The Astrophysical Journal Supplement Series, 184:218–229, 2009 October doi:10.1088/0067-0049/184/2/218

THE zCOSMOS 10k-BRIGHT SPECTROSCOPIC SAMPLE*

Simon J. Lilly[1], Vincent Le Brun[2], Christian Maier[1], Vincenzo Mainieri[3], Marco Mignoli[4], Marco Scodeggio[5], Gianni Zamorani[4], Marcella Carollo[1], Thierry Contini[6], Jean-Paul Kneib[2], Olivier Le Fèvre[2], Alvio Renzini[7], Sandro Bardelli[4], Micol Bolzonella[4], Angela Bongiorno[8], Karina Caputi[1], Graziano Coppa[4], Olga Cucciati[9], Sylvain de la Torre[2], Loic de Ravel[2], Paolo Franzetti[5], Bianca Garilli[5], Angela Iovino[9], Pawel Kampczyk[1], Katarina Kovac[1], Christian Knobel[1], Fabrice Lamareille[6], Jean-Francois Le Borgne[6], Roser Pello[6], Yingjie Peng[1], Enrique Pérez-Montero[6], Elena Ricciardelli[7], John D. Silverman[1], Masayuki Tanaka[3], Lidia Tasca[2], Laurence Tresse[2], Daniela Vergani[4], Elena Zucca[4], Olivier Ilbert[10], Mara Salvato[11], Pascal Oesch[1], Umi Abbas[2], Dario Bottini[5], Peter Capak[11], Alberto Cappi[4], Paolo Cassata[2], Andrea Cimatti[11], Martin Elvis[13], Marco Fumana[5], Luigi Guzzo[9], Gunther Hasinger[8], Anton Koekemoer[14], Alexei Leauthaud[15], Dario Maccagni[5], Christian Marinoni[16], Henry McCracken[16], Pierdomenico Memeo[5], Baptiste Meneux[8], Cristiano Porciani[1], Lucia Pozzetti[4], David Sanders[10], Roberto Scaramella[18], Claudia Scarlata[11], Nick Scoville[11], Patrick Shopbell[11], and Yoshiaki Taniguchi[19]

[1] Institute for Astronomy, ETH Zurich, Wolfgang-Pauli-strasse 27, 8093 Zurich, Switzerland
[2] Laboratoire d'Astrophysique de Marseille, 38, rue Frédéric Joliot-Curie, 13388 Marseille Cedex 13, France
[3] European Southern Observatory, Karl-Schwarzschild-Strasse 2, 85748 Garching, Germany
[4] INAF Osservatorio Astronomico di Bologna, Via Ranzani 1, 40127 Bologna, Italy
[5] INAF – IASF Milano, Via E. Bassini 15, 20133 Milano, Italy
[6] Laboratoire d'Astrophysique de Toulouse/Tarbes, Université de Toulouse, 14 avenue Edouard Belin, 31400 Toulouse, France
[7] Dipartimento di Astronomia, Universita di Padova, Vicolo dell'Osservatorio 5, 35122 Padova, Italy
[8] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße, 85748 Garching, Germany
[9] INAF Osservatorio Astronomico di Brera, via Brera 28, 20121 Milano, Italy
[10] University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822-1839, USA
[11] Department of Astronomy, Caltech, MC 105-24, 1200 East California Blvd., Pasadena, CA 91125, USA
[12] Dipartimento di Astronomia, Università degli Studi di Bologna, Via Ranzani 1, 40127 Bologna, Italy
[13] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[14] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[15] Berkeley Lab & Berkeley Center for Cosmological Physics, University of California, Lawrence Berkeley National Lab, 1 Cyclotron Road, MS 50-5005, Berkeley, CA 94720, USA
[16] Centre de Physique Theorique, Case 907 – 13288 Marseille cedex 9, France
[17] Institut d'Astrophysique de Paris, UMR7095 CNRS, Université Pierre & Marie Curie, 75014 Paris, France
[18] INAF – Osservatorio Astronomico di Roma, Osservatorio Astronomico di Roma, Via di Frascati 33, 00040 Monte Porzio Catone, Italy
[19] Ehime University, 2-5 Bunkyo-cho, Matsuyama 790-8577 Japan
Received 2009 March 19; accepted 2009 August 21; published 2009 September 15

ABSTRACT

We present spectroscopic redshifts of a large sample of galaxies with $I_{AB} < 22.5$ in the COSMOS field, measured from spectra of 10,644 objects that have been obtained in the first two years of observations in the zCOSMOS-bright redshift survey. These include a statistically complete subset of 10,109 objects. The average accuracy of individual redshifts is 110 km s^{-1}, independent of redshift. The reliability of individual redshifts is described by a Confidence Class that has been empirically calibrated through repeat spectroscopic observations of over 600 galaxies. There is very good agreement between spectroscopic and photometric redshifts for the most secure Confidence Classes. For the less secure Confidence Classes, there is a good correspondence between the fraction of objects with a consistent photometric redshift and the spectroscopic repeatability, suggesting that the photometric redshifts can be used to indicate which of the less secure spectroscopic redshifts are likely right and which are probably wrong, and to give an indication of the nature of objects for which we failed to determine a redshift. Using this approach, we can construct a spectroscopic sample that is 99% reliable and which is 88% complete in the sample as a whole, and 95% complete in the redshift range $0.5 < z < 0.8$. The luminosity and mass completeness levels of the zCOSMOS-bright sample of galaxies is also discussed.

Key words: cosmology: observations – galaxies: active – galaxies: distances and redshifts – galaxies: evolution – large-scale structure of universe – quasars: general – surveys

Online-only material: color figure, machine-readable table

1. INTRODUCTION

The zCOSMOS project (Lilly et al. 2007, hereafter ZC-07) is a major redshift survey of galaxies in the COSMOS field using 600 hr of clear dark observing time on the VLT. The survey consists of two parts. The first, zCOSMOS-bright, will ultimately consist of spectra of about 20,000 galaxies selected to have $I_{AB} < 22.5$ across the full 1.7 deg^2 of the COSMOS field (Scoville et al. 2007). It was designed to yield a high and fairly

* Based on observations undertaken at the European Southern Observatory (ESO) Very Large Telescope (VLT) under Large Program 175.A-0839. Also based on observations with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute, operated by AURA Inc., under NASA contract NAS 5-26555, with the Subaru Telescope, operated by the National Astronomical Observatory of Japan, with the telescopes of the National Optical Astronomy Observatory, operated by the Association of Universities for Research in Astronomy, Inc. (AURA) under cooperative agreement with the National Science Foundation, and with the Canada–France–Hawaii Telescope, operated by the National Research Council of Canada, the Centre National de la Recherche Scientifique de France, and the University of Hawaii.

uniform sampling rate across most of the field (about 70%), with a high success rate in measuring redshifts (approaching 100% at $0.5 < z < 0.8$) and to have sufficient velocity accuracy (about 100 km s^{-1}) to efficiently detect cosmic structures down to the scale of galaxy groups. The second part, zCOSMOS-deep, will consist of about 10,000 spectra of $B_{AB} < 25.25$ galaxies, color-selected to have redshifts in the $1.4 < z < 3.0$ redshift range and lying in the central 1 deg^2 region of the COSMOS field.

After the first two observing seasons (2005 Spring and 2006 Spring), 83 of the 180 spectroscopic masks for zCOSMOS-bright have been observed, yielding a total of 10,644 spectra from which redshift measurements have been made or attempted. This so-called 10k sample has been used to carry out a number of science investigations, including a reconstruction of the galaxy density field to $z \sim 1$ (Kovac et al. 2009a), the production of a first group catalog (Knobel et al. 2009), several studies of galaxy properties as a function of environment and redshift (Maier et al. 2009; Bolzonella et al. 2009; Pozetti et al. 2009; Cucciati et al. 2009; Iovino et al. 2009; Kovac et al. 2009b; Tasca et al. 2009; Vergani et al. 2009; Zucca et al. 2009), studies of the correlation function (Gilli et al. 2009; Meneux et al. 2009; de la Torre et al. 2009; C. Porciani et al. 2009, in preparation), and studies of active galactic nucleus (AGN; Brusa et al. 2009; Silverman et al. 2009a, 2009b), far-IR (Caputi et al. 2008, 2009), and radio sources (Bardelli et al. 2009). Studies of the merging rate in the sample will be presented by L. de Ravel et al. (2009, in preparation) and P. Kampcyzk et al. (2009, in preparation). The redshifts have been used to calibrate photometric redshifts in Mandelbaum et al. (2008).

zCOSMOS, as with COSMOS generally, is undertaken in the spirit of a Legacy Program with prompt public release of data products. The present paper therefore describes and presents the zCOSMOS 10k sample (version 3.51). The electronic data, including one-dimensional spectra, are also available through the ESO Public Archive and IRSA. In the future, a final "20k sample" will be released following the completion of the program in 2009.

The detailed design of both parts of zCOSMOS was described in some detail in ZC-07 and will not be repeated here. Rather, the intention here is to provide updated information on this particular "10k sample" that may be of interest to potential users of the publicly released data or to those interested in the details of the spectroscopic sample that has been used in the current set of zCOSMOS science papers outlined above.

2. zCOSMOS 10k SAMPLE

2.1. Input Catalog and Target Selection

The selection criteria and input target catalog for the zCOSMOS-bright 10k sample remain unchanged from those described in ZC-07. To summarize, the input target catalog for zCOSMOS-bright is based primarily on "total" F814W magnitudes derived from the 0.1 arcsec resolution *Hubble Space Telescope* (*HST*) images (Koekemoer et al. 2007). For this purpose, a SExtractor "magauto" magnitude was used for consistency with other COSMOS photometric catalogs (e.g., Capak et al. 2007). These are supplemented, for the small regions where the *HST* data are absent or compromised by diffraction spikes, etc., by photometry from a high-resolution *i*-band CFHT image. Of the entire region targeted for zCOSMOS-bright spectroscopy, about 4% of objects were added from the Canada–France–Hawaii Telescope (CFHT) images. However, the vast majority of these are at the extreme edges of the field (where there are no ACS images), where the spectroscopic sampling rate is in any case lower. Away from these edges, the fraction of added objects is very small, being less than 0.5%. A comparison of the magnitudes for the huge number of objects in both catalogs indicated that there is a negligible photometric offset between the CFHT and *HST* photometry. It should be noted that, as a result of the arbitrary *HST* roll angle, the diffraction spikes from the two sets of data are at different position angles. Because it is based on high-resolution *HST* data, without even the effects of *HST* diffraction spikes (because of the CFHT catalog), a negligibly small area is obscured by foreground stars and no attempt at "masking" was made. The selection magnitude range is $15.0 < I_{AB} < 22.5$.

As noted in our earlier paper, the first 12 masks (7% of the program total) were generated from an earlier target catalog that had the same selection criteria but was based on the first season of *HST* images. These covered only part of the field and their processing did not include subsequent improvements such as correction for the effects of imperfect charge transfer. As would be expected, the new catalog does not exactly match the old one at the selection boundary $I_{AB} = 22.5$, because of unavoidable small variations in the photometry. As a result, 172 of the objects observed in these first 12 masks do not, in fact, appear in the newer 2006 target catalog. To deal with this small complication, all 2571 objects in the 2005 target catalog that would have disappeared in the 2006 version were transferred across into the new catalog, an augmentation of 4.5%. These objects have been assigned a distinct target ID number beginning with a "7" instead of the usual "8." Because all but the first 12 masks were designed with the newer 2006 catalog, these objects comprise only 1.6% of the present 10k sample, and they will be less than 1% of the final 20k sample when it is completed. For many purposes, this complication may be safely ignored.

Suspected Galactic stars (representing 19% of objects) were not targeted for spectroscopic observation. Their exclusion was based on a combination of their morphologies and spectral energy distributions, using criteria that were intentionally quite conservative. As a result, about 4% of zCOSMOS slits were targeted at objects that have been subsequently spectroscopically identified as stars. Conversely, 160 of the objects that were initially "excluded" as stars actually have been observed, as serendipitous spectra either in the slits of other targets (see below) or in the first 12 masks (see above). Of these 160 objects, only three (1.8%) were subsequently securely identified spectroscopically as galaxies, 129 as stars, and the remainder were unidentified. The latter are probably stars—the spectra of these secondary targets are often noisy—and the fraction of actual galaxies excluded from the spectroscopic sample because they are masquerading as stars is therefore estimated to be on the order of 0.4%. These may represent a systematic class of compact galaxies (see Drinkwater et al. 2003).

A few hundred X-ray and ultraviolet (GALEX) identifications were included in the mask designs as "compulsory" targets. In the mask-making algorithm, slits are first placed over as many as possible of these compulsory targets. Then, in a second pass, slits are placed over as many of the "random" targets as possible. As a rule of thumb, the addition of each compulsory slit costs reduces by two the number of slits that can be placed over random targets. Because of this preferential treatment, objects designated as compulsory have a greater chance of being included in the mask designs, especially for the first masks designed in each patch of sky. Many, but not all, of these compulsory objects also satisfy

the *I*-band selection criteria for the main statistically complete spectroscopic sample, i.e., $15.00 < I_{AB} < 22.5$. This means that for any analysis of the statistically complete sample, these "compulsory" targets should be included with an appropriate deweighting to correct for their higher chance of having been observed spectroscopically.

In the current 10k sample, the sampling rate of these compulsory targets is estimated (from the input catalog within the zCOSMOS survey area of 1.77 deg^2) to be 2.05 times higher than that of the other "random" targets. Although this deweighting factor is approximate, since it depends on the detailed spatial distribution of both compulsory and random targets, it should be accurate enough for most purposes. Since it is fortuitously close to two, an alternative approach is simply to remove every other object in this category. It should be noted that the compulsory targets also have a subtly different spatial distribution within the masks because they tend to be more uniformly distributed in the masks on account of their higher priority. The compulsory targets are identified in the spectroscopic catalog by a flag.

Finally, a small number of radio identifications were included in the masks as "random" targets even though they were fainter than $I_{AB} = 22.5$. These have the same sampling rate as the main sample, but are not part of the statistically complete $15 < I_{AB} < 22.5$ sample.

2.2. Observations

Observations for the 10k sample were carried out with the VIMOS spectrograph (Le Fèvre et al. 2005) mounted on the 8 m VLT/UT3 telescope during extensive Service Mode runs in 2005 and 2006. Slit masks were prepared for the four quadrants of VIMOS using the VMMPS software (Bottini et al. 2005). Observations used the MR grism with 1.0 arcsec slits, yielding a spectral resolution $R \sim 600$ at 2.5 Å pixel^{-1} and a spectral range 5550–9450 Å. All masks are observed with the slits oriented N-S. The pattern of pointings is such that, when the survey is completed, every target in a large rectangular region will have had eight opportunities to be selected for observation (see ZC07).

Data reduction for each mask was carried out independently in two institutes using the VIPGI software (Scodeggio et al. 2005). Redshift measurements were carried out using the EZ software (R. Scaramella et al. 2009, in preparation). Although an initial automated cross-correlation redshift is automatically produced, the spectra are always checked visually. We find with these spectra that the cross-correlation finds the correct redshift for 60%–70% of spectra. These are generally also the easiest spectra to recognize by eye, and it is unlikely that the process of redshift determination could be fully automated with an acceptable degree of reliability. The final redshifts are determined purely from the spectra, independently of a photometric redshift or other information about the objects.

As noted above, all one-dimensional spectra are available for download from the ESO Science Archive and IRSA.

2.2.1. Primary and Secondary Targets

While each slit is placed so as to observe one object, the so-called primary target, it is sometimes possible to obtain a spectrum for another object in the catalog in the same slit, a so-called secondary target. It should be noted that these secondary targets may have been primary targets in some other mask, or they may have been excluded from the target list as stars (see above). At present, a total of 380 objects in the 10,644 object 10k sample (3.6% of the total) have yielded an identifiable spectrum only as a secondary target. A further 135 spectra were extracted but no redshift was discernible.

The secondary spectra have a higher failure rate in redshift determination, presumably because of higher slit losses due to mis-centering of the slit on the target and less optimal sky-subtraction. Despite this, there is no evidence, however, that the precision of the redshifts of secondary targets is affected (see below). The 380 targets with only a secondary spectrum will be biased toward brighter, or otherwise "easier," targets, and it is recommended that they be removed from the sample whenever statistical completeness is required. Likewise, the secondary spectra where we failed to measure a redshift should also be discarded since they were not a "fair test."

All of the objects for which only a secondary spectrum has been obtained are nonetheless included in the published catalog. They are identified in the catalog by preceding the redshift Confidence Class (described in the next Section) by a "2."

2.2.2. Confidence Classes for the Redshifts

Inevitably, in a redshift survey of this type, there is a degree of subjectivity in the redshift measurement for some objects and range of reliability in the redshifts. As described in ZC-07, the reduction of the data through to the crucial redshift assignment is carried out independently at two different institutions, using the same software packages but with sufficient human intervention that the reductions are largely independent. This procedure already reduces the level of subjectivity. The reliability is quantified by means of a confidence parameter that is assigned to each redshift.

All redshifts are first assigned an integer Confidence Class that captures the subjectively estimated reliability of the spectroscopic redshift measurement. Relative to the description in ZC-07, there have been a couple of small modifications to the description of the integer Confidence Classes. The basic scheme used earlier ran from 0 (no redshift) to a maximum of 4 (most secure, and with a textbook spectrum) with an additional Class 9 for one-line redshifts where the line is expected to be either [O II] 3727 or Hα, i.e., where an identification with Hβ or the [O III] 4959, 5007 doublet can be rejected. A modifier of +10 is given for a broad line object. Preceding the class by a further "2" (i.e., either by adding +20 or +200 for a broad line object) indicates that the spectrum was of a secondary object in the slit, as described in Section 2.2.1. An additional Class 18 is also added for broad line objects for which only a single line is measured. We do not use 19 for these since the choice of redshifts is less clear-cut for a single broad line than it is for a single narrow emission line. The Confidence Classes are summarized in Table 1.

As before, the independent redshift measurements and integer Confidence Classes are then "reconciled" at a face-to-face meeting between representatives of the two institutes. Discrepant redshifts and significantly discrepant confidence classes are all inspected interactively and resolved to mutual agreement.

2.3. Spectroscopic Verification Rate and Velocity Accuracy

In the 10k sample, repeat observations now exist for 632 objects, more than 5% of the total. These repeats occur for three reasons: (1) a few (random) masks were observed more than once; (2) some primary targets reappear as secondary targets in different masks; and (3) a randomly chosen subset of objects observed from the 2005 catalog were reinserted into the pool of selectable targets for subsequent observing runs based on the new catalog. These repeat observations can be used to establish

Table 1
Definition of Confidence Classes

Class	Integer Confidence Classes
4	A very secure redshift with an exhibition-quality spectrum
3	A very secure redshift
2	A likely redshift about which there is some doubt
9	A securely detected line which is believed to be either [O II] 3727 or Hα 6563
1	An insecure redshift
0	No redshift measurement attempted
+10	As above but for broad line AGN, with 18 instead of 9, reflecting the greater range of possibilities
+20 or +200	As above but for a target only observed as a secondary target in a slit centered on another object
	Decimal place modifiers
.5	The spectroscopic and photometric redshifts are consistent to within $0.08(1+z)$
.4	No photometric redshift is available for some reason
.3	For Class 9 and 18 one-line redshifts only, the spectroscopic redshift is consistent only after the spectroscopic redshift is changed to the alternate redshift
.1	The spectroscopic and photometric redshifts differ by more than $0.08(1+z)$

Note. The set of objects defined as Classes 1.5, 2.4, 2.5, 9.3, 9.5 and all Class 3.x and 4.x comprise 88% of the sample (95% within $0.5 < z < 0.8$) and are estimated to be 99% reliable.

Table 2
Empirical Reliability of Integer Confidence Classes

Class	Fraction of Primary Sample	Spectroscopic Verification Rate p_i	Photo-z Consistency within $\Delta z = 0.08(1+z)$		
			ZEBRA v3.4	Ilbert et al. (2009) v3.5	
			All	All objects	Not photometrically masked[b]
Classes 3 and 4	61%	99.8%	95%	96%	98.5%
Class 9	6%	86–96%[a]	94%	94%	95%
Class 2	15%	92%	93%	93%	94%
Class 1	10%	70%	72%	72%	72%
Class 0	8%	...	...	...	...

Notes.
[a] One-line redshifts, before and after adjustment of the spectroscopic redshift to the alternate redshift on the basis of the photo-z.
[b] Photometrically unmasked regions comprise 85% of objects.

empirically the redshift reliability of the different Confidence Classes and also determine the typical velocity accuracy.

Looking first at the velocity accuracy, the 569 objects for which the redshift was in broad agreement, i.e., to within $\Delta z = 0.0066\ (1+z)$, or 2000 km s^{-1}, yield an overall velocity rms, per observation, of $\sigma_v = 110$ km s^{-1} or $\sigma_z = 0.00036\ (1+z)$. The same value is obtained if the secondary spectra are eliminated, or if repeat observations through the same mask are removed from the analysis. There is also no evidence for a significant variation in velocity error with redshift and 110 km s^{-1} may be taken as a good global estimate of the velocity uncertainty in zCOSMOS-bright.

We also use the repeat observations to calibrate the repeatability of the Confidence Classes. A more sophisticated algorithm is used to evaluate the statistical reliability of the spectroscopic redshifts than that was used for ZC-07. We assign to each Confidence Class, i, a probability p_i (to be determined) that the redshifts are correct. We then assume that the chance of getting the same redshift (within 2000 km s^{-1}) when both redshifts are wrong should be negligible. This is unlikely to be completely valid—the same mistake could be made twice—but it is probably a reasonable approximation (see below). This means that, for any pair of independent redshift measurements with Confidence Classes i and j, respectively, the probability that they agree should be just the product $p_i p_j$. By examining all pairs of measurements of common objects, we can construct two triangular matrices N and F, for which the elements N_{ij} (with $i \geqslant j$) contain the number of "trials" where one Confidence Class was i and the other j, and F_{ij} gives the fraction of these trials where the two redshifts were actually in agreement. For a given set of p_i, the probabilities of observing F_{ij} with N_{ij} trials, P_{ij}, is evaluated using standard binomial statistics and the preferred set of p_i selected through a maximum likelihood approach by maximizing the product $\Pi_{i,j \geqslant i}\ P_{ij}$. The resulting p_i are shown in Table 2.

As noted above, the single emission lines in the Class 9 objects could be either [O II]3727 or Hα 6563, since objects are only assigned this class if an identification with Hβ or the [O III] 4959, 5007 doublet is rejected. Unfortunately, with our $R \sim 600$ spectra, the splitting of the [O II] 3726, 3729 doublet cannot be used as an identifier. An initial guess for the Class 9 redshifts is made based purely on the spectrum, but we then modify the redshift to the alternative line identification if the photometric redshift (see below) is found to be inconsistent with the initial redshift but consistent with the alternative. This accounts for the two reliabilities given in Table 1 for this class. We find that the initial spectroscopic guess is in fact correct in 86% of cases and that the alternative redshift is consistent in two-thirds of the remainder. The remaining 4% photometric redshift inconsistency is approaching that of even the most secure redshifts and may well be due to photometric redshift problems (see below).

Compared with Table 2 of ZC-07, it can be seen that there are slightly more low-confidence redshifts in the new sample, but that the reliability of these lower Confidence Classes is higher. For example, even the Class 2 redshifts (which were aimed to be 75% reliable) are apparently correct more than nine times out of ten, and the 75% level is almost achieved by our least reliable Class 1. The differences from ZC-07 may reflect a gradual drift in confidence through the duration of the program or simply better statistics in the new larger sample.

2.4. Consistency with Photo-z Estimates

A very extensive and impressive set of deep photometric data now exists over the COSMOS field, permitting the estimation of photometric redshifts with very high accuracy for both galaxies and AGNs (see, e.g., Ilbert et al. 2009; Salvato et al. 2009). The photo-z and spectroscopic redshifts are highly complementary in both directions.

Once validated with our most secure redshifts, the photo-z provide a further calibration of the reliability of our less secure Confidence Classes. They also allow resolution of the ambiguous Class 9 and 18 "single-line" redshifts. The photometric redshifts can also be used to indicate individually which of the less secure spectroscopic redshifts (within the less secure classes) are likely to be correct and which are likely to be incorrect. This information is encapsulated in a decimal place that is added to the Confidence Class of each redshift (see below). Finally, the photo-z give information on the (approximate) redshifts and rest-frame colors of the roughly 12% of galaxies for which an incorrect spectroscopic redshift, or no redshift at all, was obtained, i.e., on the biases that will be present within the spectroscopic sample.

The quality of the photo-z in COSMOS has continuously improved. In this paper, we consider two sets of photo-z. The first was obtained by running the ZEBRA code (Feldmann et al. 2006) on the 11-band photometry (e.g., Capak et al. 2007; Sanders et al. 2007) that was available to us in early 2007. These photo-z were used in the construction of v3.4 of the zCOSMOS bright sample, which has been used for our current science analyses referenced above. More recently, the photo-z produced by Ilbert et al. (2009) and Salvato et al. (2009) using 30-band photometry have become available. These show a gain in individual precision, from $\sigma_z/(1+z) \sim 0.02$ to $\sigma_z/(1+z) \sim 0.01$, and also allow an improved

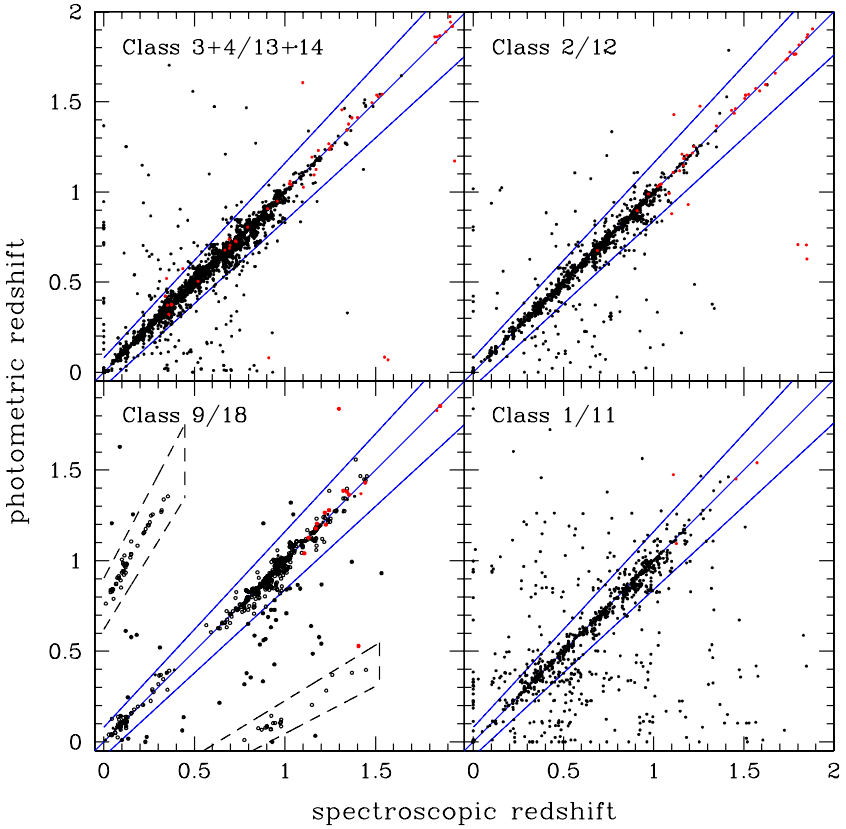

Figure 1. Spectroscopic vs. photometric redshifts for the 10k sample, split by the integer Confidence Class of the spectroscopic redshift (based on the spectrum alone). The blue lines indicate a region of agreement within $\Delta z = 0.08(1+z)$. The dashed lines for Class 9 indicate the "alternative" redshifts obtained by misidentifying Hα and [O II] 3727. Broad line objects, which have a Confidence Class increased by 10 (with $9 \rightarrow 18$), are indicated by red symbols.

(A color version of this figure is available in the online journal.)

treatment of stars and AGNs. However, the fraction of outliers, or "catastrophic failures," remains comparable, being mainly driven by photometric problems, such as galaxies in masked areas or those with close companions on the sky.

Figure 1 shows the comparison of spectroscopic and photometric redshifts for objects in different spectroscopic Confidence Classes. The fraction of objects for which the redshift difference is less than $\Delta z/(1+z) = 0.08$ is shown in Table 2, both for these photo-z and those based on the earlier ZEBRA photo-z.

About 4% of the objects in our most secure Confidence Classes 3 and 4 have discrepant photo-z. This fraction can be reduced to 1.5% by avoiding the 15% or so of objects that lie in the photometrically "masked" areas of Ilbert et al. (2009). These areas have incomplete or otherwise less reliable photometry. The fraction with discrepant photo-z can be reduced still further by eliminating objects with bimodal redshift likelihood functions (see Ilbert et al. 2009 for a discussion). All the cases with discrepant spectroscopic and photometric redshifts in our secure Classes 3 and 4 have been individually examined. This confirms that in, all cases, it is the photometry that is "at fault." Outside of the masked areas, this is often due to close neighboring objects. These discrepancies clearly represent a ceiling in photo-z performance that can be raised toward 100% by sufficiently aggressive cleaning of the photometric catalog.

If we now look at the galaxies in our less secure spectroscopic Confidence Classes 1, 2, and 9, we find that there is very good agreement between the fraction of objects with consistent spectroscopic and photometric redshifts and the "repeatability rate" determined internally from the spectroscopy alone, as discussed above. This agreement is shown in Table 2.

Together with the excellent performance of the photo-z on our most secure spectroscopic redshifts, this good agreement between "spectroscopic repeatability" and "photometric consistency" suggests that the photo-z can be used to identify, object by object, which of the less reliable spectroscopic redshifts are likely to be correct and which are probably wrong. We have therefore introduced an additional decimal place modifier to the Confidence Class to indicate whether the spectroscopic redshift is consistent with the photo-z to within $\Delta z = 0.08(1+z)$. A decimal place of 0.5 indicates consistency and one of 0.1 represents inconsistency. A decimal place of 0.4 represents an inadequate test due to lack of a photo-z. The special case of the 10% of Class 9 single line redshifts that become consistent only after the redshift is changed to the "alternate" value is indicated by a 0.3 decimal. To illustrate the meaning of this scheme, a Class 3.5 redshift is virtually certain to be right, as are the rare Class 3.1 redshifts, since in this case the photo-z is almost certainly wrong and the discrepancy was anyway examined individually. On the other hand, while a Class 2.5 redshift is very likely to be correct, a Class 2.1 redshift is likely to be wrong, since the photometric redshifts are normally reliable (>96%). A Class 2.4 redshift, in the absence of photo-z information, has the default Class 2 chance of being correct of 92%.

For many purposes, we can define a usable (i.e., reasonably secure) galaxy redshift to be all the Class 3.x, 4.x redshifts, plus the Classes 1.5, 2.4, 2.5, 9.3, and 9.5. Running the set of these through the spectroscopic verification program described

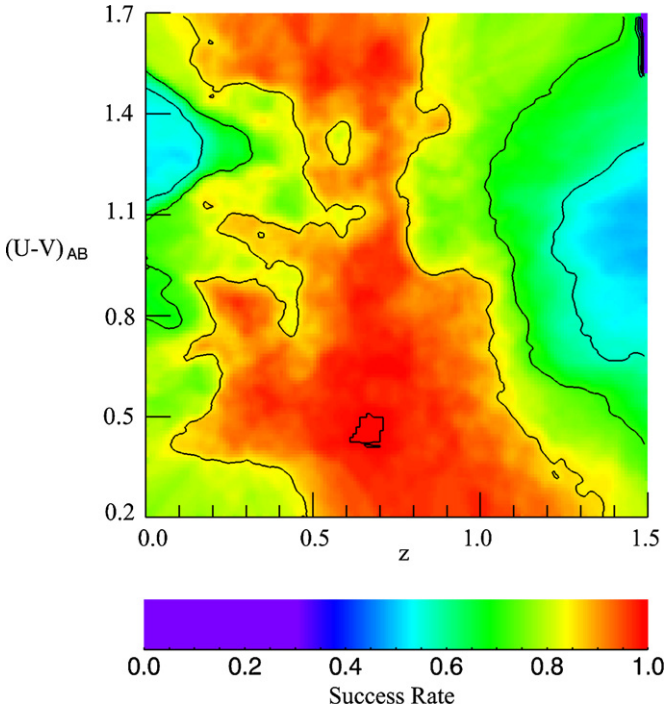

Figure 2. Fraction of spectra yielding a successful redshift measurement as a function of redshift and rest-frame color (derived from the photo-z). The U in this analysis is a relative red u passband used at CFHT.

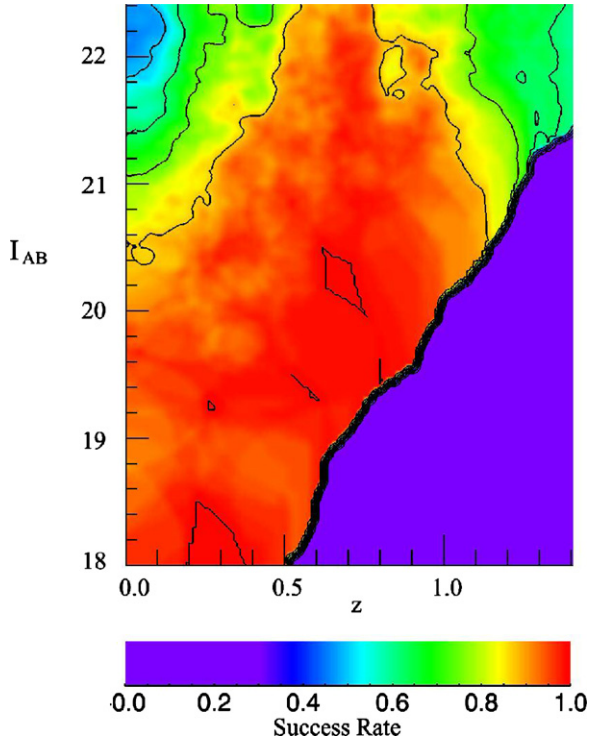

Figure 3. Fraction of spectra yielding a successful redshift measurement as a function of redshift (derived from the photo-z) and I_{AB}.

earlier indicates that the redshifts in this set of objects will be individually 99% reliable. These objects comprise 88% of the overall sample. In the main $0.5 < z < 0.8$ redshift interval, this completeness rises to 95% (see Section 2.4).

Defining a "failed" redshift to be all Class 0, plus Classes 1.1, 1.4, 2.1, and 9.1, the failures amount to 12% of the 10k sample (increasing to 13% if the spectroscopically confirmed stars are removed). In future, a reexamination of these spectra in the light of the photo-z may enable us to recover accurate redshifts for these objects, but this has not yet been done. It is important to stress that the inconsistency with the photo-z makes it very likely that these redshifts are incorrect. A maximum of 4% of galaxies with secure spectroscopic redshifts have inconsistent photo-z, so the chance of an insecure spectroscopic redshift with a discrepant photo-z being correct is correspondingly small. At face value, we would expect most (formally 50%–90%) of these inconsistent insecure redshifts to be incorrect. However, they are retained in the published catalog for completeness, *but we would strongly recommend that they not be used for scientific analysis.*

While most of the scientific analyses of the 10k sample referenced above used the v3.4 of the catalog, the newest photo-z of Ilbert et al. (2009) and Salvato et al. (2009) were used for the released version v3.5 of the 10k sample. The photo-z, of course, enter only in the construction of the decimal place reliability modifiers and in the resolution of the single line Class 9 and 18 ambiguous redshifts, and in practical term, the main differences between v3.4 and v3.5 is in the decimal place reliability of the star and AGN redshifts, and in occasional changes between the decimal places as objects moved in or out of the photo-z consistency criteria. The only objects to actually change their spectroscopic redshifts were a handful of Class 9 or 18 objects that changed their redshifts because of the role of the photo-z in resolving the single-line redshift.

2.5. Overall Redshift Success Rate as Functions of Redshift, Color, and Magnitude

Based on the apparent success of the photo-z in reproducing the redshift of those galaxies with secure redshifts, the photo-z of those for which we failed (as defined in Section 2.4) to secure a redshift can be used to examine the biases in the spectroscopic sample. Figures 2 and 3 show the fractional success rate of securing a redshift as functions of redshift, and rest-frame color and I_{AB} magnitude, respectively. The broad features of these diagrams can be easily understood in terms of the observability of features important for redshift determination within the spectral window 5550–9650 Å, specifically the exclusion of the emission and absorption features between Hβ 4861 and 5175 at very low redshifts and the migration of [O II] 3727 and the 4000 Å break region into progressively worse regions of OH emission at $z > 1.2$. As noted above, in the important redshift range $0.5 < z < 0.8$, the success rate is 95%.

2.6. Spatial Sampling and Correction Methods Thereof

The final "20k sample" will have a relatively uniform spatial sampling with observations with 180 separate masks on a grid of 90 field centers, giving most objects eight opportunities to be included in the mask design. However, the current set of 83 VIMOS masks provides a highly nonuniform coverage. Figure 4 shows the overall distribution of the 10k sample on the sky, and Figure 5 shows a map of the spatial sampling rate obtained by a median filtering of the calculated sampling rate on 2 arcmin scales. The complex pattern arises because slightly less than a half of the final number of the masks have been observed so far and because of the nonuniform pattern of slits within each mask.

The latter effect is shown in Figure 6, which shows the location of all the slits contributing to the 10k sample in the four VIMOS quadrants. This distribution is highly nonuniform:

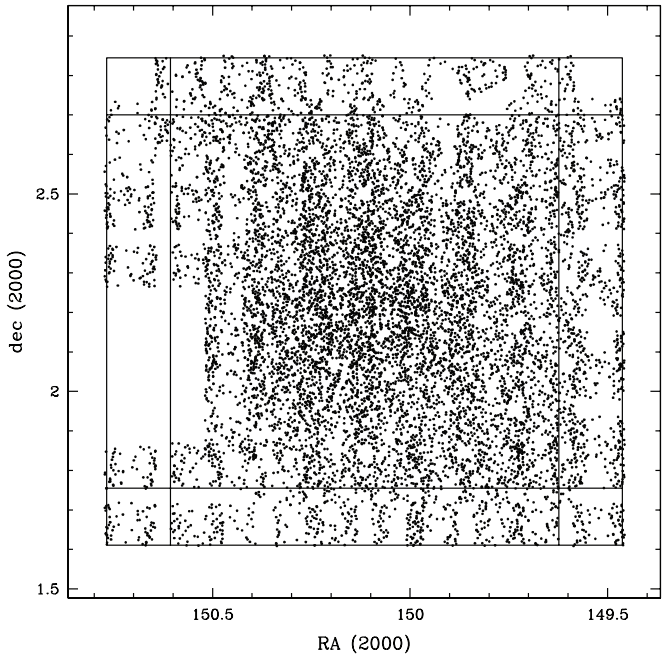

Figure 4. Spatial distribution of spectroscopic targets in the 10k sample.

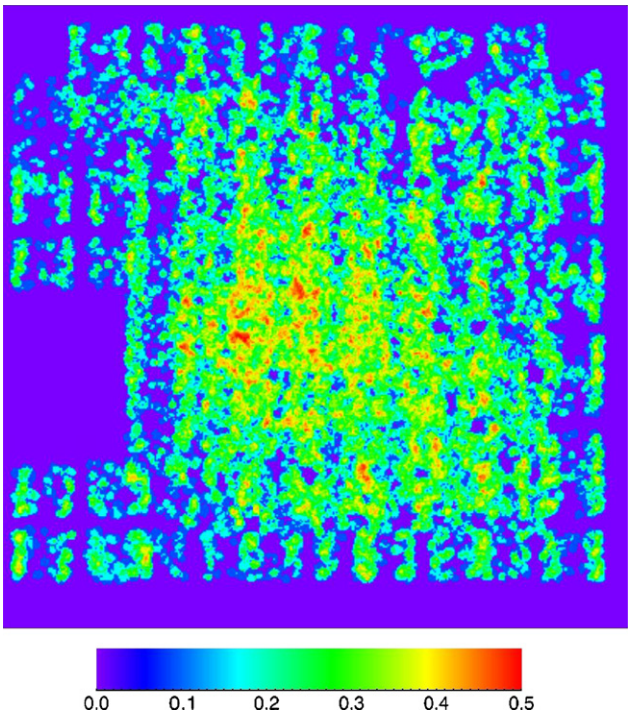

Figure 5. Map of the spatial sampling rate obtained by dividing the number of objects in the 10k sample (see Figure 4) by the number in the input target catalog. The map is adaptively smoothed. Axes are the same as in Figure 4.

because individual spectra have a length of about 1550 pixels, equivalent to 2/3 of the spatial field of view of the spectrograph, the automated software placing the slits tends to populate two strips along the top and bottom of each mask, since in this way two objects at the same declination can be observed (note that the full spectral range for both objects is still obtained because of the larger dimension of the detector in the dispersion direction). Different VIMOS pointings are offset by an amount equal to the quadrant dimensions, so the sparsely sampled regions in one mask broadly coincide with the peaks in the other masks. However, the final sampling will not be completely uniform because of this pattern. A further complication is that very occasionally one of the VIMOS quadrants fails (e.g., the grism fails to insert correctly) and the observation of this particular pointing is not generally repeated just to replace one of the four quadrants.

There is also a significant variation in the spectroscopic success rate from quadrant to quadrant and even within quadrants. This is shown in Figure 7. Quadrant 1 is so far the best with a 90% success rate, while Quadrant 2 is the worst with only 84% of spectra yielding a secure redshift.

The factors affecting the spatial sampling of the 10k sample are sufficiently complex that it is almost impossible to model them in their entirety. The same is likely to be true of the final 20k sample, and indeed in most redshift surveys using multi-slit spectrographs. For many purposes, it may be adequate to simply use an empirical sampling map, as constructed in Figure 5. However, two new analysis algorithms have been developed for zCOSMOS to specially account for this spatial sampling, without the need for modeling. The first is to combine spectroscopic redshifts and photo-z likelihood functions $L(z)$ for those objects without spectroscopic redshifts in order to optimally generate the density field (ZADE, see Kovac et al. 2009a) and the second is a new method to construct the spatial correlation function from such nonuniformly sampled data (C. Porciani & S. Lilly 2009, in preparation). In both cases, careful simulations on mock catalogs indicate that the new methods are superior to simple correction by a sampling map, and the interested reader is referred to these for more details.

3. THE 10k SAMPLE

The latest v3.51 of the 10k catalog is given in Table 3. The table includes all 10,644 objects for which a spectrum has been extracted, even if no redshift was measurable. Table 1 includes both primary and secondary targets and it also includes objects that were intended to be excluded as stars but which were observed anyway.

Figure 8 shows the overall redshift distribution of objects with secure redshifts. There is prominent spiking from the structure in the field. This plot emphasizes the importance of cosmic variance even in fields as large as COSMOS which at $z \sim 1$ spans a transverse dimension of approximately 100 Mpc. The full galaxy density field is presented by Kovac et al. (2009a) and a group catalog by Knobel et al. (2009).

3.1. A Statistically Complete Subsample

Users of the catalog who wish to construct a statistically complete subsample of the 10k sample should perform the following four operations:

1. All objects observed *only* as a secondary target should be excluded. These are indicated by a Confidence Class beginning with a "2" in Column 5 of Table 3.
2. All objects that were "forbidden" but which were observed anyway should be excluded. These are indicated by a "0" in Column 7 of Table 3.
3. All objects that were observed as "compulsory" should be downweighted by a factor of 2.05. These are indicated by a "2" in Column 7 of the table.
4. All objects outside of the range $15 < I_{\rm AB} < 22.5$ in Column 6 of the table should be excluded.

3.2. Galaxian Properties in the Sample

As a flux-limited sample, the range of galaxian luminosities varies with redshift. Figure 9 shows the $M_{\rm B,AB}$ as a function

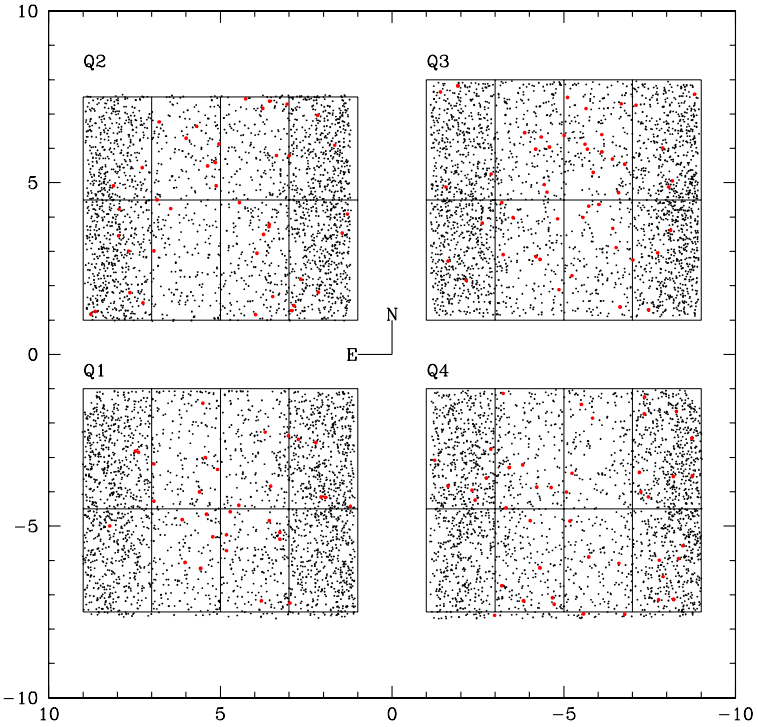

Figure 6. Location of all slits in the 10k sample within the VIMOS focal plane. Axes are in units of arcmin on the sky. The slits have a nonuniform pattern to maximize the number of slits. Compulsory (X-ray selected) targets are indicated by red and have a more uniform distribution on account of their higher priority in the mask design process.

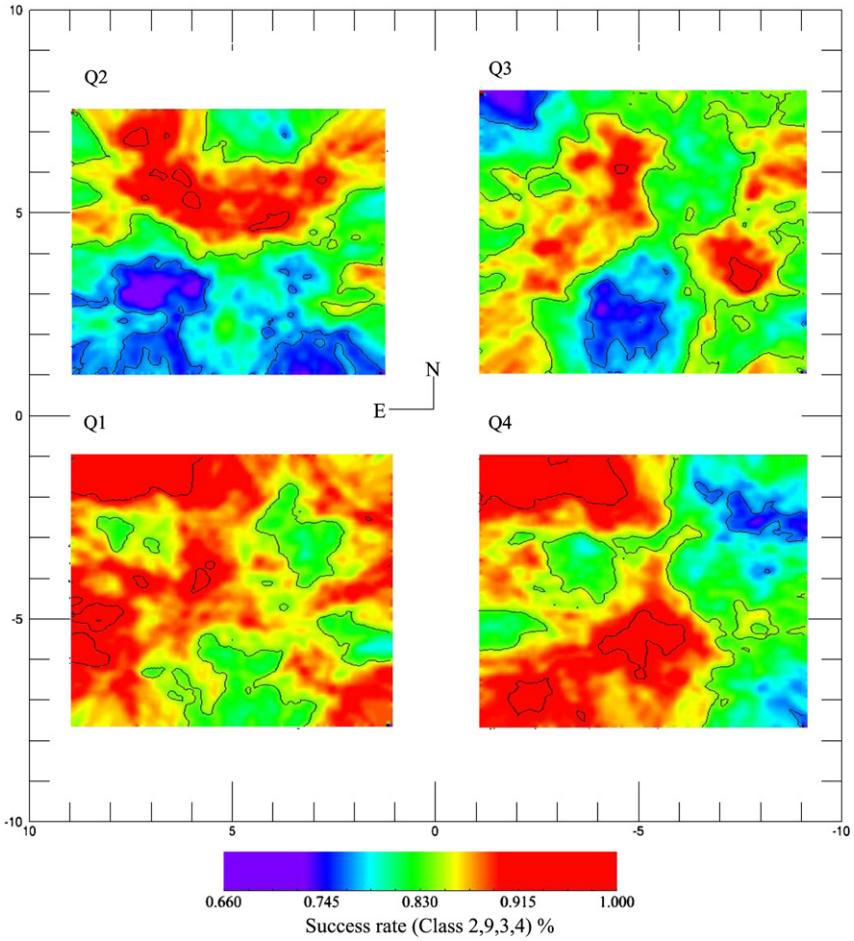

Figure 7. Average success rate in measuring a redshift as a function of position within the VIMOS field of view.

Table 3
zCOSMOS-bright 10k Sample

ID	R.A. (J2000)	Decl. (J2000)	z[a]	Confidence Class[b]	I_{AB}	Mask Priority[c]
701831	149.978638	2.473719	9.9999	0.0	22.27	1
701855	149.967194	2.522962	0.2639	4.5	22.49	1
701867	149.942169	2.552188	0.6683	2.5	22.46	1
701872	149.968689	2.561477	9.9999	0.0	22.50	1
701873	149.968887	2.563699	0.1821	9.1	22.46	1
701911	149.849625	2.522367	0.3756	4.5	22.48	1
701916	149.855301	2.541329	0.5012	4.5	22.47	1
701926	149.857544	2.588878	0.1516	9.1	22.46	1
702163	150.064880	2.651896	9.9999	20.0	22.48	1
702398	150.335876	2.053984	0.8378	2.5	23.02	1
702408	150.162262	1.991348	0.1240	4.5	19.29	1
800074	150.767700	1.612734	0.9361	3.5	21.65	1
800075	150.766983	1.617270	0.1721	4.5	20.37	1
800091	150.757111	1.625080	0.2148	4.5	21.73	1
800159	150.719849	1.617681	0.7932	4.1	22.34	1
800216	150.687683	1.623919	1.3207	9.3	22.28	1
800229	150.677124	1.613463	0.0000	1.5	21.32	1
800246	150.669510	1.610566	0.6230	3.5	21.85	1
800248	150.668030	1.616499	0.3725	4.5	19.99	1
800249	150.667923	1.618669	0.0000	23.4	16.27	0
800270	150.653290	1.625360	0.4808	4.5	19.36	1
800377	150.586746	1.614795	0.9817	2.1	22.24	1
800413	150.566101	1.609985	1.0167	9.5	22.30	1
800415	150.563858	1.619524	0.3661	4.5	22.07	1
800454	150.535828	1.609854	0.9299	1.5	21.75	1
800468	150.524902	1.614850	0.0000	4.5	18.12	1
800507	150.506027	1.623860	0.2061	4.5	20.28	1
800509	150.505997	1.612080	0.2061	4.5	18.15	1
800514	150.503265	1.625479	0.1733	4.5	20.58	1
800557	150.478561	1.619601	9.9999	0.0	22.23	1
800559	150.477417	1.615959	0.9235	2.5	21.88	1
800685	150.411606	1.625469	0.7465	2.5	21.80	1
800696	150.405518	1.619081	0.0997	2.5	18.14	1
800717	150.391556	1.614257	0.5897	3.5	22.33	1
800816	150.339005	1.615391	0.6163	22.1	18.67	1
800821	150.338730	1.614216	0.2270	3.5	19.86	1
800827	150.337265	1.608831	0.2202	3.5	21.99	1

Notes.
[a] A redshift of $z = 9.9999$ means no redshift identification was made.
[b] See text and Table 1 for explanation. The set of objects defined as Classes 1.5, 2.4, 2.5, 9.3, 9.5 and all Class 3.x and 4.x comprise 88% of the sample (95% between $0.5 < z < 0.8$) and are estimated to be 99% reliable. The redshifts with Classes 1.1, 2.1, and 9.1 (and 21.1, 22.1, and 29.1) are likely to be incorrect and should not be used, even though they are retained in the catalog for completeness.
[c] The mask priority indicates the priority for insertion in the slit masks. "1" indicates a randomly selected object from the main catalog. "2" indicates a "compulsory target" that has a roughly twice higher chance of having been observed. "0" indicates an object that was not intended to be observed (e.g., because it was thought to be a star) but in fact was observed, usually as a secondary object in the slit of another target.
(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

of redshift. For many applications, it is desirable to construct samples that approximate "volume limited" samples, in which the same galaxies would appear at all redshifts.

Our current understanding of galaxy evolution has not yet reached the level of sophistication where this is possible, even in principle. Furthermore, practical complications, such as the merging of galaxies or the differential build-up of their stellar populations through star formation, may make such a goal in any case illusory. However, stellar populations with a wide range of star formation histories follow a similar luminosity evolution in the rest-frame B band (see, e.g., Figure 16 of Lilly et al. 1998). This is because for declining star formation rates, the luminosity evolution at longer wavelengths primarily reflects the passive evolution of a dominant older population. This can easily reach one magnitude of luminosity evolution at a redshift of unity. Accordingly, a simple luminosity selection over a wide redshift range is inappropriate and we recommend the use of an Absolute Magnitude selection cut that increases in luminosity by 1.0 magnitude per unit redshift interval, i.e., $M_{B,cut} = M_{B,0} - z$ as a first-order correction for the effects of luminosity evolution of individual galaxies to produce an approximation to a volume-limited sample. However, it should be appreciated that such a correction is at best approximate.

An alternative approach would be to construct mass-limited samples of galaxies. These also face the conceptual ambiguities discussed above, i.e., nonconservation of number through

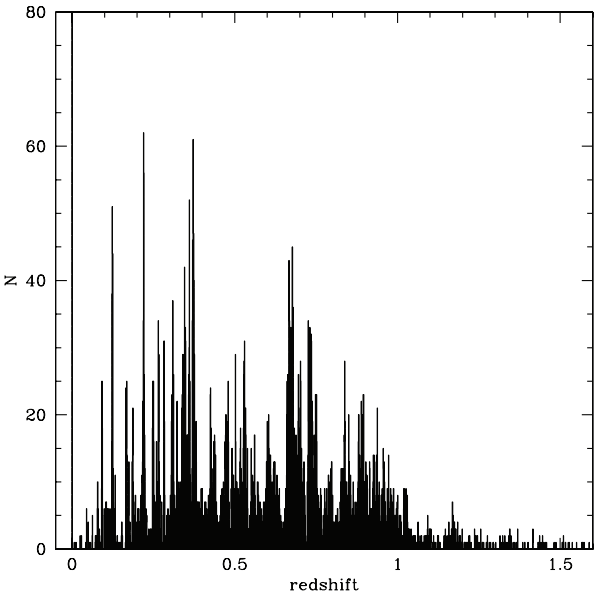

Figure 8. Redshift distribution of extragalactic objects in the zCOSMOS-bright 10k sample with secure redshifts, binned in intervals $\Delta z = 0.001$, which is larger than the redshift uncertainty by a factor of about three at $z = 0$ and of 1.5 at $z \sim 1$. The redshift distribution shows structure on a large range of scales. See Kovac et al. (2009a) for a full reconstruction of the density field.

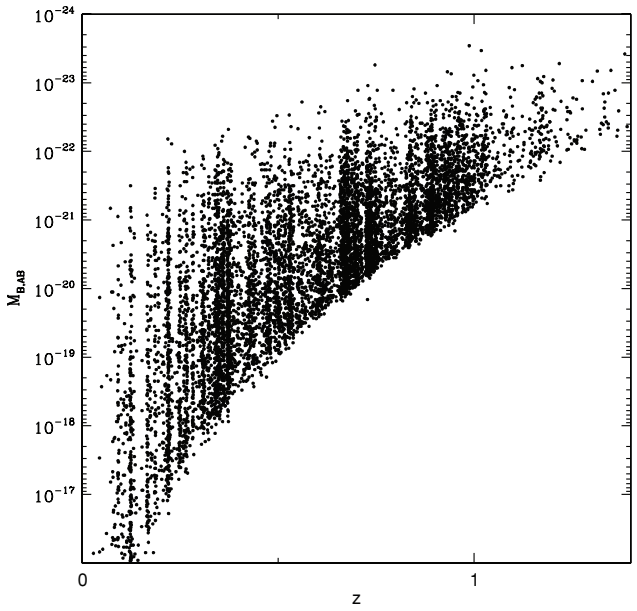

Figure 9. Absolute magnitudes in the rest-frame *B* band of the 10k sample.

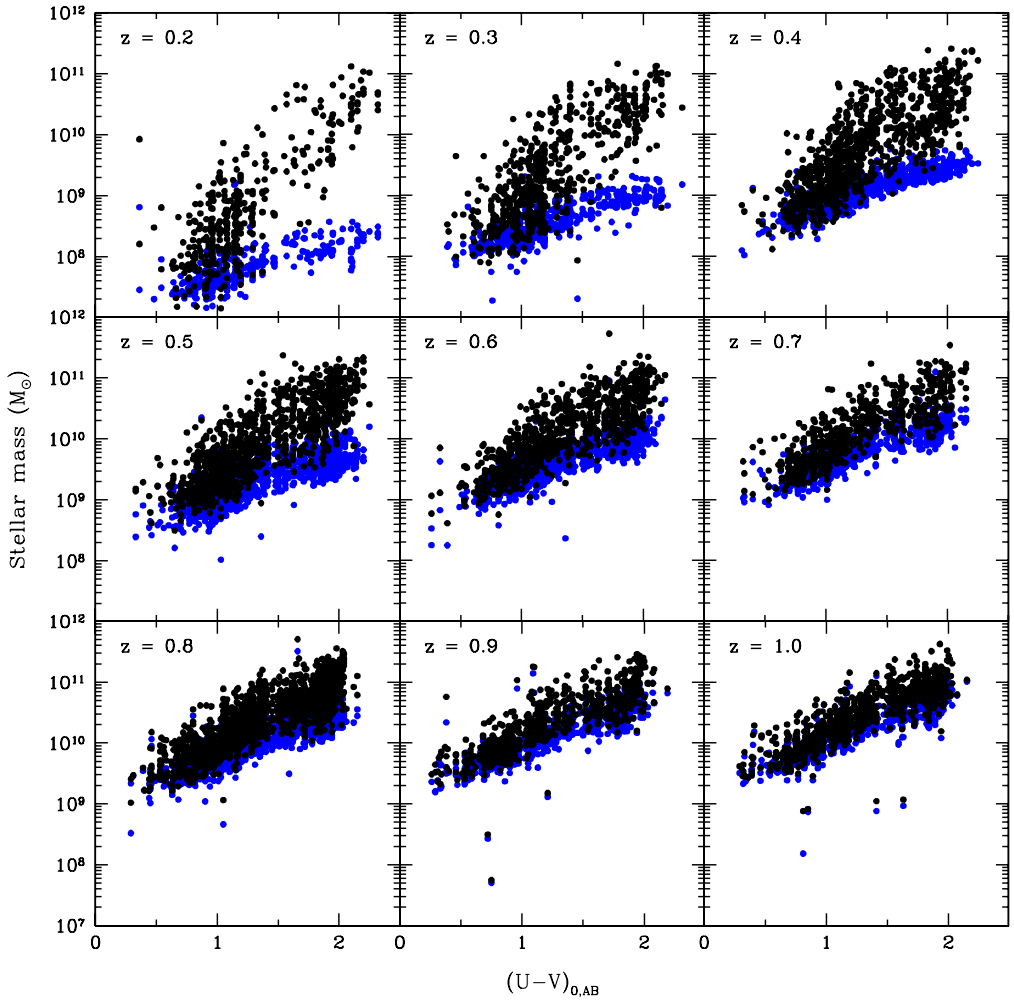

Figure 10. Stellar masses of zCOSMOS galaxies from Bolzonella et al. (2009) in bins of $\Delta z = 0.1$, centered on the indicated redshifts (black points). Blue points show the masses of these galaxies if they had the limiting magnitude at $I_{AB} = 22.5$ and effectively shows the mass limits of the zCOSMOS-bright survey at different redshifts. The U filter in this plot is a standard U and is bluer than that used in Figure 2.

merging and differential mass assembly, with the additional complication that the mass-to-light ratio of a galaxy depends on the stellar populations and thus on the rest-frame colors. A straight flux-limited sample reaches to substantially lower masses for blue star-forming galaxies than for quiescent red galaxies, and the sample will only be complete (i.e., representative of all galaxy types) for relatively high masses. This is illustrated in Figure 10, which shows the stellar masses of the galaxies in the 10k sample computed as in Bolzonella et al. (2009). The black dots show the actual masses of the galaxies in the 10k sample, while the underlying blue dots show the masses that the galaxies would have if they had the limiting magnitude $I_{AB} = 22.5$ (and the same stellar population), which gives a crude indication of the mass completeness of the sample as a function of redshift and rest-frame $(U - V)$ color.

4. SUMMARY

This paper presents the current set of spectroscopic redshifts that have been measured in the COSMOS field from 10,644 spectra taken so far in the zCOSMOS-bright program. A careful attention has been paid to characterize the properties of the spectroscopic sample and to quantify the reliability and accuracy of the individual redshifts, both through repeat measurements and through the comparison with high-quality independently estimated photometric redshifts.

The redshifts are individually accurate at a 1σ level of about 110 km s^{-1} enabling the identification, reported elsewhere, of groups and other manifestations of large-scale structure in the Universe.

Some 88% of the galaxies observed so far in zCOSMOS-bright have a spectroscopic redshift that is individually secure at the 99% level. In the key $0.5 < z < 0.8$ redshift range, the success rate for obtaining reliable redshift measurements increases to 95%.

The redshift distribution shows strong spiking at all redshifts, emphasizing the presence of "cosmic variance" even in rather large survey fields such as COSMOS.

Efficient execution of such a large program in Service Mode on the VLT requires the efforts of many individual observatory staff at ESO, both in Garching and on Cerro Paranal, and we gratefully acknowledge the contribution of these many and regrettably sometimes anonymous individuals to our project. The zCOSMOS program builds on many of the hardware and software tools built by the VIMOS/VVDS team, and we acknowledge here the contribution of those in the VIMOS/VVDS team who are not part of the zCOSMOS project. Finally, we acknowledge with appreciation the contributions of those many individuals not listed as authors who have worked, directly or indirectly, to produce the superb imaging data on which the COSMOS survey is based.

REFERENCES

Bardelli, S., et al. 2009, A&A, submitted
Bolzonella, M., et al. 2009, arXiv: 0907.0013
Bottini, D., et al. 2005, PASP, 117, 996
Brusa, M., et al. 2009, ApJ, 693, 8
Capak, P., et al. 2007, ApJ, 172, 99
Caputi, K., et al. 2008, ApJ, 680, 939
Caputi, K., et al. 2009, ApJ, 691, 91
Cucciati, O., et al. 2009, A&A, submitted
de la Torre, S., et al. 2009, A&A, submitted
Drinkwater, M. J., Gregg, M. D., Hilker, M., Bekki, K., Couch, W. J., Ferguson, H. C., Jones, J. B., & Phillipps, S. 2003, Nature, 423, 519
Feldmann, R., et al. 2006, MNRAS, 372, 565
Gilli, R., et al. 2009, A&A, 494, 33
Ilbert, O., et al. 2009, ApJ, 690, 1236
Iovino, A., et al. 2009, A&A, submitted (arXiv:0909.1951)
Knobel, C., et al. 2009, ApJ, 697, 1842
Koekemoer, A. M., et al. 2007, ApJS, 172, 196
Kovac, K., et al. 2009a, arXiv:0903.3409
Kovac, K., et al. 2009b, ApJ, submitted, arXiv:0909.2032
Le Fèvre, O., et al. 2005, A&A, 439, 845
Lilly, S. J., et al. 1998, ApJ, 500, 75
Lilly, S. J., et al. 2007, ApJS, 172, 70 (ZC-07)
Maier, C., et al. 2009, ApJ, 694, 1099
Mandelbaum, R., et al. 2008, MNRAS, 386, 781
Meneux, B., et al. 2009, arXiv:0906.1807
Pozetti, L., et al. 2009, A&A, submitted (arXiv:0907.5416)
Salvato, M., et al. 2009, ApJ, 690, 1250
Sanders, D., et al. 2007, 172, 86
Scodeggio, M., et al. 2005, PASP, 117, 1284
Scoville, N. Z., et al. 2007, ApJ, 172, 1
Silverman, J. D., et al. 2009a, ApJ, 695, 171
Silverman, J. D., et al. 2009b, ApJ, 696, 396
Tasca, L., et al. 2009, A&A, 503, 379
Vergani, D., et al. 2009, A&A, submitted (arXiv:0909.1968)
Zucca, E., et al. 2009, A&A, submitted