Astronomy & Astrophysics manuscript no. ldf_photoz
November 18, 2020
©ESO 2020

The LOFAR Two-metre Sky Survey Deep Fields - Data Release 1

IV. Photometric redshifts and stellar masses[⋆,⋆⋆]

K. J. Duncan[1,2], R. Kondapally[1], M. J. I. Brown[3], M. Bonato[4,5,6], P.N. Best[1], H. J. A. Röttgering[2], M. Bondi[4], R. A. A. Bowler[7], R. K. Cochrane[8], G. Gürkan[9], M. J. Hardcastle[10], M. J. Jarvis[7,11], M. Kunert-Bajraszewska[12], S. K. Leslie[2], K. Małek[13,14], L. K. Morabito[15], S. P. O'Sullivan[16], I. Prandoni[6], J. Sabater[1], T. W. Shimwell[2,17], D. J. B. Smith[10], L. Wang[18,19], and A. Wołowska[12]

(Affiliations can be found after the references)

arXiv:2011.08204v1 [astro-ph.GA] 16 Nov 2020

ABSTRACT

The Low Frequency Array (LOFAR) Two-metre Sky Survey (LoTSS) is a sensitive, high-resolution 120-168 MHz survey split across multiple tiers over the northern sky. The first LoTSS Deep Fields data release consists of deep radio continuum imaging at 150 MHz of the Boötes, European Large Area Infrared Space Observatory Survey-North 1 (ELAIS-N1), and Lockman Hole fields, down to rms sensitivities of ∼32, 20, and 22 μJy beam^{-1}, respectively. In this paper we present consistent photometric redshift (photo-z) estimates for the optical source catalogues in all three fields - totalling over 7 million sources (∼ 5 million after limiting to regions with the best photometric coverage). Our photo-z estimation uses a hybrid methodology that combines template fitting and machine learning and is optimised to produce the best possible performance for the radio continuum selected sources and the wider optical source population. Comparing our results with spectroscopic redshift samples, we find a robust scatter ranging from 1.6 to 2% for galaxies and 6.4 to 7% for identified optical, infrared, or X-ray selected active galactic nuclei (AGN). Our estimated outlier fractions ($\left|z_{\rm phot} - z_{\rm spec}\right| / (1 + z_{\rm spec}) > 0.15$) for the corresponding subsets range from 1.5 to 1.8% and 18 to 22%, respectively. Replicating trends seen in analyses of previous wide-area radio surveys, we find no strong trend in photo-z quality as a function of radio luminosity for a fixed redshift. We exploit the broad wavelength coverage available within each field to produce galaxy stellar mass estimates for all optical sources at $z < 1.5$. Stellar mass functions derived for each field are used to validate our mass estimates, with the resulting estimates in good agreement between each field and with published results from the literature.

Key words. galaxies: distances and redshifts – galaxies: active – radio continuum: galaxies

1. Introduction

Combining extremely high sensitivity with a wide field-of-view, the Low Frequency Array (LOFAR; van Haarlem et al. 2013) offers an unprecedented capability for performing large statistical surveys of the radio sky. The LOFAR Surveys Key Science Project (Rottgering 2010) is undertaking a set of tiered surveys over the northern sky at 120-168 MHz. The first release of data from the all-sky LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017) presented 424deg^2 to an average sensitivity of 71μJy beam^{-1} at 150 MHz[1]. Complementary to the wide area LoTSS data, the LoTSS Deep Fields First Data Release (LoTSS DR1) reaches radio continuum sensitivities comparable to or deeper than the deepest surveys currently available, and over orders of magnitude wider areas (> 50 deg^2; see Tasse et al. (2020) and Sabater et al. (2020) - Paper I and Paper II, respectively). Located in some of the best-studied northern extragalactic survey fields - Boötes, European Large Area Infrared Space Observatory Survey-North 1 (ELAIS-N1, or EN1 hereafter), and the Lockman Hole (LH) - the LoTSS Deep Field data reach a current rms sensitivity of ∼32, 20, and 22 μJy beam^{-1} at 150 MHz, respectively, sufficient to detect radio-quiet AGN and extremely star-forming galaxies out to the highest redshifts ($z > 5$).

Extracting the maximum scientific value from the LOFAR radio continuum observations requires robust identification of the host galaxies of radio sources, alongside knowledge of the source redshifts, to extract intrinsic physical properties for both the radio sources (e.g. physical size, luminosity) and their host galaxies. By design, the LOFAR deep fields are located in regions of the sky that contain extensive ancillary imaging data, from ultraviolet (UV) all the way to far-infrared (FIR). Kondapally et al. (2020; hereafter Paper III) present new multi-wavelength optical to mid-IR photometry catalogues for all three fields alongside careful cross-identification with the LOFAR radio source population (with host identifications for $\gtrsim$ 97% of radio sources). However, the wide range of intrinsic host properties in radio continuum selected samples, including both active galactic nuclei (AGN) and extreme star-forming galaxies, means that even with extensive ancillary data, deriving reliable photometric redshifts (photo-zs) can be non-trivial (Norris et al. 2019).

⋆ The redshift catalogues are available online at https://lofar-surveys.org/, as part of this data release.
⋆⋆ Catalogues are available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/
[1] Formally, the central frequency of the LoTSS Deep Fields data is 144 MHz in Boötes and Lockman Hole, and 146 MHz in ELAIS-N1. However, throughout this paper we will refer to the LoTSS frequency colloquially as 150 MHz.

In Duncan et al. (2018a, hereafter D18a), we demonstrated that photo-z estimates for radio continuum sources obtained by combining multiple template-fitting estimates can be both more precise and more reliable than those using just one single template library. However, there remained key subsets of the AGN population for which template-based photo-zs could not produce satisfactory results (e.g. IR or X-ray selected AGN at $1 < z < 3$).

In Duncan et al. (2018b, hereafter D18b), we built upon the previous template-fitting method of D18a by incorporating additional machine-learning-based photo-z estimates trained for different AGN subsets. When combined together within a hierarchical Bayesian (HB) combination framework, the resulting consensus photo-z estimates improve on the performance of either individual method (specifically template-fitting or machine learning, see also Cavuoti et al. 2017; Fotopoulou & Paltani 2018, for other successful approaches). This combined 'hybrid' approach was successfully applied to shallow optical and mid-IR photometry over very wide fields as part of LoTSS DR1 (Shimwell et al. 2019; Williams et al. 2019). The resulting photo-zs - presented in Duncan et al. (2019b, hereafter D19) - provided precise and reliable photo-zs ($\approx 3\%$ scatter and $< 2\%$ outlier fraction; OLF) out to $z \sim 0.8$ for LOFAR radio sources with host-dominated optical-infrared spectral energy distributions (SEDs; specifically, sources for which there is no strong evidence for a dominant AGN contribution to the SED at X-ray, optical, or IR wavelengths). Beyond $z \sim 0.8$, the shallow optical data leads to incomplete samples and highly uncertain redshift estimates for this population in LoTSS DR1. Thanks largely to the extensive training samples available for machine learning estimates, reliable photo-z estimates for quasar-like sources in D19 extended out to $z \lesssim 3$.

In the case of the Boötes field, a test field for both D18a and D18b, we have already demonstrated that our hybrid approach can produce high quality photo-zs across a broad range of source types and redshifts. However, Boötes is currently unique among the LoTSS Deep Fields with respect to its larger sample of high quality spectroscopic redshifts both for calibration of template fits and for training of machine learning photo-z estimates. Neither the EN1 nor LH fields contain the representative samples of AGN spectroscopic redshifts necessary to apply the hybrid method of D18b with the same effectiveness as achieved in Boötes. Until large samples of spectroscopic redshifts provided by the forthcoming WEAVE-LOFAR survey (Smith et al. 2016), a modified approach is therefore required if we wish to maximise the quality of photo-zs across all three LOFAR Deep Fields and fully exploit the extraordinary sample of faint radio continuum sources they provide.

In this paper we present photometric analysis of the matched aperture multi-wavelength optical to mid-IR data in each of the three LoTSS Deep Fields. The primary aim of the paper is to provide optimal photo-z estimates for all sources in the fields to enable the exploitation of the deep radio continuum observations. Additionally, given that all three fields have deep near-infrared (NIR) or mid-infrared (mid-IR) photometry required for robust SED fitting, we provide stellar mass estimates for the subset of the optically detected sources for which we can make reliable measurements - enabling a wide range of science and providing a valuable reference sample for studies of the faint LOFAR radio population.

The remaining sections of this paper are set out as follows. In Section 2 we summarise the properties of the multi-wavelength photometry and spectroscopic samples used for photo-z analysis in this study. Next, in Section 3 we outline the key changes to the hybrid photo-z method used in this analysis. In Section 4 we analyse the precision and accuracy of the resulting photo-zs across all three fields as a function of redshift, magnitude, source type and radio luminosity. Section 5 then presents the method used for deriving stellar mass estimates for the optical sample, along with derivation of mass completeness limits and tests demonstrating their overall quality and limitations. In Section 6 we then briefly explore the physical properties of the LoTSS Deep Fields radio source population based on our derived redshifts and stellar masses. Finally, Section 7 presents a summary of our work. Throughout this paper, all magnitudes are quoted in the AB system (Oke & Gunn 1983) unless otherwise stated. We also assume a Λ Cold Dark Matter cosmology with $H_0 = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.

2. Data

2.1. Photometry

Optical to mid-IR imaging, photometry and radio source identification for each of the LOFAR deep fields are presented in Paper III, to which we refer the reader for full details. In summary, available deep photometry in the EN1 and Lockman Hole fields has been mosaicked onto a common pixel and flux scale. For each field, forced aperture photometry in all bands has then been performed based on detections in two stacked χ^2 S/N images (optical to NIR and mid-IR) and the resulting sets of catalogues have been merged to produce a consistent multi-wavelength catalog. In the Boötes field, existing forced aperture photometry catalogues based on detections in I and *Spitzer* Infrared Array Camera (IRAC; Fazio et al. 2004) 4.5μm images have been merged following the same procedure as used for the other two fields. Aperture corrections based on either curve of growth analysis for sources in the field (EN1 and LH) or the analytic point spread function (PSF; Boötes) have then been calculated to provide total flux estimates in each filter.

For all three fields we make use of $3''$ apertures for optical to NIR bands and $4''$ apertures for IRAC (due to the lower resolution in the IRAC imaging). As described in Paper III, the galactic extinction ($E(B-V)$) for each source has been calculated based on the Milky Way extinction map of Schlegel et al. (1998), queried using the `dustmaps` Python package (Green 2018). Filter-dependent extinction factors are then calculated by convolving the respective filter response curves with the Milky Way dust extinction law of Fitzpatrick (1999). In Figure 1 we visually summarise the wavelength coverage and sensitivity of the resulting photometry catalogues used for the photo-z analysis in this paper.

In addition to the processing steps above that are presented in Paper III, for the photo-z estimation and SED fitting work in this paper we include further processing as follows. Due to the combination of large areas, the number of unique bands and depth of the photometry, small numbers of spurious datapoints are inevitable. Therefore, as a final additional step before photo-z estimation, we automatically filter the photometric datasets for spurious

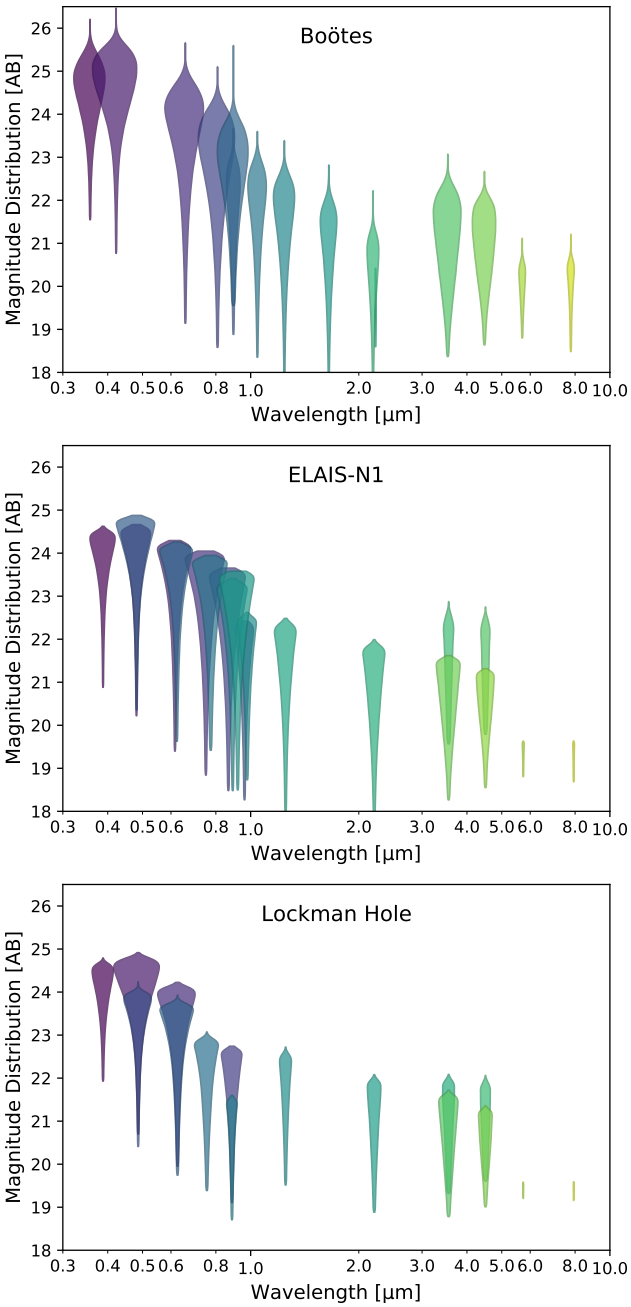

Fig. 1. Illustration of the wavelength coverage and depths for the photometric datasets used for photo-z estimation in the three LoTSS Deep Fields. The violin plot for each filter shows the distribution of magnitudes for 5σ sources. The widths of the violin plots are indicative of the area available at a given depth within each field and are scaled by the fraction of sources with 5σ detections compared to the total size of the catalog. For example, the deeper *Spitzer* IRAC photometry of the smaller SERVS regions (at 3.6 and 4.5μm) are evident in the EN1 and LH plots.

datapoints. Specifically, we filter for unphysical colours or extreme datapoints that are likely to be caused by artefacts such as stellar diffraction spikes, cosmic rays, cross-talk etc. Measurements that lie > 2.5 mag above or < 1 mag below those at *both* adjacent filters at shorter and longer wavelengths, indicative of extreme excess or deficit, are excluded. The large magnitude cuts are chosen to not exclude strong emission or absorption line sources. Similarly, measurements that result in a > 5 magnitude jump or drop in colour between consecutive filters are also excluded. These cuts are designed to be conservative, excluding clearly unphysical colours whilst not affecting large but physical colours such as Lyman break or strong emission line features. We note that the full optical catalogues released in Paper III have not been processed in this way as users may wish to apply different cuts. However, the exact photometric catalogues used for photo-z estimation (including outlier identification) are available through the LOFAR Surveys Data Release site[2].

There likely remain a number of spurious measurements for which more detailed analysis would be required for identification (e.g. iterative fits excluding individual photometry points, c.f. Chung et al. 2014). However, given the large number of available bands and the overall high quality of imaging available in these fields, we do not expect that our photo-z estimates (or science results derived from them) are significantly affected by any remaining anomalous datapoints.

2.2. Multi-wavelength classifications

We broadly classify all sources in the photometric samples using the following additional criteria:

Infrared AGN are identified using the *Spitzer* IR colour and and monotonically increasing mid-IR SED criteria presented by Donley et al. (2012). As we are primarily concerned with identifying only the robustly selected AGN sources, we also require > 5σ detections in all four IRAC bands. For the subset of the optical catalogues which satisfy our primary quality cuts, we find 0.7, 0.3 and 0.3% of the catalogues are found to satisfy the IR AGN criteria in Boötes, EN1, and LH, respectively.

Optical AGN: were identified through cross-matching the optical catalogue with the Million Quasar catalogue compilation of optical AGN, which is primarily based on Sloan Digital Sky Survey (SDSS; Alam et al. 2015) and other literature catalogues (Flesch 2015). Additionally, sources flagged as AGN based on spectroscopic observations are included in the optical AGN sample. For the subset of the optical catalogues which satisfy our primary quality cuts, we find consistently that $\approx 0.1\%$ of the optical sources in all three fields satisfy the optical AGN selection criteria.

X-ray AGN: In the Boötes field where deep X-ray observations are present over a large area, X-ray AGN were identified by cross-matching the positions of sources in our catalogue with the X-Böotes *Chandra* survey of NDWFS (Kenter et al. 2005). We calculate the X-ray-to-optical flux ratio, $X/O = \log_{10}(f_X/f_{\rm opt})$, based on the I band magnitude following Brand et al. (2006) For a source to be selected as an X-ray AGN, we require that an X-ray source has $X/O > -1$ or an X-ray hardness ratio > 0.8 (Bauer et al. 2004). In total, we find 2811 X-ray AGN in Boötes based on these criteria. In EN1 and LH, bright X-ray sources were identified based on the Second *ROSAT* All-Sky Survey (2RXS; Boller et al. 2016) and the *XMM-Newton* Slew Survey (XMMSL2)[3], as in LoTSS DR1. X-ray sources were matched to their IR counterparts using the published AllWISE (Wide-field Infrared Survey Explorer, WISE; Wright et al. 2010) cross-matches of Salvato et al. (2017), with

[2] https://lofar-surveys.org/releases.html
[3] https://www.cosmos.esa.int/web/xmm-newton/xmmsl2-ug

Table 1. Total spectroscopic redshift samples available for photo-z training and/or validation and the size of the corresponding photometric catalogue used in this analysis. We note that these numbers represent the total contained within the full catalogues and do not account for any cuts applied to the catalogues for analysis (e.g. flagging and restrictions to regions with specific photometric coverage).

Field	$N_{\mathrm{spec}-z}$ (Galaxies)	$N_{\mathrm{spec}-z}$ (AGN)	N_{Total}
Boötes	19 143	2 714	2 214 329
EN1	3 419	593	2 105 993
LH	4 787	1 182	3 041 794

sources then matched to the deep fields photometric dataset using the corresponding AllWISE source positions.

We note here that as in previous works, these broad sample selections are designed to identify clear AGN dominated SEDs for the purposes of optimising the photo-z analysis. The classifications are not intended to be complete samples of the AGN population within the fields. Subsequent studies will combine the multi-wavelength photometry with radio and FIR information to provide robust source classifications for the LOFAR detected population (see Best et al. 2020; Paper V).

2.3. Spectroscopic redshift samples

Spectroscopic redshifts for sources in Boötes are taken from a compilation of observations within the field comprising primarily the results of the AGN and Galaxy Evolution Survey (AGES; Kochanek et al. 2012) spectroscopic sample, with additional redshifts provided by a large number of smaller surveys in the field including Lee et al. (2012, 2013, 2014), Stanford et al. (2012), Zeimann et al. (2012, 2013) and Dey et al. (2016). Included in this sample are a number of proprietary redshifts used for spectroscopic redshift training only (M. Brown, *private communication*). In the release catalogues provided in this paper, we incorporate publicly available spectroscopic redshifts compiled as part of the Herschel Extragalactic Legacy Project (HELP, PI: S. Oliver)[4].

Our spectroscopic redshift samples for the EN1 and LH fields are also based on the HELP compilations. The majority of spectroscopic redshifts in these compilations originate from the SDSS spectroscopic sample (Alam et al. 2015), with additional data from a number of smaller spectroscopic follow-up campaigns (namely Berta et al. 2007; Swinbank et al. 2007a,b; Lacy et al. 2013).

For all three fields, the spectroscopic samples are matched to the optical photometry catalogues using a simple nearest neighbour match with a maximum radius of 1″. Table 1 summarises the spectroscopic redshift samples available for each field and the corresponding redshift and magnitude distributions are presented in Fig. 2 and Fig. 3, respectively. The disparity in currently available spectroscopic samples between Boötes and the other two deep fields is evident, with much larger samples of low-redshift galaxies available and significant AGN samples that extend to higher redshift and fainter magnitudes than available in EN1 or LH. However, we can see from Fig. 3 that the available spectroscopic sources are biased towards brighter magnitudes in all fields. Nevertheless, a key goal of this work is to provide consistent photo-z estimates across the three fields in all parameter space - ideally of comparable quality.

[4] http://hedam.lam.fr/HELP/

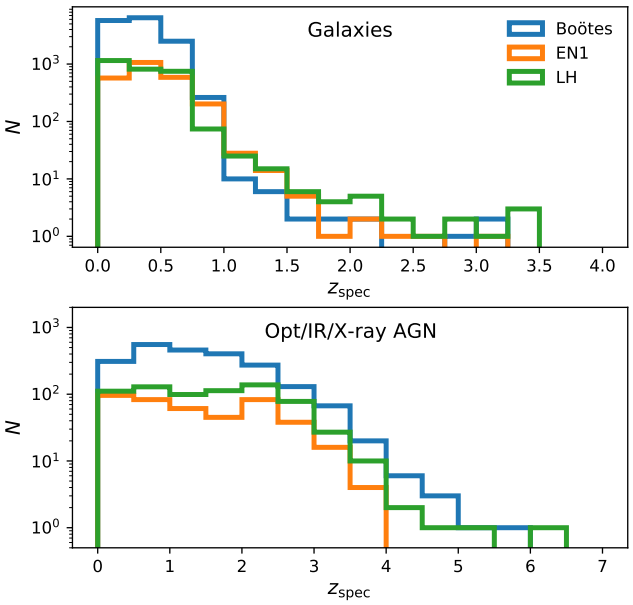

Fig. 2. Redshift distributions for the spectroscopic redshift training and test samples available in each of the fields (see Table 1 for total numbers). While Boötes has a much greater number of sources with z_{spec} available, these are largely limited to $z < 1$ for galaxies but extend over a greater redshift range for the known AGN.

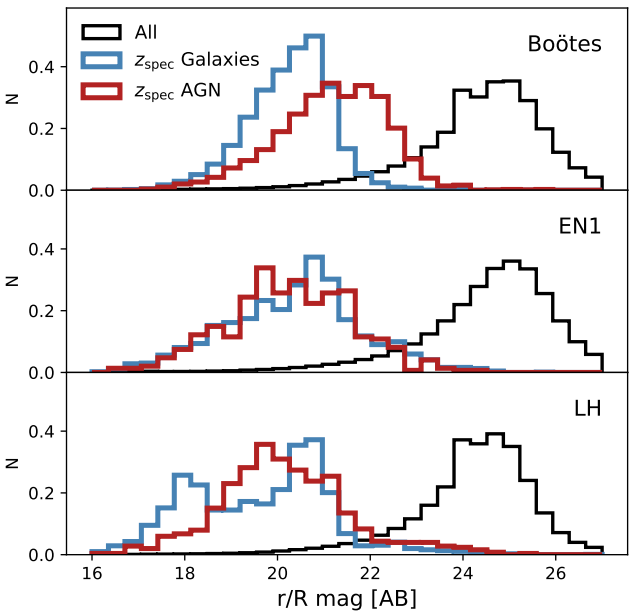

Fig. 3. Normalised optical magnitude distributions for the galaxy (blue histogram) and AGN (red histrogram) spectroscopic redshift samples in comparison to the full parent photometric sample. For EN1 and LH we plot the r-band magnitude distribution, while for Boötes we show the nearest equivalent wavelength, R. The offset in the magnitude distributions highlights the importance of high quality template estimates in the regime where spectroscopic training samples are not available.

3. Photometric redshift estimates

We estimate photo-zs for the full matched aperture optical catalogues presented in Paper III following a modified version of the hybrid approach presented in D18a and D18b and applied to LoTSS DR1 in D19. For full details of the motivation behind the hybrid approach, as well as details regarding the HB combination method we refer the reader to those papers. In this paper we focus solely on the modifications or changes to the method presented in D19 that are specific to this work.

3.1. Template fitting

Due to the limited availability of spectroscopic training samples in two of the deep fields (EN1 and Lockman Hole) and the scientific focus on higher redshift where spectroscopic redshifts will always be limited, the primary modifications to our hybrid photo-z method are designed to maximise the accuracy and reliability of our template-fitting based estimates for the AGN population. Our aim is to reduce the potential need for machine learning estimates and provide consensus photo-zs for AGN in EN1 and LH that are of comparable quality to those produced in Boötes (D18b). For the template fitting photo-z estimates, we use updated versions of all three template libraries that either improve the wavelength coverage or broaden the representation of different AGN SED types. The three updated libraries now used for template fitting are:

Updated EAZY *models:* In previous work, the 'default' EAZY template set (version 1.3) was found to produce photo-z estimates with the smallest scatter and OLF for star-forming or quiescent galaxy populations (D18a). Here we make use of the new template set derived from the Flexible Stellar Population Synthesis code (FSPS; Conroy et al. 2009; Conroy & Gunn 2010). While designed to reproduce the same representative combination of stellar emission dominated galaxy SEDs, the revised FSPS templates now include dust reprocessed emission in the mid-IR. The emission from dust extends the range of rest-frame wavelengths that can reliably be used to constrain the photo-z estimate, which is particularly valuable in the deep field datasets that includes photometry out to $8\,\mu$m.

Extended Atlas Library: The second template set, the 'Atlas of Empirical SEDs' Brown et al. (2014) was recently extended to incorporate a wide range of the AGN and quasi-stellar object (QSO) populations not previously present in the original library (Brown et al. 2019). Additionally, the new Atlas of AGN SEDs incorporates a range of combinations of Seyfert type AGN spectra and different underlying host stellar populations (with varying AGN to host contributions). As illustrated in Brown et al. (2019), the inclusion of a fully representative range of AGN SEDs (and combinations of AGN + host galaxy) leads to significant improvement in the photo-z statistics for AGN dominated galaxies - the area for which machine learning estimates were found to provide the greatest improvement in the hybrid method (D18b).

Revised 'XMM-COSMOS' Team templates: Finally, we also make use of a new iteration of the 'XMM-COSMOS' galaxy and AGN SEDs (Polletta et al. 2007; Salvato et al. 2008, 2011) as presented in Ananna et al. (2017). Optimised for significantly larger survey fields than previous iterations (comparable to the total LOFAR Deep Fields coverage), the implementation in Ananna et al. (2017) includes more luminous quasar SED types than previous versions - as would be expected in the larger survey volumes than the deep pencil-beam surveys for which the Salvato et al. (2011) library was optimised.

The implementation of all three libraries within our template fitting method is the same as their corresponding versions in D18a and D18b. One key exception is the incorporation of new template library specific rest-frame model uncertainties within the fitting process that we outline below.

3.1.1. Photometric zero-point offsets

The inclusion of small magnitude offsets, or zero-point offsets, to the observed photometry of some datasets has been shown to improve photo-z estimates from template fitting (e.g. see Dahlen et al. 2013). While typically small ($\lesssim 10\%$), these additional offsets can often substantially reduce the overall scatter or OLF for photo-z estimates. In Hildebrandt et al. (2012), detailed comparison between the resulting photo-z performance for different levels of photometry precision indicates that zero-point flux offsets serve largely to correct for PSF effects. Given the nature of the photometry used in the work, we would therefore expect that the inclusion of zero-point offsets during template fitting will be beneficial.

Zero-point offsets for all template sets are derived for each photometric dataset following the method outlined in D18a. In summary, for 50% of the spectroscopic redshift subset (with 50% retained for validation/testing), the template set is fit to the observed photometry with the redshift fixed to the true redshift and the corresponding zeropoint offset is then calculated from the median offset between the observed and fitted flux values for sources with $S/N > 3$ in that band.

As previously seen for the Boötes photometry in D18a, the inclusion of the zero-point offsets during template fitting leads to substantial improvement in photo-z quality (as tested using the 50% of spectroscopic sources not included in the derivation of zero-point offsets). Similar improvements are observed for both EN1 and LH. For reference, we present the derived zero-point offsets for all template library and field combinations in Appendix B. Within a given field, we find good agreement between the largest zero-point corrections derived for the different template sets (e.g. the UKIRT Infrared Deep Sky Survey, UKIDSS, K-band data in EN1 and LH are consistently found to require a correction of $\sim 10\%$). We conclude that for the datasets employed in this analysis, the zero-point corrections are likely correcting for small offsets in the photometry itself, rather than large systematic errors in the templates themselves (we note that the three different template sets encompass both model and empirically derived templates).

3.1.2. Template model uncertainties

Incorporating an estimate of the rest-frame model uncertainties during photo-z template fitting has been shown to significantly improve the accuracy of the resulting estimates (Brammer et al. 2008; Dahlen et al. 2013). As part of our revised photo-z template fitting procedure we therefore estimate the rest-frame uncertainties for each template

library from the spectroscopic training sample following the method outlined by Brammer et al. (2008). For the Boötes field (chosen due to its large available spec-z sample), for each template set we fit the templates to the spectroscopic sample (including the zero-point flux correction) while fixing the fit to the known redshift and then measuring the distribution of flux residuals.

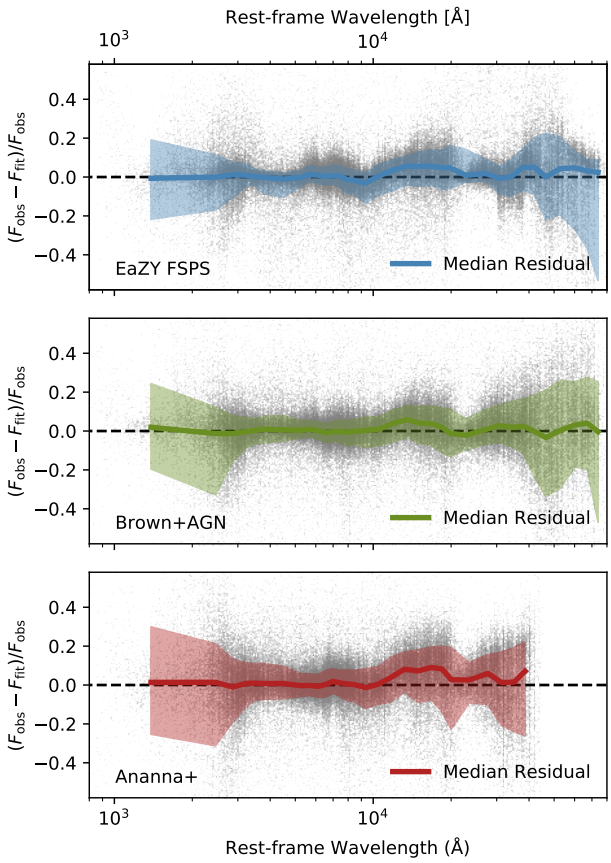

Fig. 4. Rest-frame residuals for the template fits using each of the three SED libraries employed in this work. Background grey points correspond to individual data points (i.e. one for each fitted filter per source) while the solid coloured lines correspond to the median residual within a given wavelength bin and the shaded region corresponds to the 1σ distribution (16 and 84th percentiles).

The measured rest-frame residuals for each template library and the corresponding 1σ ranges are illustrated in Fig. 4. The resulting template error functions for each library are then shown in Fig. 5. Consistent with previous measurements of the photo-z template errors (Brammer et al. 2008), we find that rest-frame UV and mid-IR wavelengths have the highest model uncertainties. We note however that model uncertainties estimated by this method are representative of the average uncertainty across the full template set given the spectroscopic sample available. As demonstrated by Brown et al. (2019), the model uncertainties for known AGN samples can be substantially larger at certain rest-frame wavelengths. However, a full Bayesian template fitting framework that can account for variance associated with each individual template (see e.g. Leistedt et al. 2019) is not practical for this work.

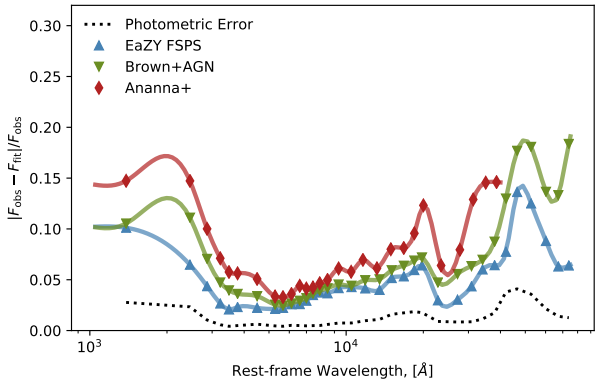

Fig. 5. Rest-frame model uncertainties derived for each of the template libraries used during our analysis (solid coloured lines). The model uncertainties are calculated by subtracting in quadrature the scaled average fractional error (black dotted line) from the median absolute rest-frame residuals (coloured symbols). However, the contribution of the average error is negligible.

In our implementation of the rest-frame model uncertainties for this study, we also identify and correct for a subtle but important statistical error in the version of the Brammer et al. (2008) photo-z code used. Specifically, by default EAZY simply stores the best-fit χ^2 at each step in the chosen redshift grid ($\hat{\chi}^2(z)$). The redshift likelihood is then typically taken as $\propto \exp(-\hat{\chi}^2(z)/2)$ (e.g. Dahlen et al. 2013; Finkelstein 2014), which may then be convolved with a redshift prior (e.g. a magnitude prior). However, when redshift-dependent model errors are included in the fitting, the simplifying assumption of dropping the normalisation in the Gaussian likelihood term is no longer valid since it is no longer a constant. This has the effect of the $P(z)$ becoming biased towards redshift ranges where the template error itself is maximised. For bright sources with strong colour features, this has minimal effect, but for fainter sources this can significantly bias the $P(z)$. Correctly accounting for the redshift-dependent model errors is trivial, with the maximum likelihood at a given redshift, $\hat{\mathcal{L}}(z)$, being:

$$\hat{\mathcal{L}}(z) = \prod_i^N \frac{1}{\sqrt{2\pi\sigma_i(z)}} \exp\left(-\frac{(f_{i,\mathrm{obs}} - \hat{f}_{i,\mathrm{t}}(z))^2}{2\sigma_i(z)^2}\right), \quad (1)$$

where for a given filter, i, $\sigma_i(z)$ is the total error (i.e. $\sigma_i(z)^2 = \sigma^2_{i,\mathrm{obs}} + \sigma^2_{i,\mathrm{model}}(z)$), $f_{i,\mathrm{obs}}$ the observed flux and $\hat{f}_{i,\mathrm{t}}(z)$ the model flux of the best-fitting template at that redshift. Taking the logarithm of both sides, the log-likelihood can be separated into two terms:

$$\ln \hat{\mathcal{L}}(z) = -\frac{1}{2}\sum_i^N \ln 2\pi\sigma_i(z) - \frac{1}{2}\sum_i^N \left(\frac{(f_{i,\mathrm{obs}} - \hat{f}_{i,\mathrm{t}})^2}{\sigma_i(z)^2}\right), \quad (2)$$

where the right-hand term can be rewritten as

$$\ln \hat{\mathcal{L}}(z) = -\frac{1}{2}\sum_i^N \ln 2\pi\sigma_i(z) - \frac{\hat{\chi}^2(z)}{2}, \quad (3)$$

with $\hat{\chi}^2(z)$ the best-fit χ^2 at each redshift step of the template fitting (as calculated and stored by EAZY). When deriving the full photo-z posterior, we therefore calculate the first term of the right-hand side of Eq. 3 for every source after the fact and incorporate alongside magnitude priors.

For all three fields, we use an IRAC 4.5μm prior derived from the Boötes spectroscopic sample following the method outlined in Section 5.1.1 of D18a. The use of the magnitude prior for the photo-z estimates was found to improve overall statistical performance at $z < 2$ (reduced scatter), the key area of interest for deep fields.

3.2. Gaussian process redshift estimates

Following previous iterations of our hybrid redshift methodology, we incorporate machine learning estimates derived using the Gaussian process redshift code GPz (Almosallam et al. 2016b,a). As in D18b, when training the GPz classifiers, we employ the magnitude and colour-based weighting scheme (based on the method presented in Lima et al. 2008) to benefit from GPz's cost-sensitive learning features. We train GPz using 25 basis functions and allowing variable covariances for each basis function (i.e. the 'GPVC' of Almosallam et al. 2016b). Finally, we also follow the practices outlined in Section 6.2 of Almosallam et al. (2016b) and allow pre-processing of the input data to normalise or decorrelate the features (also known as 'sphering' or 'whitening').

A key change to our method from previous implementations is the choice of input magnitudes. While the use of standard logarithmic magnitudes has produced good photo-z results in our previous efforts (D19), they are however not suited to datasets containing large numbers of non-detections (i.e. flux measurements with S/N < 2) or low S/N measurements. In multi-wavelength forced photometry catalogues such as those employed here where depth varies significantly between filters (Fig. 1), non-detections in individual bands are inevitable and therefore a number of sources may have zero or negative flux measurements and hence are undefined in standard logarithmic magnitudes. In many cases, most notably high-redshift galaxies, those flux limits provide valuable colour (and hence redshift) information.

We therefore make use of *asinh* magnitudes (Lupton et al. 1999, sometimes referred to as 'luptitudes') for GPz analysis, since they are able to incorporate zero or negative flux measurements and therefore allow us to train using all available measurements.[5] For a given flux, f (with a flux zeropoint $f_0 = 3631$ Jy), we define the *asinh* magnitudes as

$$m = \frac{-2.5}{\log(10)} \times \sinh^{-1}\left(\frac{f/f_0}{2b}\right) + \log(b), \tag{4}$$

where the softening parameter, b, for each band is derived from the median flux uncertainty on $\approx 5\sigma$ sources across the field (specifically, $4.95 - 5.05\sigma$).[6] For high S/N measurements, the choice of flux zeropoint ensures that the *asinh* magnitudes are equal to traditional AB magnitudes. As demonstrated in Buchs et al. (2019), photo-z estimates from *asinh* magnitudes are typically not sensitive to the details of the softening parameter used if the data are of similar depth across the field. For individual photometric bands within the deep fields optical datasets, this assumption is valid due to the relatively homogeneous depth within our datasets (the heterogeneity within the fields is largely a result of different spatial coverage of different bands).

Due to the differing range and sources (i.e. telescopes or surveys) of optical to mid-IR photometric imaging across the three fields, GPz must be trained separately for each of the corresponding spectroscopic datasets. The primary GPz classifier for each field is trained on optical sources that do not satisfy any of the AGN selection criteria - corresponding to the significant majority of both the training sample and photometric catalog. In Boötes, we also use additional GPz classifiers trained on the IR, optical and X-ray selected AGN subsets. However, due to the lack of comparable quality X-ray imaging, in the EN1 and LH fields we use additional IR and optical AGN GPz estimates only. As illustrated in D18a, sources can satisfy multiple AGN selection criteria. The resulting spectroscopic training samples for the different AGN selections are as follows: the IR, Optical and X-ray AGN sample sizes are 1936, 1750 and 1307 (2714 in total) for Boötes, while for EN1 and LH the IR and Optical AGN sample sizes are 346 and 563 (593), and 579 and 1152 (1182), respectively. For all training, each input sample was split at random into training, validation and test samples consisting of 80% (training), 10% (validation), and 10% (test) of the full sample, respectively.

The photometric bands used for GPz training for each subset and field are listed in Table 2. The exact choice of bands used was based on balancing the maximum number of bands (and hence information) with the fraction of each field covered by the respective bands. We note that the relative merit of including specific bands in the GPz estimates varies between fields due to the variation in depth and coverage of different bands (for example, less improvement was gained by including IRAC 5.8 and 8.0μm photometry for the galaxies sample in EN1 - see Table 2).

3.3. Calibration of photo-z uncertainty

As outlined in Section 3.1.2, a key goal of the modifications to our template fitting method is to incorporate model uncertainties. Nevertheless, additional calibration of the uncertainties on the resulting photo-z is still necessary. To quantify the over- or under-confidence of our photo-z estimates, we follow the method outlined in Section 3.3.1 of D18b (and originally proposed by Wittman et al. 2016) and calculate the distribution of threshold credible intervals, c, where the spectroscopic redshift intersects the redshift posterior. For perfectly accurate estimates of the uncertainties, the cumulative distribution of credible intervals, $\hat{F}(c)$, should follow a straight 1:1 relation, i.e. a quantile-quantile (or $Q - Q$) plot. Curves that fall below this 1:1 relation indicate that there is overconfidence in the photo-z errors (i.e. the $P(z)$s are too sharp) while curves that fall above indicate under-confidence.

Following D18a, we scale the uncertainties on the template fitting estimate for a source i, such that

$$P(z)_{\mathrm{new},i} \propto P(z)_{\mathrm{old},i}^{1/\alpha(m_i)} \times P(z|m_i), \tag{5}$$

where $\alpha(m)$ is a magnitude-dependent function following:

$$\alpha(m) = \begin{cases} \alpha_\eta & m \leq m_\eta \\ \alpha_\eta + \kappa \times (m - m_\eta) & m > m_\eta, \end{cases} \tag{6}$$

with $\alpha(m)$ being constant (α_η) below some characteristic apparent magnitude, m_η, and following a simple linear relation above this magnitude (Ilbert et al. 2009). For Boötes,

[5] We reiterate that magnitudes are used only for GPz estimates. Template fitting is performed using flux measurements.

[6] The corresponding magnitude uncertainties are then given by $\sigma_m = \frac{-2.5}{\log(10)} \times \frac{(\sigma_f/|f|)}{\sqrt{(1+(2b/(f/f_0))^2)}}$.

Table 2. Photometric bands used for GPz estimates in each field. See Kondapally et al. (2020) for details on depths and spatial coverage. For EN1 and LH, IRAC photometry is taken from SWIRE only.

Subset	Boötes	EN1	LH
Galaxies	u, $B_{\rm w}$, R, I, J, H, $K_{\rm s}$, 3.6μm, 4.5μm, 5.8μm, 8.0μm	u, $g_{\rm PS}$, $r_{\rm PS}$, $i_{\rm PS}$, $z_{\rm PS}$, $y_{\rm PS}$, J, K, 3.6μm, 4.5μm	u, g, r, z, 3.6μm, 4.5μm, 5.8μm, 8.0μm
AGN (IR, Opt, X-ray*)	u, $B_{\rm w}$, R, I, J, H, $K_{\rm s}$, 3.6μm, 4.5μm, 5.8μm, 8.0μm	u, $g_{\rm PS}$, $r_{\rm PS}$, $i_{\rm PS}$, $z_{\rm PS}$, $y_{\rm PS}$, J, K, 3.6μm, 4.5μm, 5.8μm, 8.0μm	u, g, r, z, 3.6μm, 4.5μm, 5.8μm, 8.0μm

EN1, and LH we use the I, $r_{\rm PS}$ and r-band optical *asinh* magnitudes, respectively for calculating the magnitude dependence of the error scaling and assume a characteristic magnitude of $m_\eta = 16$. The parameters α_η and κ are then fit using the EMCEE Markov chain Monte Carlo fitting tool (MCMC; Foreman-Mackey et al. 2013) to minimise the Euclidean distance between the measured and ideal $\hat{F}(c)$ distributions.

When calibrating the uncertainties produced by GPz, we calculate the threshold credible interval following $c_i = \mathrm{erf}\left(|z_{i,\rm spec} - z_{i,\rm phot}|/\sqrt{2}\sigma_i\right)$ (D19) and scale the uncertainties as a function of magnitude in the same manner as for the template estimates:

$$\sigma_{{\rm new},i} = \sigma_{{\rm old},i} \times \alpha(m_i), \tag{7}$$

where $\alpha(m_i)$ follows the same functional form as Eq. 6. As above, the parameters for $\alpha(m_i)$ are optimised through MCMC minimisation of the difference between the measured and ideal $\hat{F}(c)$ distributions.

As highlighted in Fig. 3, the available spectroscopic sample is biased towards bright optical magnitudes. Even though our photo-z uncertainties are calibrated as a function of magnitude, the biased sample could result in improved photo-z uncertainty accuracy for the brightest sources at the expense of the faint population. In order to prevent the optimisation of the photo-z uncertainties being dominated by the most populous magnitude ranges, we calibrate the uncertainties using a magnitude balanced subset of the total spec-z population following the approach presented in D19. For both the AGN and galaxy samples separately, a subset of each training sample is created by randomly selecting up to 750 sources in magnitude bins (of width = 1 mag) over the range covered by the spectroscopic redshift subset. Calibration of the uncertainties is then done on two-thirds of this subsample, with the other one-third retained for testing. We note that this balancing of the spec-z sample is most important for the Boötes field where the larger samples of bright magnitude selected spec-zs could significantly bias the optimisation.

3.4. Hierarchical Bayesian combination

To produce the final consensus redshift prediction for a given source, we use the HB combination method outlined by D18a (based on the method presented in Dahlen et al. 2013) and subsequently extended to hybrid GPz + template estimates in D18b. In summary, the HB combination produces a consensus redshift prediction, $P(z)$ from a set of n individual predictions while marginalising over the probability that any individual $P(z)$ is incorrect. Hyperparameters for the fraction of measurements that are bad, $f_{\rm bad}$, and the relative covariance between the different estimates β, are optimised using training data to ensure that the posterior redshift distributions more accurately represent the redshift uncertainties.

As in D19, GPz estimates are evaluated on the same redshift grid as used during the template fitting procedure. If a source does not have a photo-z estimate for a given GPz estimator (either through not satisfying the selection criteria for a given subset or lack of observations in a required band) it is assumed to have a flat redshift posterior for that specific estimator. GPz therefore contributes no additional information to the consensus HB estimates for these sources.

For the application in this work, we assume $0 \leq f_{\rm bad} \leq 0.05$ and $0 \leq f_{\rm bad} \leq 0.2$ for the galaxy and AGN subsets, respectively, and a flat prior on the redshift distribution for 'bad' estimates. However, one further improvement to our method in this work is that we optimise the hyperparameter β (i.e. the degree of covariance between different photo-z estimates) as a function of magnitude. This change is necessary to ensure that the resulting consensus photo-z posteriors provided accurate uncertainties across all magnitude ranges. Assuming a constant β for all magnitudes results in underestimates of the uncertainties for the brightest and faintest sources.

In the following section, we present analysis of the consensus photo-z estimates for each field.

4. Photometric redshift properties

In this work we have estimated photo-zs for all sources within the optical catalogues for which estimates could be made, including parts of the fields with very limited photometric coverage (including regions outside the area where radio to optical cross-identifications have been performed in Paper III). However, due to the heterogeneous photometry coverage within the deep fields, the resulting photo-z quality will also vary across the field. In the following analysis we examine the photo-z quality within the core regions of each field where photometric data is relatively homogeneous. These areas broadly conform to those where GPz estimates are available and where radio source cross-matching has been performed for the LoTSS Deep Field data (Paper III). Specifically, we apply the following cuts in each field based on the flags in the input photometric catalogues of Paper III: In Boötes, we require that `FLAG_DEEP` $\neq 0$ to remove duplicate or masked sources from the I-band detected catalogue sources. In EN 1, we require `FLAG_OVERLAP` ≥ 6 to restrict the analysis to the region that includes both NIR and *Spitzer* mid-IR photometric coverage. Due to the smaller coverage of the NIR data relative to the optical and mid-IR coverage, in LH we do not require NIR data when analysing the photo-z in this field - restricting the analysis to regions with SpARCS optical coverage

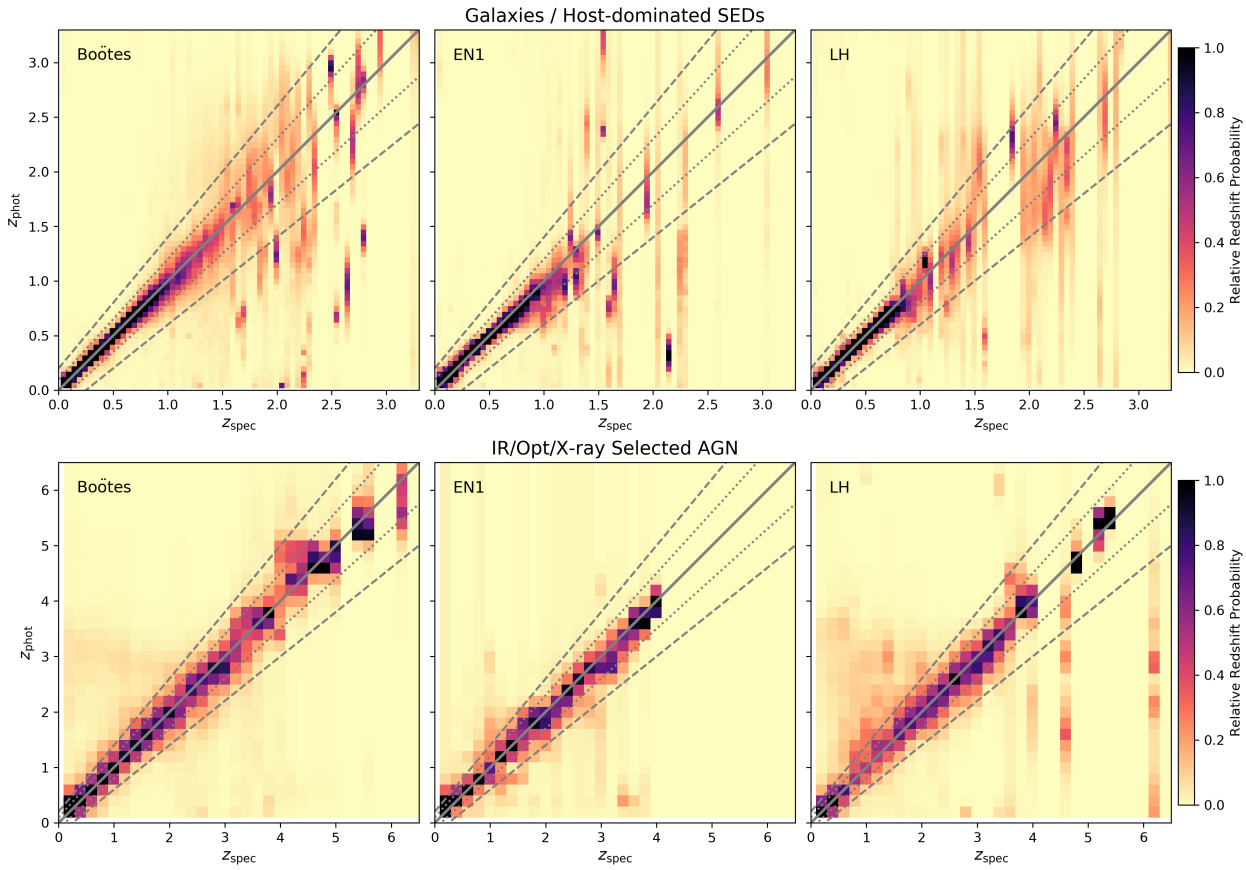

Fig. 6. Stacked redshift posterior estimates for the galaxy (or host-dominated) population (top) and for the combined AGN selected population (bottom; IR, X-ray or optically selected) for each of the three fields. The spectroscopic redshift samples available become particularly sparse at $z > 1$ for the EN1 and LH fields. Grey dotted and dashed line correspond to $\pm 0.1\times(1+z_{\rm spec})$ and $\pm 0.2\times(1+z_{\rm spec})$, respectively. The colour scale is illustrative and represents the summed $P(z)$ (normalised such that $\int P(z)\mathrm{d}z = 1$) for each $z_{\rm spec}$ bin, more precise photo-z therefore result in darker peaks.

(FLAG_OVERLAP ≥ 6) which overlaps with the *Spitzer* imaging.

Additionally, in all fields we require FLAG_CLEAN = 1 to exclude sources within the optical bright star mask. Compared to the initial combined total of $\approx 7.2\times10^6$ catalogue sources available across the three fields, these cuts reduce the number of sources with the most robust photo-z estimates to a combined $\approx 5\times10^6$ sources.

Fig. 6 presents a qualitative illustration of the final consensus redshifts after the error calibration for all input estimates and the tuning of the Bayesian combination hyperparameters. We show the stacked redshift posteriors as a function of spectroscopic redshift for each field, with the spectroscopic samples separated into the galaxy or host-dominated (top row) and AGN subsets (bottom row). From Fig. 6 we can see that the photo-z quality for galaxies at $z < 1$ is excellent for all three fields. At higher redshifts, the spectroscopic samples become much more limited (especially outside of Boötes). However, from the limited number of spec-zs available in this regime, it is clear that the performance deteriorates between $1 \lesssim z \lesssim 1.5$ for the host-dominated population, and gets substantially worse beyond $z \sim 1.5$. Examining the AGN population posteriors, we find that photo-z performance looks generally good out to $z > 3$ in all fields - albeit with significantly larger scatter than the host-dominated population. Few spectroscopically confirmed sources at $z > 4$ exist in the EN1 and LH but we find that our photo-z estimates in Boötes provide accurate estimates out to $z > 6$ (including the confirmed radio-loud quasar at $z = 6.1$ McGreer et al. 2006). Given that hybrid photo-z estimates in this regime are found to be dominated by the template estimates (see D18b), we would expect the performance in the EN1 and LH datasets to therefore be comparable in quality (or potentially better in the case of EN1). Nevertheless, we caution against the use of the photo-zs above $z \sim 4$ without careful analysis of the individual estimates.

4.1. Overall photo-z statistics

The $z_{\rm spec}$ versus $z_{\rm phot}$ distribution shown in Fig. 6 allows us to qualitatively assess the photo-z performance. However, a more quantitative analysis is required to enable comparison between fields and for defining appropriate selection criteria for future science exploitation. When calculating photo-z statistics, we use the median of the primary redshift peak ($z_{\rm 1,median}$) following D19: see Section 4.1 of that paper for details on how this value is defined. For our measure of robust scatter, we then use the normalised median absolute deviation, $\sigma_{\rm NMAD}$, defined as:

$$\sigma_{\rm NMAD} = 1.48 \times \mathrm{median}(|\Delta z| / (1 + z_{\rm spec})), \quad (8)$$

Table 3. Photo-z quality statistics for the galaxy and AGN spectroscopic redshift samples in each deep field. The samples (of size N) are cut to have $z_{\rm phot} < 1.5$ for the galaxy subset and $z_{\rm phot} < 4$ for the AGN subset to allow a more direct comparison between fields. The robust scatter, $\sigma_{\rm NMAD}$, and OLF, are defined in Equations 8 and 9, respectively.

Galaxies / Host-dominated			
	N	$\sigma_{\rm NMAD}$	OLF
Boötes	15200	0.016	0.018
ELAIS-N1	2570	0.02	0.016
Lockman Hole	2743	0.017	0.015
	AGN		
	N	$\sigma_{\rm NMAD}$	OLF
Boötes	2215	0.07	0.177
ELAIS-N1	442	0.064	0.224
Lockman Hole	701	0.077	0.223

where $\Delta z = z_{\rm 1,median} - z_{\rm spec}$. Similarly, we define the OLF as sources where

$$|\Delta z| / (1 + z_{\rm spec}) > 0.15, \quad (9)$$

as is common in the literature (e.g. Dahlen et al. 2013).

In Table 3 we present the overall $\sigma_{\rm NMAD}$ and OLF statistics derived for the galaxy and identified AGN subsets. To ensure a fair comparison across the three fields, for the galaxy subset we limit the statistics to those sources with $z_{\rm 1,median} < 1.5$ while for the AGN subsample we limit to $z_{\rm 1,median} < 4.0$. We note that these cuts are defined using $z_{\rm 1,median}$ (as opposed to $z_{\rm spec}$) to better represent the performance expected for science samples that will be defined based on their photo-z alone.

We find that all three fields exhibit very similar $\sigma_{\rm NMAD}$ and OLF metrics for both subsets, with less than 2% scatter and OLF for galaxies or host-dominated sources. As already seen in Figure 6 the photo-z quality for the AGN subset is significantly worse, with OLF reaching $\approx 20\%$ or higher. Since the available spectroscopic redshifts are biased to typically brighter galaxies and lower redshifts, these statistics are not fully representative of the true photo-z accuracy and reliability for the average galaxy population.

In Fig. 7 and 8 we present the resulting robust scatter and OLF as a function of redshift and magnitude, respectively. The resulting statistics provide a quantitative confirmation of the photo-z quality visible in Fig. 6, where we find that our photo-z estimates have excellent scatter $< 0.04\times(1+z)$ and OLF $\sim 5\%$ for the galaxy/host-dominated population at $z < 1$ across all three fields. For all three fields, there is a deterioration in the photo-z scatter and OLF with increasing redshift. However, in LH and EN1, not enough spectroscopic redshifts are available above $z \sim 1$ redshift to assess the quality beyond this point. In Boötes, where spec-zs are available in this regime, we find that the trend in the statistics verifies the visual interpretation of Figure 6 - whereby the measured scatter deteriorates above $z \sim 1$ before becoming substantially worse at $z > 1.5$.

In contrast to the strong correlation with redshift observed for the galaxy subset, the measured $\sigma_{\rm NMAD}$ for the identified optical, IR and X-ray AGN samples exhibits negligible evolution with redshift. We also find that the OLF appears to decline as a function of redshift. This trend is likely driven by biases within the spectroscopic sample (with the $z > 2$ training and test samples are typically dominated by optically bright QSOs) and the presence of the strong Lyman break feature redshifting into the u-band at $z \sim 3$, leading to improved template fitting estimates (see also D18b, for further discussion).

We find that the measured OLF for known AGN at a given redshift or 4.5μm magnitude in EN1 and LH is higher than for Boötes. We note that this difference is likely driven by the demographics of the respective spectroscopic redshift samples available. The EN1 and LH spec-z samples are dominated by QSOs while the Boötes sample includes a large number of sources explicitly selected on X-ray or IR AGN criteria. In D19, photo-z performance for AGN selected as optical QSOs was measured to be significantly worse than for those selected based on the other AGN criteria.

4.1.1. Comparison with existing literature photo-z estimates

In D18b we compared the quality of photo-z produced by our hybrid methodology with existing estimates in the literature, finding that our results produced better statistics than redshifts presented by Brodwin et al. (2006), with the substantial benefits of the accurate redshift posteriors. Thanks to the improved template fitting methodology in this study, the results presented in this data release represent a further improvement in photo-z quality available for this field. Specifically, when calculating photo-z quality statistics for the same subset of spectroscopic sources as used in Table 3, the photo-z presented in D18b have $\sigma_{\rm NMAD} = 0.032$ and OLF= 0.028 for galaxies and $\sigma_{\rm NMAD} = 0.12$ and OLF= 0.31 for the known AGN. The results presented in this paper therefore represent up to a factor of ~ 2 improvement in redshift quality averaged over the available spectroscopic population (we note that the degree of improvement in parameter space not probed by the existing spec-z sample could differ substantially from this, both positively and negatively).

In the EN1 and LH fields, there are a number of recently published photo-z catalogues to which we can directly compare our results. As part of the Hyper Suprime-Cam Subaru Strategic Program Second Public Data-release (HSC PDR2; Aihara et al. 2019), photo-z estimates derived from the HSC optical photometry (g,r,i,z,y) are available for the EN1 field for much of the area covered by the LoTSS Deep Field optical catalogues. The HSC PDR2 photo-z estimates presented by Nishizawa et al. (2020, see also Tanaka et al. 2018) are based on two different approaches: empirical estimates using the 'Direct Empirical Photometric method' (DEmP Hsieh & Yee 2014), and template based estimates following the method presented in (Tanaka 2015, Mizuki).

To directly compare our results with the HSC PDR2 estimates, we cross-match the catalogues with a maximum separation of 1″and calculate the photo-z quality statistics for all three estimates (this work, HSC DeMP and HSC Mizuki) for the subset of spectroscopic sources with measurements in both our results and at least one of the HSC catalogues. The resulting statistics, presented in Table 4, demonstrate that the estimates presented in this work are again comparable to or better than the existing literature. While the overall statistics for the galaxy/host-dominated AGN population offer only a small improvement on the excellent scatter and OLF of the HSC estimates, our estimates for the known AGN population offer significantly better precision and reliability. This improvement is even starker when restricting the analysis to the higher redshift AGN

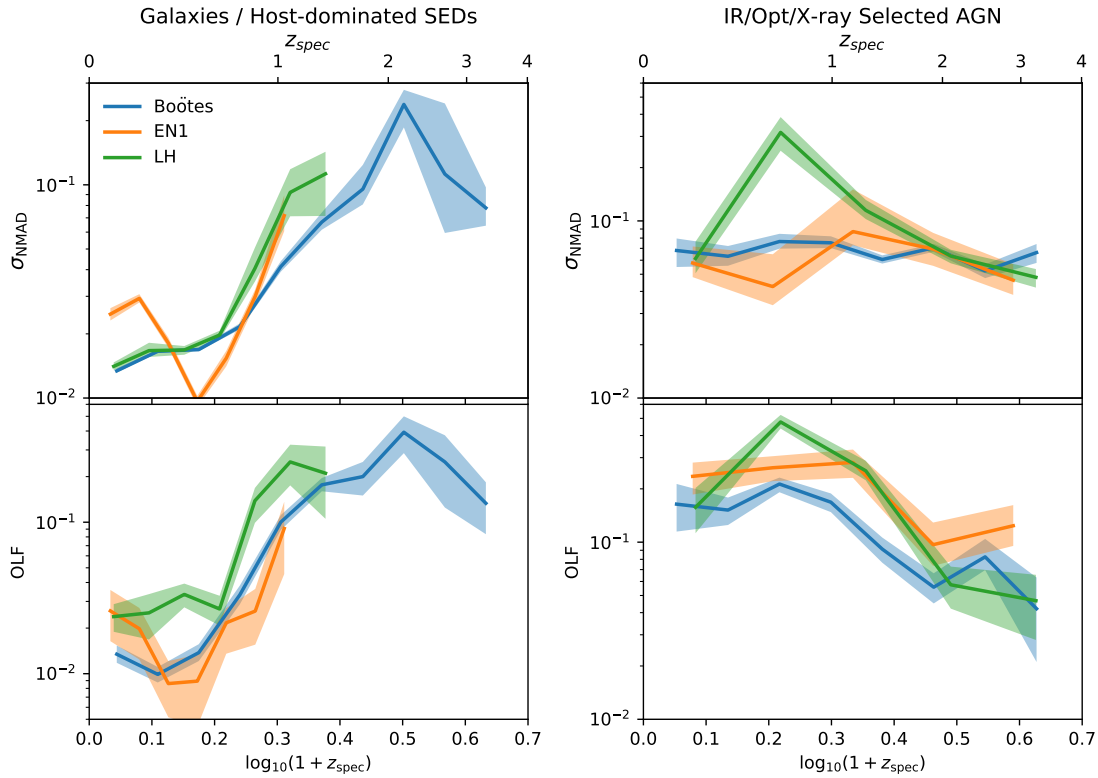

Fig. 7. Robust scatter ($\sigma_{\rm NMAD}$; upper panels) and OLF (lower panels) for the consensus photo-z estimates ($z_{\rm 1,median}$) as a function of spectroscopic redshift. Shaded regions illustrate the statistical uncertainties on the respective metrics derived from bootstrap resampling.

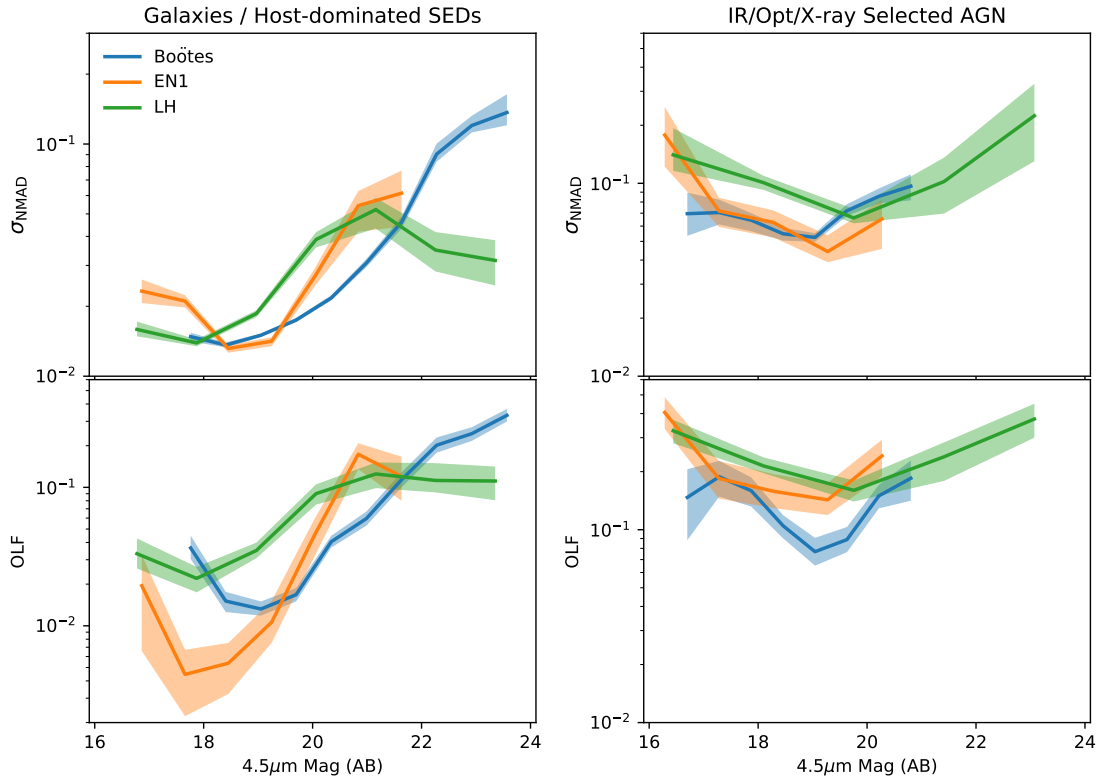

Fig. 8. Robust scatter ($\sigma_{\rm NMAD}$; upper panels) and OLF (lower panels) for the consensus photo-z estimates ($z_{\rm 1,median}$) as a function of 4.5 μm magnitude. Shaded regions illustrate the statistical uncertainties on the respective metrics derived from bootstrap resampling. The magnitude range plotted for each field is representative of the magnitude distribution of the spectroscopic sample and is not indicative of the relative photometric depths (see Fig. 1).

Table 4. Photo-z quality statistics for the galaxy and AGN spectroscopic redshift samples in EN1 in comparison to literature values from the HSC PDR2 (Nishizawa et al. 2020). The samples (of size N) are cut to have $z_{\rm phot} < 1.5$ for the galaxy subset and $z_{\rm phot} < 4$ for the AGN subset as in Table 3 and are also limited . For the AGN subset, we also show in parentheses the statistics corresponding to limited to sources at $z_{\rm spec} > 1$.

Galaxies / Host-dominated			
	N	$\sigma_{\rm NMAD}$	OLF
This paper	2439	0.019	0.016
HSC DeMP	2439	0.019	0.027
HSC Mizuki	2414	0.025	0.031
AGN ($z > 1$)			
	N	$\sigma_{\rm NMAD}$	OLF
This paper	420 (247)	0.058 (0.058)	0.193 (0.158)
HSC DeMP	420 (247)	0.069 (0.144)	0.317 (0.445)
HSC Mizuki	318 (164)	0.09 (0.188)	0.349 (0.457)

population ($z > 1$, for which we provide statistics in Table 4 in parentheses), where our results have lower $\sigma_{\rm NMAD}$ and OLF by up to a factor of ~ 3 compared with those available from HSC PDR2.

Additionally, Pforr et al. (2019) present template based photo-z estimates for 18 deg^2 of multi-wavelength optical photometry in the Spitzer Extragalactic Representative Volume Survey (SERVS; Mauduit et al. 2012) over five fields, including subsets of the EN1 and LH fields. Pforr et al. (2019) quote robust scatter and outlier statistics calculated using the same definitions as in this paper for their full spectroscopic test sample (and the subset with the best available photometry), finding $\sigma_{\rm NMAD} = 0.042\,(0.037)$ and OLF $= 0.105\,(0.028)$ in the EN1 field and $\sigma_{\rm NMAD} = 0.067\,(0.03)$ and OLF $= 0.205\,(0.048)$ in LH. As the photo-z estimates in Pforr et al. (2019) are primarily designed for galaxies (with no AGN templates included in the fitting), we can compare these values with the statistics for the galaxy population presented in Table 3. It is clear that the results presented in this work represent a substantial improvement in overall photo-z quality.

Our improved photo-z precision and reliability for the EN1 and LH fields highlights the benefits of not just our photo-z methodology but also the advantages of the full aperture matched photometry across the UV to mid-IR regime provided by Paper III. In particular, the availability of NIR and mid-IR photometry is crucial for providing reliable estimates for the AGN population at higher redshifts (a key area of scientific interest for the LoTSS Deep Fields). Furthermore, the additional colour information provided by forced photometry across all available photometric bands is vital, offering huge improvements in photo-z estimates or SED fitting when compared to the cross-matched catalogues employed by Pforr et al. (2019, see also Nyland et al. 2017).

4.1.2. Photo-z properties for the LOFAR population

Finally, we explore the quality of the consensus photo-z estimates as a function of their radio properties. Due to the larger spectroscopic sample required to bin in multiple properties, we explore the statistics only in the Boötes field. However, given the similar performance and redshift or magnitude trends observed across the three fields we would expect any observed trends to hold for all fields. Figure 9 presents the $\sigma_{\rm NMAD}$ and OLF as a function of both spectroscopic redshift and 150MHz radio luminosity. When converting from observed flux density to rest-frame radio luminosity, we assume an average spectral slope of $\alpha = -0.7$ for all sources - consistent with the typical slope observed in previous studies (Calistro Rivera et al. 2017)

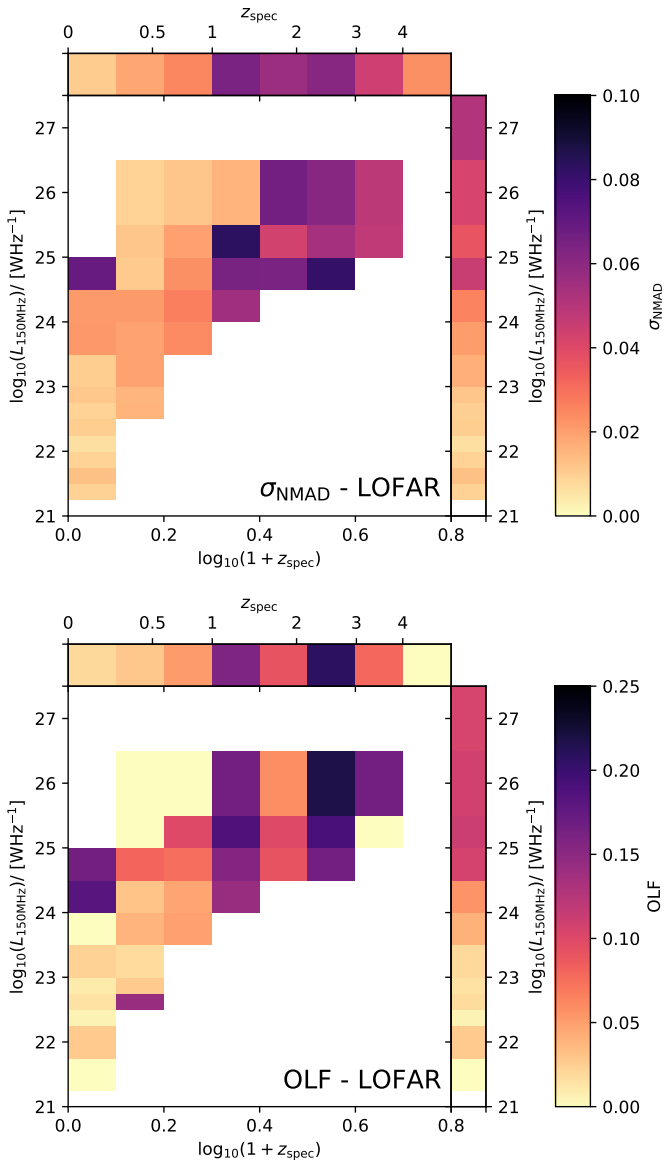

Fig. 9. Robust scatter ($\sigma_{\rm NMAD}$; top) and OLF (bottom) for the consensus photo-z estimate as a function of spectroscopic redshift and 150 MHz radio continuum luminosity in the Boötes field. The top and side panel show the trends averaged over all redshifts and luminosities, respectively. For a cell to be plotted we require a minimum of five galaxies - for some redshifts (or luminosities) we only have the number statistics available to plot the statistics averaged over all luminosities (or redshift).

In D19 we observed a clear evolution in the scatter and OLF of the radio source population with $z_{\rm spec}$, whereby the photo-z properties of the highest redshift sources are significantly worse than for sources with similar radio luminosity at low redshift. However, within a given spectroscopic redshift bin, we found no evidence for any significant trend with radio luminosity in either the scatter or OLF.

Although exhibiting a noisier evolution than observed in D19, our photo-z estimates follow the same overall trend as a function of redshift and radio power. Specifically, for a fixed redshift we do not observe a strong correlation in $\sigma_{\rm NMAD}$ or OLF as a function of radio luminosity. Averaged over all redshifts, the measured scatter and OLF increase with increasing radio power. However, this trend is clearly driven by the redshift distribution of spectroscopic sources available at a given radio power.

4.2. Accuracy of the photo-z uncertainties

In Fig. 10 we illustrate the accuracy of the final calibrated redshift posteriors for the AGN and galaxy subsets. Shown in both plots are the cumulative distribution ($\hat{F}(c)$) of threshold credible intervals, c, for bins of apparent optical magnitude (coloured lines). As outlined in Section 3.3, uncertainty calibration was performed as a function of I band for Boötes and the deepest available r-band for EN1 and LH. For all three LoTSS Deep fields, the uncertainties for the galaxy spectroscopic sample are very well calibrated, with the measured uncertainties lying close to the desired 1:1 relation across the full magnitude range.

For the AGN population, the uncertainties are also generally very well calibrated across the available magnitude ranges. However, in Boötes we find that the wings of our photo-z posteriors overestimate the true uncertainties for the very brightest ($I < 19$) and faintest ($I > 23$) AGN population. In these regimes the available spectroscopic training sample is much more limited and the optimisation of the uncertainties is therefore still dominated by the more numerous AGN types. When compared to the results obtained in D19, we find that the photo-z uncertainties presented in this work are more accurate (with $\hat{F}(c)$ closer to the desired 1:1 trend). However, we caution that this does not necessarily mean the photo-z posteriors at a given magnitude are more precise.

As with the training of the machine learning estimates, we note that the large spectroscopic training samples available in the forthcoming WEAVE-LOFAR spectroscopic survey will allow for significant improvements in the calibration of the photo-z uncertainties in future studies. Nevertheless, the analysis presented in this section verifies that the photo-zs provided for the LoTSS Deep Field release are of a high quality and suitable for scientific exploitation.

5. Stellar mass estimates

Thanks to the extensive panchromatic photometry available in each field (including deblended photometry for mid-IR to FIR), detailed physical properties can be derived for the LOFAR source population through SED fitting. Future papers will exploit this information to derive robust source classifications, AGN accretion modes and star-formation rates for the faint radio source population, however running such codes for the full optical catalogues ($> 10^6$ sources) is impractical. Nevertheless, estimating the stellar masses of the full optically selected population is critical for understanding how radio AGN affects galaxy evolution. Here, we present stellar mass estimates (and rest-frame optical colours) using a simpler grid-based SED fitting approach that scales to the massive samples available across the three deep fields.

5.1. Spectral energy distribution fitting

Stellar masses are estimated using the PYTHON-based SED fitting code previously used by Duncan et al. (2014) and Duncan et al. (2019a). Composite stellar populations are generated using the stellar population synthesis models of Bruzual & Charlot (2003) for a Chabrier (2003) initial mass function (IMF), with the following assumptions:

Star-formation histories: Recent studies have shown that the widely used assumption of exponentially declining star-formation histories can lead to biases in the ages and star-formation histories derived from photometric SED fits, resulting in biased stellar mass estimates (see e.g. Pacifici et al. 2016; Carnall et al. 2019a; Leja et al. 2019, and references therein). Carnall et al. (2018) show that the more complex double power-law parametrisation provides sufficient flexibility to accurately describe the star-formation histories of a wide range of possible formation and quenching mechanisms. As a compromise between the tractability of fitting large samples and the optimal prior assumption on star-formation histories, we therefore define a grid of SFH based on the double power-law model with the priors on the range of power-law slopes and turnover ages taken from Carnall et al. (2019b). Specifically, we use a logarithmically spaced grid of 7 values from $0.1 \leq \alpha_i \leq 1000$ for both power-law slopes ($i \in \{1, 2\}$) and 7 turnover ages from 0.1 to 1 times the age at observation (totalling 343 star-formation history models for every age step).

For a given source the time since the onset of star-formation is defined as the time since a fixed formation redshift of $z_{\rm f} = 20$ at the closest redshift step in a grid from $0 < z < 1.5$.

Stellar metallicity: We assume constant stellar metallicities at fixed values of $Z \in \{0.1, 0.4, 1.0\}\, Z_\odot$.

Nebular emission: Due to the redshift range being probed and the available constraints on the rest-frame NIR SEDs of our target sample, nebular emission lines are not expected to have a significant effect on the inferred stellar masses (unlike at $z > 3$ where they can have a significant effect, e.g. Stark et al. 2013; Schenker et al. 2013). Nevertheless, a simple prescription for nebular emission is included in the model SEDs allowing for escape fractions of $f_{\rm esc} \in \{0., 0.2\}$. Details of the assumed emission line ratios for Balmer and metal lines, as well as the nebular continuum prescription can be found in Duncan et al. (2014). As in previous studies we make the simplifying assumption that the gas-phase stellar metallicity equals the stellar metallicity.

Dust attenuation: We incorporate dust attenuation following the two-component dust model of Charlot & Fall (2000), which includes contributions from birth clouds (for stellar populations younger than 10^7 years) and an additional screen component for all ages. The wavelength dependence on the extinction curve and fraction of attenuation from the ISM are fixed at $A(\lambda) \propto \lambda^{-0.7}$ and 0.3, respectively - as originally presented by Charlot & Fall (2000, to which we refer the reader for formal definitions).

After convolution of the model spectral energy distributions with the photometric filters of each field at each redshift in the grid, the model grid is fit to the observational datapoints for all optical catalogue sources with a photo-z $z_{1,\rm median} < 1.5$ – the regime for which photo-z can be considered reliable for the galaxy population. Fitting to the observed photometry is done using a simple least-

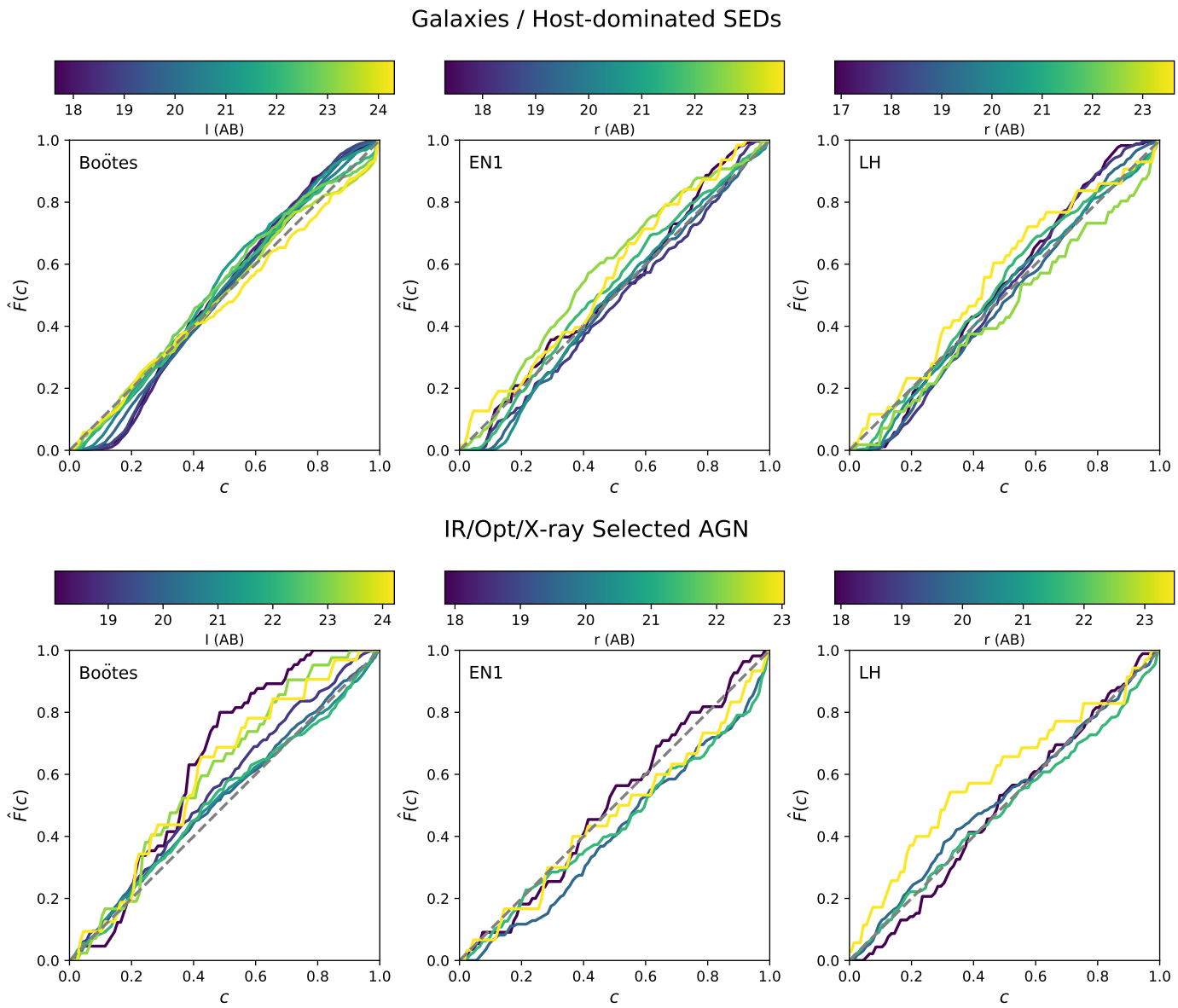

Fig. 10. Panels illustrate the cumulative distribution of threshold credible intervals, c, ($\hat{F}(c)$) plot for the final consensus photo-z estimates for the galaxy (or host-dominated AGN) population (top) and the optical, infrared and X-ray selected AGN population (bottom) within each of the deep fields. Coloured lines represent the distributions in bins of apparent optical magnitude (I-band for Bootes, r-band for EN1 and LH). Lines that fall above the 1:1 relation illustrate under-confidence in the photo-z uncertainties (uncertainties overestimated) while lines under illustrate over-confidence (uncertainties underestimated).

squares fit to all available photometric bands for a given source (see Section 2.5 of Duncan et al. 2019a, for details). However, rather than using simply the normalisation (and hence stellar mass) of only the best-fit individual model, we marginalise over the full set of stellar population parameters (star-formation history, dust attenuation, metallicity) to derive a likelihood weighted distribution for the inferred stellar mass. This marginalisation implicitly assumes a flat prior on the respective stellar population parameters. For the resulting catalog, we then take the median and 1σ (16 and 84th percentiles) of the stellar mass distribution as the estimate and corresponding uncertainty for a given source. We note however that this uncertainty does not reflect the full statistical uncertainty on the stellar masses as we do not account for the uncertainty in redshift.

In addition to the stellar mass estimate for each source, we also derive the rest-frame magnitude for all photometric filters in a given field (regardless of whether that filter was used during fitting) based on the best-fitting template. Future papers that incorporate the full panchromatic SED (including deblended 24 to 500μm observations) will provide robust estimates of the star-formation rates and other key physical properties within the LoTSS Deep Field source population.

5.2. Flux zero-point and total flux corrections

During testing and calibration of the stellar mass estimates, we found that additional flux zeropoint corrections and model uncertainties similar to those used during photo-z estimation are required for accurate and reliable mass es-

timates (see also Tanaka 2015). Offsets to the flux zero-points for stellar mass estimates are calculated following a similar approach to that applied during photo-z template fitting. The subset of optical sources with robust spectroscopic redshifts are first fit with no flux offsets applied and with no additional errors included to account for model uncertainty. Corrections to each observed filter are then calculated based on the median ratio between the observed best-fit model fluxes in that band. The sample is then re-fit incorporating the initial zero-point corrections to derive iterative changes to the corrections, with the process repeated a final time. After the third iteration of zero-point corrections, we then estimate the residual model uncertainties for the SED fitting model grid following the same method as in Section 3.1.2. The resulting corrections are included in Appendix B for reference.

When the final derived zero-point corrections are applied during fitting, we observe a significant improvement in the distribution of best-fit χ^2 values for the full photometric samples. For example in EN1 we find a median reduction per source of $\Delta\chi^2 = -0.68$, and a reduction in the average χ^2 for the whole sample by ≈ 2. Furthermore, we also find an increased agreement between the observed stellar mass functions (SMFs) between fields and with those available in the literature. We note that when fitting stellar masses using the zeropoint corrections derived from the photo-z estimates (specifically the combined template sets of Brown et al. 2014, 2019), we find no systematic offset in the resulting stellar masses - with a median offset of $\lesssim 0.03$ dex, significantly smaller than the typical uncertainty on individual estimates.

However, due to the use of fixed apertures (and aperture corrections) within the optical photometry catalogues, we find that we underestimate the mass of the most massive galaxy population at lower redshifts (which have sizes significantly larger than those of the faint galaxies or stars used for the aperture corrections). To estimate corrections from aperture to total fluxes for brighter resolved sources we cross-match the optical photometry in the deep fields to the model-fitting photometry of the Legacy Surveys (Dey et al. 2019). For sources detected in Legacy Surveys imaging, we calculate the ratio between the LOFAR optical catalogue (Paper III) aperture corrected flux and the model-fitting photometry based total flux in the available bands (g, r, z for EN1 and LH and z for Boötes). After accounting for any small global offset between the two flux measurements based on the median flux ratio for the source types used for aperture corrections, the resulting flux ratio can be used to estimate an approximate aperture to total flux conversion for each source. In the faint limit where a source has a S/N less than 5σ or is undetected in Legacy Surveys imaging, we assume that sources in the LOFAR optical catalogue are unresolved and the aperture corrected fluxes provide a reliable estimate of the total flux. In the stellar mass catalogue presented in this data release we provide the stellar mass and rest-frame magnitude values as measured, alongside total flux corrections where available.

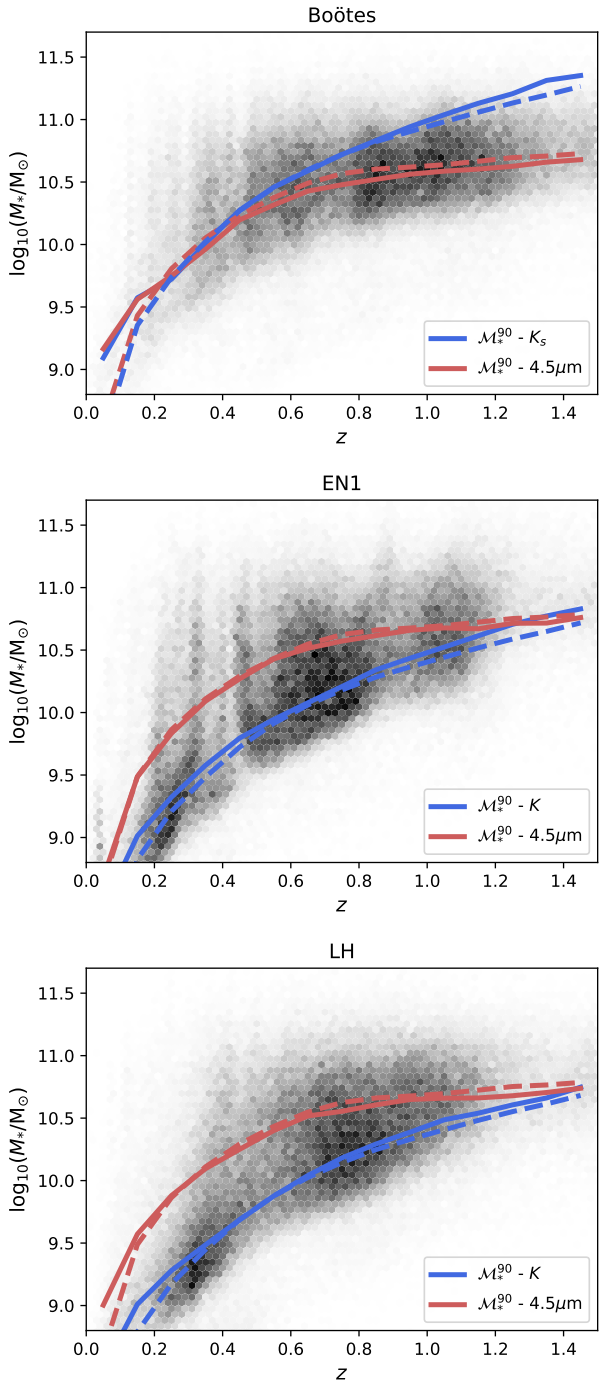

Fig. 11. Observed stellar mass distribution as a function of redshift for the three LoTSS Deep Fields. The background density plot shows the mass distribution of sources brighter than the 90% magnitude limit of the most sensitive reference band available in that field (4.5μm for Boötes and K for EN1 and LH). Solid lines represent the 90% mass completeness limits, $\mathcal{M}_\star^{90}$, derived empirically (Pozzetti et al. 2010) and dashed lines show the corresponding limits derived from model stellar populations. The blue completeness curves are derived based on the K or K_s magnitude limit, while red curves are derived from the widest area 4.5μm limit (i.e. SWIRE for EN1 and LH).

5.3. Stellar mass completeness

When performing studies of different galaxy or AGN populations across redshift and environments in flux limited surveys, understanding the associated completeness limits is essential for minimising biases and defining appropriate volume limited samples. We are therefore interested in the mass completeness limits of our surveys in order to enable reliable studies in mass-selected samples.

We estimate the detection completeness limit in the K (or K_s) and IRAC 4.5μm bands empirically by fitting a power law distribution to the observed number counts for $> 5\sigma$ sources in a regime where the catalogues are known to be complete. The 90% and 50% completeness limits are then defined as the magnitude at which the observed number counts are 0.9 and 0.5× the expected number counts predicted by the power law distributions. Although the true magnitude completeness limits will vary as a function of intrinsic source size and morphology, completeness estimates derived from simple power-law fits have been shown to be in good agreement with completeness estimates derived from detailed simulations (through injection and detection of fake sources within the imaging, see e.g. Guo et al. 2013). We estimate the 90% magnitude completeness limits to be K or Ks = 20.44, 21.78, and 21.87 for Boötes, EN1, and LH, respectively. While for the IRAC 4.5μm band we estimate 90% magnitude completeness limits of 21.33, 21.19, and 21.18 respectively, based on the SDWFS observation in Boötes and the SWIRE observations in the other two fields.

To estimate the mass completeness limit associated with the measured detection completeness limit, we derive the limiting stellar mass-to-light ratio as a function of redshift in two ways. Firstly, following the empirical method presented by Pozzetti et al. (2010), in small redshift bins we take the 20% faintest sources above the 90% magnitude completeness limit. The measured stellar masses for this sample are scaled to the magnitude limit ($\log_{10}(\mathcal{M}_{\rm lim}) = \log_{10}(\mathcal{M}) + 0.4(m - m^{90}_{\rm lim})$) and the mass completeness limit derived from 95th percentile of the scaled $\mathcal{M}_{\rm lim}$ mass distribution. This measurement is repeated for both the observed stellar masses and K (K_s) and IRAC 4.5μm magnitude limits in each field.

Our second derivation of the mass completeness limit is based on a maximally old stellar population representative of the highest mass-to-light ratio expected within the sample. Using the same stellar populations models as used for SED fitting, we model a dust-free exponentially declining star-formation history with an e-folding time of 50 Myr and a formation redshift of $z = 20$. For each redshift bin, the stellar population is convolved with the K (K_s) and IRAC 4.5μm filter response curves with the stellar population age set by the time since the formation redshift. The mass completeness limit corresponding to the maximally old population is then derived by scaling the observed template flux at each redshift bin to the corresponding magnitude completeness limit.

In Figure 11 we find that the two approaches for estimating the limiting mass-to-light ratio at each redshift yield almost identical results. Selecting based on either near- or mid-IR, all three fields are complete to $\log_{10}(\mathcal{M}^{90}_{\star}/M_{\odot}) \approx 10.7$ at $1.4 < z \leq 1.5$ - sufficient to probe below the knee of the galaxy SMF at this redshift. At $z < 1$, the deeper NIR observations provided by UKIDSS Deep Extragalactic Survey (DXS) mean that the EN1 and LH fields are complete to significantly lower masses if using K band to select samples. We note that by restricting observations to the *Spitzer* SERVS region of the EN1 and LH fields (≈ 1 magnitude deeper than SWIRE), mass complete samples at $z \sim 1.5$ could likely reach ~ 0.4 dex lower in mass at the expense of total volume probed.

5.4. Stellar mass validation

To validate our stellar mass estimates and verify the consistency both between the three LOFAR deep fields and with the literature, we estimate the galaxy SMF in each field. After masking regions affected by bright stars and restricting to regions with both optical and NIR or mid-IR coverage, the resulting survey areas in each field are 8.63, 6.5 and 7.24 deg^2 for Boötes, EN1, and LH, respectively (we note that these areas correspond to the region where our selection criteria have been satisfied, rather than the full areas quoted in Paper III). For the purposes of this comparison, we construct our estimate of the galaxy SMF through simple calculation of the volume density in bins of stellar mass. We do not include additional corrections to weight for the reduced volume for sources near the detection limit (i.e. the classical $1/V_{max}$ method) or for the effects of Eddington bias on the high mass bins.

For each redshift range, we limit to the 90% mass completeness limit based on the deepest available band over a large area, specifically the UKIDSS DXS K for the EN1 and LH fields, and the IRAC 4.5μm band for Boötes. We exclude known IR, X-ray and optical AGN that are likely to result in poor SED fits or biased stellar mass estimates. Additionally, we apply a cut on the best-fit χ^2 to exclude sources with χ^2/N (where N equals the number of filters used during fitting) to exclude the 5% worst fits indicative of unidentified stars or AGN within the sample.

Our derived SMF estimates for four redshift bins ($0.15 < z \leq 0.3$, $0.3 < z \leq 0.5$, $0.5 < z \leq 1.0$, and $1.0 < z \leq 1.5$) are presented in Figure 12 alongside a selection of published SMFs from the literature that probe comparable redshifts (Pozzetti et al. 2007; Muzzin et al. 2013; Leauthaud et al. 2016; Wright et al. 2018; Leja et al. 2020; McLeod & et al. 2020, in prep.). We plot error bars for our binned SMF estimates that represent the uncertainty only from Poisson noise and the approximate cosmic variance based on the volume probed for each field. Our cosmic variance uncertainties are calculated following the prescription and code presented in Moster et al. (2011), with the standard deviation on the number counts at a given mass interpolated from the mass ranges produced by the code.

As noted above, we find that total flux corrections are necessary to produce a broad agreement between the three fields and with observed SMFs and those of the literature. However, due to the different assumptions used for aperture corrections between Boötes (analytic corrections based on the PSF), EN1 and LH (empirical corrections), we include total flux corrections at $z < 0.5$ for the EN1 and LH fields, while for Boötes we include corrections out to $z < 1$. We find that corrections are not required at higher redshift, with the three fields providing results that are consistent with the population becoming largely unresolved.

At $z > 0.3$, where the fields probe a large representative volume, our simple galaxy SMFs are in excellent agreement both with published literature values and with self-consistent results across all three fields. In the lowest redshift bin, $0.15 < z \leq 0.3$, we observe an increased scatter between the SMF estimates of the three different fields, with EN1 having a consistently higher normalisation than the other two fields. Given the significantly smaller volume probed in this bin ($\sim 3.4\times$ less than at $0.3 < z \leq 0.5$), larger cosmic variance between fields is to be expected. While the estimated cosmic variance uncertainties do not fully

account for the offset between fields in all redshift bins, the overall normalisation and shape of our observed SMFs are still broadly consistent with the range published in the literature.

Finally, as with the photo-zs, we can compare our stellar mass estimates on a source by source basis with those available in the literature. Specifically, we compare our EN1 masses with those presented in the HSC PDR2 (Nishizawa et al. 2020, see Sec. 4.1.1), which assume the same Chabrier (2003) IMF as used in this work. For sources with $\log_{10}(\mathcal{M}_\star/M_\odot) > 9$ in our catalog, we find a median offset of 0.09 dex between our stellar masses and those of the HSC-DeMP (Hsieh & Yee 2014) methodology, with a corresponding robust scatter of 0.28 dex. For the HSC-Mizuki template based estimates (Tanaka 2015), we find a median offset of just 0.01 dex and robust scatter of 0.32 dex. We note that the typical difference between our estimates and either of the HSC estimates is smaller than the difference between the two HSC estimates themselves when calculated for the same galaxy sample (a median offset of 0.14 dex).

Given the good agreement and consistency between our three fields at the redshifts of interest, we are therefore confident, that with appropriate sample selection and quality cuts, our stellar mass estimates are suitable for robust quantitative studies across all three fields. Furthermore, the large samples and volumes probed by the combined deep fields dataset over a wide range of cosmic history offer reliable reference samples for studying radio properties as a function of history, mass, and environment.

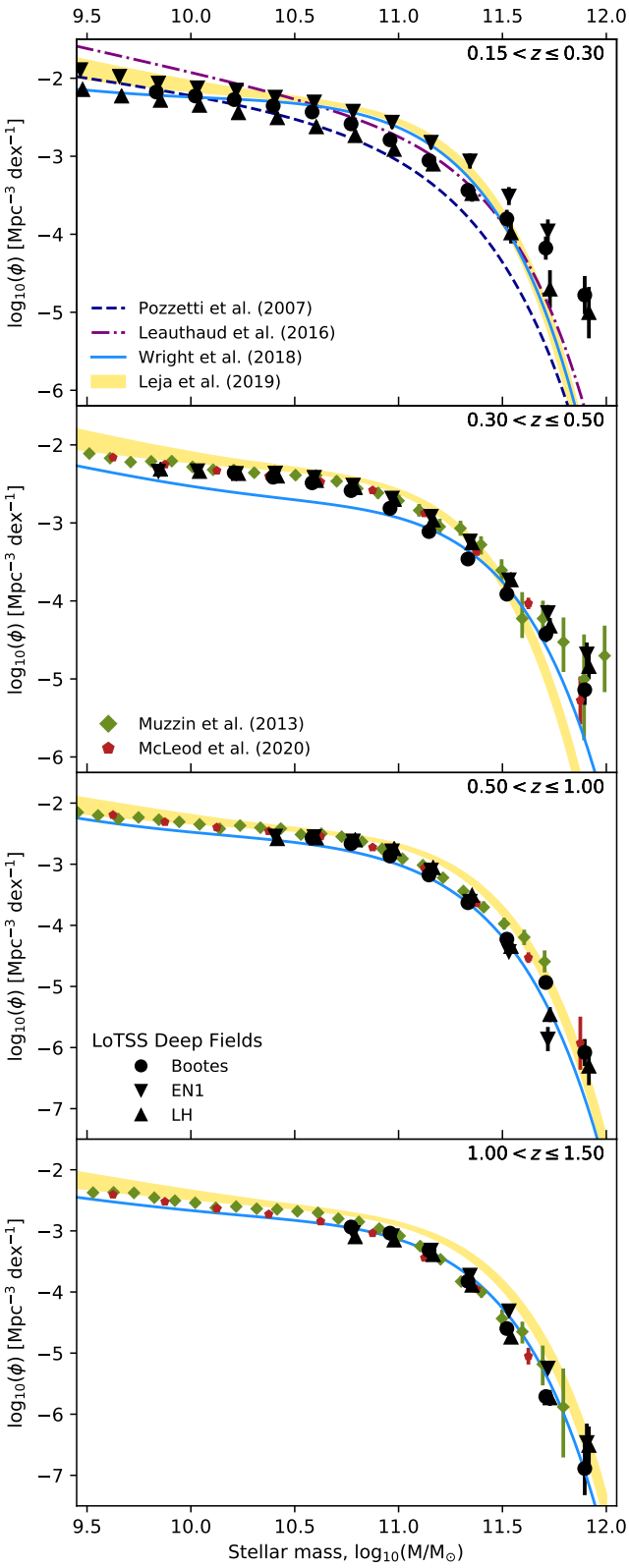

Fig. 12. Estimated galaxy SMFs in each LoTSS deep field based on the stellar mass estimates presented in this data release (black data points), with a sample of results from the literature over the same redshift ranges. For the SMF estimates from this work, we only plot only the mass ranges that are estimated to be 90% complete (see Section 5.3. Total flux corrections to the inferred stellar masses are applied for all fields at $z < 0.5$ and for Boötes at $z < 1$.

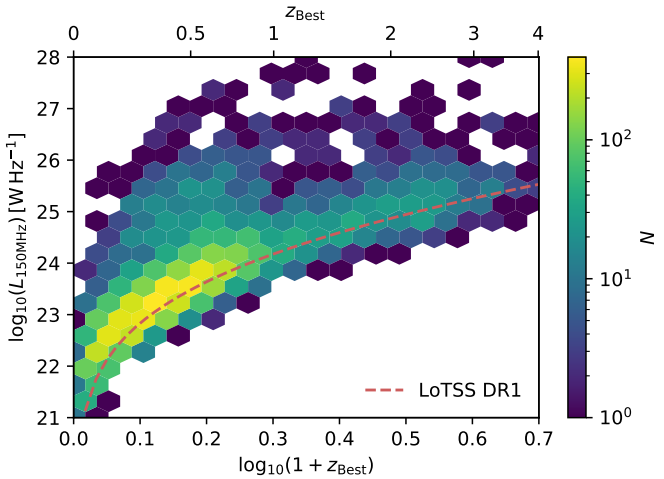

Fig. 13. Distribution of LoTSS Deep Field sources as a function of redshift, $z_{\rm Best}$ (Z_BEST in our catalogues), and radio power $L_{150\rm MHz}$ assuming a constant spectral index of $\alpha = -0.7$. The colour scale illustrates the combined number of sources across the three deep fields with a logarithmic scale. The 5σ luminosity limit for the wide area LoTSS DR1 (Shimwell et al. 2019) data is plotted in red for comparison.

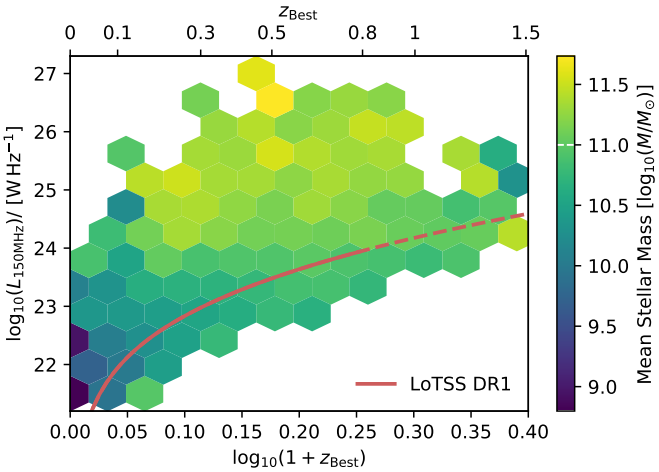

Fig. 14. Average stellar mass (taken as the mean in $\log_{10}$ space) of the LoTSS Deep radio population as a function of redshift, $z_{\rm Best}$ (Z_BEST in our catalogues), and radio power $L_{150\rm MHz}$ - assuming a constant spectral index of $\alpha = -0.7$. The 5σ luminosity limit for the wide area LoTSS DR1 (Shimwell et al. 2019) data is plotted in red for comparison (with the solid line representing the redshift range with high quality photo-z for the comparable galaxy population D19). The typical $M^\star$ for the galaxy SMF at these redshifts is marked on the colour scale for reference (white dashed line).

6. Radio source properties

The combination of exquisite radio continuum observations and extensive multi-wavelength data covering > 20 deg^2 provided by the LoTSS Deep Fields data release offers the potential for a wide range of studies of the radio population - ranging from studies of the cosmic star-formation history through to obscuration free studies of the black hole accretion history (e.g. Smolčić et al. 2017b; Novak et al. 2017). In addition to the SED fitting presented in this paper, extensive detailed physical modelling of the radio population that incorporates additional FIR information (e.g. Calistro Rivera et al. 2016) can offer further insight into the physical properties of the faint radio source population. Building upon the multi-wavelength datasets presented in Paper III and the photo-zs presented here, Best et al. (2020; Paper V) combine results from multiple SED fitting tools to derive consensus source classifications for the radio population. Further papers will utilise these classifications and the associated detailed SED modelling to study various open questions regarding AGN and galaxy evolution.

We therefore do not present a detailed analysis of the LoTSS Deep Field radio source population in this study. However, to demonstrate the scientific potential offered by the LoTSS Deep Fields release, we present some basic properties of the radio source population derived from the catalogues presented in this work. In Figure 13, we show the overall distribution of the LoTSS Deep Field sources as a function of both redshift (based on our best available estimate, $z_{\rm Best}$, where spectroscopic redshifts are used when available and photo-zs otherwise) and of the inferred low-frequency radio luminosity, $L_{150\rm MHz}$.

The colour scale in Figure 13 illustrates the total number of sources per cell across all three fields, with no additional correction for radio source completeness. However, we do apply a cut on the measured photo-z, restricting the sample to $z_{\rm Best} < 4$ and also exclude sources affected by bright stars. Clearly illustrated by this figure is the broad dynamic range offered by the LoTSS Deep Fields, which offer large samples of low power radio sources ($\log_{10}(L_{150\rm MHz}/\rm W\ Hz^{-1}) < 25$) while also probing a sufficient volume to detect statistical samples of the high power radio source population.

In Figure 14, we present a different view of the radio source population, showing the mean log stellar mass in cells of redshift and radio luminosity. We limit the analysis to $z_{\rm Best} < 1.5$, where stellar mass estimates are available, and exclude sources with identified AGN components likely to bias the stellar mass estimates. At low redshift and low radio power, the average stellar mass is significantly below $M^\star$ at these redshift ranges ($M^\star \approx 10^{11}\rm M_\odot$, see Figure 12); this is consistent with the expectation that the radio continuum population at these luminosities is dominated by star-forming galaxies (Best et al. 2005; Smolčić et al. 2017a). As radio luminosity increases, the radio population is hosted by increasingly massive galaxies. This apparent evolution is likely driven by two trends: the first being the transition from the regime where star-forming galaxies dominate the radio source population to where AGN dominate, and the second being the strong correlation between stellar mass and radio AGN activity for radio AGN (Sabater et al. 2019).

We note the important caveat that this analysis excludes the radio AGN that are hosted by sources with significant dust torus emission (as would be detected by IRAC colour criteria) as well as radio-quiet quasars. The trends observed are therefore not representative of the full evolution in radio AGN host properties over this redshift range. Nevertheless, Figure 14 illustrates the potential of the catalogues presented in this work, which when combined with additional source classification and complex SED fitting can offer a detailed picture of the evolution of the radio source population.

7. Summary

In this paper we present details of photo-z and stellar mass estimates produced for the LoTSS Deep Fields DR1. Photo-zs are estimated for all optical sources within the three Deep Fields (Boötes, ELAIS-N1 and LH), totalling over 5 million estimates across a combined ~ 25 deg^2 after appropriate optical quality cuts. Building on previous work our photo-z method combines multiple template fitting and empirical training based estimates to produce a consensus redshift prediction with well-calibrated photo-z uncertainties.

Based on the available spectroscopic training and test sample in each field, the resulting consensus photo-zs have robust scatter ranging from $\sigma_{\rm NMAD} = 0.016$ to 0.02 for galaxies and/or host dominated AGN sources, and from $\sigma_{\rm NMAD} = 0.064$ to 0.07 for identified AGN sources. Our estimated OLF for the corresponding subsets range from 1.5 to 1.8% and 18 to 22%, respectively. Similar to previous studies we find that the photo-z quality is a function of both optical magnitude and spectroscopic redshift.

Exploring the photo-z quality of the LoTSS Deep Field radio source populations we find that there is no strong trend in photo-z quality as a function of radio luminosity (for a fixed redshift), reproducing trends observed in previous studies (Duncan et al. 2019b). However, there is clear deterioration in photo-z quality as a function of redshift for a given radio luminosity that we attribute to selection effects in the spectroscopic sample and/or intrinsic evolution within the radio population.

In the redshift range for which we find our photo-z estimates to be reliable for host-dominated SEDs ($z < 1.5$), we exploit the extensive wavelength coverage to produce consistent stellar mass estimates across the three fields, as well as estimates of the corresponding stellar mass completeness. From simple estimations of the galaxy SMFs within each field and comparison with the literature, we validate our stellar masses provide reliable and self consistent estimates suitable for statistical studies across all three fields.

The catalogue presented in this work builds upon both the LOFAR radio data presented in Paper I & II, and the optical catalogues and radio-optical cross identification presented in Paper III. All data produced in this work are made available for public release to enable the full exploitation of the LoTSS Deep Field survey for a wide range of scientific goals.

Acknowledgements

KJD and HJAR acknowledge support from the ERC Advanced Investigator programme NewClusters 321271. RK acknowledges support from the Science and Technology Facilities Council (STFC) through an STFC studentship via grant ST/R504737/1. PNB and JS are grateful for support from the UK STFC via grant ST/R000972/1 MB acknowledges support from INAF under PRIN SKA/CTA FORECaST and from the Ministero degli Affari Esteri della Cooperazione Internazionale - Direzione Generale per la Promozione del Sistema Paese Progetto di Grande Rilevanza ZA18GR02. RAAB acknowledges support from the Glasstone Foundation. MJH acknowledges support from the UK STFC via grant ST/R000905/1. MJJ acknowledges support from the UK Science and Technology Facilities Council [ST/N000919/1] and the Oxford Hintze Centre for Astrophysical Surveys which is funded through generous support from the Hintze Family Charitable Foundation. MKB and AW acknowledge support from the National Science Centre, Poland under grant no. 2017/26/E/ST9/00216. K.M. has been supported by the National Science Centre (UMO-2018/30/E/ST9/00082). IP acknowledges support from INAF under the SKA/CTA PRIN "FORECaST" and the PRIN MAIN STREAM "SAuROS" projects.

This paper is based (in part) on data obtained with the International LOFAR Telescope (ILT) under project codes LC0 015, LC2 024, LC2 038, LC3 008, LC4 008, LC4 034 and LT10 01. LOFAR (van Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and which are collectively operated by the ILT foundation under a joint scientific policy. The ILT resources have benefitted from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, MIWF-NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands, The Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland.

The Pan-STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation Grant No. AST-1238877, the University of Maryland, Eotvos Lorand University (ELTE), the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation.

This work is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France and the University of Hawaii. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. RCSLenS data processing was made possible thanks to significant computing support from the NSERC Research Tools and Instruments grant program. This work is based in part on observations made with the Spitzer Space Telescope, which was operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA.

References

Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, Publ. Astron. Soc. Jpn, 71, 114
Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12
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
Ananna, T. T., Salvato, M., LaMassa, S., et al. 2017, ApJ, 850, 66
Bauer, F. E., Alexander, D. M., Brandt, W. N., et al. 2004, AJ, 128, 2048
Berta, S., Lonsdale, C. J., Siana, B., et al. 2007, A&A, 467, 565
Best, P. N., Kauffmann, G., Heckman, T. M., et al. 2005, MNRAS, 362, 25
Boller, T., Freyberg, M. J., Truemper, J., et al. 2016, arXiv, A103
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brand, K., Brown, M. J. I., Dey, A., et al. 2006, ApJ, 641, 140
Brodwin, M., Brown, M. J. I., Ashby, M. L. N., et al. 2006, ApJ, 651, 791
Brown, M. J. I., Duncan, K. J., Landt, H., et al. 2019, MNRAS, 489, 3351
Brown, M. J. I., Moustakas, J., Smith, J. D. T., et al. 2014, ApJS, 212, 18
Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000
Buchs, R., Davis, C., Gruen, D., et al. 2019, MNRAS, 489, 820
Calistro Rivera, G., Lusso, E., Hennawi, J. F., & Hogg, D. W. 2016, ApJ, 833, 98
Calistro Rivera, G., Williams, W. L., Hardcastle, M. J., et al. 2017, MNRAS, 469, 3468
Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019a, ApJ, 873, 44
Carnall, A. C., McLure, R. J., Dunlop, J. S., et al. 2019b, MNRAS, 490, 417
Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379
Cavuoti, S., Tortora, C., Brescia, M., et al. 2017, MNRAS, 466, 2039
Chabrier, G. 2003, The Publications of the Astronomical Society of the Pacific, 115, 763
Charlot, S. & Fall, S. M. 2000, ApJ, 539, 718
Chung, S. M., Kochanek, C. S., Assef, R., et al. 2014, ApJ, 790, 54
Conroy, C. & Gunn, J. E. 2010, Astrophysics Source Code Library, ascl:1010.043
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Dahlen, T., Mobasher, B., Faber, S. M., et al. 2013, ApJ, 775, 93
Dey, A., Lee, K.-S., Reddy, N., et al. 2016, ApJ, 823, 11
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142
Duncan, K. J., Brown, M. J. I., Williams, W. L., et al. 2018a, MNRAS, 473, 2655
Duncan, K. J., Conselice, C. J., Mortlock, A., et al. 2014, MNRAS, 444, 2960
Duncan, K. J., Conselice, C. J., Mundy, C., et al. 2019a, ApJ, 876, 110
Duncan, K. J., Jarvis, M. J., Brown, M. J. I., & Rottgering, H. J. A. 2018b, MNRAS, 477, 5177
Duncan, K. J., Sabater, J., Rottgering, H. J. A., et al. 2019b, A&A, 622, A3
Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, ApJS, 154, 10
Finkelstein, S. L. 2014, eprint arXiv:1410.5439, 810, 71
Fitzpatrick, E. L. 1999, The Publications of the Astronomical Society of the Pacific, 111, 63
Flesch, E. W. 2015, Publ. Astron. Soc. Aust., 32, e010
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306
Fotopoulou, S. & Paltani, S. 2018, A&A, 619, A14
Green, G. M. 2018, JOSS, 3, 695
Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24
Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, MNRAS, 421, 2355
Hsieh, B. C. & Yee, H. K. C. 2014, ApJ, 792, 102
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Kenter, A., Murray, S. S., Forman, W. R., et al. 2005, ApJS, 161, 9
Kochanek, C. S., Eisenstein, D. J., Cool, R. J., et al. 2012, ApJS, 200, 8
Kondapally, R. & et al. 2020, A&A, subm.
Lacy, M., Ridgway, S. E., Gates, E. L., et al. 2013, ApJS, 208, 24
Leauthaud, A., Bundy, K., Saito, S., et al. 2016, MNRAS, 457, 4021
Lee, K.-S., Alberts, S., Atlee, D., et al. 2012, ApJ: Letters, 758, L31
Lee, K.-S., Dey, A., Cooper, M. C., Reddy, N., & Jannuzi, B. T. 2013, ApJ, 771, 25
Lee, K.-S., Dey, A., Hong, S., et al. 2014, ApJ, 796, 126
Leistedt, B., Hogg, D. W., Wechsler, R. H., & DeRose, J. 2019, ApJ, 881, 80
Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019, ApJ, 876, 3
Leja, J., Speagle, J. S., Johnson, B. D., et al. 2020, ApJ, 893, 111
Lima, M., Cunha, C. E., Oyaizu, H., et al. 2008, MNRAS, 390, 118
Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, AJ, 118, 1406
Mauduit, J. C., Lacy, M., Farrah, D., et al. 2012, Publications of the Astronomical Society of the Pacific, 124, 714
McGreer, I. D., Becker, R. H., Helfand, D. J., & White, R. L. 2006, ApJ, 652, 157
McLeod, D. J. & et al. 2020, MNRAS, in prep.
Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18
Nishizawa, A. J., Hsieh, B.-C., Tanaka, M., & Takata, T. 2020, arXiv, arXiv:2003.01511
Norris, R. P., Salvato, M., Longo, G., et al. 2019, Publications of the Astronomical Society of the Pacific, 131, 108004
Novak, M., Smolčić, V., Delhaize, J., et al. 2017, A&A, 602, A5
Nyland, K., Lacy, M., Sajina, A., et al. 2017, ApJS, 230, 9
Oke, J. B. & Gunn, J. E. 1983, ApJ, 266, 713
Pacifici, C., Kassin, S. A., Weiner, B. J., et al. 2016, ApJ, 832, 79
Pforr, J., Vaccari, M., Lacy, M., et al. 2019, MNRAS, 483, 3168
Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81
Pozzetti, L., Bolzonella, M., Lamareille, F., et al. 2007, A&A, 474, 443
Pozzetti, L., Bolzonella, M., Zucca, E., et al. 2010, A&A, 523, A13
Rottgering, H. J. A. 2010, in Proc. ISKAF2010, 50–
Sabater, J., Best, P. N., Hardcastle, M. J., et al. 2019, A&A, 622, A17
Sabater, J., Best, P. N., Tasse, C., & et al. 2020, A&A, submitted (LoTSS SI)
Salvato, M., Buchner, J., Budavári, T., et al. 2017, arXiv, 4937
Salvato, M., Hasinger, G., Ilbert, O., et al. 2008, ApJ, 690, 1250
Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, ApJ, 742, 61
Schenker, M. A., Ellis, R. S., Konidaris, N. P., & Stark, D. P. 2013, ApJ, 777, 67
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Shimwell, T. W., Rottgering, H. J. A., Best, P. N., et al. 2017, A&A, 598, A104
Shimwell, T. W., Tasse, C., Hardcastle, M. J., et al. 2019, A&A, 622, A1
Smith, D. J. B., Best, P. N., Duncan, K. J., et al. 2016, eprint arXiv:161102706,
Smolčić, V., Delvecchio, I., Zamorani, G., et al. 2017a, A&A, 602, A2
Smolčić, V., Novak, M., Delvecchio, I., et al. 2017b, A&A, 602, A6
Stanford, S. A., Brodwin, M., Gonzalez, A. H., et al. 2012, ApJ, 753, 164
Stark, D. P., Schenker, M. A., Ellis, R., et al. 2013, ApJ, 763, 129
Swinbank, A. M., Bower, R. G., Smith, G. P., et al. 2007a, MNRAS, 376, 479
Swinbank, A. M., Edge, A. C., Smail, I., et al. 2007b, MNRAS, 379, 1343
Tanaka, M. 2015, ApJ, 801, 20
Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, Publ. Astron. Soc. Jpn, 70, S9
Tasse, C., Shimwell, T. W., Hardcastle, M. J., & et al. 2020, A&A, submitted (LoTSS SI)
van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
Williams, W. L., Hardcastle, M. J., Best, P. N., et al. 2019, A&A, 622, A2
Wittman, D., Bhaskar, R., & Tobin, R. 2016, MNRAS, 457, 4005
Wright, A. H., Driver, S. P., & Robotham, A. S. G. 2018, MNRAS, 480, 3491
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Zeimann, G. R., Stanford, S. A., Brodwin, M., et al. 2013, ApJ, 779, 137
Zeimann, G. R., Stanford, S. A., Brodwin, M., et al. 2012, ApJ, 756, 115

[1] SUPA, Institute for Astronomy, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK
[2] Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, The Netherlands
[3] School of Physics and Astronomy, Monash University, Clayton, Victoria 3800, Australia
[4] INAF-Istituto di Radioastronomia, Via Gobetti 101, I-40129, Bologna, Italy
[5] Italian ALMA Regional Centre, Via Gobetti 101, I-40129, Bologna, Italy
[6] INAF-Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, I-35122, Padova, Italy
[7] Astrophysics, Department of Physics, Keble Road, Oxford, OX1 3RH, UK
[8] Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, USA
[9] CSIRO Astronomy and Space Science, PO Box 1130, Bentley WA 6102, Australia
[10] Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK
[11] Department of Physics & Astronomy, University of the Western Cape, Private Bag X17, Bellville, Cape Town, 7535, South Africa
[12] Institute of Astronomy, Faculty of Physics, Astronomy and Informatics, NCU, Grudziadzka 5, 87-100 Toruń, Poland
[13] National Centre for Nuclear Research, ul. Pasteura 7, 02-093 Warszawa, Poland
[14] Aix Marseille Univ. CNRS, CNES, LAM, Marseille, France
[15] Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham, DH1 3LE, UK
[16] School of Physical Sciences and Centre for Astrophysics & Relativity, Dublin City University, Glasnevin, D09 W6Y4, Ireland
[17] ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, 7991 PD, Dwingeloo, The Netherlands
[18] SRON Netherlands Institute for Space Research, Landleven 12, 9747 AD, Groningen, The Netherlands
[19] Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, the Netherlands

Appendix A: Catalogue description

The contents of the catalogue added by this work are as follows:

- `Z_BEST` - Best available redshift estimate
- `Z_BEST_SOURCE` - The source of the best available redshift, `Z_BEST` where 1 corresponds to spectroscopic redshift and 0 corresponds to the photo-z presented in this work.
- `Z_SPEC` - Literature Spectroscopic Redshift
- `Z_SOURCE` - Source of the spectroscopic redshift
- `Z_QUAL` - Spectroscopic redshift quality, where flag $Q = 3$ means probable, $Q \geq 4$ means reliable. Lower reliability redshifts have not been included.
- `AGN_ZSPEC` - Spectroscopic AGN/QSO flag where provided
- `z1_median` - Median of the primary redshift peak above 80% HPD CI
- `z1_min` - Lower bound of the primary 80% HPD CI peak|
- `z1_max` - Upper bound of the primary 80% HPD CI peak
- `z1_area` - Integrated area of the primary 80% HPD CI peak
- `z2_median` - Median of the secondary redshift peak (if present) above 80% HPD CI
- `z2_min` - Lower bound of the secondary 80% HPD CI peak
- `z2_max` - Upper bound of the secondary 80% HPD CI peak
- `z2_area` - Integrated area of the secondary 80% HPD CI peak
- `nfilt_eazy` - Number of filters included in EAZY template fit|
- `nfilt_atlas` - Number of filters included in Atlas+AGN template fit
- `nfilt_ananna` - Number of filters included in Ananna et al. template fit
- `chi_r_best` - χ^2 / nfilt for best-fit galaxy or AGN template (any library)
- `chi_r_stellar` - χ^2 / nfilt for best-fit stellar template
- `stellar_type` - Stellar type of best-fit stellar template

Also included for all sources are the multi-wavelength AGN classifications used during photo-z estimation.

- `AGN` - Sources flagged by any one of optAGN, IRAGN or XrayAGN
- `optAGN` - Flag indicating whether source is included in Million Quasar catalogue compilation (Flesch 2015), where 1 means a source is included. Sources flagged as AGN based on their spectroscopic redshifts are also flagged.
- `IRAGN` - Source satisfies Donley et al. (2012) IR AGN selection criteria.
- `XrayAGN` - Source has X-ray counterpart.

For the Boötes field, where sources have been matched to the X-Böotes *Chandra* survey of NDWFS (Kenter et al. 2005), we provide the additional associated values:

- `XrayFlux_0.5-2` [10^{-14} erg/cm^2/s] XBoötes Soft X-ray Flux
- `XrayHardness` - XBoötes X-ray Hardness Ratio

For sources with either `Z_SPEC` or `z1_median` < 1.5, we provide our estimate of the galaxy stellar mass as well as the rest-frame magnitudes for the individual best-fitting SED model. The catalogue columns associated with the additonal SED fitting are:

- `zmodel` - Model grid redshift used in stellar mass fit
- `chi_best` - χ^2 for the single best-fit model SED (for `Z_BEST`)
- `Mass_median` - 50th percentile of the marginalised stellar mass posterior (for `Z_BEST`), in units of $\log_{10}(M/M_\odot)$
- `Mass_l68` - 16th percentile of the marginalised stellar mass posterior (for `Z_BEST`), in units of $\log_{10}(M/M_\odot)$
- `Mass_u68` - 84th percentile of the marginalised stellar mass posterior (for `Z_BEST`), in units of $\log_{10}(M/M_\odot)$
- `Nfilts` - Number of photometric bands included in stellar mass fit.
- `ap_to_model_[x]` - Estimated to aperture to total flux correction derived from the model fitting photometry of the Legacy Survey in a band, x.[7]
- `ap_to_model_err_[x]` - Statistical uncertainty on `ap_to_model_[x]` derived from the combined flux uncertainties of the LoTSS Deep Field optical and Legacy Surveys flux measurement.

Finally, for each photometric band used during stellar mass estimation we provide the rest-frame magnitude in that filter for the best-fit SED template:

- `[x]_rest` Rest-frame magnitude in a given band, x, for best-fit SED (for `Z_BEST`)

[7] For EN1 and LH, we provide corrections for the g, r, and z bands calculated using either PS1 or RCSLens photometry, respectively. In Boötes we provide corrections based on the $z_{\rm Subaru}$ and z bands.

Appendix B: Zeropoint flux corrections

The inclusion of the zero-point offsets during template fitting has been demonstrated to lead to substantial improvement in photo-z quality. As outlined in Section 3.1, zero-point offsets for the photo-z estimates for each template library and dataset were derived following the method outlined in D18a. Similarly, zero-point corrections for stellar mass estimates are calculated following the method described in Section 5.2. Here we provide the multiplicative correction factors for all template sets in each of the LoTSS Deep Fields.

Table B.1. Flux zero-point corrections applied to the Boötes photometry during photo-z fitting or stellar mass estimates. Values quoted are multiplicative corrections applied to the observed fluxes and flux uncertainties. The three photo-z template sets consist of the 'EAZY' (Brammer et al. 2008), 'Brown' (Brown et al. 2014, 2019) and 'Ananna' (Ananna et al. 2017) libraries as described in Section 3.1. If a filter was not included in the SED fitting for a specific template set, we do not provide a corresponding value.

Boötes				
Filter	Template Set			
	EAZY	Brown	Ananna	Masses
u	0.979	0.957	0.946	0.999
B_W	1.027	1.029	1.0	0.95
R	0.986	0.972	0.97	0.882
I	0.969	0.976	0.958	0.898
z	0.933	0.945	0.927	0.879
$z_{\rm Subaru}$	1.028	1.038	1.02	0.917
Y	1.003	1.008	0.992	0.971
J	0.99	0.978	0.99	0.991
H	1.044	1.053	1.096	1.076
K	0.797	0.815	0.856	0.895
K_s	1.004	1.024	1.061	1.082
3.6μm	1.002	0.989	1.0	1.079
4.5μm	1.058	0.994	1.006	1.015
5.8μm	0.99	0.958		1.0
8.0μm	0.98	0.922		1.0

Table B.2. Flux zero-point corrections applied to the EN1 photometry during photo-z fitting or stellar mass estimates. Values quoted are multiplicative corrections applied to the observed fluxes and flux uncertainties. The three photo-z template sets consist of the 'EAZY' (Brammer et al. 2008), 'Brown' (Brown et al. 2014, 2019) and 'Ananna' (Ananna et al. 2017) libraries as described in Section 3.1. If a filter was not included in the SED fitting for a specific template set, we do not provide a corresponding value.

EN1				
Filter	Template Set			
	EAZY	Brown	Ananna	Masses
u	1.023	1.033	1.0	1.047
g	1.052	1.041	1.02	0.918
r	1.048	1.024	1.0	0.88
i	1.019	1.014	0.993	0.879
z	0.998	1.006	0.993	1.092
y	0.992	1.007	0.994	0.738
$g_{\rm HSC}$	0.951	0.945	0.926	
$r_{\rm HSC}$	0.972	0.95	0.929	
$i_{\rm HSC}$	0.992	0.987	0.97	
$z_{\rm HSC}$	0.933	0.944	0.93	
$y_{\rm HSC}$	0.94	0.955	0.94	
NB921$_{\rm HSC}$	0.925	0.939	0.925	
J	1.203	1.182	1.165	1.129
K	1.088	1.061	1.087	1.072
3.6μm SERVS	0.944	0.946	0.946	0.952
4.5μm SERVS	1.0	0.96	0.98	0.917
3.6μm SWIRE	0.955	0.953	0.956	0.959
4.5μm SWIRE	1.014	0.961	0.978	0.951
5.8μm SWIRE	1.0	1.0		1.0
8.0μm SWIRE	1.16	1.057		1.0

Table B.3. Flux zero-point corrections applied to the LH photometry during photo-z fitting or stellar mass estimates. Values quoted are multiplicative corrections applied to the observed fluxes and flux uncertainties. The three photo-z template sets consist of the 'EAZY' (Brammer et al. 2008), 'Brown' (Brown et al. 2014, 2019),and 'Ananna' (Ananna et al. 2017) libraries as described in Section 3.1. If a filter was not included in the SED fitting for a specific template set, we do not provide a corresponding value.

LH				
Filter	Template Set			
	EAZY	Brown	Ananna	Masses
u	1.115	1.102	1.107	1.121
g	0.989	0.978	0.94	0.904
r	1.008	0.987	0.946	0.911
z	0.99	1.006	0.992	0.981
$g_{\rm RCS}$	0.921	0.915	0.879	0.873
$r_{\rm RCS}$	0.963	0.947	0.904	0.889
$i_{\rm RCS}$	0.931	0.941	0.906	0.9
$z_{\rm RCS}$	0.919	0.935	0.921	0.927
J	1.117	1.11	1.08	1.16
K	1.113	1.071	1.109	1.126
3.6μm SERVS	0.962	0.952	0.969	0.933
4.5μm SERVS	1.0	0.94	1.0	0.908
3.6μm SWIRE	1.0	0.977	1.0	0.954
4.5μm SWIRE	1.019	0.967	1.0	0.941
5.8μm SWIRE	1.04	0.977		1.0
8.0μm SWIRE	1.2	1.039		1.0