Deep *ugrizY* imaging and DEEP2/3 spectroscopy: a photometric redshift testbed for LSST and public release of data from the DEEP3 Galaxy Redshift Survey

Rongpu Zhou [1,2][*] Michael C. Cooper [3] Jeffrey A. Newman[1,2]
Matthew L. N. Ashby,[4] James Aird [5] Christopher J. Conselice,[6] Marc Davis,[7,8]
Aaron A. Dutton [9] S. M. Faber,[10] Jerome J. Fang,[11] G. G. Fazio,[4]
Puragra Guhathakurta,[10] Dale Kocevski,[12] David C. Koo,[10] Kirpal Nandra,[13]
Andrew C. Phillips,[10] David J. Rosario,[14] Edward F. Schlafly,[15] Jonathan R. Trump,[16]
Benjamin Weiner,[17] Christopher N. A. Willmer[18] and Renbin Yan[19]

Affiliations are listed at the end of the paper

ABSTRACT

We present catalogues of calibrated photometry and spectroscopic redshifts in the Extended Groth Strip, intended for studies of photometric redshifts (photo-z's). The data includes *ugriz* photometry from Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) and *Y*-band photometry from the Subaru Suprime camera, as well as spectroscopic redshifts from the DEEP2, DEEP3, and 3D-HST surveys. These catalogues incorporate corrections to produce effectively matched-aperture photometry across all bands, based upon object size information available in the catalogue and Moffat profile point spread function fits. We test this catalogue with a simple machine learning-based photometric redshift algorithm based upon Random Forest regression, and find that the corrected aperture photometry leads to significant improvement in photo-z accuracy compared to the original SEXTRACTOR catalogues from CFHTLS and Subaru. The deep *ugrizY* photometry and spectroscopic redshifts are well suited for empirical tests of photometric redshift algorithms for LSST. The resulting catalogues are publicly available at http://d-scholarship.pitt.edu/36064/. We include a basic summary of the strategy of the DEEP3 Galaxy Redshift Survey to accompany the recent public release of DEEP3 data.

Key words: catalogues – surveys – galaxies: distances and redshifts.

1 INTRODUCTION

Redshift is a crucial observable in the study of galaxies and cosmology. Spectroscopic redshifts are accurate, but the observations required are much more expensive than photometric measurements. Modern imaging surveys can measure the photometry of a huge number of objects very efficiently, but only a very small fraction will have observed spectra. For such surveys, redshifts must be estimated from broad-band photometry, and the large number of photometric redshift (photo-z) measurements compensates for their inaccuracy. The availability of large imaging data sets has made photometric redshift estimates an increasingly important component of modern extragalactic astronomy and cosmology studies.

The Large Synoptic Survey Telescope (Ivezic et al. 2009; LSST Science Collaboration 2009) will rely on photometric redshifts to achieve many of its science goals. For 10 yr, LSST will survey the sky in six filters to a depth unprecedented over such a wide area. The resulting data set should provide important clues to the nature of dark matter and dark energy, detailed information on the structure of the Milky Way, a census of near-Earth objects in the Solar system, and a wealth of information on variable and transient phenomena. In this paper, we present catalogues with robust spectroscopic redshift measurements and well-calibrated photometry in the Extended Groth Strip (EGS) with filter coverage and depths similar to the LSST *ugrizy* system. The LSST Science Requirements Document[1] specifies that for galaxies with $i < 25$

---

[*] E-mail: rongpu.zhou@pitt.edu

[1] www.lsst.org/scientists/publications/science-requirements-document

the LSST data should be capable of delivering a root-mean-square (RMS) error in redshift smaller than $0.02(1 + z)$ with a rate of $>3\sigma$ outliers below 10 per cent. The data set we have assembled will be useful for assessing if current photometric redshift algorithms can meet these requirements, and for improving them if not.

A previous paper, Matthews et al. (2013), matched redshifts from the DEEP2 Galaxy Redshift Survey (Newman et al. 2013) to photometry from the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS; Hudelot et al. 2012) and the Sloan Digital Sky Survey (SDSS Gunn et al. 1998; Alam et al. 2015). This works builds on that effort by adding DEEP3 (Cooper et al. 2011, 2012) and 3D-HST (Brammer et al. 2012; Momcheva et al. 2016) redshifts and *Y*-band photometry, and using Pan-STARRS (Chambers et al. 2016; Magnier et al. 2016) instead of SDSS for photometric calibration. We also have developed a method for calculating corrected aperture photometry from the CFHTLS catalogues, and we perform tests with a simple photometric redshift algorithm to demonstrate the superiority of this photometry for measuring galaxy colours.

The structure of this paper is as follows. Section 2 describes the data sets that we used to produce the final catalogues. We use spectroscopic redshifts from the DEEP2 and DEEP3 surveys, as well as grism redshifts from 3D-HST. The photometry in the *ugriz* bands is from CFHTLS. Additionally, *Y*-band imaging was obtained from SuprimeCam at the Subaru Telescope (Miyazaki et al. 2002); photometry based on these images was derived using SEXTRACTOR (Bertin & Arnouts 1996). In Section 3, we describe the methods used to bring the CFHTLS, Subaru *Y*-band, and Pan-STARRS1 catalogues to a common astrometric system, based on those employed by Matthews et al. (2013). We describe our photometric zero-point calibration methods in Section 4 and the techniques used to produce corrected aperture photometry in Section 5. In Section 6, we describe the resulting matched catalogues, which are being released in concert with this paper. In Section 7, we present tests of these catalogues using photometric redshifts measured via Random Forest regression. We provide a summary in Section 8.

2 DATA SETS

In this section, we describe the spectroscopic and imaging data sets used to construct the catalogues presented in this paper.

2.1 Spectroscopy

The first spectroscopic sample included in our catalogues comes from the DEEP2 Galaxy Redshift Survey, which is a magnitude-limited spectroscopic survey performed using the DEIMOS spectrograph at the Keck 2 telescope. Galaxy spectra were observed in four fields, with targets lying in the magnitude range $R_{AB} < 24.1$. Field 1 (corresponding to the EGS) applied no redshift pre-selection, though objects expected to be at higher redshift received greater weight in targeting. In the remaining three fields, DEEP2 targeted only objects expected to be in the redshift range of $z > 0.75$. Only Field 1 is used for this paper. Details of DEEP2 are given in Newman et al. (2013).

The second spectroscopic sample included constitutes the public data release of spectra from the DEEP3 Galaxy Redshift Survey (Cooper et al. 2011, 2012), which was primarily intended to enlarge the DEEP2 survey within the EGS field to take advantage of the wealth of multiwavelength information available there. This

release is distributed at http://deep.ps.uci.edu/deep3/home.html. We describe DEEP3 in more detail in Appendix A to accompany this data release.

We also incorporate grism redshift data from the 3D-HST survey (Brammer et al. 2012; Momcheva et al. 2016), which measures redshift down to $JH_{IR} = 26$. The 3D-HST sample reaches higher redshifts than DEEP2 or DEEP3. The 3D-HST grism redshifts are derived using a combination of grism spectra and photometric data, and proper selection is needed to ensure a set of robust redshifts. The selection criteria used are described in Section 7.

2.2 Photometry in *ugriz* bands

For the *ugriz* bands, we used the CFHTLS-T0007 (Hudelot et al. 2012) catalogues of photometry from CFHT/MegaCam. We utilize data from the CFHTLS Deep field D3 as well as the seven pointings in the Wide field W3 which overlap with DEEP2/3 and 3D-HST. The list of pointings may be found in Table 2. The CFHTLS Wide field sample reaches 5σ depths of $u \sim 24.7$, $g \sim 25.4$, $r \sim 24.8$, $i \sim 24.3$, and $z \sim 23.5$. The CFHTLS Deep data reaches 5σ depths of $u \sim 27.1$, $g \sim 27.5$, $r \sim 27.2$, $i \sim 26.9$, $i_2 \sim 26.6$, and $z \sim 25.8$, where i_2 is the replacement filter for the *i*-band filter. This filter was named *y* in the CFHTLS catalogues, but within this paper and in our catalogues we refer to this filter as i_2 to avoid confusion with the *y* band in the LSST *ugrizy* filter system. The default photometry from CFHTLS is the Kron-like elliptical aperture magnitude MAG_AUTO. We also have calculated a set of corrected aperture magnitudes as described below, which we designate as MAG_APERCOR in catalogues. See Section 5 for details of the aperture correction procedure applied.

We have utilized an internal version of the Pan-STARRS1 (PS1) catalogue (Chambers et al. 2016; Magnier et al. 2016) to calibrate the photometric zero-points for the *griz* and *Y* bands. For the CFHTLS *u* band, we have used the Deep field photometry as the standard against which we calibrate the Wide field data, as described in Section 4.3.

2.3 *Y*-band data

In addition to the *ugriz* bands which are included in CFHTLS, LSST will obtain data in the *y* band. To obtain photometry of comparable depth in a similar filter, we used the *Y*-band filter available for SuprimeCam on the Subaru Telescope (Miyazaki et al. 2002) over the course of two nights to cover a portion of the DEEP2 EGS field. The wavelength coverage of this filter is slightly redder and narrower than the LSST *y*-band filter, but it is otherwise similar. The *Y*-band observations consist of two pointings centred on RA = $14^h17^m58^s.2$, Dec. = $+52°36'4''.0$ and RA = $14^h22^m28^s.0$, Dec. = $+53°24'58''.0$, with exposure times of 234 and 9 min, respectively. The unequal exposure times were not planned, but rather a result of the onset of poor weather conditions. The 5σ depth of the two pointings are 25.0 and 23.4 mag, respectively, and the seeing full width at half-maximum (FWHM) values were 0.662 and 0.632 arcsec, respectively. A mosaic was created using the Subaru/Suprime SDFRED2 pipeline (Ouchi et al. 2004). The initial astrometry for the mosaic was determined using Astrometry.net (Lang et al. 2010). We then used SEXTRACTOR (Bertin & Arnouts 1996) to detect sources and obtain a photometric catalogue. Slightly different SEXTRACTOR parameters were used for the two pointings to account for differences in depth and seeing. The parameters are listed in Appendix B. An initial 'guess' of the image zero-point was used for SEXTRACTOR. We determine a more accurate zero-point later in the calibration procedure as described in
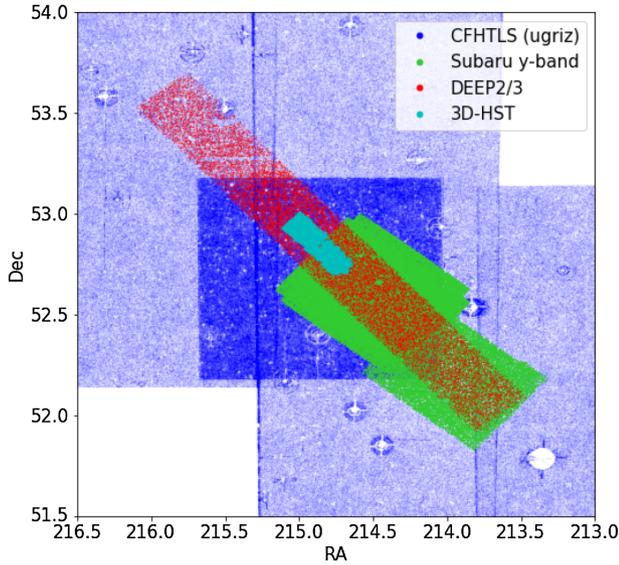
Figure 1. Sky coverage of the catalogues used in this paper (Pan-STARRS1 and SDSS are not shown as they cover the entire region). The region covered by DEEP2 and DEEP3 is shown in red, 3D-HST is in cyan, the CFHTLS imaging pointings included in our catalogues are shown in blue, and the two rectangular pointings of Subaru *Y*-band imaging are shown in green (the deeper of the two pointings partially overlaps with 3D-HST).

Section 4. A subset of the detected sources was visually inspected to optimize the parameters, enabling us to minimize false detections and to ensure that nearby and overlapping sources are de-blended properly.

SEXTRACTOR requires a weight map for processing. To create the weight map, we set the BACK_SIZE parameter to 16, and the resulting BACKGROUND_RMS check image was used as the weight map (in other words, the RMS of the background evaluated over 16 pixel boxes was used as a weight map). To avoid false detections near the image boundary, objects within 15 pixels of the image boundary were not used. The astrometry was further corrected by cross-matching to SDSS (cf. Section 3). Besides the default MAG_AUTO photometry, we also produced aperture photometry (MAG_APER) with aperture diameters ranging from 9 to 56 pixels in 1 pixel spacing (the pixel size of SuprimeCam is 0.2 arcsec). The MAG_APER photometry and half-light radius were used to calculate the corrected aperture magnitudes as described in Section 5.

The sky coverage of the data sets incorporated in this work is shown in Fig. 1.

3 ASTROMETRIC CORRECTION

To avoid false matching between catalogues, we applied astrometric corrections to CFHTLS, the *Y*-band catalogue, and PS1 to make them each match the SDSS coordinate system before cross-matching the catalogues. The astrometric offsets required varied spatially for each of these data sets. There was no significant offset between DEEP2/3 positions (which were previously remapped to match SDSS coordinates) and SDSS. For 3D-HST, a constant RA and Dec. offset were needed to match SDSS but no spatial variation in offsets was needed.

In order to derive astrometric corrections for the CFHTLS, Subaru *Y*-band, and PS1 catalogues to match SDSS, we have applied the

Table 1. The mean and RMS of $RA_{CFHTLS} - RA_{SDSS}$, $RA_{Subaru} - RA_{SDSS}$, and $RA_{PS1} - RA_{SDSS}$. The values before correction are listed as plain text and the values after correction are in *italic* font. The astrometric corrections applied are described in Section 3.

Pointing	$RA - RA_{SDSS}$ (arcsec)		$Dec. - Dec._{SDSS}$ (arcsec)	
	Mean	σ	Mean	σ
CFHTLS D3	0.071	0.303	−0.023	0.180
	0.003	*0.267*	*0.002*	*0.155*
CFHTLS W3-0-1	0.107	0.286	0.016	0.157
	0.002	*0.257*	*0.000*	*0.150*
CFHTLS W3-1-2	0.058	0.271	0.042	0.163
	0.002	*0.258*	*0.001*	*0.152*
CFHTLS W3-0-3	0.125	0.281	−0.011	0.155
	0.004	*0.243*	*−0.001*	*0.148*
CFHTLS W3 + 1-2	0.075	0.269	−0.027	0.155
	0.001	*0.252*	*0.000*	*0.147*
CFHTLS W3-0-2	0.107	0.284	−0.007	0.158
	0.001	*0.259*	*0.000*	*0.151*
CFHTLS W3 + 1-1	0.094	0.266	0.007	0.150
	0.002	*0.243*	*0.000*	*0.146*
CFHTLS W3-1-3	0.033	0.252	−0.003	0.157
	0.003	*0.244*	*0.000*	*0.147*
Subaru *Y*-band	−0.042	0.285	−0.165	0.296
	−0.001	*0.259*	*0.000*	*0.151*
PS1	0.020	0.285	−0.022	0.171
	−0.001	*0.264*	*0.000*	*0.153*

same methodology as described in Matthews et al. (2013). In this paper, we give only a brief outline of these techniques; we refer the reader to this prior work for details. We describe the correction of CFHTLS for sake of example.

The correction is done separately for each pointing from CFHTLS. First we cross-match CFHTLS to SDSS with a search radius of 1.0 arcsec. If more than one match is found, the nearest match is kept. The differences in RA and Dec. (ΔRA and ΔDec.) are calculated for every matched object. The matched objects are binned according to their RA and Dec., with a bin size of 1.2 arcmin × 1.2 arcmin. This bin size was chosen because smaller bins did not significantly reduce the residuals and could lead to problems with overfitting. Within each bin, the mean value of the ΔRA and ΔDec. are calculated using the robust Hodges–Lehmann estimator (Hodges & Lehmann 1963). For bins that have fewer than three objects, values from the neighbouring bins are used. A 3 × 3 boxcar average is performed to smooth ΔRA and ΔDec., and we perform bivariate spline interpolation on the smoothed ΔRA and ΔDec. grid to obtain the functions ΔRA(RA, Dec.) and ΔDec.(RA, Dec.). For each object in the CFHTLS catalogues we then evaluate ΔRA(RA, Dec.) and ΔDec.(RA, Dec.) to determine the offsets at its position, and subtract them from the CFHTLS coordinates. The same method is used to correct the astrometry of PS1 and the *Y*-band catalogue, with the only difference being the bin sizes used (4 arcmin × 4 arcmin and 1.7 arcmin × 1.7 arcmin, respectively, for PS1 and Subaru). Table 1 lists the mean and standard deviation of ΔRA and ΔDec. for each catalogue before and after these corrections.

4 PHOTOMETRIC ZERO-POINT CALIBRATION

The CFHTLS photometry is in the AB system but has systematic zero-point offsets that must be corrected. We also need to determine the *Y*-band zero-point. PS1 has *grizy* photometry that is well calibrated (Magnier et al. 2016), so it is well suited to use as a standard for improving the calibration of most bands used in this work. The calibration of CFHTLS *u* band must be handled differently, however, since this filter is not observed by PS1. Our methods for *u*-band calibration are described in Section 4.3.

4.1 Pan-STARRS1 catalogue

The PS1 catalogue contains columns corresponding to the mean flux, median flux, and flux error in each band for all objects. For convenience we convert the mean flux and flux error to AB magnitude and magnitude error via standard error propagation. To eliminate false detections, we require that an object has at least three 'good' detections (nmag_ok≥1) in the six bands. The PS1 photometry has been found to have small zero-point offsets compared to the standard AB system (Scolnic et al. 2015); we have shifted the PS1 *grizy* magnitudes by +20, +33, +24, +28, and +11 mmag (*griz* offsets from table 3 of Scolnic et al. 2015; *y*-band offset from private communication from Dan Scolnic), respectively, to match to the AB system.

4.2 Zero-point calibration of *grizY* bands

The filter throughputs and overall system responses vary between different telescopes even for the same nominal band, so in general the measured fluxes of the same source should differ between catalogues. However, if the filter responses are sufficiently similar and the source spectrum is nearly flat over the filter wavelength range, the brightness measured from the two telescopes should be approximately the same, as the colour measured between any two instruments/filters should be zero for a flat spectrum source (by the definition of the AB system). Such flat-spectrum sources can be approximated by observed objects with zero colour in the AB system; the magnitudes measured from two telescopes should be the same for these objects if all photometry is properly calibrated to AB. Based on this idea, we calculated the zero-point offset between PS1 and other photometry by performing a linear fit of magnitude difference as a function of colour for stars that are found in a given pair of catalogues:

$$g_c - g_p = a_{0,g} + a_{1,g} * (g_p - r_p), \tag{1a}$$

$$r_c - r_p = a_{0,r} + a_{1,r} * (r_p - i_p), \tag{1b}$$

$$i_c - i_p = a_{0,i} + a_{1,i} * (i_p - z_p), \tag{1c}$$

$$i2_c - i_p = a_{0,i} + a_{1,i} * (i2_p - z_p), \tag{1d}$$

$$z_c - z_p = a_{0,z} + a_{1,z} * (i_p - z_p), \tag{1e}$$

$$Y_s - y_p = a_{0,y} + a_{1,y} * (z_p - y_p), \tag{1f}$$

where $a_{0,m}$ is the zero-point offset, and the subscripts c, s, and p stand for CFHTLS, Subaru and PS1, respectively. As noted previously, the variable *i*2 in equation 1d represents the magnitude from the replacement filter for the CFHTLS *i* band, which was slightly different from the original *i*-band filter. It is labelled as the *y* band

in CFHTLS catalogues, but we relabel it *i*2 here to avoid confusion with the Subaru *Y* band.

In order to perform these fits, we have cross-matched the PS1 catalogue to CFHTLS and Subaru with a search radius of 1.0 arcsec. To avoid objects with large photometric errors in PS1, we require the PS1 magnitude errors to be smaller than 0.05 mag in both bands used for a given fit. Only stars that are not saturated or masked are used for calculating the offsets. For *griz* bands, we require the 'flag' value in the CFHTLS catalogue be 0 ('star' and 'not saturated or masked') and the SEXTRACTOR flag in each band to be smaller than 3, providing an additional rejection of saturated objects.

To select stars for the *Y* band, we used the star/galaxy classifier 'CLASS_STAR' from SEXTRACTOR, selecting those objects with CLASS_STAR > 0.983. There are a number of objects with much larger size that are misclassified as stars, and we removed them by applying a cut on the half-light radius: $r < 0.44$ arcsec for the deep pointing and $r < 0.41$ arcsec for the shallow pointing. We also removed saturated objects by requiring the SEXTRACTOR flag be smaller than 3 and applying a cut on MAG_AUTO to reject the brightest objects, corresponding to MAG_AUTO > 17.0 for the deep pointing and MAG_AUTO > 15.0 for the shallow pointing.

To avoid influence from outliers, we applied robust linear fitting using the PYTHON package STATSMODELS and used Huber's T as an M-estimator with the tuning constant $t = 2MAD$, where MAD is the median absolute deviation between the data and the fit. The zero-point calculation is done separately for each pointing in the CFHTLS Wide field, and separately for the two *Y*-band pointings. Fig. 2 shows the linear fit of equations (1a) to (1f) using the MAG_AUTO photometry for the CFHTLS Deep field and the Subaru deep pointing. The coefficients from the linear fits are listed in Table 2 for CFHTLS and Table 3 for the Subaru *Y* band. The a_0 in Table 3 corresponds to the offset between the initial zero-point value for the *Y*-band image and the zero-point of PS1.

So far we have assumed that the zero-point offset is uniform in each pointing. That might not be the case, and we also tried correcting for any spatial variations of the zero-point offset. To do this, we used a fixed value of the slope a_1 from the previous fit, and calculated the zero-point offset a_0 for each matching star. For example, the *g*-band offset for the *j*th object is calculated as follows:

$$a_{j,0,g} = g_{j,c} - g_{j,p} - a_{1,g} * (g_{j,p} - r_{j,p}). \tag{2}$$

After obtaining the zero-point offsets for each object, we obtained the spatial variation of the zero-point offset $a_{0,m}$(RA, Dec.) by fitting the zero-point offset to a second-order bivariate polynomial of RA and Dec. Then we obtained the calibrated magnitudes: $m' = m - a_{0,m}$(RA, Dec.). To test if the spatial correction actually improves the photometry, we calculated the median absolute deviation (MAD) of $a_{0,m}$ before and after spatial zero-point correction. Here, we randomly select 75 per cent of all objects to calculate the bivariate polynomial fit, and apply the correction on the other 25 per cent. We repeat this procedure many times to find the statistical distribution of the difference in MAD before and after correction. For corrections to be statistically significant, we require that MAD should be smaller after correction at least 95 per cent of the time. Only one pointing in CFHTLS met this requirement in one band (*z* band). Thus, we conclude that there is no significant improvement by applying spatially varying zero-point corrections, so uniform corrections were applied instead.
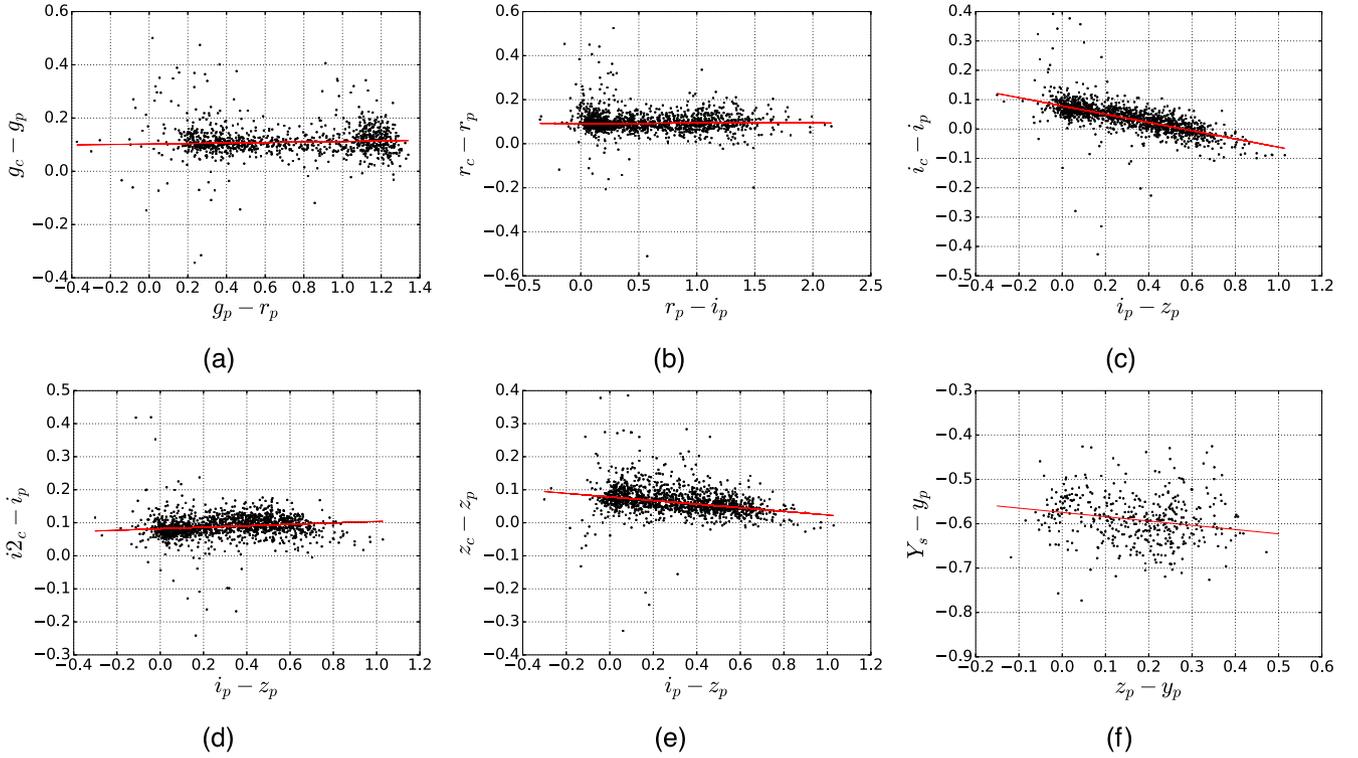
Figure 2. Panels (a–e) show difference in magnitude between CFHTLS Deep field D3 (subscript c) and PS1 (subscript p) plotted as a function of colour. Panel (f) shows the same plot for Subaru *Y* band (subscript s) from the deep pointing. Only stars are used. The red lines are the linear fits described by equations (1a)–(1f). The intercepts correspond to the zero-point offsets between the two systems, and are listed in Tables 2 and 3.

Table 2. Coefficients in equations (1a)–(1f) for CFHTLS. The coefficient a_0 corresponds to the zero-point offset between CFHTLS and Pan-STARRS, and is subtracted from the CFHTLS magnitudes to obtain calibrated values.

Pointing	Method	*g* band		*r* band		*i* band		*i2* band		*z* band	
		a_0	a_1	a_0	a_1	a_0	a_1	a_0	a_1	a_0	a_1
D3	MAG_AUTO	0.055	0.004	0.038	0.000	0.039	−0.128	0.044	0.029	0.038	−0.048
	MAG_APERCOR	0.017	0.005	0.000	0.001	0.004	−0.145	0.010	0.013	−0.002	−0.062
W3-0-1	MAG_AUTO	0.074	0.024	0.042	0.024	0.023	−0.110	–	–	0.035	−0.029
	MAG_APERCOR	0.048	−0.007	0.002	0.012	−0.008	−0.143	–	–	0.001	−0.064
W3-1-2	MAG_AUTO	0.081	0.015	0.036	0.011	0.042	−0.122	–	–	0.062	−0.036
	MAG_APERCOR	0.039	−0.001	0.011	0.001	0.000	−0.150	–	–	0.012	−0.067
W3-0-3	MAG_AUTO	0.071	0.006	0.036	0.005	0.032	−0.128	–	–	0.061	−0.036
	MAG_APERCOR	0.033	0.000	−0.004	−0.002	−0.001	−0.158	–	–	0.006	−0.061
W3 + 1-2	MAG_AUTO	0.062	−0.005	0.064	−0.002	0.025	−0.124	–	–	0.053	−0.040
	MAG_APERCOR	0.032	−0.007	0.020	−0.010	−0.006	−0.143	–	–	0.006	−0.061
W3-0-2	MAG_AUTO	0.053	0.012	0.067	0.017	0.030	−0.127	–	–	0.060	−0.031
	MAG_APERCOR	0.019	0.010	0.027	0.001	0.001	−0.149	–	–	0.013	−0.068
W3 + 1-1	MAG_AUTO	0.067	0.008	0.055	0.005	0.018	−0.119	–	–	0.058	−0.010
	MAG_APERCOR	0.031	0.000	0.015	−0.004	−0.003	−0.144	–	–	0.003	−0.060
W3-1-3	MAG_AUTO	0.065	0.000	0.056	0.001	0.025	−0.112	–	–	0.027	−0.031
	MAG_APERCOR	0.028	−0.006	0.015	−0.009	−0.013	−0.142	–	–	−0.016	−0.060

4.3 Calibration of the *u* band

Because there is no *u* band in PS1, the zero-point calibration of CFHTLS *u* band is done differently. We tried using SDSS *u* band as the standard photometry, but we encountered difficulties with this approach. First, the SDSS *u* band is significantly bluer (by

∼270 Å) than the CFHTLS *u* band; as a result the slope a_1 is large and our assumptions are less valid. Secondly, there are not many stars near zero colour in *u* − *g*, and the stars that do have colours near zero exhibit large scatter. What is worse, SDSS photometry is not exactly in the AB system. For the *u* band, it is estimated

Table 3. Coefficients in equation (1f) for Subaru Y-band photometry. The coefficient a_0 corresponds to the zero-point offset between initial zero-point value for the Y-band image and PS1. These offsets are subtracted from the Y-band magnitude to obtain calibrated values.

Pointing	Method	a_0	a_1
Deep	MAG_AUTO	-0.584	-0.097
	MAG_APERCOR	-0.646	-0.101
Shallow	MAG_AUTO	-0.653	-0.145
	MAG_APERCOR	-0.695	-0.142

that $u_{\mathrm{SDSS}} = u_{\mathrm{AB}} + 0.04$ mag with uncertainties at the 0.01–0.02 mag level.[2] Because of these problems, we have instead assumed that the CFHTLS Deep field u band is well calibrated based on the tests done for the SNLS survey (Hudelot et al. 2012), and calibrate the u-band zero-point of Wide field pointings by requiring that their $u - g$ versus $g - r$ stellar locus matches that from the Deep field. According to Hudelot et al. (2012), the calibration accuracy is at the 2 per cent level in the u band for the Deep field. Although this uncertainty in the absolute calibration remains, the procedure we have followed ensures that all the pointings at least have a uniform zero-point offset from the AB system, ensuring consistent photometry for calculating photometric redshifts.

Because not all of the CFHTLS Wide pointings overlap with the CFHTLS Deep pointing, direct calibration of the u band by cross-matching Wide and Deep objects is not feasible. Thus, we resort to an indirect calibration approach. Specifically, if all pointings are calibrated in the u, g, and r bands, their $u - g$ versus $g - r$ stellar loci should be the same. Since g and r are already calibrated, the only shift in the stellar locus should be in the $u - g$ direction, and correspond to variations in the u-band zero-point. To tie the u-band zero-point of Wide field pointings to the Deep field, we therefore need to find the relative shift in the $u - g$ direction between stellar loci in the Deep field and a Wide field pointing.

To do this, we first selected stars in the range $0.4 < g - r < 0.8$ and $u - g > 0.7$, where the stellar locus is roughly a straight line (the second cut removes outliers that are much bluer in the $g - r$ colour range). The colours of the selected stars in the Deep field were fitted to a linear function. With the same colour cuts, we fitted the stars in the Wide field pointings with a slope fixed at the Deep field value, so that the only variable is the intercept. Fig. 3 shows the $u - g$ versus $g - r$ stellar loci and linear fits for the Deep field and one of the Wide field pointings. The differences in the intercept between the Wide field pointings and the Deep field are the u-band zero-point offsets, and they are listed in Table 4.

4.4 Correction for dust extinction

The original CFHTLS $ugri(i2)z$ photometry is not corrected for Galactic extinction, nor are the PS1 magnitudes used for the photometric zero-point calibration. After zero-point calibration, we applied extinction corrections to the $ugri(i2)z$ and Y-band photometry. We followed the procedure described in Schlafly & Finkbeiner (2011), and calculated $A_b/E(B - V)_{\mathrm{SFD}}$, where A_b is the total extinction in a specific band and $E(B - V)_{\mathrm{SFD}}$ is the SFD reddening value (Schlegel, Finkbeiner & Davis 1998). We assumed

[2] http://www.sdss.org/dr12/algorithms/fluxcal/#SDSStoAB

a Fitzpatrick (1999) extinction law with $R_V = 3.1$ and used the total transmission curves of each filter for the calculation. With $A_b/E(B - V)_{\mathrm{SFD}}$, we calculated A_b using $E(B - V)_{\mathrm{SFD}}$ from the SFD dust map and applied corrections. Although the DEEP2/3 footprint is relatively small, there is a small spatial variation in $E(B - V)$ across the field, ranging from 0.006 to 0.022 with a median of 0.010. Thus, we correct for this spatial variation using the SFD map. Table 5 shows these $A_b/E(B - V)_{\mathrm{SFD}}$ values and median A_b for each band.

5 CORRECTED APERTURE PHOTOMETRY

The MAG_AUTO from SExtractor is commonly used as the default photometry in extragalactic astronomy, and it is provided in our data set. However, it is not optimal for photometric redshift calculation for several reasons. First, it uses a relatively large aperture in order to capture most of the flux from the source, but larger apertures also lead to larger background noise. Secondly, even though a large aperture is used, it still cannot capture all the flux – in our analysis typically ∼95 per cent of the total flux of a point source is captured by MAG_AUTO. Thirdly, the fraction of flux captured by MAG_AUTO might be different for objects with different sizes or images with different point spread functions (PSFs). To address these problems, we developed a method to calculate the corrected aperture photometry for both point sources and extended objects. This method utilized the aperture magnitudes at different apertures provided within the public CFHTLS catalogues, and therefore it did not require any reprocessing of the CFHTLS images. The corrected aperture magnitude is labelled 'MAG_APERCOR' in our catalogues. The MAG_APERCOR photometry is calibrated the same way as MAG_AUTO (as described in Section 4), and its zero-point offsets are listed in Tables 2–4.

Here, we summarize the techniques used for calculating 'MAG_APERCOR'. Details can be found in Appendix C. Our methods are similar to the aperture correction method described in Gawiser et al. (2006). In that work, it is assumed that all objects have a Gaussian light profile with a width calculated from the half-light radius. However, actual light profiles typically have more extended 'wings' – i.e. more flux at large radius – than Gaussian profiles do. In our work, instead of a Gaussian profile, we have used the more flexible Moffat profile (cf. equation C1), which has two free parameters, though we still assume that all objects have circularly symmetric light profiles that only depend on the half-light radius. This method essentially measures the flux in a small aperture ($r_0 = 0.93$ arcsec for $ugriz$ and $r_0 = 0.9$ arcsec for Y band) and extrapolates to infinity using the Moffat profile, the parameters of which are obtained by fitting the curve of growth (the fraction of included flux as a function of aperture radius). The aperture corrections for stars and galaxies are determined slightly differently, and the Y band is also treated differently since Y-band imaging is not available for all objects. The steps of the aperture correction for *galaxies* in band b (which could be any band except Y) in pointing x are as follows:

(i) Bin the objects in pointing x by their r-band half-light radius (FLUX_RADIUS from SExtractor).

(ii) For each r-band radius bin, find the averaged b-band curve of growth and fit the Moffat profile to that curve.

(iii) From the resulting best-fitting parameters, obtain the correction factor ApCorr = Flux(∞)/Flux(r_0) for each radius bin.
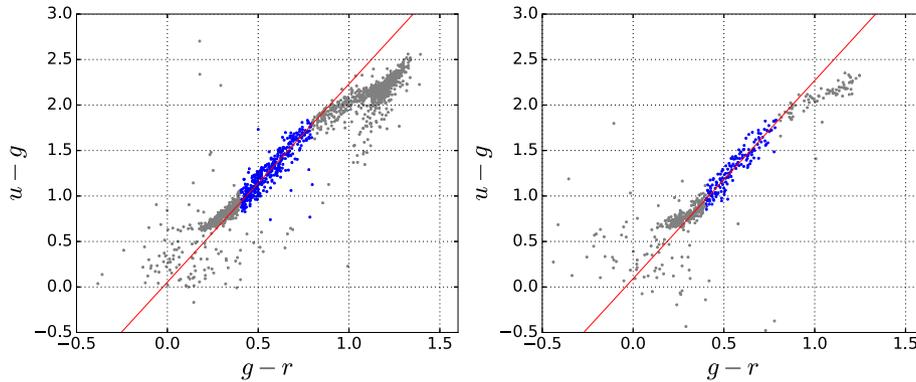
Figure 3. Stellar loci of $u - g$ versus $g - r$, using MAG_AUTO photometry. Left-hand panel: Stellar locus and linear fit of the Deep field. The red line shows a linear fit to the points in blue. The grey points are not used for the fit. The slope of the fit is used for the Wide field pointings. Right-hand panel: Wide field pointing W3-0-1; the red line has the same slope as in Deep field, and the difference in the intercept corresponds to the zero-point offset.

Table 4. The u-band zero-point offsets of the Wide field pointings relative to the Deep field. These offsets are subtracted from the Wide field u-band magnitude to obtain calibrated values.

Pointing	W3-0-1	W3-1-2	W3-0-3	W3 + 1-2	W3-0-2	W3 + 1-1	W3-1-3
MAG_AUTO	0.03	0.03	0.02	0.12	0.01	0.08	0.01
MAG_APERCOR	0.06	0.04	0.02	0.02	− 0.02	0.03	0.00

Table 5. The values of $A_b/E(B - V)_{SFD}$ in each band listed here were calculated using the procedure described in Schlafly & Finkbeiner (2011). The median A_b values are calculated for the set of DEEP2 and DEEP3 objects with spectroscopy.

Band	u	g	r	i	$i2$	z	y
$A_b/E(B - V)_{SFD}$	4.010	3.191	2.249	1.647	1.683	1.295	1.039
Median A_b	0.038	0.031	0.022	0.016	0.016	0.012	0.010

(iv) Interpolate and extrapolate the relation between the correction factor ApCorr and the mean r-band half-light radius $R_{1/2,r}$ to obtain the continuous function ApCorr$_{x,b}(R_{1/2,r})$.

(v) Use ApCorr$_{x,b}(R_{1/2,r})$ and the aperture magnitude of aperture radius r_0 to obtain the corrected aperture magnitude.

For stars, the procedure is the same except that they are not binned by radius, since the stars should effectively all have the same light profile set by the PSF; as a result, they are all placed in a bin together.

Although we can reduce background noise by choosing a small aperture, any errors in half-light radius will propagate into the total photometric error via the correction factor, and this can be a big problem for bands that have low S/N. For this reason, instead of using the SEXTRACTOR radius measurement in each band to assign the correction factor, we calculate the correction factor as a function of r-band half-light radius. In this way, we can obtain u-band MAG_APERCOR photometry even for objects with no valid radius measurement in the u band. Although the absolute photometry can be affected by any r-band radius error, the colours are not affected as much because all bands use the same r-band radius for aperture correction and thus the magnitudes are all biased in the same direction. The one exception is the Y band, for which we use the Y-band half-light radius to determine aperture corrections, as in some cases r measurements may not be available or may be noisy. The use of a matched radius makes MAG_APERCOR well suited for calculating photometric redshifts. A comparison of the photo-z performance using MAG_AUTO and MAG_APERCOR is presented in Section 7.

6 COMBINED CATALOGUES

We cross-matched the CFHTLS, Subaru Y-band catalogue and DEEP2/3 catalogues using a search radius of 1 arcsec. CFHTLS Wide field pointings were first combined into a single catalogue. For objects that appear in multiple pointings, we only kept the values from the objects that have the smallest r-band MAG_APER error. Then the Wide field combined catalogue was combined with the Deep field, keeping only the Deep field value if there is overlap. The combined CFHTLS catalogue was then matched to the Subaru Y-band catalogue. This final combined catalogue is matched to the DEEP2/3 catalogue, and all DEEP2/3 objects and columns are kept, with additional columns from CFHTLS and Subaru Y-band added. DEEP2/3 provides a quality flag, 'zquality'. Objects with secure redshifts can be selected by requiring zquality ≥ 3 (see Newman et al. 2013).

Similarly, we produced a 3D-HST grism redshift catalogue containing photometry from CFHTLS *ugriz* and Subaru Y band, as well as DEEP2/3 redshifts where available. To select objects with accurate grism redshifts, we require that either of the following criteria is met:

```
1.
((z_grism_u68-z_grism_l68)/(z_phot_u68-
z_phot_l68)<0.1)
  & ((z_grism_u68 - z_grism_l68) < 0.01)
  & (z_best_s != 0)
  & (use_phot == 1)
  & (z_max_grism > z_phot_l95)
```
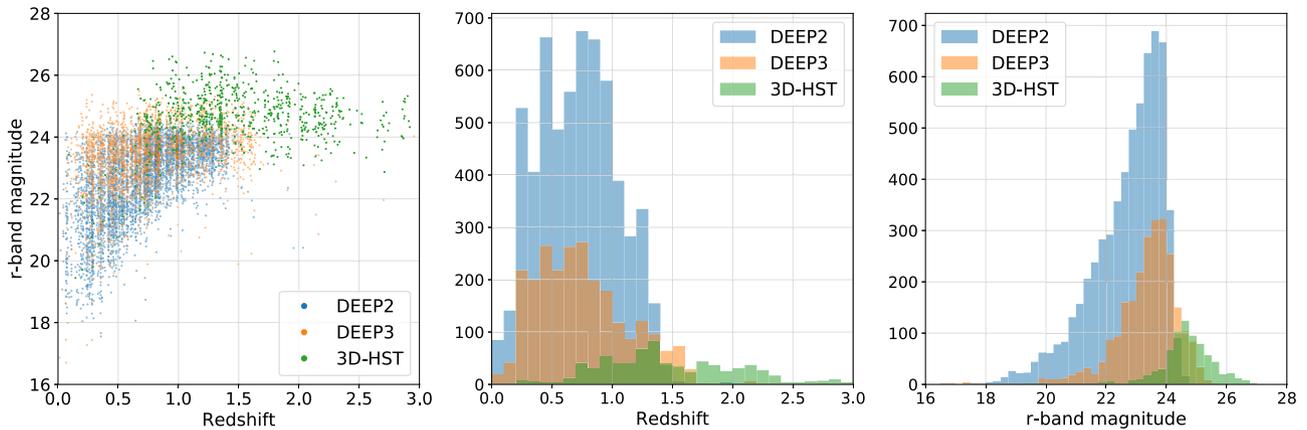
Figure 4. Left-hand panel: *r*-band magnitude versus redshift for objects in DEEP2, DEEP3, and 3D-HST with *ugrizY* photometry. If an object appears in both DEEP2/3 and 3D-HST, only the DEEP2/3 object is plotted. The large-scale structure is clearly visible. The middle panel and the right-hand panel show the redshift distributions and the *r*-band magnitude distributions, respectively.

```
& (z_max_grism < z_phot_u95)
& (z_max_grism > 0.6)
OR
2.
(z_grism_err < 0.025)
& (use_zgrism == 1),
```

where all names are quantities provided in the 3D-HST catalogue. We have compared the grism redshifts selected using the above criteria with DEEP2/3 redshifts; the normalized median absolute deviation between the spectroscopic and grism redshifts of the resulting sample is < 0.3 per cent, and the fractions of objects with larger than $0.10(1 + z)$ or larger than $0.02(1 + z)$ redshift difference are 3 per cent and 11 per cent, respectively. For convenience, we added a flag 'use_zgrism1' to the catalogue, and objects that meet the above criteria are assigned the flag value 1; otherwise this flag value will be 0.

Before cross-matching, the CFHTLS Deep and Wide catalogues include 603 852 and 1415 859 objects, respectively, and the *Y*-band catalogue includes 94 014 objects. The combined DEEP2/3 catalogue from the aforementioned cross-matching procedures includes 8479 objects with *ugrizY* photometry and secure DEEP2/3 redshifts, and the combined 3D-HST catalogue provides an additional 741 objects with accurate grism redshifts. Fig. 4 shows the distribution of *r*-band magnitude (MAG_APERCOR) and redshift for objects with *ugrizY* photometry and secure redshift measurements.

The columns in the catalogues are structured in the following way: the first columns listed are those from the relevant spectroscopic/grism catalogue; then the columns from CFHTLS with MAG_APERCOR and its errors are provided; and finally the *Y*-band columns are given. In the 3D-HST catalogue, we also include columns of DEEP2/3 redshift and other values. Description of the DEEP2/3 columns can be found at http://deep.ps.uci.edu/deep3/zt ags.html and are described in Newman et al. (2013). Description of the 3D-HST columns can be found in table 5 of Momcheva et al. (2016). Description of the CFHTLS columns can be found at http://terapix.iap.fr/cplt/T0007/doc/T0007-docsu22.html. The *Y*-band columns follow the same naming convention as CFHTLS. Note that the '*y*' variant of the CFHTLS *i* band is relabelled *i*2 in our catalogues to limit confusion. In the catalogues, 99 indicates non-detection and −99 indicates the object/quantity is not observed. Columns of principal interest are described in Table 6.

MAG_APERCOR has two sources of error: image noise and uncertainty in aperture correction. We note that errors in colours cannot be obtained by simply adding up the two kinds of errors in quadrature due to covariances between how magnitudes were determined in each band; colour errors will be smaller than one would expect if measurements in each filter were assumed to be independent. More details of how to use the errors in MAG_APERCOR can be found in Appendix C.

We also provide the photometry-only catalogues of CFHTLS Wide, CFHTLS Deep, and *Y* band. These catalogues contain calibrated MAG_AUTO and MAG_APERCOR photometry, but are not matched to any other data set.

7 PHOTOMETRIC REDSHIFT TESTS

In this section, we describe the photo-*z* tests performed on the catalogues. In general, there are two classes of method for calculating the photometric redshifts. One is the template-fitting method, in which the redshift is obtained from the best fit to the photometry (in the chi-squared sense) determined using known template SEDs. The other is the empirical method, in which a with spectroscopic redshifts is used to train an empirical relation between photometry and redshift (typically via machine learning algorithms), and the empirical relation is then applied to new photometric data to estimate the redshift. Here, we use a machine learning algorithm called random forest regression (Breiman 2001) which is included in the PYTHON package SCIKIT-LEARN (Pedregosa et al. 2011). Random forest is an ensemble learning method based on decision trees. A simple decision tree is trained by minimizing the sum of squared errors, and it tends to fit the noise in the data (i.e. overfitting). The overfitting results in reduced accuracy when the algorithm applied to new data. Random forest addresses this problem in two ways. First, a large number of new samples are created by bootstrapping the original training sample, and separate decision trees are trained using each sample. Secondly, instead of all the features (colours in our case), a random subset of the features may be used at each tree split to reduce the correlation between the trees. Although overfitting can occur in individual trees, the effect is reduced by using subsets of features and averaged out by combining the predictions from all the trees. In our analysis using a subset of features did not significantly improve the results, and thus all available features were used at each split.

Table 6. Description of some of the principal columns included in our matched catalogues. The last three columns are DEEP2/3 values added to the 3D-HST catalogue.

Column name	Description
u, g, . . .	MAG_AUTO magnitude in *u* band, *g* band, . . .
uerr, gerr, . . .	MAG_AUTO magnitude error in *u* band, *g* band, . . .
u_apercor, g_apercor, . . .	MAG_APERCOR magnitude in *u* band, *g* band, . . .
uerr_aper, gerr_aper, . . .	MAG_APERCOR magnitude error from image noise in *u* band, *g* band, . . .
uerr_apercor, gerr_apercor, . . .	MAG_APERCOR magnitude error from correction uncertainty in *u* band, *g* band, . . .
r_radius_arcsec	*r*-band half-light radius in arcsec
y_radius_arcsec	*Y*-band half-light radius in arcsec
cfhtls_source	Source of the *ugri(i2)z* photometry: 0 = Deep field; 1 = Wide field; −99 = not observed
subaru_source	Source of the *Y*-band photometry: 0 = deep pointing; 1 = shallow pointing; −99 = not observed
ra_deep2, dec_deep2	Right ascension and declination from DEEP2/3
ra_cfhtls, dec_cfhtls	Right ascension and declination from CFHTLS after astrometric correction
ra_subaru, dec_subaru	Right ascension and declination from the Subaru *Y*-band data after astrometric correction
sfd_ebv	$E(B - V)$ from Schlegel, Finkbeiner, and Davis (1998) dust map
zhelio	(In DEEP2/3 catalogue) DEEP2/3 heliocentric redshift
zquality	(In DEEP2/3 catalogue) DEEP2/3 redshift quality flag
use_zgrism1	(In 3D-HST catalogue) our grism redshift quality flag: 0 = less accurate; 1 = accurate
z_max_grism	(In 3D-HST catalogue) 3D-HST grism redshift
z_deep2	(In 3D-HST catalogue) DEEP2/3 heliocentric redshift
z_err_deep2	(In 3D-HST catalogue) DEEP2/3 redshift error
zquality_deep2	(In 3D-HST catalogue) DEEP2/3 redshift quality flag

Both DEEP2/3 and 3D-HST data were employed to train and assess the performance of the algorithm. The selection of DEEP2/3 and 3D-HST redshifts is described in Section 6. For objects that appear in both DEEP2/3 and 3D-HST, the DEEP2/3 redshift values are used. To avoid training and testing on the same data set, we applied the K-fold cross-validation method: the data set is first randomly divided into five subsets. Then one subset is selected as the testing set and the other four subsets are combined as a training set for optimizing the random forest, and this procedure is repeated five times so that the entire data set has been used as the testing set in the end. The estimated photometric redshift derived for a given object when it was in the testing set is then compared with the spectroscopic/grism redshift (from now on simply spectroscopic redshift or $z_{\rm spec}$ for convenience) and the redshift difference $\Delta z = z_{\rm phot} - z_{\rm spec}$ is calculated. Two quantities are used to evaluate the photo-*z* performance here: the normalized median absolute deviation $\sigma_{\rm NMAD} = 1.48\,{\rm MAD}$, where ${\rm MAD} = {\rm median}(|\Delta z|/(1 + z_{\rm spec}))$, and the outlier fraction η which is defined as the fraction of objects with $|\Delta z| > 0.15/(1 + z_{\rm spec})$.

For consistent S/N in the photometry, the CFHTLS Wide field and Deep field are tested separately, and in both cases the *Y*-band photometry from both the deep and shallow pointing are used. Valid photometry in all six bands (*ugrizY*) is required. We have tested the photometric redshift performance for both MAG_AUTO and MAG_APERCOR photometry. The five colours $u - g$, $g - r$, $r - i$, $i - z$, $z - y$, and *i*-band magnitude are used as the input.

Fig. 5 shows the photo-*z* results using the CFHTLS Wide field photometry, and Fig. 6 shows the results with CFHTLS Deep field photometry. We find that using the MAG_APERCOR photometry, we achieve photo-*z* accuracy $\sigma_{\rm NMAD} = 0.018$ and outlier fraction of 4.7 per cent in the CFHTLS Deep field, and $\sigma_{\rm NMAD} = 0.039$ and 6.3 per cent outliers in the CFHTLS Wide field. This represents a significant improvement over MAG_AUTO: $\sigma_{\rm NMAD}$ is reduced by 28 per cent in CFHTLS Wide and 27 per cent in CFHTLS Deep, and there is also a significant reduction in the outlier fraction. The

scatter in Δz is larger at $z_{\rm spec} > 1.4$ for both MAG_AUTO and MAG_APERCOR photometry and in both the Deep and Wide areas. This is due to both the small number of training objects in this redshift range, as well as the lack of available features (e.g. the 4000 Å break) in the optical.

As an additional validation of the MAG_APERCOR photometry, we have performed similar photo-*z* tests using the CFHTLS photometry from the 3D-HST photometric catalogues (Skelton et al. 2014). In that work, the objects were detected with HST imaging, and forced photometry of these objects were performed on the CFHTLS Deep *ugriz* images with an aperture of 1.2 arcsec. We performed photo-*z* tests using the *ugriz* photometry from Skelton et al. (2014) and redshifts from DEEP2/3 and 3D-HST, and for comparison we ran the same test using the CFHTLS Deep MAG_AUTO and MAG_APERCOR photometry in *ugriz* bands for the same objects. We find that the Skelton et al. (2014) *ugriz* photo-*z*'s have very similar accuracy to the MAG_APERCOR photo-*z*'s, with the former having 2 per cent smaller $\sigma_{\rm NMAD}$ and 17 per cent fewer outliers. Both significantly outperform the MAG_AUTO photo-*z*'s, with the Skelton et al. (2014) *ugriz* photo-*z*'s having 37 per cent smaller $\sigma_{\rm NMAD}$ and 47 per cent fewer outliers than MAG_AUTO.

The CFHTLS Deep field and the Subaru *Y* band have depth similar to LSST 10 yr data. Therefore, this test also demonstrates that in the magnitude and redshift range of DEEP2/3, at least, it is possible for LSST to achieve the goal of $0.03(1 + z)$ photo-*z* accuracy as specified by the Science Requirements Document of the LSST Dark Energy Science Collaboration (The LSST Dark Energy Science Collaboration 2018).

8 SUMMARY

In this work, we have presented a set of new catalogues with improved *ugrizY* photometry and spectroscopic or grism redshifts in the EGS. We calibrated CFHTLS *ugriz* photometry and Subaru *Y*-
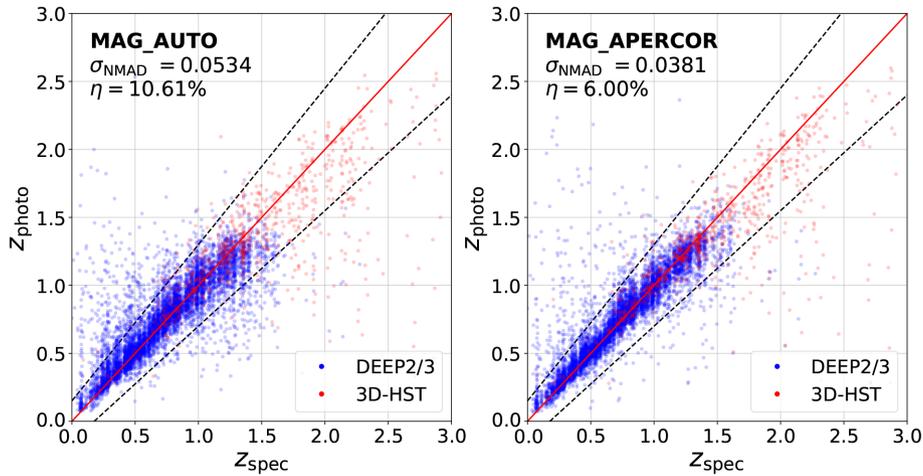
Figure 5. Photometric redshift versus spectroscopic or grism redshift using CFHTLS Wide field *ugriz* and Subaru *Y*-band photometry. The red solid line corresponds to $z_{\text{photo}} = z_{\text{spec}}$. The dashed lines mark the boundary separating the outliers. The MAG_APERCOR photometry produces photo-*z*'s with significantly better accuracy than MAG_AUTO.
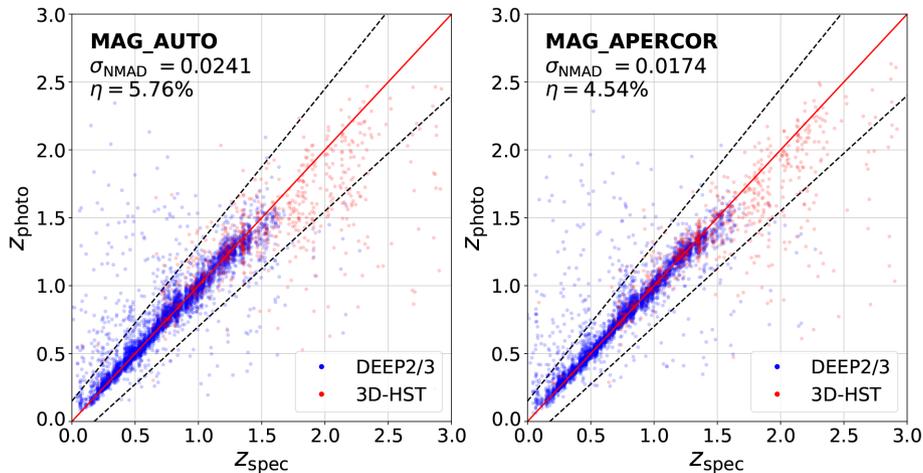

Figure 6. Same as Fig. 5, but using CFHTLS Deep field photometry instead.

band photometry and also produced corrected aperture magnitudes. We combined the *ugrizY* photometry with DEEP2/3 and 3D-HST redshifts. The *ugrizY* photometry has depth similar to the LSST 10 yr stack, and the catalogues will be useful for LSST photo-*z* tests. All data is publicly available.

We have implemented a random forest photo-*z* algorithm on our data set, and found the photo-*z* accuracy to be ∼2 per cent or better for the available spectroscopic sample in the deepest region, where the photometry has LSST-like depth. We also found significant improvement in photo-*z* accuracy from the corrected aperture magnitude, indicating that our corrections provide a real improvement in the measurement of galaxy colours (as they tighten the colour–redshift relation).

REFERENCES

Alam S. et al., 2015, ApJS, 219, 12
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Brammer G. B. et al., 2012, ApJS, 200, 13
Breiman L., 2001, Mach. Learn., 45, 5
Chambers K. C. et al., 2016, preprint (arXiv:1612.05560)
Conselice C. J. et al., 2007, MNRAS, 381, 962
Cooper M. C. et al., 2011, ApJS, 193, 14
Cooper M. C. et al., 2012, MNRAS, 419, 3018
Fitzpatrick E. L., 1999, PASP, 111, 63
Gawiser E. et al., 2006, ApJS, 162, 1
Gunn J. E. et al., 1998, AJ, 116, 3040
Gwyn S. D. J., 2008, PASP, 120, 212
Hodges J. L., Lehmann E. L., 1963, Ann. Math. Stat., 34, 598
Hudelot P. et al., 2012, VizieR Online Data Catalog, 2317
Ivezic Z. et al., 2009, American Astronomical Society Meeting Abstracts, #213. p. 366
Lang D., Hogg D. W., Mierle K., Blanton M., Roweis S., 2010, AJ, 139, 1782
LSST Science Collaboration, 2009, preprint (arXiv:0912.0201)
Magnier E. A. et al., 2016, preprint (arXiv:1612.05242)
Matthews D. J., Newman J. A., Coil A. L., Cooper M. C., Gwyn S. D. J., 2013, ApJS, 204, 21
Miyazaki S. et al., 2002, PASJ, 54, 833
Momcheva I. G. et al., 2016, ApJS, 225, 27
Newman J. A. et al., 2013, ApJS, 208, 5
Ouchi M. et al., 2004, ApJ, 611, 660
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Scolnic D. et al., 2015, ApJ, 815, 117
Skelton R. E. et al., 2014, ApJS, 214, 24
The LSST Dark Energy Science Collaboration, 2018, preprint (arXiv:1809.01669)
Willner S. P. et al., 2006, AJ, 132, 2159
Willner S. P. et al., 2012, ApJ, 756, 72

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

3D-HST_Terapix_Subaru_v1.fits
3D-HST_Terapix_Wide_Subaru_v1.fits
DEEP2_uniq_Terapix_Subaru_v1.fits
DEEP2_uniq_Terapix_Wide_Subaru_v1.fits

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

APPENDIX A: THE DEEP3 GALAXY REDSHIFT SURVEY

The DEEP3 Galaxy Redshift Survey was a Large Multi-Annual Program allocated 25.5 nights of time on the DEIMOS spectrograph at the Keck 2 telescope to measure redshifts and properties of galaxies in the EGS. The combination of DEEP2 and DEEP3 provides roughly 18 000 redshifts in the portion of the EGS overlapping the greatest amount of multiwavelength data, including multiband imaging from *HST* and *Spitzer* and deep ACIS imaging with *Chandra*.

DEEP3 includes observations of 56 DEIMOS slitmasks, tiling the central portion of the EGS and building upon the 120 slitmasks observed in the EGS as part of DEEP2. Observations for DEEP3 began in 2008 April and continued until 2011 May. In total, DEEP3 targeted ∼7500 sources, yielding ∼5000 secure redshifts. Here, we provide details regarding the target selection, observations, and data reduction for DEEP3. The first public version of the DEEP3 redshift catalogue as well as sky-subtracted one- and two-dimensional spectra of each target are available at http://deep.ps.uci.edu/deep3/home.html.

A1 DEEP3 target samples

The DEEP3 targeting strategy differs in a number of respects from the target selection strategy used by DEEP2 in the EGS, which was described by Newman et al. (2013). First, at highest priority a set of objects were targeted based upon their unusual multiwavelength properties (e.g. X-ray or far-IR sources) in AEGIS imaging. A list of the various multiwavelength sources observed and the bits used to identify objects from each sample in the DEEP3 redshift catalogue can be found in Table A1. These objects were restricted to comprise only a small fraction of the overall sample to make sure that clustering measurements for the overall sample are not strongly affected.

Secondly, the ∼35 per cent of $R_{AB} < 24.1$ galaxies which were unable to be targeted by DEEP2 due to slit collisions were assigned the next highest priority in maskmaking for DEEP3, providing in combination with DEEP2 a uniform sample of more than 90 per cent of all $R_{AB} < 24.1$ in the DEEP3 area that can be used for measurements of environment statistics and galaxy clustering. Thirdly, at lowest priority a 'faint extension' of targets with $24.1 < r_{AB} < 25.5$ in CFHTLS imaging were targeted; the resulting sample was expected to be systematically incomplete but still yield a number of useful redshifts (in the end, roughly 40 per cent of $R > 24.1$ targets in DEEP3 have secure redshift measurements).

A major difference between the strategies of DEEP2 and DEEP3 was the use of the 600-line grating on DEIMOS for the latter, instead of the 1200-line grating that was used in DEEP2. The added spectral coverage to the blue from using a lower resolution grating enables enhanced studies of line ratios, metallicities, AGN properties, K + A galaxy signatures, and Mg II wind absorption compared to DEEP2. Kinematic measurements of small-linewidth galaxies are not possible at this lower resolution, but these are already abundantly available in DEEP2. Tests prior to the start of DEEP3 found no reduction in redshift success using a lower resolution, despite the greater difficulty in resolving the [O II] doublet; in the end, DEEP3 obtained secure (ZQUALITY 3 or 4) for 69 per cent of galaxy targets with $R_{AB} < 24.1$, versus 73 per cent in DEEP2.

The DEEP3 spectra cover a broader wavelength range than those from DEEP2, spanning 4550–9900 Å (with a central wavelength of 7200 Å). The GG455 order-blocking filter was used to limit flux blueward of 4550 Å, and each slitmask was observed for approximately 1 h, depending on the observing conditions (i.e. transparency and/or seeing). Typical slitlengths were ∼4–8 arcsec, with a standard 1 arcsec slitwidth. A standard DEEP3 exposure consists of three 1200 s subexposures, which are used to remove cosmic rays and are then co-added to make a total exposure of 1 h.

A2 DEEP3 maskmaking and tiling strategies

The sky region covered by DEEP3 corresponds to the central 50 per cent (in the long direction) of the DEEP2 region of EGS, as shown in Fig. A1. This region corresponds to the intersection of the most important multiwavelength surveys in the field, including coverage with *Spitzer* IRAC and MIPS, *HST*/ACS, *Chandra*, and *GALEX*. VLA 20 cm data are poorer in the lower part of the strip owing to interference by the bright source 3C295, providing additional reason to avoid the southern end of DEEP2 for this project.

DEEP3 masks were spaced 1.5 arcmin apart, rather than 1 arcmin as in DEEP2, to match the density of targets for the program (since the majority of $R_{AB} < 24.1$ galaxies were already targeted by DEEP2). DEEP3 masks cover a strip that is 15 arcmin wide with DEIMOS, as DEEP2 did, even though the region covered by *Spitzer* and *HST* is only 10 arcmin wide. This is needed in order to create an overhang region that extends at least 2.5 arcmin beyond the prime imaging area in all directions. This buffer zone allows us to measure environmental densities for all objects in the prime zone free of edge effects. Without it, only half of the 10 arcmin wide zone would be suitable for environmental studies. The strip covered by DEEP3 masks is 1 deg long, versus 2 deg for DEEP2.

Targets were placed on masks using a modified version of the maskmaking algorithms described in Newman et al. (2013). For multiwavelength-selected objects, a wide variety of selection algorithms were used (cf. Table A1). The priority of objects from this table was used as the selection weight (W as defined in Newman et al. 2013) for them.

For $R_{AB} < 24.1$ objects, the selection is similar but not identical to that used for DEEP2 in the Groth Strip. As before, objects were required to meet the magnitude limits of the DEEP2 survey ($18.25 < R_{AB} < 24.1$; to have at least 20 per cent probability of being a galaxy ($p_{gal} > 0.2$, as defined in Newman et al. 2013); and to have no imaging pixel flags set in the R band. Unlike in DEEP2, however, objects on either side of the DEEP2 colour selection cuts were treated identically for DEEP3, and objects with non-detections in the B or I band or with low surface brightnesses were included in the sample. DEEP2-like objects received a magnitude-based target selection weight W_R (again, as defined in Newman et al. 2013) given by $min(0.75 \times 10^{-0.4*(R-24.1)}, 1)$; this function falls from 1 at $R \leq 23.8$ to 0.75 at $R = 24.1$; this is the same functional form used for higher redshift objects in the EGS in DEEP2. Objects which were previously observed by DEEP2 but received non-secure redshifts in visual inspection ($Q = 2$) were included in the sample for DEEP3, but with W_R lowered by a factor of two

Table A1. Table of all target samples included in DEEP2. For each sample, we specify the corresponding bit in EGSFLAGS; the nature of the sample; the individual who provided it; the priority assigned in target selection (as input to the target selection procedure described in Newman et al. (2013); the number of targets within the overall DEEP3 footprint; the number of targets whose spectra were obtained; and the fraction of objects in the catalogue which were targeted.

| Bi	Target class	Contact	Median priority	# in mask area	# with spectrum	# with secure z	Fraction targeted	Fraction of targets with secure z
2^0	Spitzer/MIPS 70 μm sources, priority 1	Mark Dickinson	4×10^{14}	300	204	144	0.680	0.706
2^1	Spitzer/MIPS 70 μm sources, priority 2	Mark Dickinson	2×10^{14}	83	46	32	0.554	0.696
2^2	Spitzer/MIPS 24 μm sources, priority 1	Mark Dickinson	4×10^{14}	86	47	27	0.547	0.574
2^3	Spitzer/MIPS 24 μm sources, priority 2	Mark Dickinson	2×10^{14}	20	15	9	0.750	0.600
2^4	Chandra sources	Kirpal Nandra	8×10^{8}	205	141	72	0.688	0.511
2^5	$z < 2$ Massive Galaxies from AEGIS (Conselice et al. 2007)	Christopher Conselice	5×10^{11}	441	255	192	0.578	0.753
2^6	VLA 20 cm sources (Willner et al. 2012)	Robert Ivison	8×10^{14}	125	93	49	0.744	0.527
2^7	Bright Akari/IRC 15 μm sources	Myungshin Im	1×10^{14}	65	55	39	0.846	0.709
2^8	Faint Akari/IRC 15 μm sources	Myungshin Im	5×10^{5}	91	44	16	0.484	0.364
2^9	Spitzer IRS targets	Jiasheng Huang	2×10^{14}	4	3	0	0.750	0.000
2^{10}	Spitzer/IRAC Power-law AGN candidates	Jiasheng Huang	2.5×10^{8}	76	49	18	0.645	0.367
2^{11}	VLA 6 cm sources (Willner et al. 2006)	Steven Willner	4×10^{14}	90	73	38	0.811	0.521
2^{12}	Spitzer/IRAC-identified AGN	David Rosario	1×10^{14}	24	22	9	0.917	0.409
2^{13}	New strong lens systems	Leonidas Moustakas	2×10^{15}	3	2	2	0.667	1.000
2^{14}	Spitzer IRS object	Christopher Willmer	Not in area	N/A	N/A	N/A	N/A	N/A
2^{15}	Dual AGN candidates	Brian Gerke	2×10^{15}	7	7	7	1.000	1.000
2^{16}	DEEP2 objects – previously untargeted	Jeffrey Newman	8.2×10^{11}	7454	4420	3181	0.593	0.720
2^{17}	DEEP2 objects – previously targeted	Jeffrey Newman	0.5	2595	1205	605	0.464	0.502
2^{18}	DEEP3 faint extension	Jeffrey Newman	0.0595	30868	1346	539	0.044	0.400
2^{19}	DEEP2 strong lens reobservations	Jeffrey Newman	1×10^{14}	12	12	12	1.000	1.000
2^{20}	SNLS supernova hosts – high priority	Saul Perlmutter	2000	3	1	1	0.333	1.000
2^{21}	SNLS supernova hosts – low priority	Saul Perlmutter	1×10^{14}	42	29	23	0.690	0.793
2^{22}	AEGIS-X sources	Kirpal Nandra	8×10^{14}	141	96	63	0.681	0.656
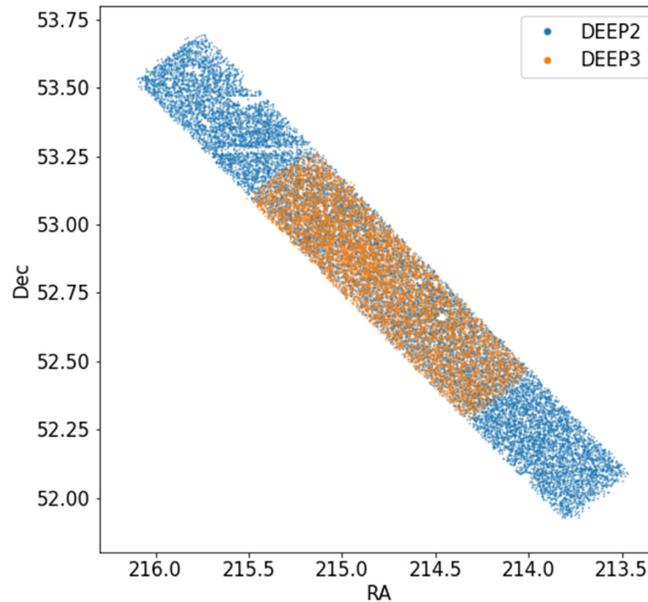
Figure A1. Sky coverage DEEP2 and DEEP3. The points are all the targets in the surveys, regardless of whether successful redshifts were obtained.

(so the maximum possible W_R for such objects was 0.5, instead of 1 for an unobserved DEEP2-like object). Unlike in DEEP2, W_R was not multiplied by the galaxy probability from star–galaxy separation, so the overall selection priority is $W = W_R$ for this sample.

For the 'faint extension' of $R_{AB} > 24.1$ objects, the selection procedure was modified since CFHTLS data was used. Specifically, the CFHT Sextractor MAG_AUTO r magnitudes from the 2008A Megapipe CFHTLS catalogues produced by Stephen Gwyn (http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/megapipe/cfhtls/index.html; Gwyn 2008) were used to select targets for DEEP3. Eligible faint extension targets had $R_{AB} > 24.1$ in the DEEP2 photometric catalogues and $r > 23.5$ in CFHTLS, or a non-detection in the DEEP2 catalogues and $r > 24.22$ in CFHTLS (reflecting the average offset between DEEP2 R and CFHTLS r), $r < 25.62$ in CFHTLS, and no r-band pixel flags set in the CFHTLS imaging. These objects were given a weight $W_R = 0.2 \times \min(0.25 \times 10^{-0.2*(r-25.5)}, 1)$; this function falls from 0.090 at $r = 24.22$ to 0.047 at $r = 25.62$.

Maskmaking then proceeded via the same procedure used by DEEP2, with the central region (in the wavelength direction) of each mask populated in a first mask, and outer portions populated second; there were only two minor differences in the procedure. First, for DEEP3, a minimum slit length of 4 arcsec, rather than 3 as in DEEP2, was used.

Secondly, in cases where multiple objects conflicted with each other such that they could not all be observed simultaneously, the target to be observed is chosen randomly. For DEEP2, this was done by generating a random value between 0 and 1 for each object and choosing the one with highest random value. For DEEP3, this behaviour was altered to ensure selection of high-priority targets. Specifically, for objects with weight $W > 1$, the object weight is multiplied by a random number uniformly distributed between 0.75 and 1; for objects with weights between 0.25 and 1, the object is assigned a random number uniformly distributed between 0 and 1 with no multiplication by weight; and for objects with weights below 0.25, the object weight is multiplied by a random number between 0 and 1. Apart from these minor differences, maskmaking proceeded as in DEEP2.

A3 Data reduction and catalogues

The DEIMOS data were reduced using a version of the DEEP2 DEIMOS `spec2d` pipeline slightly modified to improve handling of 600-line grating data, yielding sky-subtracted 1D and 2D spectra for each object. Redshifts were then measured using the DEEP2 `spec1d` Redshift Pipeline, with each redshift inspected by eye by at least one individual and assigned a quality code. The quality code system used is the same as DEEP2. ZQUALITY = −2 indicates a spectrum with data so poor for instrumental reasons that it was effectively not observed. ZQUALITY = −1 is used for stars. ZQUALITY = 1 indicates a spectrum with such poor signal to noise that it is unlikely a redshift could be recovered, and ZQUALITY = 2 indicates that a reliable redshift could not be established for reasons specified in the COMMENT field. Finally, ZQUALITY = 3 indicates a secure redshift (> 95 per cent probability of being correct), and ZQUALITY = 4 indicates highly secure cases (> 99 per cent probability of being correct). More details on the DEEP2 code used and the basic properties included in redshift catalogues may be found in Newman et al. (2013).

The DEEP3 redshift catalogue adds a new tag (or column) for each object, EGSFLAGS, which has no analogue in the DEEP2 redshift catalogue. This tag provides information about which objects belong to which input target list. Unlike DEEP2, which employed a single set of selection cuts on $R_{AB} < 24.1$ galaxies, DEEP3 has targeted a variety of sources pulled from a variety of input catalogues provided by collaborators. Table A1 shows the breakdown of the target list according to the flag values (and associated target lists). Many objects will have been eligible for targeting based on multiple reasons; e.g. a source might be both a 'FIDEL 24 µm priority 1' source and a 'DEEP2 previously untargeted' object. In such cases, all of the relevant flags are set – for example, a *Chandra* source which is also a power-law AGN candidate will have both the 2^4 and 2^{10} bits set, corresponding to an EGSFLAGS value of 1040. In other words, the EGSFLAGS value is

an integer value containing the bitwise OR of all of the flag values pertaining to a given object. As can be seen from the table, the fractions of objects selected varied from survey to survey both due to varying target priorities (as listed in the table) and varying sky coverage; for instance, many *Chandra* sources fell at the ends of the slitmasks and thus were not able to be assigned a slit.

APPENDIX B: *Y*-BAND SEXTRACTOR PARAMETERS

Source catalogues in the *Y* band were obtained by running SEXTRACTOR on the *Y*-band images. The SEXTRACTOR parameters used for the deep pointing are listed in Appendix B1. For the shallow pointing, only a few parameters were altered; these are listed at the end of the table. The 'PHOT_APERTURES' parameters specify the aperture diameters of the MAG_APER photometry, which we use to compute the MAG_APERCOR photometry. Note that SEXTRACTOR (version 2.19.5) cannot produce more than 30 aperture magnitudes, so we had to separate the apertures into two parameter files (but with the same maximum aperture size to ensure the same set of detections) and run them separately.

B1 SEXTRACTOR parameters

```
SExtractor parameters for the deep pointing

#---------------------------- Extraction ----------------------------------

DETECT_TYPE CCD # CCD (linear) or PHOTO (with gamma correction)
DETECT_MINAREA 3 # min. # of pixels above threshold
DETECT_MAXAREA 6400

DETECT_THRESH 2.0 # <sigmas> or <threshold>,<ZP> in mag.arcsec-2
ANALYSIS_THRESH 2.0 # <sigmas> or <threshold>,<ZP> in mag.arcsec-2
THRESH_TYPE RELATIVE

FILTER Y # apply filter for detection (Y or N)?
FILTER_NAME gauss_2.5_5x5.conv # name of the file containing the filter

DEBLEND_NTHRESH 64 # Number of deblending sub-thresholds
DEBLEND_MINCONT 0.001 # Minimum contrast parameter for deblending

CLEAN Y # Clean spurious detections? (Y or N)?
CLEAN_PARAM 1.0 # Cleaning efficiency

#---------------------------- WEIGHTing ----------------------------------

WEIGHT_GAIN N # If true, weight maps are considered as gain maps.
WEIGHT_TYPE MAP_RMS # type of WEIGHTing: NONE, BACKGROUND,
 # MAP_RMS, MAP_VAR or MAP_WEIGHT
WEIGHT_IMAGE weight_maps/BACKGROUND_RMS_SIZE_16.FITS # weight-map filename

#---------------------------- FLAGging ----------------------------------

FLAG_IMAGE edge_flag.fits # filename for an input FLAG-image
FLAG_TYPE MOST # flag pixel combination: OR, AND, MIN, MAX
 # or MOST

#---------------------------- Photometry ----------------------------------

PHOT_APERTURES 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24,
25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44,
45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56
 # MAG_APER aperture diameter(s) in pixels
PHOT_AUTOPARAMS 2.5, 3.5 # MAG_AUTO parameters: <Kron_fact>,<min_radius>
PHOT_PETROPARAMS 2.0, 3.5 # MAG_PETRO parameters: <Petrosian_fact>,
 # <min_radius>
PHOT_AUTOAPERS 20.0,20.0 # <estimation>,<measurement> minimum apertures
 # for MAG_AUTO and MAG_PETRO
PHOT_FLUXFRAC 0.2, 0.5, 0.8 #Fraction of FLUX AUTO defining each element of
```

```
the FLUX RADIUS vector.

SATUR_LEVEL 36000.0 # level (in ADUs) at which arises saturation
SATUR_KEY SATURATE # keyword for saturation level (in ADUs)

MAG_ZEROPOINT 31.2 # magnitude zero-point
MAG_GAMMA 4.0 # gamma of emulsion (for photographic scans)
GAIN 1 # detector gain in e-/ADU
GAIN_KEY GAIN # keyword for detector gain in e-/ADU
PIXEL_SCALE 0 # size of pixel in arcsec (0 = use FITS WCS info)


#---------------------- Star/Galaxy Separation ----------------------------


SEEING_FWHM 0.648 # stellar FWHM in arcsec
STARNNW_NAME default.nnw # Neural-Network_Weight table filename


#--------------------------- Background ---------------------------------


BACK_TYPE AUTO # AUTO or MANUAL
BACK_VALUE 0.0 # Default background value in MANUAL mode
BACK_SIZE 128 # Background mesh: <size> or <width>,<height>
BACK_FILTERSIZE 5 # Background filter: <size> or <width>,<height>
BACKPHOTO_TYPE LOCAL
BACKPHOTO_THICK 24


#-------------------- Memory (change with caution!) ---------------------


MEMORY_OBJSTACK 3000 # number of objects in stack
MEMORY_PIXSTACK 9000000 # number of pixels in stack
MEMORY_BUFSIZE 1024 # number of lines in buffer

The following parameters are for the shallow pointing:
DETECT_MINAREA 5
SATUR_LEVEL 280000.0
SEEING_FWHM 0.625
```

APPENDIX C: APERTURE CORRECTION PROCEDURES

Two assumptions are made in determining our aperture corrections. The first is that all objects have a circular symmetry and their light profiles can be described by a Moffat profile (described in more detail below). The second assumption is that in each band in each pointing, the parameters describing the Moffat profile only depend on the half-light radius and that they are smooth functions of this quantity. Under these assumptions, we can measure the flux in a small aperture and use the Moffat profile appropriate for a given object's half-light radius to extrapolate the total flux. We perform aperture corrections separately for each band in each pointing so that we can account for differences between seeing in each image.

The Moffat light profile is described by the equation

$$I(r; \alpha, \beta) = \frac{\beta - 1}{\pi \alpha^2} \left[1 + \left(\frac{r}{\alpha} \right)^2 \right]^{-\beta}, \tag{C1}$$

where I denotes the flux density and r is the angular distance from the centre of the source. There are two free parameters: α determines the width of the profile and β determines its shape. If β is small, the light profile includes more flux at larger radii (larger 'wings'), while $\beta \to \infty$ corresponds to a Gaussian profile. In this formula, the light profile is normalized so that the total flux is 1. The fraction of flux inside radius r is then

$$\mathrm{frac}(r) = \int_0^r 2\pi x I(x) \mathrm{d}x = 1 - \alpha^{2(\beta-1)} \left(\alpha^2 + x^2 \right)^{1-\beta}. \tag{C2}$$

A measurement of the half-light radius from SEXTRACTOR is provided by CFHTLS. In principle, we can determine α by solving equation (C2) for the case $I(R_{1/2}; \alpha, \beta) = 1/2$, where $R_{1/2}$ is the half-light radius, leaving only one free parameter, β. However, we found that the 'half-light' radius measured by SEXTRACTOR does not capture exactly half of the total flux, so we treat α as a free parameter as well. In the rest of this section, we use $R_{1/2}$ and the word radius to refer to the SEXTRACTOR-measured half-light radius rather than the value derived from the Moffat fit.

One set of α and β is enough to characterize the light profiles of stars since they have essentially the same light profile (i.e. the PSF). Galaxies have different light profiles, so we divide galaxies into radius bins and find the optimal α and β for each bin. The bin sizes are

0.0558 arcsec for *u* and *z* bands, 0.0372 arcsec for *g*, *r*, *i* bands, and 0.03 arcsec for *Y* band. The smallest bin is set by the PSF (stars) and the largest bin has a radius of 1.1–1.2 arcsec. We use the CFHTLS 'flag' column for star–galaxy separation.

To avoid large radius errors in bands with low S/N, and also to reduce errors in colours (e.g. $u - g$) by ensuring consistent treatment of radii, we binned objects according to their *r*-band radii when determining the aperture correction for each CFHTLS passband. For the *Y*-band aperture correction the *Y*-band radius was used for binning as many objects are not detected in *r*. For each radius bin, we compute the average curve of growth of flux as a function of radius by simply averaging the curve of growth of the individual objects within that bin.

CFHTLS provides SEXTRACTOR aperture magnitudes (MAG_APER) for aperture radii ranging from 5 to 30 pixels in 1 pixel spacing; we use these magnitudes for the curve of growth calculations. For the *Y* band, we also produced similar SEXTRACTOR aperture magnitudes; see Appendix B for details of the *Y*-band aperture magnitudes.

Only objects with relatively high S/N must be used for calculating the curve of growth to avoid background contamination, so we require the MAG_AUTO error be smaller than these limits: [0.02, 0.01, 0.01, 0.01, 0.01, 0.01] for *u*, *g*, *r*, *i*, *i*2, *z* in the CFHTLS Deep field, [0.05, 0.05, 0.05, 0.04, 0.05] for *u*, *g*, *r*, *i*, *z* in CFHTLS Wide fields, 0.02 for the *Y*-band deep pointing and 0.05 for *Y*-band shallow pointing. We also exclude saturated, masked or blended objects by requiring the CFHTLS 'flag' value to be ≤1 and the SEXTRACTOR flag (in *r* band or *Y* band) to be 0. Fig. C1 shows examples of the curve of growth fits.

We then obtain α and β by fitting equation C2 to the measured curve of growth for a given radius bin by least squares. Once we know α and β, we can measure the flux of each object in a small aperture r_0, and extrapolate to infinity to obtain the total flux. Essentially, we have then determined the aperture correction factor for a given radius bin:

$$\text{ApCorr} = \text{frac}(\infty)/\text{frac}(r_0) = \left\{ \frac{\beta - 1}{\pi\alpha^2} \left[1 + \left(\frac{r_0}{\alpha} \right)^2 \right]^{-\beta} \right\}^{-1}, \qquad (C3)$$

where α and β are fit separately for each bin.

For the *ugriz* bands, we choose the aperture radius $r_0 = 5$ pixels (0.93 arcsec), because among available apertures this choice yielded the highest signal-to-noise photometry for all but the brightest objects. For *Y* band, we choose a similar aperture radius of $r_0 = 4.5$ pixels (0.9 arcsec).

After obtaining ApCorr for each radius bin, we calculate ApCorr as a function of radius by linear interpolation to determine the correction for each individual object. To obtain the correction factor for objects larger than the largest radius bins, we must extrapolate $\text{ApCorr}(R_{1/2})$ to larger radii. To do this, we use the α and β from the largest radius bin to calculate the actual fraction of light within the SEXTRACTOR 'half-light' radius, and assume that this fraction is the same for all objects of larger radii; we then keep β fixed and use the SEXTRACTOR 'half-light' radius to estimate α and obtain ApCorr. Fig. C2 shows the correction factor ApCorr as a function of radius.

Finally, we use the function $\text{ApCorr}(R_{1/2})$ to obtain the total flux from the aperture flux within aperture radius r_0 for every object in the catalogue.

C1 Error estimation

Assuming that our model of the star and galaxy light profiles is correct, the corrected aperture magnitude MAG_APERCOR should have two sources of error: photometric errors in the aperture magnitude which were measured by SEXTRACTOR, and the error in the correction factor ApCorr which we multiplied by. In the catalogue and in this paper, we label MAGERR_APER (uerr_aper, gerr_aper, etc.) as the photometric error from SEXTRACTOR, and MAGERR_APERCOR (uerr_apercor, gerr_apercor, etc.) as the statistical uncertainty in the correction factor.

Here, we assume that the error in $\text{ApCorr}(R_{1/2})$ is only due to the error in the radius $R_{1/2}$, and the correction factor itself has negligible error if the radius is accurate. SEXTRACTOR does not provide the error in the radius, so we can only estimate this quantity indirectly. For *ugriz* bands where the *r*-band radius is used, we assume that the *i*-band radius error σ_{R_i} is the same as the *r*-band radius error σ_{R_r}, and since they are independent measurements, we can estimate σ_{R_r} from the scatter of $f_{i,r} = R_i/R_r$ about its mean value, so that

$$\frac{\sigma_{R_r}}{R_r} = \frac{\sigma_{f_{i,r}}}{\sqrt{2}\,\bar{f}_{i,r}}. \qquad (C4)$$

Here, $\bar{f}_{i,r}$ in the denominator is the average value of $f_{i,r}$. The radius error increases with decreasing S/N, so we calculate $\sigma_{f_{i,r}}$ for objects in *r*-band magnitude bins, and we obtained the fractional radius error σ_{R_r}/R_r as a function of magnitude. Similarly, we can assume that $\sigma_{R_g} = \sigma_{R_r}$, and calculate σ_{R_r}/R_r using $f_{g,r} = R_g/R_r$. We find that the fractional radius errors from *g* band and *i* band are consistent, and therefore we simply use the average of the two results as the final fractional radius error. Given the resulting estimate of the fractional radius error, we calculate MAGERR_APERCOR for each object via propagation of errors:

$$\text{MAGERR\_APERCOR} = \frac{\sigma_A}{A} = \frac{1}{A}\frac{dA}{dR_r} R_r \frac{\sigma_{R_r}}{R_r}, \qquad (C5)$$

where *A* is short for ApCorr. Similarly, in the *Y* band, we match the objects to CFHTLS, and estimate σ_{R_Y} and MAGERR_APERCOR from the scatter of $f_{z,Y} = R_z/R_Y$.

In cases where one wishes to estimate the uncertainty in the total magnitude of an objects, the net error in MAG_APERCOR is

$$\sigma_{\text{MAG\_APERCOR}} = \sqrt{(\text{MAGERR\_APER})^2 + (\text{MAGERR\_APERCOR})^2}. \qquad (C6)$$

Since the *r*-band radius is used for aperture correction for all of *ugriz*, the correction error MAGERR_APERCOR is correlated and mostly cancels out when we calculate colours involving the *ugriz* bands. For example, the error in $u - g$ colour is

$$\sigma_{u-g} = \sqrt{\text{UERR\_APER}^2 + \text{GERR\_APER}^2 + (\text{UERR\_APERCOR} - \text{GERR\_APERCOR})^2}. \qquad (C7)$$
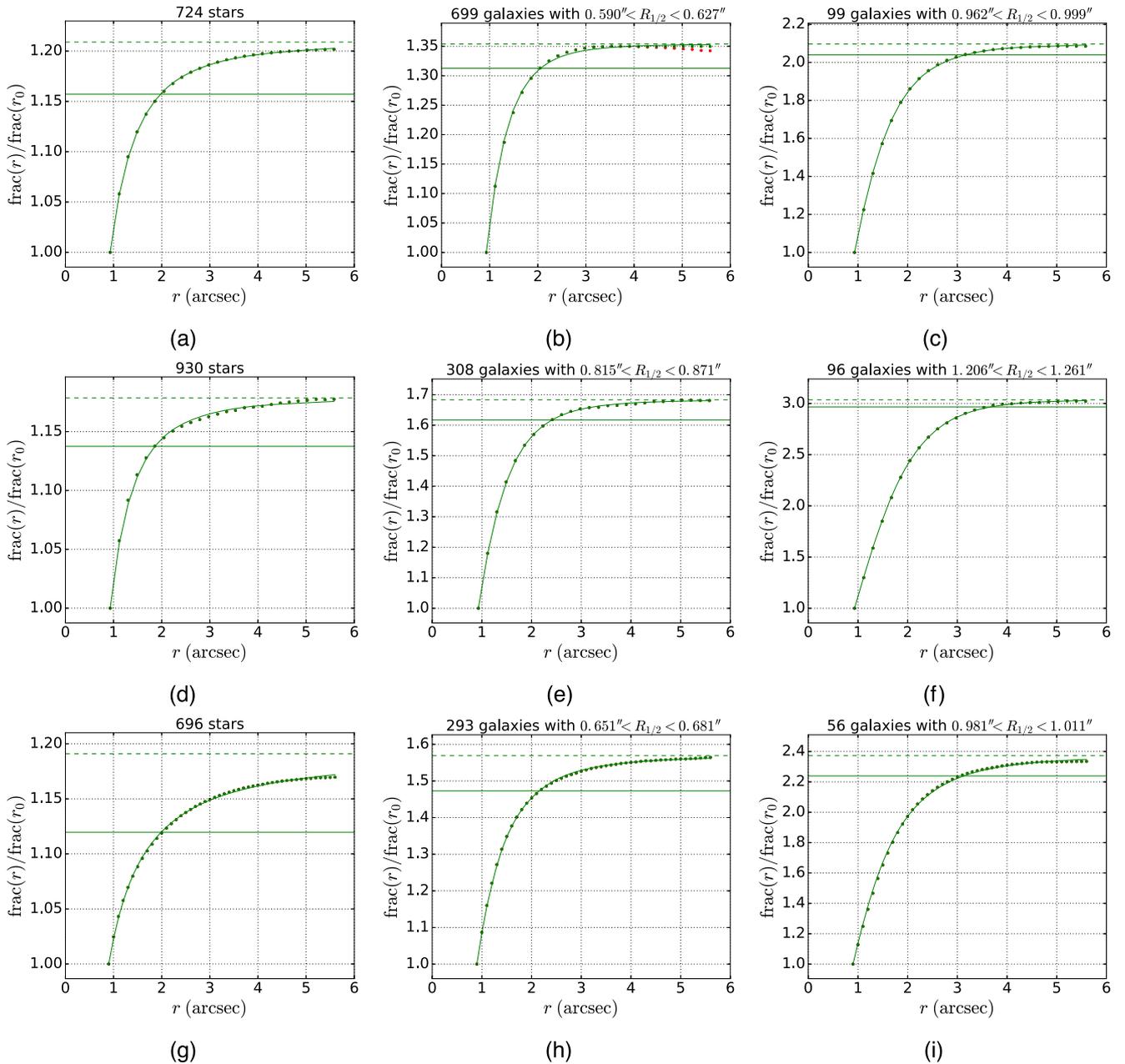
Figure C1. Examples of curve of growth and its Moffat fit. The Y-axis is the ratio of the flux in aperture radius r to the flux in the fixed aperture radius r_0. Points are the observed flux ratio for each radius bin. The solid curve is the Moffat fit. The solid horizontal line is the ratio of the flux in MAG_AUTO to the flux in the fixed aperture radius r_0, and the dashed horizontal line is the predicted flux ratio for an infinitely large aperture. Panels (a), (b), and (c) show CFHTLS D3 i band. Panels (d), (e), and (f) show CFHTLS W3-0-2 u band. Panels (g), (h), and (i) show the Subaru Y-band deep pointing. In panel (b), the flux ratio decreases at large apertures (red points) due to non-zero background, and it is corrected by extrapolating using the maximum flux ratio (green points). Such non-zero background might carry a different sign, showing as large increase of flux ratio at large apertures, although in this case it is hard to distinguish between flux from the source and the flux from the background, and no correction is applied. We tried to minimize the effects of imperfect background subtraction by selecting bright objects (with smaller photometric error) for the fit.

The Y-band aperture correction did not use r-band radius, and the error in $z - Y$ is

$$\sigma_{z-y} = \sqrt{\text{ZERR\_APER}^2 + \text{YERR\_APER}^2 + \text{ZERR\_APERCOR}^2 + \text{YERR\_APERCOR}^2}. \tag{C8}$$

Similar formulae may be used to determine the net uncertainty in any colour derived from these passbands.
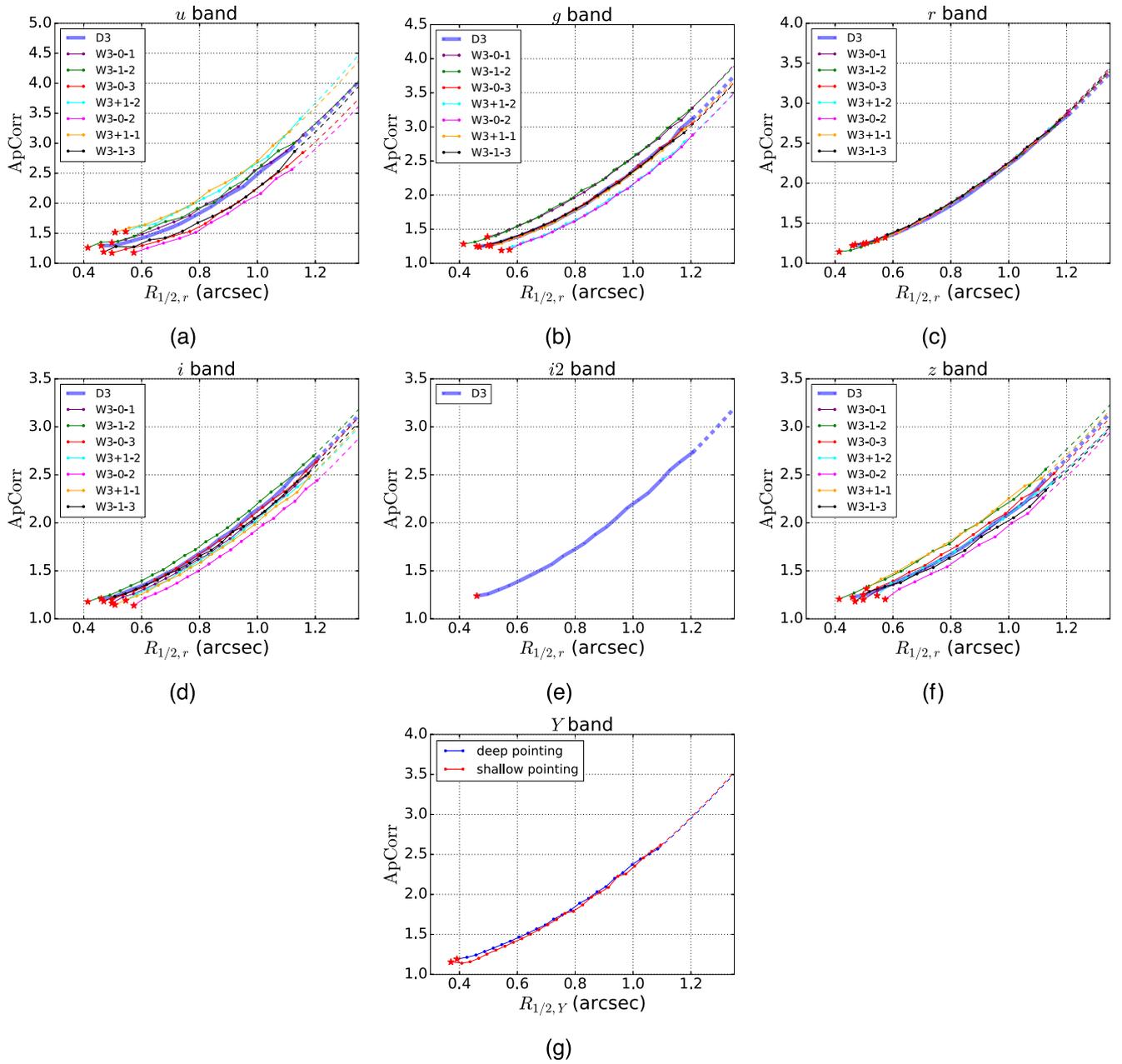
Figure C2. Each plot shows the correction factor ApCorr in one band as a function of half-light radius, overplotting all pointings . In panels (a–f), the thick line is the Deep field D3 and the thin lines are the Wide field W3 pointings. Panel (g) shows the *Y* band. The correction factor and radius of stars are plotted as the star marker. The dashed line is the extrapolation for objects larger than the radius bins.

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

[2]*Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), Pittsburgh, PA 15260, USA*

[3]*Center for Cosmology, Department of Physics and Astronomy, University of California, 4129 Reines Hall, Irvine, CA 92697, USA*

[4]*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*

[5]*Department of Physics and Astronomy, University of Leicester, University Road, Leicester LE2 3RJ, UK*

[6]*Centre for Astronomy and Particle Theory, University of Nottingham, Nottingham NG7 2RD, UK*

[7]*Department of Astronomy, University of California, Berkeley, CA 94720, USA*

[8]*Department of Physics, University of California, Berkeley, CA 94720, USA*

[9]*New York University Abu Dhabi, PO Box 129188, Abu Dhabi, United Arab Emirates*

[10]*UCO/Lick Observatory and Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA*

[11]*Astronomy Department, Orange Coast College, Costa Mesa, CA 92626, USA*

[12]*Department of Physics and Astronomy, Colby College, Waterville, ME 04961, USA*

[13]*Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstraße, D-85748 Garching, Germany*
[14]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK*
[15]*Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA*
[16]*Department of Physics, University of Connecticut, 2152 Hillside Road, Unit 3046, Storrs, CT 06269, USA*
[17]*MMT/Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA*
[18]*Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA*
[19]*Department of Physics and Astronomy, University of Kentucky, 505 Rose Street, Lexington, KY 40506-0055, USA*

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