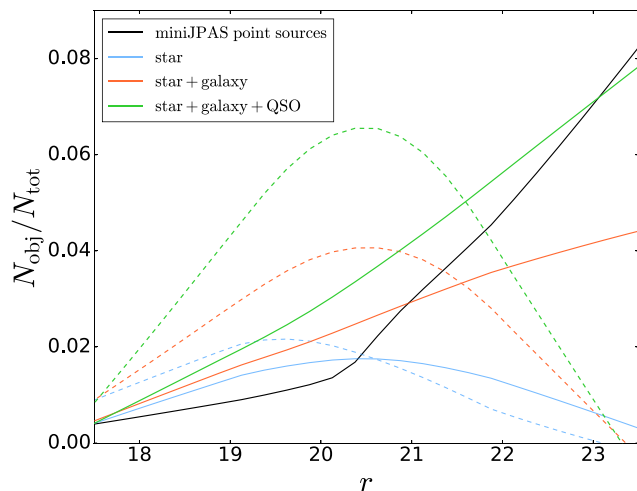
[4]https://github.com/kbarbary/extinction.

Figure 2. Original (dashed lines) and shifted (solid coloured lines) *SDSS* r-band magnitude distributions of a sample containing equal numbers of stars, galaxies, and quasars (10^4 of each). As a comparison, the black solid line shows the magnitude distribution of miniJPAS point-like sources. The y-axis corresponds to the number of objects in each magnitude bin divided by the total number of objects in the corresponding sample. The total number of objects are 10k, 20k, 30k, and 11 419 for the green, orange, blue, and black curves, respectively.

Following Abramo et al. (2012), our procedure consists in generating new fainter objects from each original spectrum. This means that we fix the spectroscopic identification (spectrum, redshift and/or type), and assign new (randomly generated) fainter r-band magnitude values for each 'new object'. Starting from a given sample of Superset spectra from Table 1, we divide them into r-band magnitude and redshift (stellar type) bins, where we consider bins of size 0.5 mag and 0.2 (0.1) in redshift for quasars (galaxies). For each object in a given bin, we generate N new objects in all magnitude bins that are fainter than the original one. The value of N is mainly informed by the luminosity function for each source and the total number of synthetic photospectra that we want to generate. In our analyses N $\sim$ 20 for quasars and galaxies, while for stars this number can vary from 40 to 120 new objects per bin (depending on the spectral type).

In Fig. 2, we illustrate how the original *SDSS* r-band magnitude distributions are shifted to fainter magnitudes (by fixing their redshifts and stellar types) in a sample containing equal numbers of stars, galaxies, and quasars (10^4 of each). As we can see, the assigned fainter magnitudes in the augmented sample are more representative of the magnitude distribution of miniJPAS point-like sources (see Section 2.3.5). A more technical explanation about the magnitude shifts is available as supplementary material of MNRAS.

Finally, we map this augmented set of spectra into the number counts specified by the corresponding luminosity functions (described in Section 3.2). The objects selected in this way constitute our final sample of parent spectra.

In the procedure outlined above we assume a weak spectral dependence on luminosity. This might not always be true in the case of AGNs due to the Baldwin effect (Baldwin 1977), the anticorrelation between the continuum luminosity and the rest-frame equivalent widths of UV emission lines (such as Ly α and C IV). This assumption of no luminosity evolution might not be very realistic for non-active galaxies in general, given that at the same redshift, fainter galaxies tend to be less massive; and thus
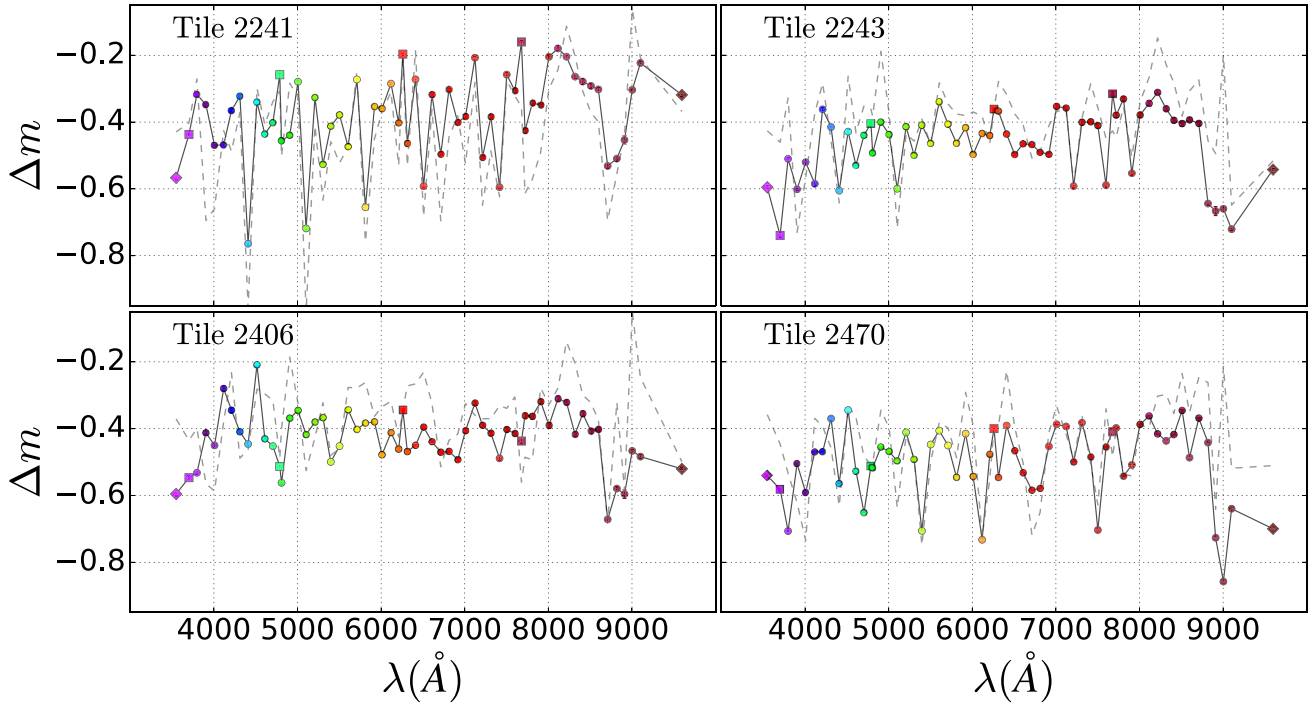
Figure 3. Comparison between the aperture corrections alone (dashed lines) and the total magnitude offsets (coloured symbols) computed for `APER_3` magnitudes as a function of the tile. Narrow, intermediate, and broad-bands are represented by circles, diamonds, and squares, respectively.

younger, bluer, and with stronger emission lines. This effect is partially mitigated by the fact that our selection of galaxy spectra was obtained independently of the spectral type (see Section 3.2.2), i.e. without explicitly considering the relative frequencies of red and blue galaxies separately. Therefore, although the current version of the galaxy mocks might not be the most appropriate one for e.g. galaxy evolution studies, for the moment it corresponds to the best available sample of main contaminants to the miniJPAS quasar population.

2.3.3 miniJPAS sample

The miniJPAS catalogue contains some quality cut flags warning whether each source has an issue in one or more filters that may impair or invalidate the photometry. The `FLAGS` column comprises the SExtractor flags and indicates close neighbours, blending, saturation, truncation, and so on. The `MASK_FLAGS`, in turn, informs if the object is outside the window frame, whether it is a bright star or is located near one, and if it has a nearby artefact. To ensure good photometry, we only selected miniJPAS sources with `FLAGS = 0` and `MASK_FLAGS = 0` in all bands. This so-called non-flagged subsample contains 46 440 sources with measured photometry in 60 or less bands.

As part of the calibration process (López-Sanjuan et al. 2019b), the miniJPAS photometry is already corrected for atmospheric extinction. Hence, we only need to correct the miniJPAS data for Galactic extinction, where the extinction coefficients per passband per tile are obtained based on the procedure outlined in López-Sanjuan et al. (2019b) and provided in the `miniJPAS.MWExtinction` table.

The miniJPAS catalogue provides different types of photometric measurements in all bands. Since our ultimate goal is to classify point-like sources, in our analyses we employ `APER_3` magnitudes (i.e. apparent magnitudes computed within a 3 arcsec-diameter aper-

ture). This choice is made to guarantee high-accuracy photometric redshifts for quasars, and will be discussed in more detail in an upcoming paper (Queiroz et al., in preparation).

Given that a 3 arcsec-aperture misses part of the total light emitted by the source, an aperture correction $\Delta m_{\mu(a)}^{3''}$ is applied to these magnitudes. To derive the aperture corrections per passband per tile, we use the non-variable, non-saturated stars from the miniJPAS-Superset cross-matched sample. These coefficients are computed based on the `APER_6` corrections ($\Delta m_{\mu(a)}^{6''}$) available in the `miniJPAS.TileImage` table, as follows:

$$\Delta m_{\mu(a)}^{3''} = \text{median} \left[m_{\mu(i,a)}^{6''} + \Delta m_{\mu(a)}^{6''} - m_{\mu(i,a)}^{3''} \right], \qquad (1)$$

where $m_{\mu(i,a)}$ is the observed aperture magnitude of the ith star measured in band μ in tile a. These aperture corrections are shown as dashed lines in Fig. 3.

After applying the aperture corrections to the `APER_3` magnitudes, we need to recalibrate the photometry. For this, we compute the offsets $\delta m_{\mu(a)}$ between the corrected magnitudes and the synthetic magnitudes obtained using the prescription from Section 3.1:

$$\delta m_{\mu(a)} = \text{median} \left[m_{\mu(i,a)}^{3''} - m_{\mu(i,a)}^{synt} \right], \qquad (2)$$

where $m_{\mu(a)}^{synt}$ is a synthetic magnitude in band μ for the ith star.

The total magnitude offsets per passband per tile are then given by

$$\Delta m_{\mu(a)} = \Delta m_{\mu(a)}^{3''} + \delta m_{\mu(a)}, \qquad (3)$$

and are shown as coloured symbols in Fig. 3. As we can see, these offsets are non-negligible, and can be as large as 0.8 mag in absolute value for some filters. These offsets are already internally available to the J-PAS collaboration, and will be publicly released together with the mock catalogues (as a separate table) upon publication.

The fact that the miniJPAS observations were performed in groups of seven filters, and carried out with different sky conditions reflects in different luminosity properties for each tile. Moreover, the reddest filters were observed last, when the AEGIS field reached the lowest elevations (i.e. highest airmass measurements). This means that the non-flagged sample contains only 2423 sources with observations in all 60 bands. In other words: the majority of the miniJPAS sources have at least one non-detection (ND) in one of the bands, a feature that needs to be incorporated into the simulated fluxes in order to build realistic mocks.

In the miniJPAS catalogue, non-detections[5] are assigned with magnitude values of 99.0: either (i) they correspond to magnitudes fainter than the limiting sensibility of the detector for that specific band (i.e. they have low signal-to-noise ratios), or (ii) they are related to negative fluxes. We opted to treat these two instances of NDs separately by adopting the following conventions for the magnitudes and uncertainties:

(i) if S/N < 1.25: $[m_{\mu(i)}, \sigma_{m, \mu(i)}] = (99.0, m_{\mu, \lim}^{3'', 5\sigma})$, where S/N is the signal-to-noise ratio in band μ for the ith source, and $m_{\mu, \lim}^{3'', 5\sigma}$ is the targeted minimum depth defined in Benitez et al. (2014);

(ii) if $F_{\lambda, \mu(i)} < 0$: $[m_{\mu(i)}, \sigma_{m, \mu(i)}] = (-99.0, 99.0)$, where $F_{\lambda, \mu(i)}$ is the APER_3 flux in units erg s^{-1} cm^{-2} Å^{-1}.

2.3.4 miniJPAS spectroscopic sample

After re-processing and recalibrating the data, we built a miniJPAS spectroscopic sample by cross-matching the non-flagged sample with the Superset sample within a radius of 1 arcsec using the TOPCAT software (Taylor 2005). This cross-match resulted in a set of 117 quasars, 40 galaxies, and 115 stars.

We also performed the cross-match with DEEP3 (Cooper et al. 2011, 2012), a dedicated spectroscopic campaign focused on the Extended Groth Strip, within a radius of 1.5 arcsec. The DEEP3 survey does not cover the whole optical wavelength range (spanning only 4550–9900 Å), and was designed to map galaxies down to a limiting magnitude of $R \sim 24.4$. This means that more noisy spectra cannot be reliably visually inspected. In particular, many of the sources classified as AGNs by DEEP3 at low redshifts seem to be actually galaxies with a very small AGN component, low-luminosity type-Is, and even type-IIs. So in the case of quasars we only selected DEEP3 sources that are identified as AGNs at redshifts $z \geq 1.5$. In addition, we also applied the following selection criteria: ZQUALITY $\geq$3 for quasars and galaxies, and ZQUALITY $=-1$ for stars (an indicator of redshift quality); RCHI2 $\geq$0.6 (reduced χ^2 square for the redshift fit); PGAL $\geq$0.6 for galaxies and PGAL <0.5 for stars (a value between 0 and 1 indicates the probability of a source being a galaxy, for unresolved sources; while a value equal to 3 indicates a resolved galaxy); r-band magnitude in the range 17.5 $\leq r < 24.0$. The final miniJPAS-DEEP3 sample contains 15 quasars at $1.5 < z < 3.7$, 8779 galaxies at $z \leq 1.7$ (6514 at $z < 0.9$) and 37 stars.

2.3.5 miniJPAS point-like sample

The miniJPAS data base provides different complementary methods to estimate the stellarity index of each source (see Bonoli et al. 2021

for more details). In these classifications an index close to one (zero) indicates that the source is likely a star (galaxy).

In this work, we employ the Extremely Randomized Trees (ERT) machine learning classifier (Baqui et al. 2021), which uses both morphological and photometric information. To build our set of miniJPAS point-like sources, we selected objects from the non-flagged sample that were classified as stars with a probability of $\mathcal{P}_{\mathrm{ERT}} \geq 0.1$. This quality cut can properly separate extended and point sources up to $r \sim 22$. In order to maximize the selection of point-like sources, whenever $\mathcal{P}_{\mathrm{ERT}} = -99.0$ we also considered objects classified as stars by the stellar-galaxy locus classifier (SGLC; López-Sanjuan et al. 2019a) with a probability of $\mathcal{P}_{\mathrm{SGLC}} \geq 0.1$. The final miniJPAS point-like subsample contains 11 419 objects, which is used as a proxy to both select realistic S/N distributions in each band and draw the pattern of non-detections to be applied to the mocks.

Finally, we built three additional subsamples containing about 10k point sources each, that are randomly selected according to the targeted number counts provided by the luminosity functions of quasars, galaxies, and stars (see Section 3.2 for more details). These subsamples are designed to provide fairer comparisons with the corresponding test sets.

3 MOCK CATALOGUES

Machine learning techniques comprehend an ensemble of adaptive learning methods to recognize and predict some sort of pattern within a specific class of objects. On one hand, such approach has the advantage of not making any prior assumptions regarding the spectral types or their evolution; on the other, it requires representative training sets to estimate the learning model and achieve high classification performances.

To classify the miniJPAS point sources, and identify the quasar candidates, we need a large and sufficiently complete sample of spectroscopically confirmed objects with miniJPAS observations. However, in the AEGIS field we fall far short of a sufficiently large sample, and the available one is neither fair nor complete to the depth we require. Consequently, the AEGIS field cannot be used alone to properly train the ML algorithms.

To bypass this fundamental limitation (which is shared with many forthcoming photometric surveys), we generated mock catalogues of simulated photospectra of quasar targets, and developed a dedicated pipeline to include realistic features from the observations, such as the S/N in all bands, the magnitude-redshift-type distributions drawn from putative luminosity functions, as well as the pattern of non-detections. Ensuring the accuracy of these simulated photospectra is key for the generalization of the models trained on mocks to real data. For this reason, this constitutes a constant work-in-progress, in the sense that, as we acquire more information from observations over larger areas, with larger samples of spectroscopically confirmed sources, we will be able to further refine the mocks.

In Fig. 4, we present the methodology developed to build the mock catalogues. Each block contains the reference to the location in the text. Throughout the following sections we describe in more detail each of the subsequent steps of the pipeline. In Section 3.1, we describe our prescription for generating synthetic fluxes from Superset spectra, which are selected according to Section 2.3.2 and using the luminosity functions summarized in Section 3.2. Then, in Section 3.3 we discuss how to add noise to the synthetic fluxes and present our noise models. Finally, in Section 3.4 we briefly review how to add the non-detection patterns. The final mock catalogues are characterized in Section 3.5. Although explicitly applied for the

---

[5]Non-observations would be signaled with a negative value of the normalized weight map flag, but we have confirmed that they are not present in the miniJPAS catalogue.
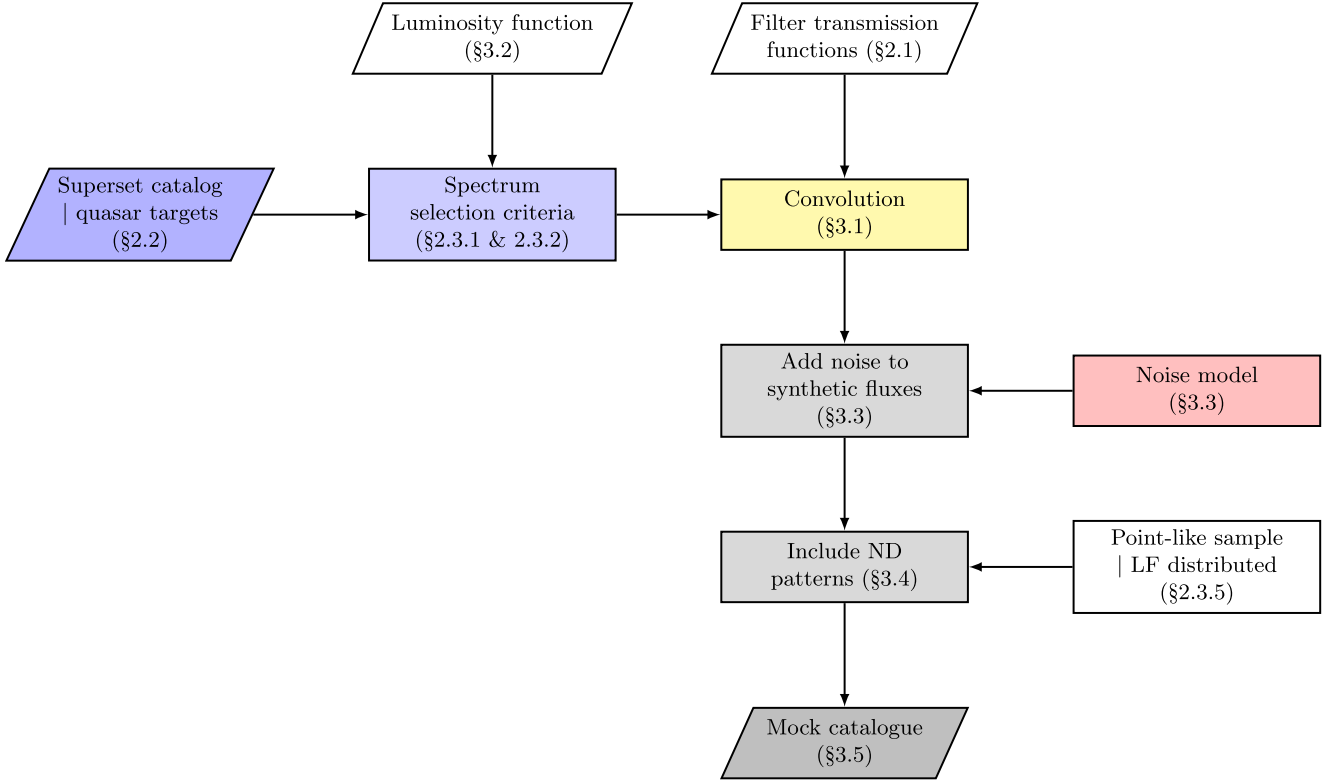
Figure 4. Pipeline to simulate photospectra with the same signal-to-noise ratio distributions expected for observations within a given photometric system.

J-PAS photometric system here, this procedure can be easily adapted for any other multiband optical survey.

3.1 Synthetic photometry

The synthetic photometry is computed using the prescription described in Díaz-García et al. (2015), based on the *HST* synphot[6] package, and in Bessell (2005); Pickles & Depagne (2010). Briefly, for modern photon-counting devices the synthetic fluxes $F_{\lambda,\mu(i)}^{synt}$ (in unit wavelength) can be obtained by

$$F_{\lambda,\mu(i)}^{synt} = \frac{\int T_\mu(\lambda)\, S_i(\lambda)\, \lambda\, d\lambda}{\int T_\mu(\lambda)\, \lambda\, d\lambda}\,, \tag{4}$$

where $S_i(\lambda)$ is the dereddened spectrum of the *i*th source, and $T_\mu(\lambda)$ is the total efficiency of the transmission curve of passband μ. We also scaled the *SDSS* spectrum to match the magnitude in the *r* band, which may correspond either to the original value or to a new randomly assigned fainter one.

The corresponding uncertainties to the synthetic fluxes are directly selected from the miniJPAS observations (point-like sample). Note that the *SDSS* spectra also have uncertainties associated with them. However, our estimates indicate that for *SDSS* the value of $<S/N>$ per pixel scales as $1/\sqrt{N_{bin}}$, with $N_{bin} \sim 145/0.87$, where N_{bin} corresponds to the ratio between the number of spectral bins necessary to compose the flux in a given narrow band and the size of each spectral bin. This implies that the *SDSS* median S/N is negligible when compared to the median S/N from the synthetic fluxes (when we consider the noise coming from the miniJPAS

observations), and hence the *SDSS* uncertainties are neglected in equation (4).

The magnitudes in the STMAG system (defined such that a source with constant flux per unit wavelength has zero colour) are computed as

$$m_{\text{ST},\mu(i)}^{synt} = -2.5 \log_{10} F_{\lambda,\mu(i)}^{synt} - 21.1\,, \tag{5}$$

and AB magnitudes can then be obtained from

$$m_{\text{AB},\mu(i)}^{synt} = m_{\text{ST},\mu(i)}^{synt} - 5 \log_{10} \lambda_{\text{pivot},\mu} + 18.692\,, \tag{6}$$

where the pivot wavelength is defined as

$$\lambda_{\text{pivot},\mu} = \sqrt{\frac{\int T_\mu(\lambda)\, \lambda\, d\lambda}{\int T_\mu(\lambda)\, d\lambda/\lambda}}\,. \tag{7}$$

Hereafter, we shall use indistinctly $m_{\mu(i)}^{synt}$ to refer to (synthetic) AB magnitudes. Finally, we assume that the synthetic photometry is centred at the corresponding effective wavelength, given by

$$\lambda_{\text{eff},\mu} = \frac{\int T_\mu(\lambda)\, \lambda\, d\lambda}{\int T_\mu(\lambda)\, d\lambda}\,. \tag{8}$$

To reach signal-to-noise ratio distributions in the mocks that resemble those of miniJPAS point sources, we still need to add a certain level of noise to the synthetic fluxes. This procedure is described in Section 3.3.

Another caveat concerns the wavelength range spanned by the *SDSS* spectra, which do not fully cover the *uJAVA* and *J1007* passbands. Since this could preclude us from including those bands in the mocks, prior to the convolution shown in equation (4) we extend the spectrum coverage by performing a template fitting. Our method consists in fitting the blue and the red parts of each spectrum

separately to ensure a smoother transition when concatenating the fluxes of the two best-fitting templates with the spectrum.

The best-fitting template is chosen by minimizing the following function:

$$\chi^2_{k(i)} = \sum_\mu \left[\frac{F^{synt}_{\lambda, \mu(i)} - \mathcal{T}_{\lambda, \mu(k)}}{\sigma^{sdss}_{F, \mu(i)}} \right]^2, \tag{9}$$

where $\mathcal{T}_{\lambda, \mu(k)}$ is the synthetic flux of the kth template scaled by the first blue (or red) filter with a valid observation, and $\sigma^{sdss}_{F, \mu(i)}$ is the uncertainty[7] obtained by error propagation of equation (4).

In the case of galaxies and stars, the best-fitting templates are chosen from a library of SED models available with the code LePhare (Arnouts et al. 1999; Ilbert et al. 2006). We use 37 (154) galaxy (stellar) templates from the COSMOS (Pickles) libraries. For galaxies, the flux densities of the templates are computed at the spectroscopic redshift. In the case of quasars, we adopt a single Vanden Berk composite spectrum (Vanden Berk et al. 2001) to which we add an adjustable amount of extinction, following the prescription of Hernán-Caballero et al. (2016).

3.2 Luminosity function

Our mock catalogues are built based on balanced training, validation and test sets containing representative relative frequencies of quasars, galaxies, and stars, drawn from putative luminosity functions. For the extragalactic sources, we assume that these number densities depend both on redshift and magnitude, without making any further assumptions on the frequencies of their sub-types. In the case of stars, besides the dependence on magnitude, their number densities also have an angular dependence. In addition, since some stellar types, such as A, F, M, and white dwarfs, are more often confused with quasars, either due to their similar colours (Richards et al. 2002) or because their continuum emission can be confused with the Lyman-break of high-redshift quasars, we also considered a dependency on the stellar spectral types.

Since the effective area surveyed by miniJPAS is small (0.895 deg^2), to have representative samples to train and assess the performance of the ML classifiers, we simulated multiple realizations of the relative frequencies expected in $\sim$1 deg^2 until we obtained the desired size for the test, validation and training sets. For instance, the training sets (which contain 100k objects each) correspond to realizations inside final areas of approximately 196, 16, and 46 deg^2 for quasars, galaxies, and stars, respectively.

Throughout this paper, we shall use indistinctly luminosity function (LF) to refer to the number counts in luminosity and type (or redshift, if applicable) for each class of object. The corresponding LFs are described in the following sections.

3.2.1 Quasar luminosity function

For quasars, we adopt the pure luminosity evolution (QLF, hereafter) function from Palanque-Delabrouille et al. (2016), which assumes that the luminosity of all quasars scales up according to some function of redshift, and it allows the bright-end and faint-end slopes to be different on either side of a pivot redshift ($z_{pivot} = 2.2$). Considering a perfect selection of objects, we find that over an area of 1/5 of the full sky (similar to what is planned for the entire J-PAS footprint),

---

[7]Note that this is the only step where we have explicitly made use of the flux uncertainties coming from the spectral bins.
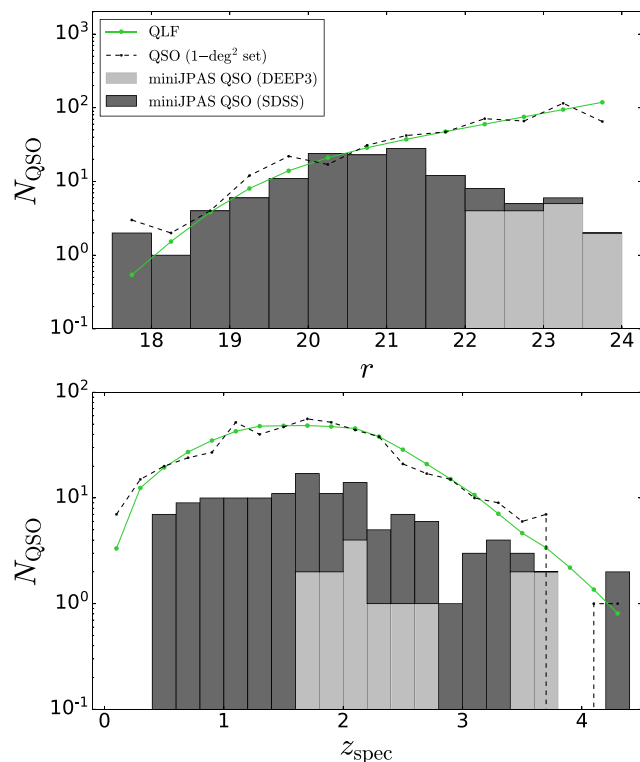

Figure 5. Number of quasars per deg^2 from the luminosity function (green solid line) and miniJPAS (grey bars) as a function of the r-band magnitude (top) and spectroscopic redshift (bottom). The miniJPAS quasars from DEEP3 (*SDSS*) are shown in light (dark) grey. As a comparison, we also show the distribution of quasars in the 1-deg^2 set (black dashed line) using noise model 11 as reference.

a flux-limited ($r < 23.5$) survey could yield more than three million quasars up to $z = 6$, which is in accordance with the estimates from Abramo et al. (2012).

To generate the mocks, we selected quasar spectra in the redshift range $0.033 \leq z \leq 4.3$ and with magnitudes between $17.5 \leq r < 24$. In Fig. 5, we show the magnitude–redshift distribution per deg^2 predicted by the QLF, and compare it with the distribution of the miniJPAS quasars, separating the contributions from the cross-matches with DEEP3 and *SDSS* Superset. The QLF predicts 510 quasars per deg^2, being 133 at high redshifts ($z \geq 2.1$). Although DEEP3 quasars are complementary to the *SDSS* sample in the faint end, we can still clearly see that the sample of spectroscopically confirmed miniJPAS quasars becomes highly incomplete at $r \gtrsim 21.5$.

In Fig. 5, we also provide the distributions of quasars in the 1-deg^2 set using noise model 11 as reference (see Section 3.3 for more details). The absence of quasars at $3.6 < z < 4.0$ is attributed to cosmic variance.

3.2.2 Galaxy luminosity function

For galaxies, instead of using a phenomenological luminosity function, we mapped the magnitude–redshift distributions of the miniJPAS spectroscopic galaxies. To ensure a fair distribution of galaxies in the bright and faint ends, we combined the contributions from both Superset and DEEP3 galaxies observed with miniJPAS.

To generate the mock catalogues, we considered spectra of galaxies at $z \leq 0.9$, and with magnitudes between $18.7 \leq r < 24$. Note that these magnitude–redshift ranges are limited by the availability of
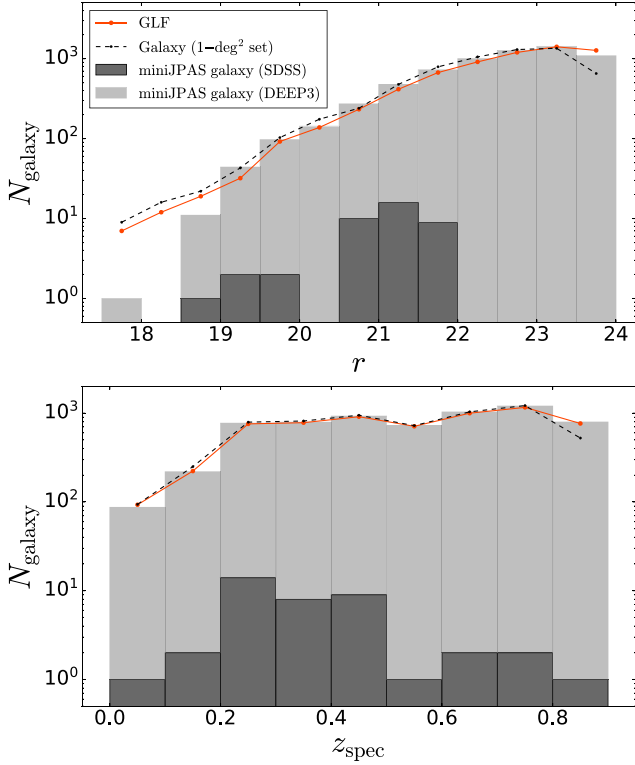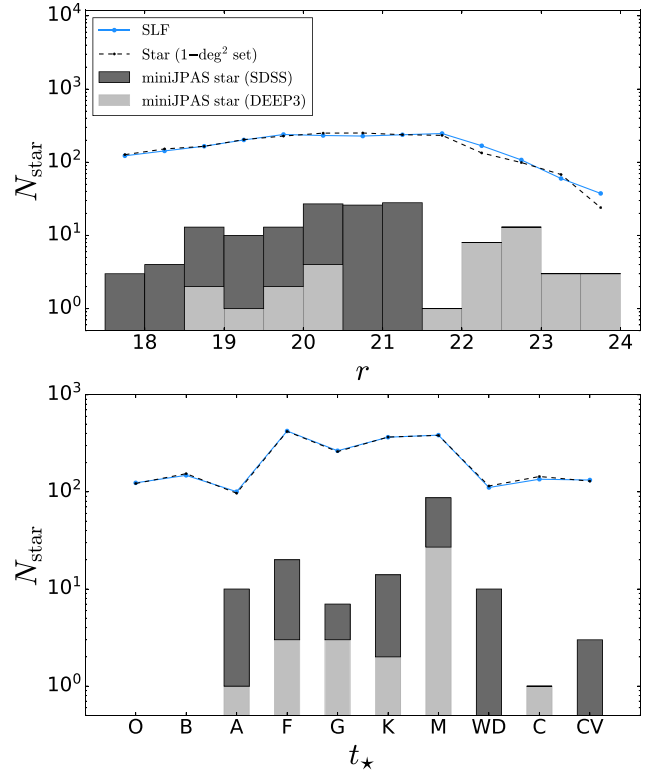
Figure 6. Number of galaxies per deg^2 from the luminosity function (orange solid line) and miniJPAS (grey bars) as a function of the r-band magnitude (top) and spectroscopic redshift (bottom). The miniJPAS galaxies from DEEP3 (*SDSS*) are shown in light (dark) grey. As a comparison, we also show the distribution of galaxies in the 1-deg^2 set (black dashed line) using noise model 11 as reference.

spectra in the Superset catalogue, and a more detailed separation in the number counts of blue and red galaxies is beyond the scope of this paper. Such galaxy luminosity function (GLF) predicts 6 410 galaxies per deg^2. In Fig. 6, we show the magnitude–redshift distributions predicted by the GLF, and compare them with the distributions of the miniJPAS galaxies. As we can see, the GLF is dominated by galaxies from DEEP3. We also compare the distributions of galaxies in the 1-deg^2 set using noise model 11 as reference.

3.2.3 Stellar luminosity function

Although the number and types of galaxies are more or less uniformly distributed across the sky, this is not true for stars: their number densities and spectral types are highly dependent on the line of sight that we are probing throughout the Milky Way.

To take this effect into account, we made use of the Besançon model of stellar population synthesis of the Galaxy (Robin et al. 2003) to compute the stellar counts per spectral type per magnitude bin in the same angular position of the miniJPAS area. We considered main-sequence stars, white dwarfs, carbon stars, and cataclysmic variables in the magnitude range $17.5 \leq r < 24$. These number counts are complemented by the number densities of miniJPAS spectroscopic stars, and we also extrapolate the relative frequencies of the following types: O, B, A, WD, C, and CV.

Such star luminosity function (SLF) predicts 2190 stars per deg^2 in the AEGIS field. In Fig. 7, we show the magnitude-type distribution of stars predicted by the SLF, and compare it with the distribution in miniJPAS and in the 1-deg^2 set. As we can see, the spectroscopic

Figure 7. Number of stars per deg^2 from the luminosity function (blue solid line) and miniJPAS (grey bars) as a function of the r-band magnitude (top) and spectral type (bottom). The miniJPAS stars from DEEP3 (*SDSS*) are shown in light (dark) grey. As a comparison, we also show the distribution of stars in the 1-deg^2 set (black dashed line) using model 11 as reference.

Table 2. Summary of the noise models tested to properly model the miniJPAS observations. Red bands are defined such that $\lambda_{\rm eff} \geq 7\,416$ Å.

Model	Description
1	$G(0, 1\sigma_\mu)$
2	$G(0, 1.5\sigma_\mu)$
3	$G(0, 2\sigma_\mu)$
4	$G(0, 2.5\sigma_\mu)$
5	$G(0, 3\sigma_\mu)$
6	$\frac{2}{3}G(0, 1\sigma_\mu) + \frac{1}{3}G(0, 2\sigma_\mu)$
7	$\frac{1}{3}G(0, 1\sigma_\mu) + \frac{1}{3}G(0, 2\sigma_\mu) + \frac{1}{3}G(0, 3\sigma_\mu)$
8	$\frac{2}{3}G(0, 1\sigma_\mu) + \frac{1}{3}G(0, 3\sigma_\mu)$
9	$\frac{2}{3}G(0, 2\sigma_\mu) + \frac{1}{3}G(0, 3\sigma_\mu)$
10	$G(0, 1\sigma_\mu)$ [blue bands]
	$G(0, 2\sigma_\mu)$ [red bands]
11	best of above (for each band)

sample of miniJPAS stars is highly incomplete in comparison with the distributions expected from SLF.

3.3 Noise model

To obtain similar S/N distributions in all bands as the ones found for miniJPAS point sources, we performed a thorough modeling of the noise profiles of the observations, and included realistic errors into the synthetic fluxes derived in equation (4), as described in the following.
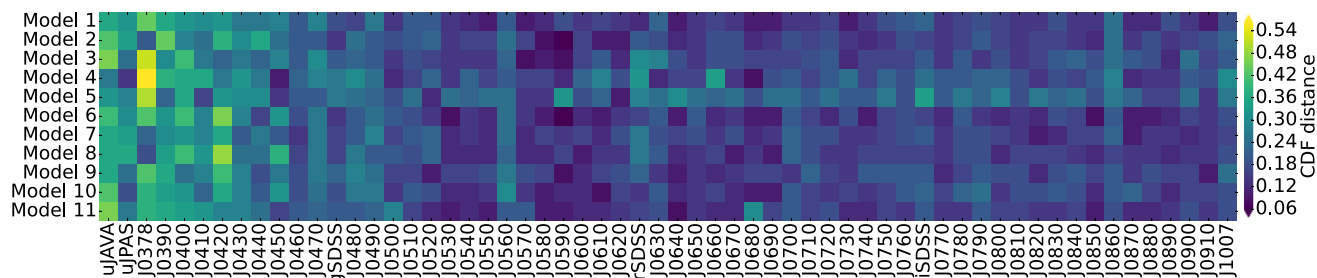
Figure 8. Goodness of the fit for each noise model as a function of the passband. The colour code represents the maximum distance between the CDFs of the S/N of the miniJPAS-Superset stars, and the S/N of the synthetic photometry obtained using a given model, where lower values correspond to a better fit. Model 11 corresponds to the best-fitting model for each band.
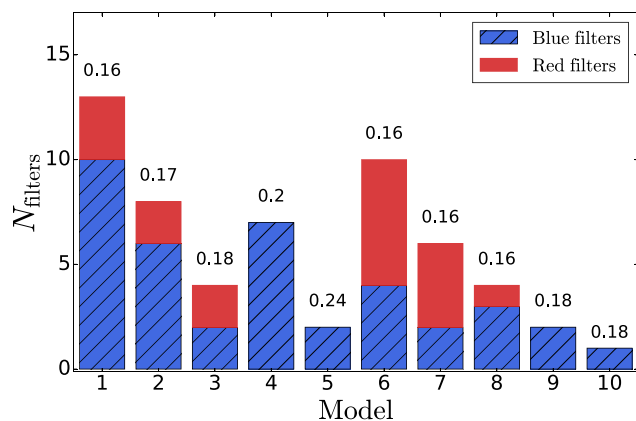

Figure 9. Histogram of the number of filters that had their noise profiles best fitted by a given model. The frequencies are separated by the contributions from the blue and red bands. The number on top of the bars correspond to the median CDF distance over all bands for a given model.

First we sorted in a consistent way the fluxes of the miniJPAS point sources in ascending order for each passband. Then, we searched in this sorted list for a value similar to the synthetic flux, and associated the corresponding observed uncertainty (a.k.a. nominal error) to it. The nominal errors in units of flux will be referred as σ_μ. Given that the nominal errors are associated to the synthetic photometry in a random way, this procedure ensures that the noise patterns from the different tiles are well represented in the mocks. Finally, the flux fluctuation was taken over a realization of a Gaussian function (or a combination of two or more Gaussians) with width proportional to the nominal error.

We tested 10 different noise models, as defined in Table 2. Such a variety of models help us to ensure a proper modelling of the noise profiles in all bands, even for the faintest miniJPAS sources. Models 1–5 correspond to single Gaussian functions with increasing widths; models 6–9 correspond to combinations of two or more Gaussian functions with different widths; model 10 samples the red bands ($\lambda_{\rm eff} \geq 7416$ Å) with a broader Gaussian than the blue bands; and model 11 corresponds to the best-fitting noise model for each band.

To select the best noise model for each band, we used the sample of non-saturated miniJPAS-Superset stars, which are expected to have almost no variability, and computed the maximum differences between the cumulative distribution functions (CDF) of the S/N of the observations, and the S/N of the synthetic photometry obtained according to model x, with x ranging from 1 to 10. The best-fitting for each band corresponds then to the model that

minimizes the difference between the CDFs, yielding model 11. In Fig. 8, we show the goodness of the fit (i.e. maximum distance between the CDFs) for each noise model as a function of the passband.

In Fig. 9, we provide the histogram of the number of filters which had their noise profile distributions best fitted by model x, with x ranging from 1 to 10. The contributions from the blue and red ($\lambda_{\rm eff} \geq 7416$ Å) bands are shown separately. The number on top of the bars corresponds to the median CDF distance over all bands for a given model; as a comparison, for model 11 this value is equal to 0.16. Although model 1 is able to reproduce the noise profile distributions of miniJPAS observations for most filters, models with larger widths are still preferable for some of the bands. These results agree with fig. A1 from González Delgado et al. (2021), which shows that the noise profile distributions of the PSFCOR magnitudes for miniJPAS extended sources are globally well fitted by a Gaussian function with standard deviation equal to 1.4. They also report that some filters present a higher dispersion in the noise distribution, and that the miniJPAS errors are particularly underestimated in the red filters.

In Fig. 10, we show the histograms of the differences between the observed and the synthetic fluxes divided by the nominal errors of the spectroscopic stars at six different bands. We compare the synthetic fluxes generated based on models 1 and 11. In the case of the miniJPAS point sources, the synthetic fluxes are computed directly from equation (4) (i.e. without adding flux fluctuations), and the sources are divided in three equal parts according to their magnitudes. As we can see, there is no clear trend on the noise distribution of the brightest and faintest objects. The absence of scattering in the *r*-band for the miniJPAS observations is interpreted as due to a narrower error distribution, which reflects the fact that this band has a significantly higher <S/N> than the narrow-bands, and was adopted as the reference band for the calibration of the miniJPAS images. The plots for the remaining 54 bands are available as supplementary material of MNRAS.

3.4 ND patterns

Following the convention adopted in Section 2.3.3, non-detections originated from low signal-to-noise detections naturally appear in the mocks whenever S/N < 1.25 (case i). As for the negative fluxes, if the noise fluctuations are sampled from very wide distributions, some of the resulting synthetic fluxes become negative, and are automatically flagged as NDs (case ii). We try to avoid as much as possible any excessive degradation of valid detections in the mocks, which would typically affect more intensely fainter objects.
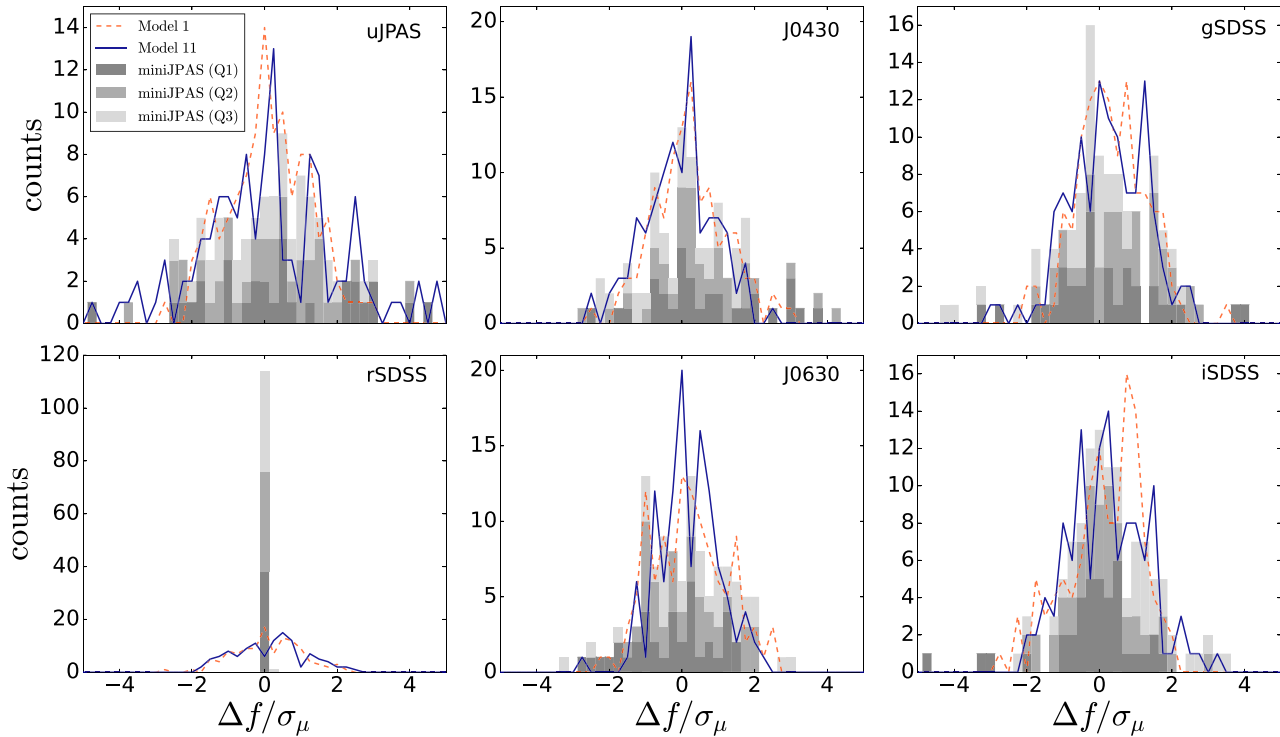
Figure 10. Histograms of the differences between the observed and the synthetic fluxes (Δf) divided by the nominal errors (σ_μ) for the miniJPAS-Superset stars. Here, we show the distributions for six different bands: *uJPAS*, *J0430*, *gSDSS*, *rSDSS*, *J0630*, and *iSDSS*. The bars correspond to the miniJPAS point sources, whose magnitudes were divided into lower (Q1), median (Q2), and upper (Q3) tertiles, shown by the different shades of grey ranging from darker to lighter, respectively. We compare noise models 1 (orange dashed lines) and 11 (blue solid lines).

3.5 Final mock catalogues

The final mock catalogues are provided in two versions: fluxes per unit wavelength and AB magnitudes with the corresponding nominal errors, so as to reproduce a real catalogue of observations. Since one does not know *a priori* the 'true' distributions of objects in each region of the sky, to avoid any biases from our putative luminosity functions our final mocks for the classification of miniJPAS sources contain balanced samples of size 10k, 10k, and 100k for the test, validation, and training sets, respectively, and for each class of object.

In Fig. 11, we provide some examples of synthetic photospectra generated with noise model 11 for galaxies, quasars, and stars. As a comparison, we also show the corresponding *SDSS* spectra, and the miniJPAS APER3 fluxes. The synthetic fluxes follow satisfactorily the miniJPAS observations, presenting some random statistical fluctuations within the expected levels, which is one of the key ingredients in our mocks.

The performances of the ML algorithms are validated on the test sets, as well as on the miniJPAS-Superset sample. The results of each classifier are also combined using a random forest algorithm (Pérez-Ràfols et al., in preparation). Besides providing balanced test samples, we also generate a sample containing the relative incidence rates of objects per deg² to allow a more direct comparison with the performance of the classifiers on the miniJPAS spectroscopic sample. The format of the final mock catalogues is shown in Fig. A.

4 MOCK VALIDATION

In this section, we validate the mock catalogues by comparing the main properties of the synthetic fluxes with the observational features

present in the miniJPAS point-like sample. Unless otherwise stated, these results correspond to the test sets generated using noise model 11.

4.1 Magnitude limits

The magnitude limits reached in each band by combining the test sets of quasars, galaxies, and stars are shown in Fig. 12. We compare these depths with the maximum magnitudes reached by the miniJPAS point-like sample (over all tiles). Note that these magnitude limits are obtained by considering only valid detections (i.e. S/N > 1.25 and $F_{\lambda,\mu}^{synt} > 0$). The depth reached in the test sets is in great accordance with the miniJPAS observations. As a comparison, we also show the J-PAS theoretical minimum depth (considering S/N = 5 within an aperture of 3 arcsec).

Although not shown here, the magnitude limits reached in the training, validation and 1-deg² sets are very similar to what we have demonstrated for the test set – apart, of course, from some fluctuations, as already expected due to the different sample size and different realizations of the luminosity function in each set.

4.2 Signal-to-noise ratio

In the left-hand panel of Fig. 13, we compare the median S/N distributions per band for the miniJPAS point sources randomly selected according to the QLF, miniJPAS-Superset quasars, and quasars from the test set. As already expected, miniJPAS-Superset quasars have typically larger S/N distributions than the mocks. On one hand, we can attribute this to the lack of a representative
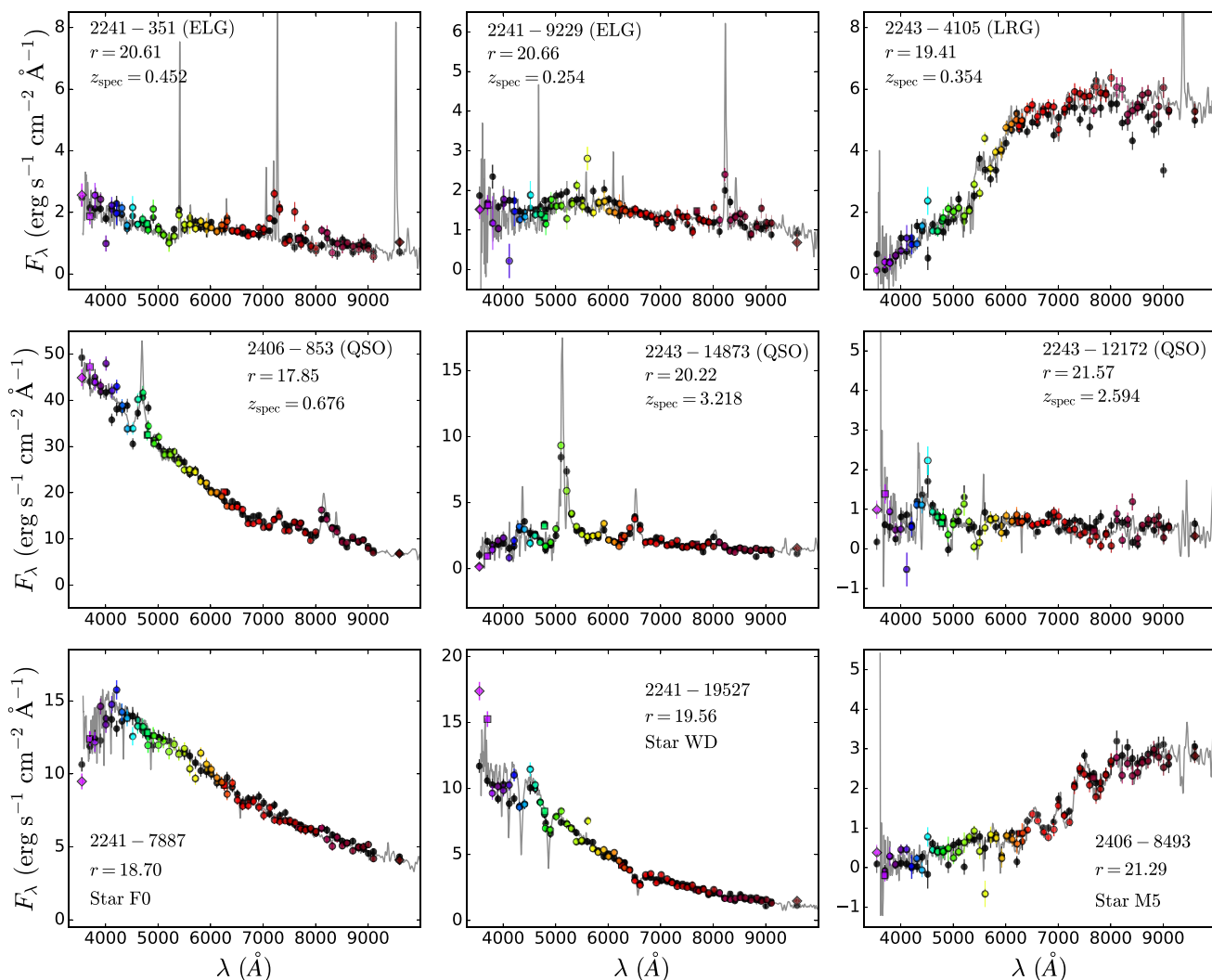
Figure 11. Simulated photospectra of starburst galaxies (top), quasars (middle), and stars (bottom). The grey solid lines correspond to the smoothed *SDSS* spectra. Coloured diamonds, squares, and dots correspond to the miniJPAS APER3 fluxes in the medium, broad, and narrow-bands, respectively. Black dots correspond to the synthetic fluxes generated using model 11 with their corresponding uncertainties. The miniJPAS objects are identified by their tile and number IDs; their *r*-band magnitudes, and spectroscopic redshifts $z_{\rm spec}$ (or stellar types $t_\star$) are also listed in the legend. The fluxes are in units erg s^{-1} cm^{-2} Å^{-1}.

sample of faint objects in the Superset sample, which increases $<S/N>$. Equivalently, the LF yields a fairer sample of faint sources in the mocks (with typically larger associated errors) which ends up dominating (and, subsequently, lowering down) the median S/N distribution. As a consequence, the signal-to-noise ratio in the test set becomes more representative of the miniJPAS point sources.

We also identify similar modulations in the S/N of adjacent filters for both the observations and simulations. Since the presence of these modulations are directly associated with the miniJPAS survey strategy (e.g. filter exposure times, net effect of sky brightness, and final number of combined images), having them in the mocks indicates that we have successfully matched the S/N distributions from the observations into the mocks. This can be better seen by comparing the median S/N achieved by the observations and the synthetic fluxes of the miniJPAS-Superset quasars.

In the middle and right-hand panels of Fig. 13, we show the median S/N distributions for the galaxy and star test sets, respectively. Again,

we observe the same patterns of modulations both in the observations and in the mocks. Note, however, that in the case of stars the median S/N distribution for the test set follows the spectroscopic sample more closely than the miniJPAS point sample. This effect is attributed to the decrease in the number counts predicted by the SLF for $r \gtrsim 21.5$ (see Fig. 7), which results in a typically brighter sample in comparison to the simulated samples of quasars and galaxies – whose luminosity functions tend to predict more objects in the faint end.

In Fig. 14, we evaluate how $<S/N>$ varies as a function of the magnitude in the test set for six different bands (*uJPAS*, *J0430*, *gS-DSS*, *rSDSS*, *J0630*, and *iSDSS*). As a comparison, we also show the subsamples of miniJPAS point sources that were randomly selected according to a given luminosity function. Results are provided for galaxies, quasars, and stars. All bands shown here present the same general trend of having a decreasing $<S/N>$ for fainter magnitudes. This demonstrates again that the mocks yield comparable signal-to-noise ratio properties as verified in a miniJPAS point-like subsample with equivalent luminosity distributions.
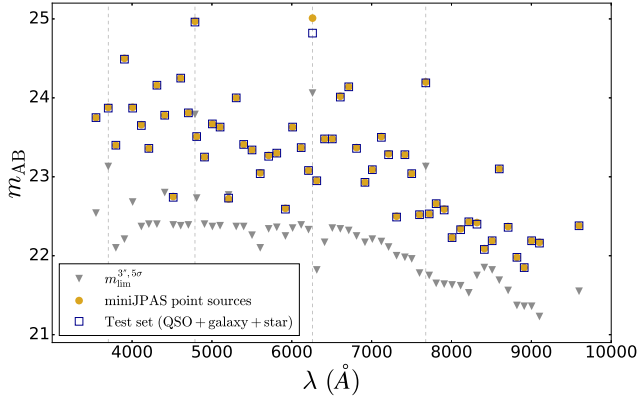
Figure 12. Maximum magnitudes reached in each filter by combining the test sets of quasars, galaxies, and stars (dark blue open squares). The yellow dots represent the depths reached by the miniJPAS point-like sample (over all tiles). As a comparison, we also show the J-PAS targeted minimum depth within an aperture of 3 arcsec (grey down-pointing triangles). The vertical grey dashed lines indicate the positions of the four broad-bands (*uJPAS*, *gSDSS*, *rSDSS*, and *iSDSS*, respectively).

4.3 Non-detections

In Fig. 15, we compare the number density of non-detections per band for the quasar test set and the subsample of miniJPAS point sources randomly selected according to the QLF. We separate the contributions from low S/N sources (upper panel) and negative fluxes (lower panel).

The pattern of non-detections is in general well reproduced in the quasar test set; in particular, we do not seem to include a too large fraction of negative fluxes. Nevertheless, the bluest bands, which correspond to filter indices < 30, seem to be affected by a large fraction of more noisy objects than the observations. These noisier quasars dominate the median signal-to-noise ratios at magnitudes $r > 22.0$. Similarly to what we found for quasars, the galaxy test set presents a large fraction of more noisy objects than the observations, but the fraction of negative fluxes is not overestimated. On the other hand, the mocks of stars result in larger fractions of objects with higher signal-to-noise ratios than the observations, and equivalently smaller fractions of negative fluxes.

An illustration of how the number of observed filters is degraded for fainter *r*-band magnitudes is provided in Fig. 16. As a comparison, we show the median number of filters with valid detections per magnitude bin for the quasar test set and the subsample of miniJPAS point sources randomly selected according to the QLF. We also divide the contributions between blue and red ($\lambda_{eff} \geq 7\,416\,\text{Å}$) bands. When compared with the miniJPAS point sources, the simulated fluxes in the blue bands present smaller fractions of non-detections in the range $20 \leq r \leq 22.5$, while presenting a strict suppression of detections in the reddest bands at the faint end ($r > 21.5$).

4.4 Quasar offsets

Quasars have long been known as intrinsically variable sources. On that account, one question that naturally arises is whether the miniJPAS quasars present any evidence of variability.

To investigate this hypothesis, we compute the median absolute differences (and corresponding standard deviations) of the miniJPAS magnitudes and synthetic magnitudes (without noise fluctuations) for Superset quasars, and compare these offsets with the ones obtained for stars. These differences are computed as a function of the tile, and the results are shown in Fig. 17. In essence, any evidence for quasar variability would be indicated by variations larger than the error bars of stars. However, we found no significant evidence of quasar variability. Hopefully, future observations will allow us to make a better assessment of quasar variability in the AEGIS field.

5 DISCUSSION

Training machine learning algorithms in the presence of simulated fluxes can be more advantageous than a training based solely on real data sets (composed of e.g. spectroscopically confirmed sources), because the latter might include spectral misclassifications and/or a non-fair distribution of redshifts and luminosities that can bias the resulting ML models, and consequently bias the classification. In particular, since the number of spectroscopically confirmed sources in the area surveyed by miniJPAS is not sufficiently large nor complete, the mock catalogues described in this work are crucial to perform the selection of the miniJPAS quasar candidates.

The results shown in this paper suggest that the mocks have successfully reached similar depths and levels of signal-to-noise ratio as the miniJPAS point sources. To further optimize the methodology
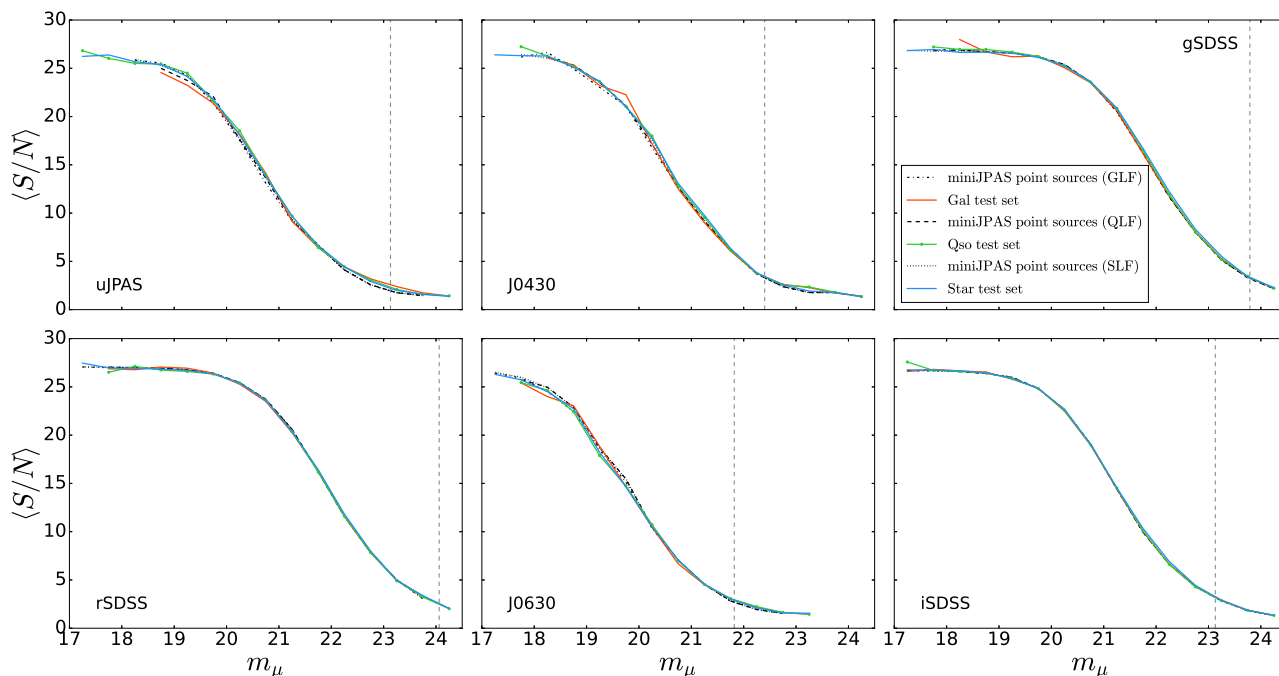

Figure 13. Median signal-to-noise ratio as a function of the filter for the subsample of miniJPAS point sources randomly selected according to the luminosity function (black lines), miniJPAS-Superset sources (grey solid lines), and test set (coloured solid lines). As a comparison, we also show the median signal-to-noise ratio obtained for the synthetic fluxes of the miniJPAS-Superset sources (grey dashed lines). We show the results for quasars (left-hand panel), galaxies (middle panel), and stars (right-hand panel).

Figure 14. Median signal-to-noise ratio as a function of the magnitude at six different bands (*uJPAS*, *J0430*, *gSDSS*, *rSDSS*, *J0630*, and *iSDSS*) for galaxies (orange solid line), quasars (green solid line), and stars (light blue solid line). As a comparison, we show the subsamples of miniJPAS point sources that were randomly selected so as to reproduce equivalent luminosity distributions for each type of source. The vertical dashed lines indicate the limiting magnitudes in each band.

developed for generating mock catalogues, we highlight in the following some future improvements:

(i) expand the library of spectra to include other types of sources that may be misrepresented in the current version. For instance, we lack a fair sample of some stellar spectral types (such as white dwarfs), galaxies brighter than $r = 18.7$ and at redshifts larger than $z = 0.9$, Lyman-break galaxies and Lyman α emitters, as well as type-II AGNs and red quasars. The inclusion of these objects in the training sets will allow us to refine the machine learning classifiers, and better assess the contaminants within the sample of quasar candidates;

(ii) better assess the luminosity priors for galaxies and stars. In particular, model the distributions of e.g. blue and red galaxies in more detail;

(iii) impose less conservative selection criteria in the miniJPAS catalogue, allowing, for instance, sources with some sort of flag, and explore different cuts in stellarity to assess how the performance of the classifiers changes;

(iv) include observations from other wavelengths (when available) – such as infrared information from the *Wide-field Infrared Survey Explorer* (*WISE*; Wright et al. 2010) and proper motions from *Gaia* (Gaia Collaboration 2016), to assist the classification. The mocks could also be supplemented with morphological parameters (coming from the modelling of miniJPAS sources) to improve e.g. the separation between unresolved galaxies and quasars. Combining all this information with the optical spectra poses an interesting challenge.

Finally, observations with the WEAVE-QSO survey (Pieri et al. 2016) will provide us with spectra of high-redshift quasars and Lyman α systems, with unprecedented spectral resolution (mostly $R = 5\,000$ but also $R = 20\,000$), allowing us to improve the simulated

fluxes (particularly at the faint end). The WEAVE-QSO spectroscopic follow-ups will enlighten us on the most critical improvements to the mock catalogues, as we will be able to better assess the performance of the ML classifiers, and confirm (or exclude) sources identified as potential quasar candidates. Moreover, J-PAS will soon start gathering data, which will further help us improve the noise models.

5.1 Additional applications

Our mock catalogues are also relevant for further interesting applications. For instance, in Queiroz et al. (in preparation) we present a novel technique to estimate the photometric redshifts (photo-zs) of the miniJPAS quasar candidates based on a best-fitting model for the quasar photospectrum in terms of the so-called eigenspectra derived from a principal component analysis. Since we do not have a large enough sample of spectroscopically confirmed miniJPAS quasars, the mocks are essential to validate the performance of this photo-z method.

Moreover, early operation of the J-PAS survey may be conducted with less than 60 filters. This means that, in addition to the classification that we perform for miniJPAS point sources, the mocks will be fundamental to forecast the accuracy with which J-PAS will be able to detect quasars, specially when we have observations in less than 60 bands. Although applied to the J-PAS photometric system, the methodology presented here can be easily adapted to build mocks for other (narrow-to-medium-band) photometric systems.

6 CONCLUSIONS

This paper is part of a series of manuscripts that aim at developing tools to classify miniJPAS sources in preparation for J-PAS, and
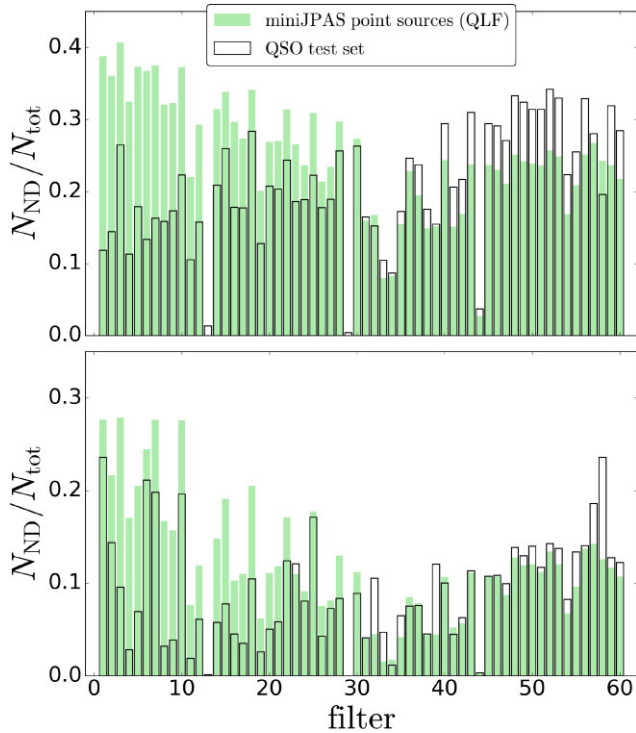
Figure 15. Non-detection fraction per band for the quasar test set (black lines), and the subsample of point sources randomly selected according to the QLF (green solid bars). The *x*-axis represents the filter indices, which are ordered according to the effective wavelength. We separate the contributions from low signal-to-noise ratio (upper panel) and negative (lower panel) fluxes. The *y*-axis corresponds to the number of non-detections in each passband divided by the total number of objects in the corresponding sample. The total number of objects are 8054 and 10 k for the green solid bars and black lines, respectively.

identify quasar candidates. Since no real data set exists at present that is sufficiently large or complete to serve that purpose, constructing mock catalogues is crucial to properly train our machine learning classifiers, and assess their performances until such a time when both our photometric data and spectroscopic follow-up data reach sufficient size.

In this paper, we present the pipeline to generate simulated photospectra of quasars, galaxies, and stars containing the same signal-to-noise ratio distributions expected for miniJPAS point-like sources. Starting from synthetic fluxes obtained by the convolution of *SDSS* spectra with the J-PAS photometric system, we show how to incorporate realistic observational features by (i) imposing that the relative incidence rates of the different classes of objects in the mocks follow the expected count numbers from putative luminosity functions; (ii) carefully modelling the noises in all bands, and adding compatible levels of noise to the synthetic photometry; and (iii) adding the patterns of non-detections (which dominate the faint end). Our results indicate that the miniJPAS fluxes in each band are best described by different noise profile distributions, but typically 1σ to 1.5σ Gaussian functions can properly fit the uncertainties in most filters.

Our final mock catalogues demonstrated the capability of correctly reaching the expected depths in all bands, and matching the signal-to-noise ratio distributions from the observations. These mock catalogues are invaluable for many scientific applications within the J-PAS collaboration, and will also be important for the whole
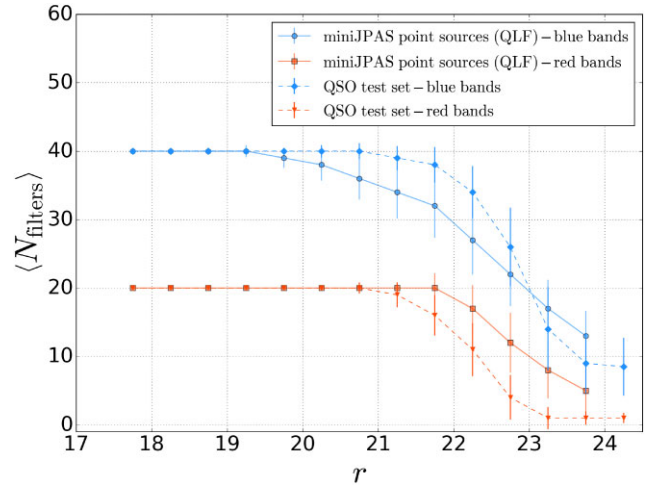
Figure 16. Median number of filters with valid detections as a function of *r*-band magnitude. We compare the number of detections present in the miniJPAS point sources randomly selected according to the QLF (solid lines) with that present in the quasar test set (dashed lines). The error bars correspond to the standard deviations. We divide the contributions between blue and red ($\lambda_{\mathrm{eff}} \geq 7\,416$ Å) bands.

astronomical community, as the procedure outlined here can be easily adapted to serve the purposes of other photometric surveys.

Once J-PAS is fully operational, we will be able to train the machine learning classifiers using directly real data. However, the mocks will not lose their crucial aspect of helping in the refinement of our methods, and avoiding misleading conclusions when analysing real observations. Even then, the mocks will perdure as our allies in the process of unraveling the cosmos in the search for quasars.
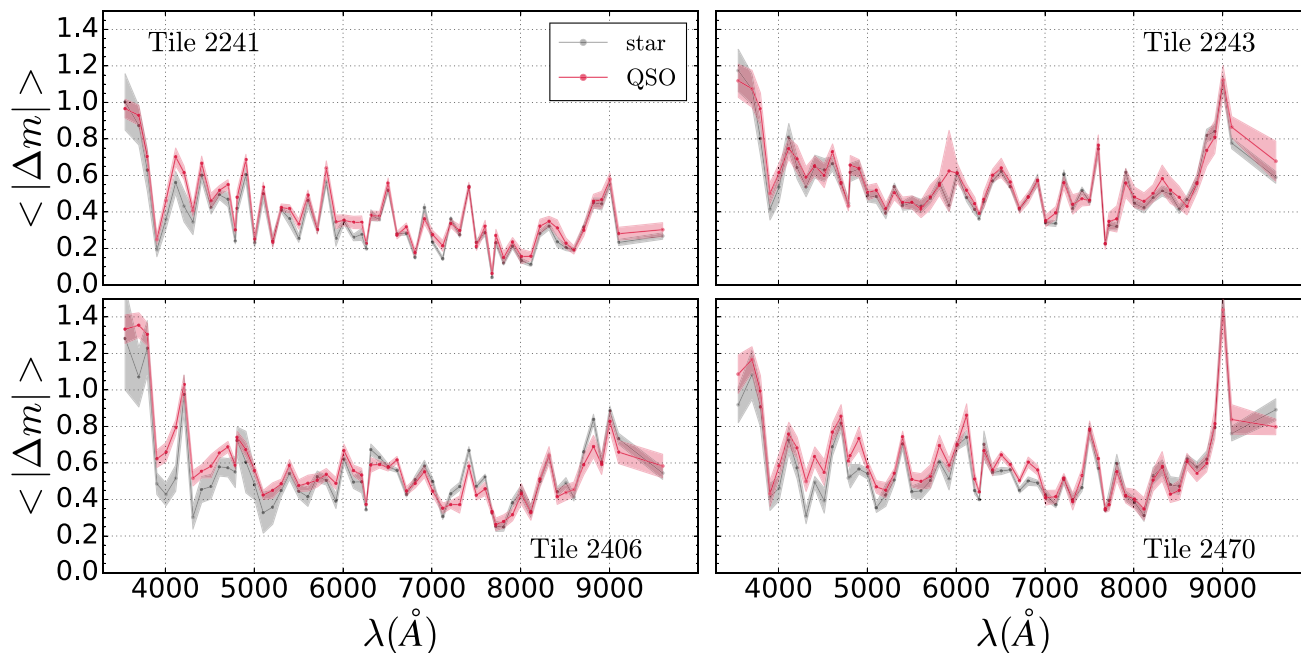
Figure 17. Median absolute differences between miniJPAS magnitudes and *SDSS* DR16Q synthetic magnitudes (without noise fluctuations) for quasars (red) and stars (grey) as a function of the tile. The solid regions correspond to the standard deviations.

DATA AVAILABILITY

The miniJPAS aperture corrections and mock catalogues developed in this work are available here.

REFERENCES

Abramo L. R., Bertacca D., 2017, Phys. Rev. D, 96, 123535
Abramo L. R. et al., 2012, MNRAS, 423, 3251
Ahumada R. et al., 2020, ApJS, 249, 3
Aihara H. et al., 2018, PASJ, 70, S8
Aihara H. et al., 2019, PASJ, 71, 114
Amendola L. et al., 2013, Living Rev. Relativ., 16, 6
Antonucci R., 1993, ARA&A, 31, 473

Aparicio Resco M. et al., 2020, MNRAS, 493, 3616

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, MNRAS, 310, 540

Astropy Collaboration, 2013, A&A, 558, A33

Astropy Collaboration, 2018, AJ, 156, 123

Baldwin J. A., 1977, ApJ, 214, 679

Ball N. M., Brunner R. J., Myers A. D., Tcheng D., 2006, ApJ, 650, 497

Baqui P. O. et al., 2021, A&A, 645, A87

Benitez N. et al., 2014, preprint (arXiv:1403.5237)

Bertin E., Arnouts S., 1996, A&AS, 117, 393

Bessell M. S., 2005, ARA&A, 43, 293

Bonoli S. et al., 2021, A&A, 653, A31

Brescia M., Cavuoti S., Longo G., 2015, MNRAS, 450, 3893

Cabayol L. et al., 2019, MNRAS, 483, 529

Cenarro A. J. et al., 2019, A&A, 622, A176

Chaves-Montero J. et al., 2017, MNRAS, 472, 2085

Clarke A. O., Scaife A. M. M., Greenhalgh R., Griguta V., 2020, A&A, 639, A84

Cooper M. C. et al., 2011, ApJS, 193, 14

Cooper M. C. et al., 2012, MNRAS, 419, 3018

Cristóbal-Hornillos D. et al., 2014, in Chiozzi G., Radziwill N. M., eds, Proc. SPIE Conf. Ser. Vol. 9152, Software and Cyberinfrastructure for Astronomy III. SPIE, Bellingham, p. 91520O

Croom S. M. et al., 2005, MNRAS, 356, 415

da Ângela J. et al., 2008, MNRAS, 383, 565

Dalton G., 2016, WEAVE: The Next Generation Spectroscopy Facility for the WHT. p. 97

Davis M. et al., 2007, ApJ, 660, L1

Dawson K. S. et al., 2013, AJ, 145, 10

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Díaz-García L. A. et al., 2015, A&A, 582, A14

Eftekharzadeh S. et al., 2015, MNRAS, 453, 2779

Eisenstein D. J. et al., 2011, AJ, 142, 72

Fadely R., Hogg D. W., Willman B., 2012, ApJ, 760, 15

Fitzpatrick E. L., Massa D., 2007, ApJ, 663, 320

Gaia Collaboration, 2016, A&A, 595, A1

Golob A., Sawicki M., Goulding A. D., Coupon J., 2021, MNRAS, 503, 4136

González Delgado R. M. et al., 2021, A&A, 649, A79

Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633

Harris C. R. et al., 2020, Nature, 585, 357

Harrison C. M., 2017, Nat. Astron., 1, 0165

Hernán-Caballero A., Hatziminaoglou E., Alonso-Herrero A., Mateos S., 2016, MNRAS, 463, 2064

Hopkins P. F. et al., 2004, AJ, 128, 1112

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90

Ilbert O. et al., 2006, A&A, 457, 841

Ivezić Ž. et al., 2019, ApJ, 873, 111

Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576

Laurent P. et al., 2017, J. Cosmol. Astropart. Phys., 2017, 017

Leistedt B., Peiris H. V., 2014, MNRAS, 444, 2

Leistedt B., Peiris H. V., Mortlock D. J., Benoit-Lévy A., Pontzen A., 2013, MNRAS, 435, 1857

López-Sanjuan C. et al., 2019a, A&A, 622, A177

López-Sanjuan C. et al., 2019b, A&A, 631, A119

Lyke B. W. et al., 2020, ApJS, 250, 8

Lynds R., 1971, ApJ, 164, L73

Margala D., Kirkby D., Dawson K., Bailey S., Blanton M., Schneider D. P., 2016, ApJ, 831, 157

Marín-Franch A. et al., 2017, in Arribas S., Alonso-Herrero A., Figueras F., Hernández-Monteagudo C., Sánchez-Lavega A., Pérez-Hoyos S., eds, Highlights on Spanish Astrophysics IX. p. 670

Martí P., Miquel R., Castander F. J., Gaztañaga E., Eriksen M., Sánchez C., 2014, MNRAS, 442, 92

Martínez-Solaeche G. et al., 2021, A&A, 647, A158

Mendes de Oliveira C. et al., 2019, MNRAS, 489, 241

Moles M. et al., 2008, AJ, 136, 1325

Nakazono L. et al., 2021, MNRAS, 507, 5847

Newman J. A. et al., 2013, ApJS, 208, 5

Odewahn S. C., Stockwell E. B., Pennington R. L., Humphreys R. M., Zumach W. A., 1992, AJ, 103, 318

Odewahn S. C. et al., 2004, AJ, 128, 3092

Palanque-Delabrouille N. et al., 2016, A&A, 587, A41

Pâris I. et al., 2017, A&A, 597, A79

Pérez-González P. G. et al., 2013, ApJ, 762, 46

Pickles A., Depagne É., 2010, PASP, 122, 1437

Pieri M. M. et al., 2016, SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. p. 259

Porciani C., Magliocchetti M., Norberg P., 2004, MNRAS, 355, 1010

Richards G. T. et al., 2002, AJ, 123, 2945

Richards G. T. et al., 2009, ApJS, 180, 67

Robin A. C., Reylé C., Derrière S., Picaud S., 2003, A&A, 409, 523

Ross N. P. et al., 2009, ApJ, 697, 1634

Ross N. P. et al., 2012, ApJS, 199, 3

Salpeter E. E., 1964, ApJ, 140, 796

Sargent W. L. W., Young P. J., Boksenberg A., Tytler D., 1980, ApJS, 42, 41

Schaye J. et al., 2015, MNRAS, 446, 521

Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103

Sevilla-Noarbe I. et al., 2018, MNRAS, 481, 5451

Shen Y. et al., 2007, AJ, 133, 2222

Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575

Taniguchi Y. et al., 2015, PASJ, 67, 104

Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29

Taylor K. et al., 2014, J. Astron. Instrum., 3, 1350010

The Dark Energy Survey Collaboration, 2005, preprint (arXiv:astro-ph/0510346)

Urry C. M., Padovani P., 1995, PASP, 107, 803

Vanden Berk D. E. et al., 2001, AJ, 122, 549

Véron-Cetty M. P., Joly M., Véron P., 2004, A&A, 417, 515

Vestergaard M., Wilkes B. J., 2001, ApJS, 134, 1

Virtanen P. et al., 2020, Nat. Methods, 17, 261

Wright E. L. et al., 2010, AJ, 140, 1868

Zel'dovich Y. B., Novikov I. D., 1964, Sov. Phys. Doklady, 9, 246

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Supplementary_material_online.pdf

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: MOCK CATALOGUES

In Table A1, we show the format of our final mock catalogues. For each type of source (quasars, galaxies, and stars) and mock version (training set, validation set, test set, and 1-deg^2 set) we provide a different catalogue. In the case of stars, for each type (O, B, A, F, G, K, M, WD, C, and CV) we associate a classification number running from 1 to 10.

The offsets per passband per tile obtained in Section 2.3.3 are also provided together with the mock catalogues as a separate table. The mock catalogues are already internally available to the

Table A1. General format of our mock catalogues, which are provided in two versions: fluxes per unit wavelength and AB magnitudes.

Column	Name	Format	Description
1	Mock	INT64	Mock version
2	id	INT64	Object identification
3–62	Filter	FLOAT[4]	Flux (or magnitude) in the 60 bands
63–122	eFilter	FLOAT[4]	Flux (or magnitude) uncertainty in the 60 bands
123	plate	INT64	*SDSS* spectroscopic plate number
124	mjd	INT64	*SDSS* spectroscopic MJD
125	fiber	INT64	*SDSS* spectroscopic fiber number
126	magr_scale	FLOAT[4]	*r*-band magnitude used to scale the *SDSS* spectrum
127	magr_synthetic	FLOAT[4]	Synthetic *r*-band magnitude
128	ztype	FLOAT[4]	Spectroscopic redshift (Stellar type)

J-PAS collaboration, and will be made publicly available upon publication.

[1]*Departamento de Astronomia, Instituto de Física, Universidade Federal do Rio Grande do Sul (UFRGS), Av. Bento Gonçalves, 9500, Porto Alegre, RS, Brazil*

[2]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, Rua do Matão, 1371, CEP 05508-090, São Paulo, Brazil*

[3]*CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies, Sorbonne Université, Université Paris Diderot, LPNHE, 4 Place Jussieu, F-75252 Paris, France*

[4]*CNRS, CNES, LAM, Aix Marseille Univ, Marseille, France*

[5]*Instituto de Astrofísica de Andalucía (CSIC), PO Box 3004, E-18080 Granada, Spain*

[6]*Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Plaza San Juan, 1, E-44001 Teruel, Spain*

[7]*Departamento de Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain*

[8]*Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain*

[9]*Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

[10]*Donostia International Physics Center, Paseo Manuel de Lardizabal 4, E-20018 Donostia-San Sebastian, Spain*

[11]*Ikerbasque, Basque Foundation for Science, E-48013 Bilbao, Spain*

[12]*Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China*

[13]*Instituto de Física de Cantabria (CSIC-UC), Avda. Los Castros s/n, E-39005 Santander, Spain*

[14]*Unidad Asociada 'Grupo de Astrofísica Extragaláctica y Cosmología', IFCA-CSIC / Universitat de València, Valencia, Spain*

[15]*Observatório Nacional/MCTI, Rua General José Cristino, 77, São Cristóvão, CEP 20921-400, Rio de Janeiro, Brazil*

[16]*Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Unidad Asociada al CSIC, Plaza San Juan 1, E-44001 Teruel, Spain*

[17]*Department of Astronomy, University of Michigan, 311 West Hall, 1085 South University Avenue, Ann Arbor, USA*

[18]*Department of Physics and Astronomy, University of Alabama, Gallalee Hall, Tuscaloosa, AL 35401, USA*

[19]*Departamento de Astronomia, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Rua do Matão, 1226, CEP 05508-090, São Paulo, Brazil*

[20]*Instruments, 44121 Pembury Place, La Canada Flintridge, CA 91011, USA*

This paper has been typeset from a TEX/LATEX file prepared by the author.