Galaxies to the redshift desert and beyond III. The photometric redshift code GAZELLE

Ralf Kotulla⋆

Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, United Kingdom

ABSTRACT

Key words:

1 INTRODUCTION

Something about the history of photo-z

Different methods: Neural nets, training sets, template codes

Downside: No physical parameters or interpretation (because interpolation), no errors, no evolution in the templates

1.1 Why yet another photometric redshift code?

Despite the wealth of publicly available redshift codes such as HYPERZ (Bolzonella et al. 2000), BPZ (Benítez 2000), ANNz (XXX missing reference XXX 2035), ZEBRA (Feldmann et al. 2006), LePhare (XXX missing reference XXX 2035), EAZY (Brammer et al. 2008) to just name the more common ones we decided to create yet another one, which we called GAZELLE. The main reason for this is that the main goal of GAZELLE is to derive not only a photometric redshift for each galaxies, but rather physical properties such as masses, star formation rates and galaxy types. While this was also possible based on results from several of these other codes using a combination of methods, we here describe an unified approach to derive these relevant parameters from one set of templates

1.2 Differences and similarities with other photo-z codes

2 THE PHOTOMETRIC REDSHIFT CODE gazelle

To make best use of the wealth of information supplied by our GALEV models (Anders & Fritze 2003; Bicker et al. 2004; Kotulla et al. 2009), we developed a new and innovative photometric redshift code that we named GAZELLE. Its working is described below.

⋆ E-mail: r.kotulla@herts.ac.uk

2.1 The object photometry catalog: Required and optional information

As in most other photometric redshift codes the first processing step is to read the input catalog. This input catalog contains **a)** an arbitrary number of columns (the number can be specified via the configuration file) that allow to identify each individual galaxy, e.g. catalog numbers, coordinates, etc.; **b)** the galaxy photometry, either as measured fluxes or magnitudes. Each photometric datapoint needs to be followed by its uncertainty, either expressed as flux error, magnitude error or signal-to-noise ratio.

After reading this input catalog, GAZELLE compares each input photometry against the detection limit of the respective filter. Detected magnitudes below the specified detection limit get flagged as suspicious, and non-detections in one or more filters for an otherwise detected object are also flagged depending on whether they lie outside the observed field-of-view (usually indicated by a magnitude of -99 mag) or unobserved because they are too faint (magnitudes of $+99$ mag). The reason for this flagging is detailed in Sect. 2.4. The format of this file hence is largely compatible with the requirements of the other codes to allow for easier inter-comparison of different codes.

2.2 The template set: Galaxy models including spectroscopic and chemical evolution

As mentioned earlier the major improvement of GAZELLE compared to all previous photometric redshift codes is that as comparison templates we use models that include the full spectroscopic and chemical evolution from the onset of star formation shortly after the Big Bang until the present day. For this reason it was inefficient to input the full spectra into GAZELLE and convolve them with the filter curves on-the-fly. Instead we decided to do this computational expensive process beforehand and use spectral energy distributions, apparent magnitudes in each of the specified filters that are generated by our GALEV models, as principle input into GAZELLE. These SEDs, one for each of the redshift range corresponding to the time-steps of our models, include both k- and e-corrections as well as the attenuation due to intergalactic neutral hydrogen clouds as shortly described above and in more detail in Kotulla et al. (2009). Supplementary input such as dust extinction for each filter as function of redshift (the SED-equivalent to extinction curves) as well as physical parameters (masses, star formation rates, metallicities, etc.) computed by GALEV are also read from their individual files and internally merged into a final catalog of SEDs and physical parameters.

2.3 The algorithm

The algorithm of GAZELLE is based upon a slightly modified χ^2 minimisation algorithm. For each (observed SED)-(model SED) combination we derive a χ^2 value by

$$\chi^2 = \sum_{i=1}^{\#\text{filters}} \left[w_i \frac{F_{\text{model}}^i - \alpha \times F_{\text{obs}}^i}{\sigma_i} \right]^2, \tag{1}$$

where F_{model}^i and F_{obs}^i are the fluxes in the i-th filter from the model grid and the observations, respectively. σ_i, the uncertainty of this SED point is given by

$$\sigma_i = \sqrt{\sigma_{i,\text{obs}}^2 + \sigma_{i,\text{model}}^2}, \tag{2}$$

where $\sigma_{i,\text{obs}}$ is the uncertainty of the observation, i.e. photometric errors. The additional term $\sigma_{i,\text{model}}$ represents uncertainties of the models to account for variations among galaxies of identical spectral types as well as uncertainties propagated from the modelling process, such as uncertainties in our knowledge of the Initial Mass Function, stellar evolution, incomplete stellar libraries, changes in metallicity due to gas infall or outflows, etc. (see Conroy et al. 2008, 2009, for more details on uncertainties of population synthesis models). We typically assume a model uncertainty of 0.1 mag independent of wavelength, although a wavelength-dependent parametrisation is supported by GAZELLE. The exact value for the model uncertainty is not crucial in the sense that it does not significantly affect the best-match solution; it does however influence the confidence ranges, with smaller model uncertainty ranges resulting in more closely confined solutions. The inclusion of this additional error term was already successfully implemented by, e.g., Anders et al. (2004) and Brammer et al. (2008).

One crucial parameter in eq. 1 is the scaling parameter α. This parameter directly influences the resulting mass and all dependent parameters such as star formation rates of the galaxy by scaling all luminosities with the same factor to match the average observed luminosity. The α parameter is computed by

$$\alpha = \frac{\sum_i w_i F_{\text{model}}^i}{\sum_i w_i F_{\text{obs}}^i}, \tag{3}$$

where w_i is a weighting factor. Unlike most χ^2 algorithms we offer, in addition to the normal weighting with all $w_i = 1$ the possibility to use $w_i = \sigma_{i,obs}^2$ in anticipation of the upcoming χ^2 computation. Tests comparing both methods showed that this alternative method yields slightly better results, in particular in cases where different filters are observed to very different depths with consequently very different photometric uncertainties.

Another crucial step on the way to accurate redshifts is the determination of the intrinsic reddening, since this is able to dramatically change both the shape and the normalisation of the SED. To find the minimum χ^2 value for one particular (observed SED)-(model SED) combination as a function of reddening we implemented a golden section search to sparsely scan the allowed reddening range and confine the best solution to within a small range (typically $\Delta E(B - V) \leqslant 0.05$ mag) of the optimum. In a last step we fit a parabola to the points near the minimum to obtain the final, best-match value.

For a typical run, each model in the comparison grid contains ≈ 800 SEDs, covering the redshift range $z = 0 - 8$ in steps of 0.01. For each galaxy we scan the full grid, i.e. we compute χ^2 values for every SED of every model.

Once all χ^2 values for each possible (observed SED)-(model SED) combination has been computed we derive raw probabilities for each SED by

$$P_{\text{raw}} = \exp\left(\frac{-\chi^2}{2}\right), \tag{4}$$

and, in a next step normalise all raw probabilities P_{raw}^j to yield

$$P_{\mathrm{norm}}^j = \frac{P_{\mathrm{raw}}^j}{\sum P_{\mathrm{raw}}^j \times \Delta z^j} \qquad (5)$$

with P_{norm} fulfilling the condition

$$\sum P_{\mathrm{norm}}(z) \times \Delta z(z) = 1. \qquad (6)$$

$\Delta z(z)$ is the width of the redshift interval corresponding to this SED point. This is necessary as consequence of a potentially inhomogenous redshift sampling of our model. Optionally we offer the possibility to first normalise all χ^2 to yield $\chi^2_{\mathrm{min}} = 1$, in analogy with the approach presented in Le Borgne & Rocca-Volmerange (2002).

We now sort the resulting array of probabilities by increasing probabilities or equivalently decreasing χ^2 values. The model SED with the highest probability determines the best-match values for all parameters, i.e. for redshift, extinction, galaxy type, masses, star formation rate, etc. To derive confidence ranges we progress down this sorted list and integrate probabilities until 0.68 (in the case of $1\,\sigma$ uncertainties) have been reached. The confidence range is determined by the minimum and maximum values for each parameter reached within this interval. This implies that the extreme values covered by the confidence ranges do not necessarily result from the same point in the model grid. We also point out that we do not make any assumptions on the distribution of values around the best-match value. This leads to generally asymmetric confidence intervals that also do not need to be gaussian distributed.

2.4 Treatment of non-detections

The χ^2 algorithm described above is not able to work with non-detections that occur frequently for a number of reasons, either because

• **a)** the galaxy intrinsically does not emit flux in the observed wavelength range or all emitted flux is absorbed on the way towards us (e.g. shorewards of the Lyman break);

• **b)** the survey is not sufficiently deep to observe the faint flux from the galaxy; or

• **c)** the galaxy happens to lie outside the field-of-view covered by the observations in one or more of the filters (for instance in the case of a wide-field optical survey that was only partially followed up in the NIR).

Similar criteria also apply to the models, that can predict the galaxy to be undetected in one or more bands, for above reasons a) and/or b). Each of these cases contains information by itself and hence deserves proper inclusion.

In all above cases where either the model and/or the observed data are not detected in one or more bands, these filters are excluded from the computation of the mass scaling parameter α.

For the following χ^2 calculation we implemented the following behaviour:

• **Reason a) and/or b) for both model AND data:** In case that both observations and model agree on a non-detection (reason a) or b) apply to both e.g. for dropouts or non-detections due to too shallow data) this filter is taken as a perfect match and assigned $\chi_i^2 = 0$.

• **Reason a) and/or b) only apply to model OR data:** In the case of disagreement, e.g. the model predicts a magnitude brighter than the detection limit but the observations did not detect it, or vice versa, the model predicts too faint a magnitude but the object was detected, GAZELLE offers two possibilities: The χ_i^2 value for this filter can be assigned a fixed value to account for this mismatch or alternatively can be computed by the difference between magnitude limit and observed (or predicted) magnitude divided by the model uncertainty.

• **Reason c):** In the case the object was not detected because it does not fall into the observed area, this data point is excluded from the χ^2 calculation.

This process ensures best use of the available data.

2.5 Treatment of bandpasses beyond NIR

While non-detections are more relevant for the shortest wavelength filters there are also potential problems with bandpasses at long wavelengths. At rest-frame wavelength beyond K-band or $\lambda \approx 2.5\mu m$ the spectra of galaxies are increasingly affected by non-stellar emission, most prominently due to dust, that is currently not yet included in our chosen template set. We therefore exclude all filters from both the mass determination as well as the χ^2 computation for which the central wavelength, as derived form the filter curve, exceeds a certain maximum rest-frame wavelength (we chose a conservative value of $\lambda_{\mathrm{max}} = 3\mu m$ as default value). Observations at longer wavelengths, for example from IRAC on board the Spitzer Space Telescope, are hence only used in the comparison with model SEDs above a certain redshift.

2.6 Implementation

The above algorithm is implemented in the software program GAZELLE. GAZELLE is programmed in C/C++ to allow easy modification and adaptations to more specific projects. However, most parameters that determine the behaviour of the algorithm (e.g. the maximum allowed wavelength, model uncertainties, etc.) can be configured via a parameter file so that GAZELLE is simple to use. For large datasets and/or large model grids GAZELLE can also be run in a MPI[1]-compliant multi-processor environment or even large CPU clusters to decrease the required wall-clock–time considerably. While GAZELLE is available on request from the author we also plan to implement a web-compatible version into the GALEV webpage (http://www.galev.org) to enable even easier access to the wider community.

2.7 Output of results

Once the algorithm described above has been applied to each SED of the model grid GAZELLE is able to derive a wealth of information about the object. The resulting output hence typically contains

---

[1] Message-Passing Interface, see http://www.mcs.anl.gov/research/projects/mpi/ or http://www.mpi-forum.org/ for further details

• The best-match photometric redshift including its respective uncertainty ranges, typically including 68% of the total probability. Other and/or further confidence limits can also be configured.

• The galaxy model containing the best-match model SED. From this model we immediately can infer the most likely past star formation history and, in the case the galaxy contained a starburst or star formation truncation event, the age since this event.

• The mass scaling factor α. Since our undisturbed galaxy models are calibrated to match the properties, in particular masses and luminosities, of a typical $1L_\star$ galaxy of their respective spectral types in the local universe we can directly infer the mass relative to the characteristic mass of galaxies of this type.

• Physical parameters inferred from the above mass scaling factor α in combination with the galaxy type. These are both stellar and gaseous masses and hence total baryonic masses, star formation rates and stellar population ages. We furthermore can derive both integrated stellar and gas-phase metallicities as well as stellar population ages. Since we compute a complete probability density distribution we also obtain confidence ranges for each one of the above parameters, fully accounting for redshift, galaxy type, mass and dust reddening uncertainties.

• Optionally the full probability distribution function (PDF) for each object from the input catalog to enable follow-up studies. Another optional output is the best-match SED, including the effects of the mass scaling and dust extinction to allow for easy presentation and visualisation of the fit.

This enables us to not only derive the redshift of a galaxy, but also its star formation history, mass assembly history and chemical enrichment history.

2.8   GAZELLE as SED-fitting code

Based upon the above described functionality GAZELLE offers several other features more applicable to studies of galaxies with available spectroscopy and/or known distances. Among those features is the ability to also include spectral indices, e.g. in the Lick-system (Trager et al. 1998), in the comparison. The main difference between the treatment of observed magnitudes and spectral indices is that the latter do not scale with mass, and hence excluded from the calculation of α. The second feature is a generalisation of the above approach to enable fitting of spectra to a grid of GALEV models.

While the internal working, i.e. the χ^2 algorithm, is the same for all operation modes, the treatment of non-detections and the scaling of the input models with the factor α changes from case to case. A detailed description of each of these analysis modes will be given with the respective applications to data.

3   THE SAMPLE

To study the variation of the inferred mass functions on the deepfield at hand we compiled a large catalog from readily available public catalogs published by the teams of each
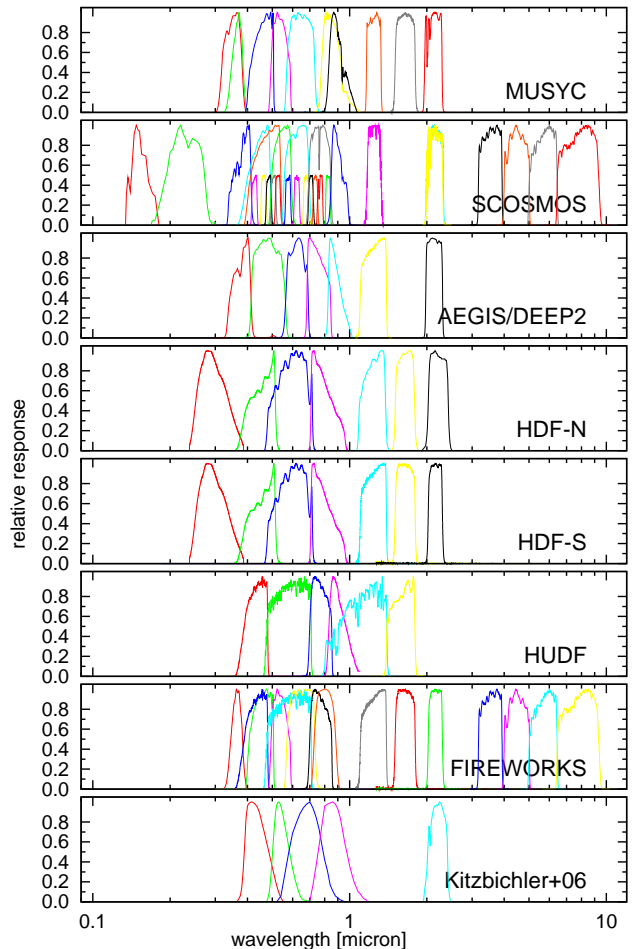

Figure 1. Response curves of the filters used in each of the survey.

respective survey. For each of the fields we used the filter functions for the according telescope and instrument to not include additional uncertainties due to colour transformations. Photometric redshifts as well as physical parameters were then derived individually for each field, so the accuracy of our photo-z compared to the true (spectroscopic) redshifts also varies from field to field as consequence of the varying filter coverage. In the following we detail the data we used for the presented analysis. A summary of field sizes, number of spectroscopic redshifts, filter sets and obtained accuracy can be found in Table **??**. For more details on data reduction, source selection, and alike we refer the interested reader to these papers.

In Fig. 1 we show the filter sets covered by each of the fields to give an overview of the covered wavelength basis and sampling of each of the SEDs.

3.1   COSMOS

The Cosmic Evolution Survey (COSMOS, Scoville et al. 2007b) covers a 2 sq.deg. equatorial field centred on $\alpha = 10^h00^m29^s$ and $\delta = +02°12'21''$, and features a very wide wavelength coverage from X-rays (Brusa et al. 2007; Cappelluti et al. 2007; Finoguenov et al. 2007; Hasinger et al. 2007), UV (Zamojski et al. 2007), optical both ground-based in broad- and narrow-band filters (Taniguchi et al. 2007) as

well as HST ([Koekemoer et al. 2007](); [Scoville et al. 2007a]()), near-infrared ([Capak et al. 2007]()), mid-infrared from Spitzer ([Ilbert et al. 2009]()) to radio ([Schinnerer et al. 2004](), [2007](); [Bondi et al. 2008]()). As such, many spectroscopic redshifts have been obtained by various groups, e.g [Lilly et al. (2007)]().

Our catalog is based on the latest (Release April 2009) i-band selected UV-optical-NIR catalog[2] including photometry in 16 broad- and 12 narrow-band filters. We restricted the catalog to sources that are also detected in the K-band to ensure the large wavelength coverage necessary to obtain reliable photo-z (our COSMOS catalog). We also matched the i+K-band selected catalog with the catalog of $[3.6\mu m]$ detected sources (yielding our SCOSMOS catalog), so that we have two slightly different source catalogs for the COSMOS field.

3.2 FIREWORKS: GOODS-CDFS

The FIREWORKS-catalog[3] ([Wuyts et al. 2008]()) contains K-band selected and aperture-matched photometry of the GOODS Chandra Deep Field South (CDFS, [Giacconi et al. 2000]()) region and covers the wavelength range from U through $24\mu m$ with in total 18 filters. For reasons outlined in Sect. **??** we do not use the $[5.8\mu m]$, $[8.0\mu m]$ and $[24\mu m]$ filters, though. The catalog furthermore contains spectroscopic redshifts for XXX sources (see [Wuyts et al. 2008](), and references therein), allowing for easy comparison of the obtained photo-z, making this catalog very convenient to use.

3.3 MUSYC: CW1255, SDSS1030, EHDFS and ECDFS

The Multiwavelength Survey of Yale-Chile (MUSYC[4]) covers a total area of $\approx$ 1sg.deg. spread out over four fields (Extended Hubble Deep Field South (EHDFS), Extended Chandra Deep Field South (ECDFS), plus an additional two fields centred on $\alpha = 10^h30^m27^s$, $\delta = 05°24'55''$ [SDSS1030] and $\alpha = 12^h55^m40^s$, $\delta = 01°07'00''$ [CW1255]). For this study we use the K-band selected deep and wide catalogs presented in [Quadri et al. (2007)](), [Blanc et al. (2008)](), and Taylor et al. (in prep.). The deep versions differ from the wide versions by a K-band detection limit $\approx$ 1 mag deeper (22.8 vs. 21.7 mag) and the additional availability of J- and H-band photometry in addition to the UBVRIzK photometry common to all fields.

3.4 SXDF/UDS

The Subaru-XMM Deep Field (SXDF[5], [Sekiguchi & et al. 2004]()) covers a 1.3 sq.deg. equatorial field centred on $\alpha = 02^h18^m00^s$, $\delta = -05°00'00''$ in four optical filters (BRiz) to great depth ([Kashikawa et al. 2004]()). This field overlaps with the Ultra Deep Survey (UDS) of the UKIRT Infrared Deep Sky Survey (UKIDSS[6], [Lawrence et al. 2007]()), adding deep near infrared data in the J- and K-bands (), and was

also covered by the Spitzer Wide-Area Infrared Extragalactic (SWIRE[7], [Lonsdale et al. 2003]()) survey, extending the wavelength range into the mid-infrared. For this study we use the public catalog presented in [Williams et al. (2009)]().

3.5 FIRES: MS1054

The Faint Infra-Red Extragalactic Survey (FIRES[8], [Franx et al. 2000](); [Rudnick et al. 2001]()) covers the HDF-South and a field centred on the intermediate redshift galaxy cluster MS1054 with extremely deep NIR J-, H-, and K-band data. The catalog we use here is described in [Förster-Schreiber et al. (2006)]() and contains the UBVRIJHK filters.

3.6 Hubble Deep Fields

The Hubble Deep Field (HDF) North covers the area of one WFPC2 footprint ($\approx$ 6 sq.arcmin) on the northern hemisphere, centred on $\alpha = 12^h36^m49^s$, $\delta = +62°12'58''$ ([Williams et al. 1996]()). For the present study we use the catalog of [Fernández-Soto et al. (1999)](), that supplements the optical HST data in UBVI with NIR-data in J, H, and K. The catalog also contains spectroscopic redshifts from 103 galaxies. The Hubble Ultra Deep Field (HUDF, **?**) was taken a few years later with the Advanced Camera for Surveys (ACS) onboard HST and observed an area of XXX arcmin of the southern sky in BVRI to even greater depth. In our study we use the catalog compiled by [Coe et al. (2006)](), also containing J- and H-band data from HST-NICMOS.

3.7 AEGIS and DEEP-2

AEGIS, the All-wavelength Extended Groth Strip International Survey, covers an elongated patch of sky centred on $\alpha = 14^h17^m00^s$ and $\delta = 52°30'00''$ with multi-wavelength data spanning the range from X-ray, UV, optical, near- and mid-infrared to radio ([Davis et al. 2007]()). It was also covered by the Deep Evolutionary Exploratory 2 Galaxy Redshift Survey (DEEP2 [Willmer et al. 2006]()), supplying spectroscopic redshifts for a large sample of $R_{AB} \leqslant 24.1$ mag galaxies. The data we use in this study was compiled from two independent catalogs covering the Optical ugriz bands observed from CFHT ([Davis et al. 2007]()) and the near-infrared J and K-bands from Palomar ([Bundy et al. 2006]()). This compilation differs from the previous catalogs in that no PSF-matching between the different filters could be performed, and we will in the following discuss the implications of this. For the matching we started with the NIR catalog, and cross-correlated it with the optical catalog, using a matching radius of $2''$. In the next step we cross-correlated the NIR+Optical catalog with the DEEP2 redshift survey (using data release 3), again using a matching radius of $2''$. The final catalog hence contains only such galaxies that are detected in the NIR and the optical and where a spectroscopic redshifts is available from DEEP2.

---

[2] http://cosmos.astro.caltech.edu/
[3] http://www.strw.leidenuniv.nl/fireworks/
[4] http://www.astro.yale.edu/MUSYC/
[5] http://www.naoj.org/Science/SubaruProject/SDS/
[6] http://www.nottingham.ac.uk/astronomy/UDS

[7] http://swire.ipac.caltech.edu/swire
[8] http://www.strw.leidenuniv.nl/ fires/

4   THE GALEV MODEL SET

For each of the aforementioned fields we used a homogenous set of model templates and, with exception of field-dependent detection limits, the same GAZELLE configuration between the fields to ensure that differences between the individual fields only reflect changes in the data and not in the configuration.

We use the following template set:

• 5 galaxy types representing the local spectral types E and Sa through Sd, using our chemically consistent (c.c.) models that include a self-consistent treatment of the chemical evolution and fully account for the increasing abundances of subsequently formed stellar populations. These c.c. models are closely calibrated to reproduce a wide range of observables in the local universe as described in Kotulla et al. (2009) and Kotulla & Fritze (2009). However, we can not a priori know whether the relation between galaxy spectral type and stellar mass holds out to arbitrary high redshifts. This could mean that galaxies at higher redshifts are more or less massive than predicted by our models, leading to different metallicities as a consequence of the mass-metallicity relation. To account for this uncertainty we also add, for each type, models with the same SFH, but with metallicities fixed to $(0.2, 0.5, 1.0) \times Z_\odot$.

• A range of starburst and post-starburst models with 16 different burst ages ranging from 0 Gyr (i.e. at the peak of star formation activity) to 10 Gyrs, (i.e. old and red galaxies as typically found for ellipticals in the local universe). This sequence naturally includes strong post-starburst at ages of $0.5 - 1.0$ Gyrs when their spectra show strong Balmer absorption lines typical for k+a type galaxies in the low-redshift universe (Dressler et al. 1999; XXX missing reference XXX 2035; Falkenberg et al. 2009a,b), but that are increasingly found in high-redshift ellipticals (XXX missing reference XXX 2035). Again we allow 3 metallicities of $(0.2, 0.5, 1.0) \times Z_\odot$ for each of these models. Using our c.c. approach for these models is not appropriate, as these starbursts are likely triggered by mergers, hence rendering our assumption of closed-box evolution invalid. However, we are currently working on implementing a multi-zone approach into GALEV, that could lift this limitation in the near future.

• A sample of stellar spectra from the Lejeune et al. (1997, 1998) library, covering the full range of effective temperatures from $2000 - 50000$ K. Adding these templates aims at finding remaining stellar contaminants that were not flagged as such based on morphological criteria.

For all galaxies we assume star formation to commence (start? XXX) at $z = 8$ and then track the full evolution from there formation to the present day. We hence naturally include evolutionary corrections, accounting for the younger ages towards higher redshifts and increasing look-back times. This naturally ensures that all galaxies are younger than the age of the universe at each redshift. It also means that the number of available templates changes with redshifts, as, for instance, older post-burst galaxy templates are restricted to lower redshifts than models of ongoing bursts or undisturbed galaxies.

All models (spectra, apparent magnitudes, k- and e-corrections for all filters of the presented fields, stellar masses, SFRs, etc as function of time and/or redshift) are freely available from our website http://www.galev.org. Additional models for different assumptions and/or in different filters can also be computed via this web-site.

4.1   Comparison of predicted colour evolution with observations

The most important factor to derive accurate photo-z from templates is to ensure the models accurately trace the evolution of the observable SED with redshift. GAZELLE in this respect differs from all other photo-z codes currently available (see Hildebrandt et al. 2010, for an overview and short description of most other codes) in that it does not limit the template set to a few templates representing galaxies at different ages and then use the redshift as free parameter, but it rather assumes a fixed relation between galaxy age and redshift. Doing so eliminates one essential free parameter, and in turn imposes stringent requirements on the underlying star formation history, but at the same time makes the derived SFH more meaningful. Several codes (e.g. EAZY, Brammer et al. 2008) interpolate between templates to improve the quality of their photo-z fits, and while these interpolation factors can basically be used to derive SFHs, smooth SFHs are likely better representations of the true SFH than such consisting of only a few individual bursts.

The physical reason for this approach is XXX * old stars in all galaxies, even low-mass dwarfs * globular clusters in all local galaxies * semi-analytical models – some SF even at high redshift.

In Fig. 2 we show observed colours of galaxies in the COSMOS field in a range of filter combinations, using both broad- and intermediate-band filters. We overplot a small selection of the models we use to derive photometric redshifts. All shown models are computed with fixed metallicity of 0.5 $Z_\odot$ and, as mentioned above, assume a formation redshift of $z_f = 8$.

Fig. 2 clearly shows that our models offer a best-possible match to the observations without requiring any additional free parameters or a-posteriori calibration (except the zero-point matching described above). We particularly note the close match to the bumps and wiggles that are caused by emission lines in the star-forming templates. Without emission lines many of the features that closely determine the photo-z accuracy, such as the P-Cygni like evolution of the starburst template in the bottom row panels, can not be explained. This was also found by Ilbert et al. (2009), who also found a significant improvement in photo-z accuracy after inclusion of emission lines into their templates.

4.2   Advantages of this model coupling

The intended close coupling between GALEV evolutionary synthesis models on the one hand and GAZELLE as photo-z code on the other hand furthermore allows us to extract more information out of each SED than "only" a redshift. We already showed that in order to match the overall normalisation of the SED we need to apply a scaling parameter α (see eq. **??**). This factor not only scales the galaxy's stellar mass as source of the emitted light, but by linking to the galaxies SFH also affects the current SFR, and other mass-dependent parameters. Furthermore from our knowledge of
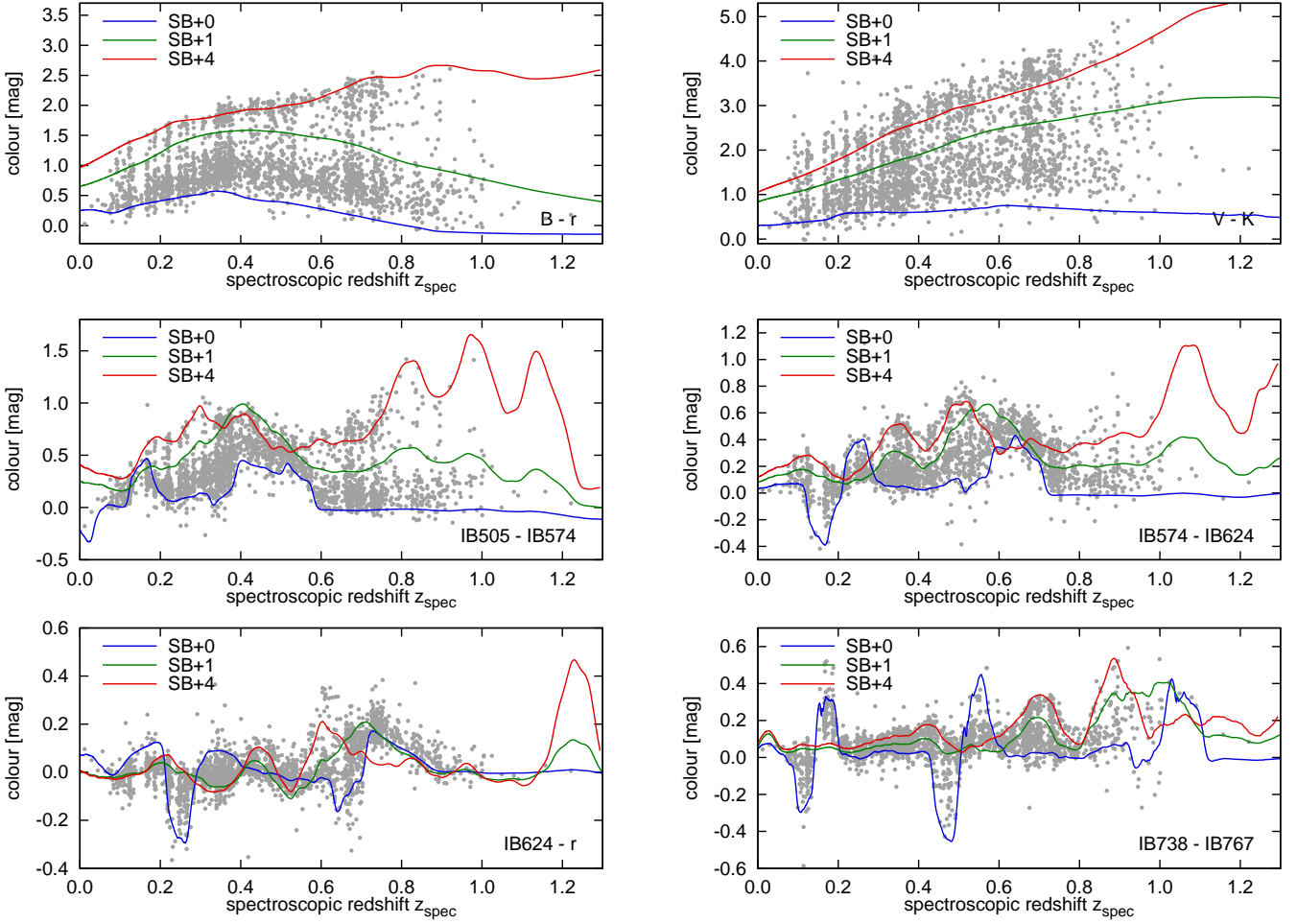
Figure 2. Comparison of the colour evolution with (spectroscopic) redshift as observed in COSMOS for a range of filter combinations: Top row: $B - r$ (left), $V - K$ (right); middle row: $IB505 - IB574$, $IB574 - IB624$; bottom row: $I624 - r$, $IB738 - IB767$. For comparison we show three of the models in our grid, a starburst at peak SFR and two post-starburst models 1 Gyr and 4 Gyrs after the SFR peak, each using a metallicity of 0.5 $Z_\odot$.

the SFH we can get a handle on mass- and light-weighted stellar ages. Mass-weighted ages in this context are derived by weighting the age of each stellar population with its mass, while light-weighted ages are obtained by weighting each age with its luminosity. As stellar population of different ages dominate different parts of the emitted spectrum (e.g. young stars dominating the emission in the UV and Blue, while NIR is dominated by older, evolved giant stars), these light-weighted ages generally depend on the filter at hand, with longer wavelength bands generally yielding older ages than bluer bands. Furthermore if we assume that galaxies evolve as closed-boxes as done in this study[9], or alternatively assume some some prescription for gas infall and outflows, we can also infer the currently available gas-mass and gaseous metallicity. Although this approach introduces spurious relations merely reflecting the initial assumptions (e.g.

a one-to-one relation between gas-mass and SFR in the case of our Sa-Sc templates) that have to be carefully considered during the analysis, it still allows to derive a wealth of physical information from the data at hand in a self-consistent manner.

Fig. 3 exemplifies this approach. Using the probability distribution derived from the SED fitting, we assign identical probabilities of each (observed SED)-(model SED) combination to each of the physical parameters of this model point, in this case the stellar mass and SFR. Confidence ranges for these paramters are then derived simultaneously with the photo-z confidence ranges by integrating probabilities from highest to lowest probability and deriving the extreme values reached within the 68% range.

This self-consistent and more realistic approach therefore also accounts for the uncertainties in the photo-z determination. This leads to generally larger uncertainties for each parameter as compared to the more common, sequential approach of first obtaining a (photometric) redshift, and then deriving masses and SFRs via SED fits *at fixed redshift* that ignores redshift uncertainties as major source of uncertainty for all other parameters.

---

[9] Note, however, that as consequence of the fraction-of-visible-mass parameter (see Kotulla et al. 2009, for details) the chemical evolution, but not the gas fraction, of our models is essentially equivalent to the model of **?** that assumes gas infall and outflows to be proportional to the star formation rate.
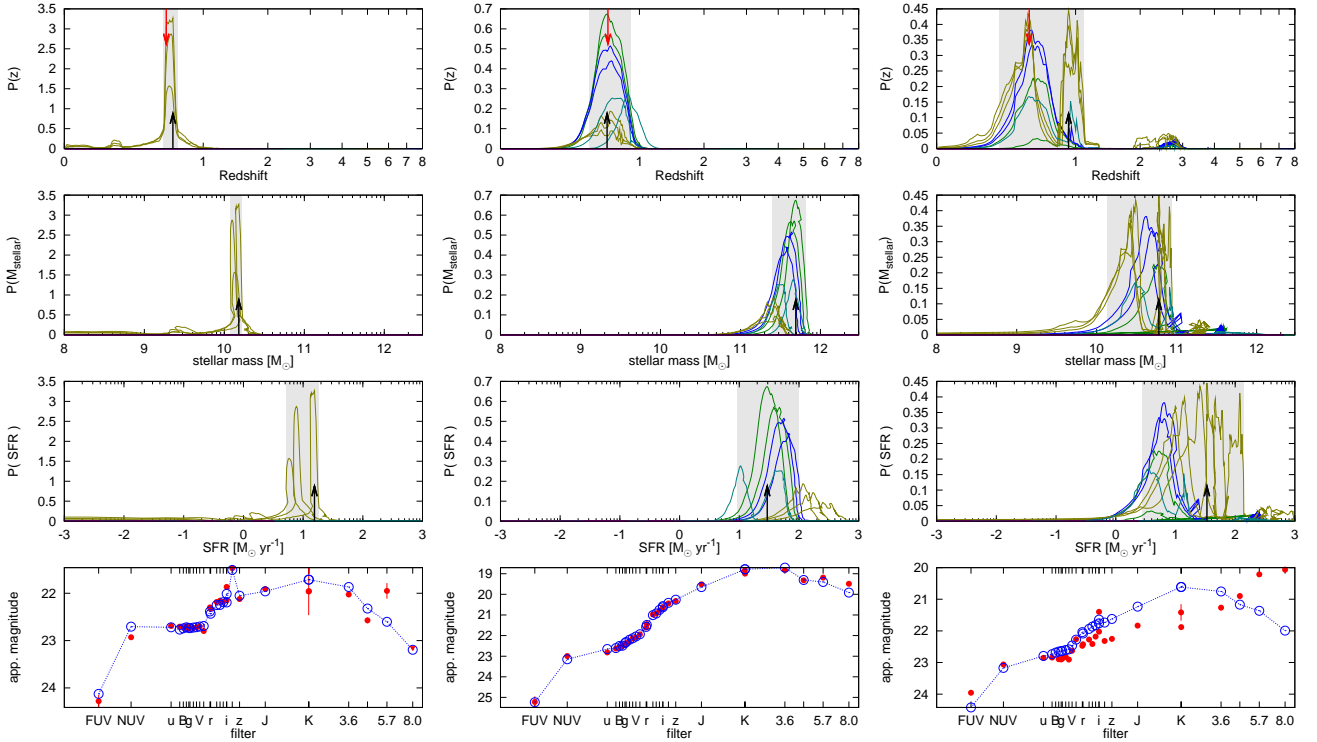
Figure 3. Examples for typical probability distribution functions for three galaxies in the COSMOS field. The top three panels show the probability distribution as function of redshift (top), stellar mass (second row) and star formation rate (third row); the bottom panel shows the observed (filled red circles) and corresponding best-fit SEDs (blue open circles, connected by lines). Black upward arrows mark the best-fit parameters (redshift, stellar mass and SFR), the downwards pointing red arrow in the top panel marks the spectroscopic redshift. The 1-σ confidence ranges are marked as gray-shaded regions.
The three galaxies are (ID [from zCOSMOS DR2 catalog], RA [deg], DEC [deg]): left: (811415, 150.242065, 1.826618); middle: (807424, 149.718628, 1.638003); right: (810153, 150.517227, 1.876145).

Fig. 3 also demonstrates three out of four possibilities of principal shapes of probability distribution functions (PDFs), two of which lead to an correct photo-z estimate. In each of the panels, line-colours indicate the galaxy template, with red lines representing our E-type model, green lines early-type spirals Sa and Sb, dark blue lines late-type spirals Sc and Sd, brown lines early burst-models with burst ages $\leqslant 500$ Myr, turquoise intermediate-age bursts with $500 <$ burst age < 1500 Myrs, and finally violett lines post-starburst models with burst ages > 1.5 Gyrs. In each sub-panel the upward facing arrows indicate the best-match value; the red, downward facing arrow in the top $P(z)$ panels marks the spectroscopic redshift. We furthermore indicate the 1σ confidence intervals with the grey-shaded regions. The bottom sub-panels give the observed photometry data and their errors with red points; the best-match model SED is given by blue open circles. Note that the lines connecting the model SED points are only meant to guide the eye.

The first case, shown in the left panel, illustrates the ideal case of having only a single template that can reproduce the observed SED. In essentially all cases this is caused by one or more strong features in the SED, either a strong break (as in the case of dropouts) or, as for this galaxy, strong emission lines that dominates the flux in at least one filter. These features can only be reproduced over a very narrow redshift range (typically $\Delta x \approx \Delta\lambda/\lambda$ of the filter that covers the feature) and hence lead to very peaked PDFs.

We remark that the SFR PDF splits up into three components that coincide in their redshift- and stellar mass PDFs. These represent three young bursts at different stages or burst ages, that despite their different physical parameters show indistinguishable observables in form of their SEDs.

The second case shown in the middle column of Fig. 3 is a more typical case where no sharp features are detected and the photo-z determination is dominated by a fit to the continuum. This also allows for dust extinction as additional degree of freedom to come into play. Dust reddening affects the spectral slope globally, but does not affect the shape and/or position of spectral lines or breaks, so we naturally expect and indeed observe a much wider PDF in these cases. The stellar mass PDF is not symmetric around the best-match value, but rather skewed due to the width of the photo-z distribution.

The third, right column in Fig. ?? is characteristic for a case where the PDF has more than one peak. Reasons for this can be found in degeneracies due to large photometric errors and/or insufficient sampling of the SED, e.g. too few filters not allowing to discriminate effects of stellar population age and dust reddening or confusing the Balmer and Lyman breaks due to lacking/too shallow UV coverage. In most of these cases the PDF still shows a local but no global probability maximum at the correct redshift. If both peaks are close in redshift space this effect is likely to average out for larger samples, as the scattering of best-match

values from the correct into the wrong peak is proportional to their relative amplitudes. For larger distances (e.g. as in the case of confused Balmer and Lyman-breaks) the effects are more severe, because nearby, low-mass and galaxies that get interpreted as high-redshift galaxies are more numerous than massive galaxies at high-redshift. However, in most of these cases the resulting photo-zs are assigned very large uncertainties to represent this degeneracy and can be excluded from any subsequent analysis.

The fourth case not shown here are "catastrophic failures", where the spectroscopic redshift does not correspond to even local maxima and lies outside the confidence ranges. This can be the case, e.g, for strong AGNs or typically late-type stars that are not represented by the template grid. In particular the latter can potentially be excluded if high-resolution imaging, e.g. from HST or using adaptive optics, are available, e.g. using the Stellarity parameter from SExtractor (Bertin & Arnouts 1996) or by comparing the FWHM to the FWHM of bona-fide points sources. However, doing so from ground-based observations alone with typical seeing of $\lesssim 1$ arcsec is less straightforward, explaining the significant fraction of objects falling into this category.

Another source of discordant photo-zs are wrong spectroscopic redshifts. ? compared spectroscopic and photometric redshifts for a sample of 140 galaxies in the HDF-N. They find for 5 out of 9 discordant redshifts the spectroscopic redshift is at fault, while only in 1/9 the photo-z is incorrect. This fraction is particularly troublesome for training-set based photo-z techniques such as ANNz (?).XXX

5   PERFORMANCE EVALUATION

To test the accuracy and reliability of our code we applied GAZELLE to a range of different surveys with available spectroscopic redshifts. The fields under investigation are the original Hubble Deep Fields North (HDF-N, Fernández-Soto et al. 1999) and South (HDF-S, Labbé et al. 2003), the Hubble Ultra Deep Field (HUDF, Coe et al. 2006), Chandra Deep Field South (CDF-S, using the FIREWORKS catalog by Wuyts et al. 2008) and the AEGIS (All-wavelength Extended Groth Strip International Survey, data from Bundy et al. 2006; Davis et al. 2007). The wavelength coverage of each of these fields is summarised in Table 1.

For each of these fields we first ran a calibration run, in which the redshift of all templates is fixed to the spectroscopic redshift of each object. This allows to derive empirical zero-point offsets that need to be applied to the data to minimise systematic offsets between the photometric and spectroscopic redshifts. Possible reasons for these zero-point shifts are imperfect calibration of the data with respect to the quoted magnitude system (as is the case, e.g., for the SDSS u- and z-filters, (Doi et al. 2010)). On the modeling side small changes to the filter response curves as consequence of different airmasses or insufficient knowledge of the response of telescope mirrors and/or detector as function of both time and wavelength can also lead to minor offsets. However, the required offsets are typically small ($\lesssim 0.05$ mag) and well within the assumed uncertainties of our models of 0.1 mag.

The resulting one-to-one comparison is shown in the left panel of Fig. 4 and also summarised in Table 1. We gener-
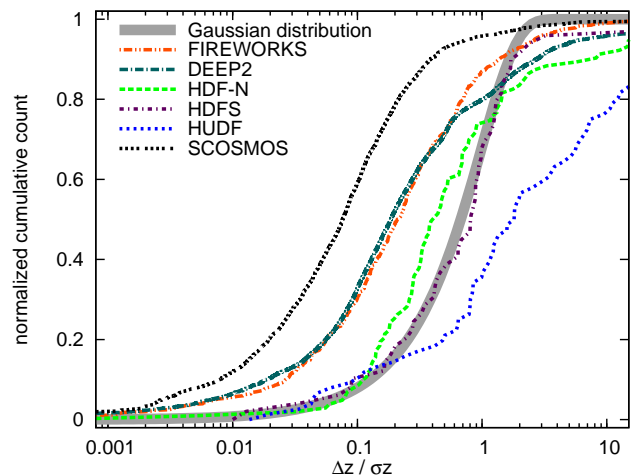

Figure 5. XXX

ally find a very good agreement between the spectroscopic and our derived photometric redshifts. Two notable cases are SCOSMOS and AEGIS/DEEP-2. SCOSMOS shows by far the best performance, which we attribute not only to the large number of filters that finely sample the SED in much the same way as a low-resolution spectra does, but also to its large wavelength coverage from FUV ($0.15\mu m$) to Mid-Infrared at $8.0\mu m$ (note, however, that in particular the $[5.8\mu m]$ and $[8.0\mu m]$ bands are mostly excluded from the fitting due to the low redshifts of the spectroscopic sample). This greatly helps to constraint dust extinctions, alleviating degeneracies of dust extinction, stellar population age and redshift. On the other end of the spectrum is our AEGIS/DEEP-2 sample, showing a relatively large number of objects with large derivations from the ideal case. We attribute this to the way the photometry sample was assembled, as described earlier in Sect. ??.

The right panel of Fig. 4 shows the distribution of photometric redshift offsets $\sigma_z = (z_{phot} - z_{spec})/(z_{spec} + 1)$ for all six fields. All fields show a minor bias of $\sigma_{z,median} \approx -0.02$ towards slightly underestimated photo-z as compared to the true spectroscopic redshift. This bias however does not significantly impact on the derived physical parameters, in particular at $z > 0.5$; small deviations in redshift do not change the overall evolutionary state of the galaxy, and the resulting error in the distance modulus ($\delta(m-M)/\delta z \leqslant XXX$) which affects masses and mass-dependent parameters is small compared to the uncertainties from different model types, i.e. star formation histories (also see Fig. 3). The dispersion of $\sigma_z \approx 0.05$ is comparable to other state-of-the-art photo-z codes on the same and/or similar datasets (Hildebrandt et al. 2010).

5.1   Distribution of photometric redshift errors

One important factor during the derivation of photo-z is to determine how robust that redshift estimate is, or what its confidence ranges are. As mentioned earlier (Sect. 2.3) GAZELLE derives confidence ranges directly from the probability distribution, accounting for potential multiple probability maxima.

In Fig. 5 we show the cumulative error distribution de-

Table 1. Surveys in comparison

Survey / Field	wavelength coverage	filters	N(z$_{\mathrm{spec}}$)	$\left\langle \frac{\Delta z}{1+z_{\mathrm{spec}}} \right\rangle$	σ_{z}	# outliers[1]
HUDF	B - J	6	86	-0.013	0.108	30
HDF-N	U - K	7	103	-0.027	0.056	17
HDF-S	U - K	7	49	-0.021	0.059	7
AEGIS/DEEP	u - K	7	5026	-0.037	0.067	1285
FIREWORKS	U - [8.0]	16	582	-0.057	0.062	117
SCOSMOS	FUV - [8.0]	16+12	2375	-0.003	0.017	148
combined	-	-	8221	-0.018	0.052	1604
Kitzbichler+06	B-K	5	112k	-0.013	0.054	14345

[1] defined as having $\Delta z \equiv |z_{\mathrm{spec}} - z_{\mathrm{phot}}| > 0.1 \times (1+z)$
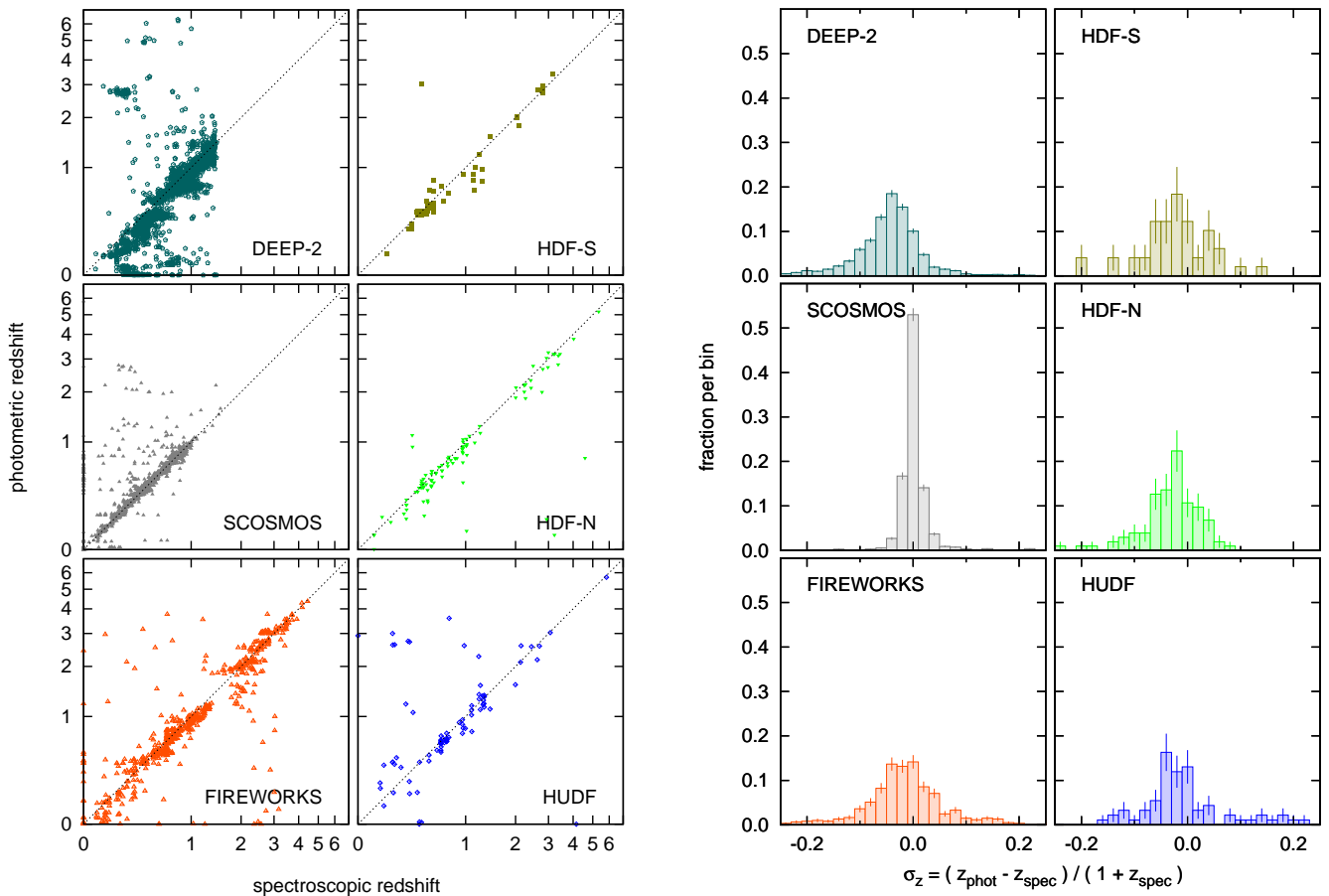
Figure 4. Comparison of spectroscopic and photometric redshifts. Direct comparison of z$_{\mathrm{phot}}$ and z$_{\mathrm{spec}}$ for the surveys listed in Table **??**. The mean deviation $\langle z_{spec} - z_{phot} \rangle = -0.03XXX$, the scatter is XXX. **Right panel:** Histogram of the deviation

rived from all galaxies with available spectroscopic redshifts. When compared to the theoretical expectations we find that for most fields the curves lie left of the gaussian distribution (shown as thicker grey line), meaning that for these we generally overestimate the uncertainties and photo-z are more accurate than could be expected from the error bars. This is most likely due to PDFs with multiple peaks and/or degeneracies intrinsic to the template set and parameter space. For instance galaxy age, redshift and dust reddening have similar impact on the observed SEDs. Only in the case of the HUDF are errors underestimated. However, this field also has the lowest number of spectroscopic redshifts and is selected differently from all the other fields, and so our result for this particular field has lower statistical significance.

We also note that at large deviations our estimated errors drop below the gaussian distribution. This originates from the fraction of "catastrophic failures" or outliers from the photo-z vs. spec-z identity. For these cases the quoted error bars are often significantly underestimated.

5.2 Systematic uncertainties in stellar mass and star formation rate due to simplified SFHs

As GAZELLE simultaneously serves as SED fitting code to derive physical parameters, we need to make sure that our assumption of smoothly varying star formation histories is capable to deal with the often complex star formation histories observed in detailed studies of individual galaxies. For this purpose we chose a large sample of galaxies from semi-analytical models. These galaxies have more realistic SFHs derived from the Millennium simulation (Springel et al. 2005), but, owing to their model nature, well known parameters, making them advantageous over real galaxies for this comparison. We use a sample of galaxies from the De Lucia & Blaizot (2007) and Bower et al. (2006) models. For all cases we correct the stellar masses and star formation rates for differences in the IMFs, following the prescription in the respective papers. In all cases we use the simulated photometry and apply the photo-z and SED fitting using identical parameters as in the real analysis of the deepfields.

The comparison with the models from De Lucia & Blaizot (2007) are shown in Fig. 6. The redshift for the four runs, i.e. $z = (0.36, 0.76, 1.50, 3.06)$, have been chosen to as closely as possible match the central redshifts for the redshift bins of the subsequent analysis, to not introduce uncertainties. We chose the photometry in the SDSS ugriz and 2MASS JHK filters, as these filters more closely represent current observations as compared to the Buser BV and Johnson RIK filters.

We find that, for galaxies more massive than $10^9 M_\odot$, stellar masses can be reproduced with a typical 1σ scatter of 0.07 dex; 99.9 per cent of all galaxies have stellar masses within $0.17 \ldots 0.24$ dex (with lower values towards higher redshifts) of the true model mass. Star formation rates for the $M_{stellar} > 10^9 M_\odot$ selected sample are typically good to 0.5 dex (1σ) but show logarithmic large scatters, mostly due to very low SFR that have very little effect on the SED. Restricting the sample to galaxies with SFR $\geqslant 1$ $M_\odot$ yr^{-1} significantly improves this situation, with scatter being reduced to 0.18 dex (1σ), and > 0.999 of all galaxies agreeing to better then 0.5 dex.

The results using the models of Bower et al. (2006) and our fits to the UBRIJHK photometry are shown in Fig. 7. These models do not give star formation rates, so we have to restrict the comparison to stellar masses. The scatter here amounts to $0.08 \ldots 0.12$ dex. In all cases essentially all galaxies have photometric stellar masses within 0.25 dex of the true value.

While these results are already more than encouraging, so far we could only fit rest-frame photometry for which we had to restrict the redshift range to match the models. This prevented us from determining the true accuracy with redshift as additional free parameter. To overcome this we also applied GAZELLE to synthetic light-cones from Kitzbichler & White (2007). These light-cones are also based on the Millennium simulation, but account for redshift by using appropriately shifted BVRIK filters. The results derived from these relatively sparse SEDs are shown in Fig. 8.

We generally find very good agreement of the true values and those derived from their SEDs.
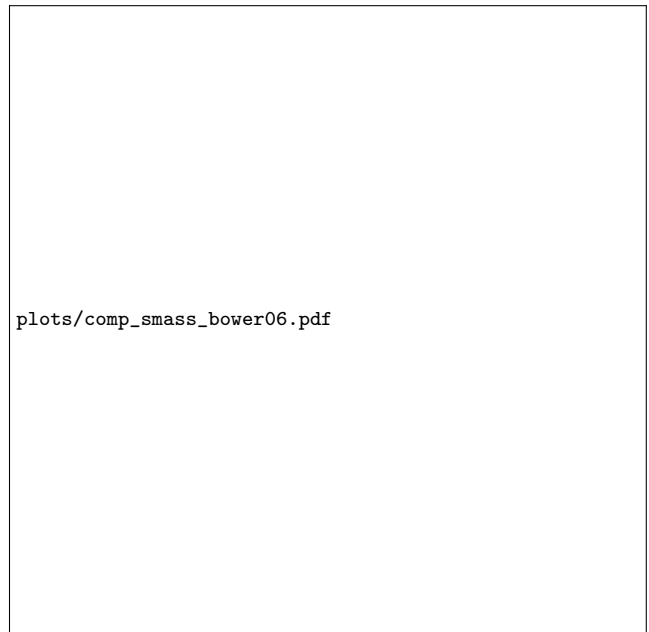
Figure 7. Comparison of stellar masses taken from the semi-analytical galaxy formation model of Bower et al. (2006) with the parameters obtained by fitting the synthetic rest-frame SEDs to a set of GALEV models using GAZELLE. Other details as in XXX Fig. 6.

Figure 8. Comparison of values obtained from the obtained photometric redshifts with the true values as given by the semi-analytical models for a 2 sq.deg. light-cone from Kitzbichler & White (2007). Top left: true and photometric redshifts; bottom left: stellar masses; top right: star formation rates; bottom right: specific star formation rates ($=$ SFR/$M_{stellar}$). The solid black line shows the identity between true and photometric values; dotted lines mark confidence ranges of $\pm 0.1(1 + z_{spec})$ and 0.3 dex ($M_{stellar}$, SFR, and sSFR).
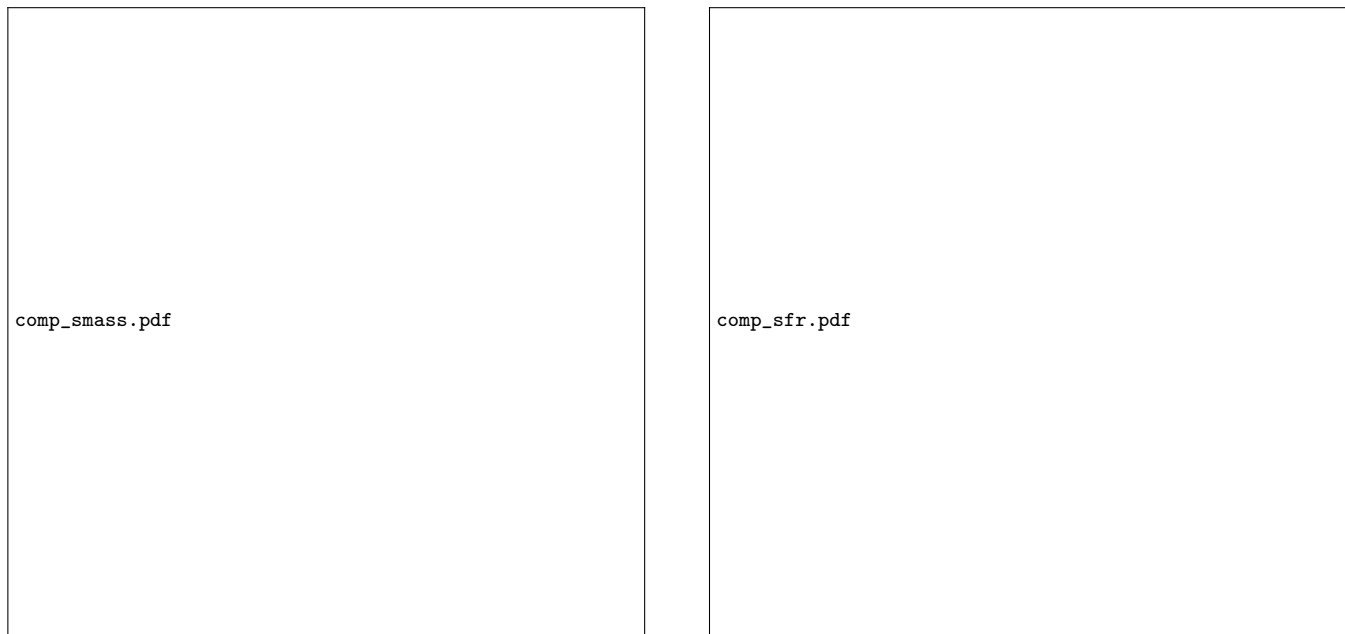
Figure 6. Comparison of stellar masses (left panels) and star formation rates (right panels) taken from a semi-analytical galaxy formation model (De Lucia & Blaizot 2007) based on the Millennium simulation with the parameters obtained by fitting the synthetic rest-frame SEDs to a set of GALEV models using GAZELLE. All data were corrected to use the same cosmology and a Salpeter IMF. Data points are shown with blue dots, perfect one-to-one agreement is shown by the solid line, and dashed lines mark 0.3 dex offsets to either side.

6   RESULTS

6.1   Redshift distribution

The first step in our comparison of the individual fields is the distribution of the derived photometric redshifts. To make the samples comparable for all subsequent analyses, we select a K-band bright sample of $m_K < 22$ mag galaxies. This brightness is reached in all samples except the HUDF that only extends to the H-band. Furthermore the HUDF catalog combines galaxies in any of the BViz frames (Coe et al. 2006) and hence should be considered carefully. Also note that the COSMOS field is actually based on a combined i-band and K-band selection. However, the i-band selection reaches significantly deeper than the K-band (26.1 vs 23.8 mag at 10σ, Capak et al. (2008, in preparation), also see Ilbert et al. 2009), which in turn is nearly 2 mag deeper than our conservative limit and hence will not likely affect the following results except in the case of extremely red $(i-K > 4)$ object.

Fig. 9 shows the resulting redshift distribution for all fields, normalised by the area of each field (top panel) and the total number of galaxies of each respective field (bottom panel).

As can be seen from this figure, the photo-z distributions of most fields roughly agree in that they show a rise at low redshifts to a maximum at $z \approx 0.5$, and a sharp decline towards higher redshifts. This agrees well with results from other groups for the same fields (Grazian et al. 2006; XXX missing reference XXX 2035), re-emphasising that both the code and our underlying model grid compare well with other solutions currently available, but with the added benefit of physical parameters from the intimate coupling to our models. A striking feature of both number- and area-normalised plots is the large scatter between the fields.

Despite our restriction to a homogeneous, $m_K < 22$ mag selected sample we still find differences of a factor of three near the peak of the distribution at $z \approx 0.5 - 0.8$. This discrepancy between the individual grows larger still with increasing redshift, but with reduced statistical significance due to the smaller number of detected galaxies at these redshifts. Possible reasons for this effect are **1)** differences in the selection of the sample; **2)** different reduction steps from the data data to the actual catalogs, including calibration uncertainties; **3)** photometric redshift errors that scatter galaxies between the individual redshift bins; **4)** real physical differences in the high-redshift galaxy populations at higher redshifts;

The first suspect was largely ruled out by our restriction to a well defined sample as outlined above. In particular the chosen magnitude limit is significantly brighter than the quoted incompleteness limits of the individual fields (see, e.g., Quadri et al. 2007 for the MUSYC fields). Furthermore the reduction steps that lead to these catalogs largely, i.e. not considering certain steps necessary to account for the different instrument characteristics (XXX missing reference XXX 2035), follow widely-accepted standard reduction procedures, so that also the second point seems unlikely to cause the large observed discrepancies.

The third point of photometric redshift uncertainties shuffling galaxies among the bins can be addressed by comparing two nearly identical fields. For this purpose we chose the MUSYC CW1255, SDSS1030, and EHDFS fields, as they were both published by Quadri et al. (2007), are based on data from the same instrument, and have photometry of equal quality in identical filters, and hence should be affected by photo-z scatter in very much the same way. However, even among these fields we still find significant $(3-5\sigma)$ differences at redshifts $z = 0.8 \dots 2.5$. At higher redshifts $z \leqslant 2.5$ the
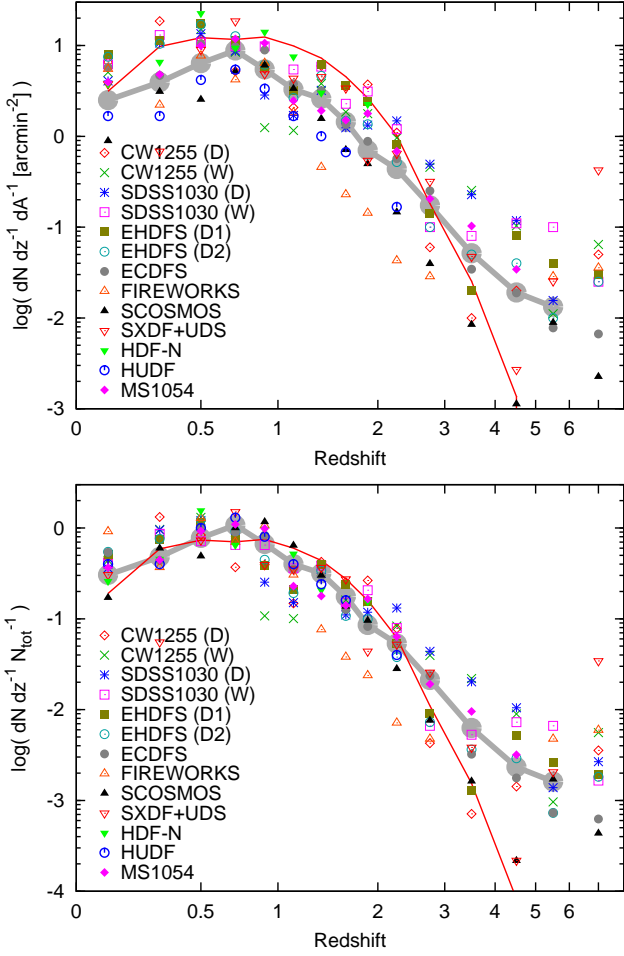
Figure 9. Projected density of galaxies per unit redshift, normalised by the area of each field in arcmin² (top panel) and by total number of galaxies in each field (bottom panel) as a function of redshift. The thick grey line marks the average of all fields (using the wide fields where both wide and deep are available); the thin, solid red line shows the redshift distribution derived from light-cones through the Millennium-simulation (Kitzbichler & White 2007).

redshift distributions from the wide catalogs agree very well, while the deep catalogs still show some differences, but due to the small number of galaxies the statistical significance of this finding is questionable.

We will discuss the evidence for the fourth point below in the light of mass functions derived from the data for each of the fields.

Another interesting feature in several fields (CW1255 D/W, SDSS1030 W) is a deviation at $z \approx 1.5 \dots 2$ in the decrease rate towards higher redshifts. This can be more clearly seen in Fig. 10 where we show the area-normalised redshift distribution of each individual field.

To test whether this is indeed a real feature or an artefact of our photo-z determination we created a large mock-catalog of simulated observations for each of our input templates. For each we started from our model SED, added dust extinction with $E(B - V) = 0 - 0.4$ mag drawn from a flat distribution, varied the mass function to represent a Schechter-function with faint-end slope $\alpha = -1.0$ and a
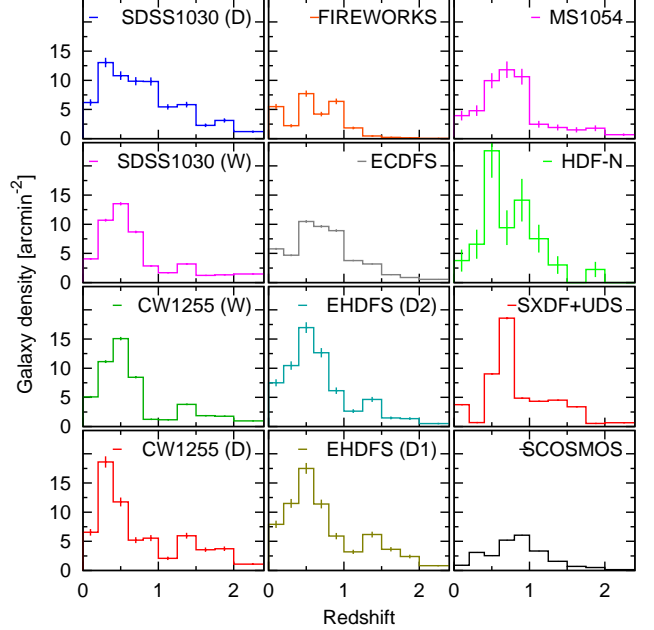

Figure 10. Redshift distribution for each individual sample, normalised by the are of each field. Poisson errors have been plotted for each individual field, but are too small in the case of SCOSMOS, ECDFS and SXDF+UDS to be recognisable.

typical mass of $3 \times 10^{11} M_{\odot}$. Finally we added photometric noise, drawn from a gaussian distribution with zero mean and a standard deviation that depends on the filter and the apparent magnitude (see Kotulla & Fritze 2009, for a more detailed description). We used the same filter set and detection limits as in the MUSYC-ECDFS field. These catalogs were then used as input catalog for GAZELLE and we determined photometric redshifts and physical parameters with parameters identical to what we used for the real data. The very last step then culled only these galaxies that fulfil the $m_K < 22$ mag criterion to make the simulation comparable to above findings. In Fig. **??** we show the comparison between true, input redshift distribution and the retrieved, photo-z distribution.

6.2 Stellar mass functions

In a next step we compare the galactic stellar mass functions for each of the fields to search for mass-dependent differences between the individual fields. This is done in four redshift intervals $(0.2, 0.5)$, $(0.5, 1.0)$, $(1.0, 2.0)$, and $(2.0, 4.0)$, each covering a timespan of ≈ 2 Gyrs in our chosen cosmology, and avoiding both the redshift extremes at very low and very high redshifts were even small redshift uncertainties have significant effect, or source densities are low, respectively.

Stellar masses for each of the galaxies is taken to be the best-match value. To account for incompleteness at low stellar masses we used the V/V_{max} technique (Schmidt 1968) and computed, for each galaxy, its space density in the volume over which the galaxy fulfils the detection- and selection criteria for our sample. To derive this maximum volume we derived the maximum redshift out to which the best-match galaxy model would be observable, hence fully and consistently accounting not only for the changing luminosity dis-

tance, but also for k- and e-correction, i.e. changes to the observed magnitude due to a shift in filter position relative to the rest-frame spectrum and due to the younger galaxy age towards higher redshifts. We then fitted each mass function with a Schechter profile (Schechter 1976). The results are shown in Fig. 11.

For each field we show the derived mass-function as data points with error bars (assuming Poisson noise) and our fitted Schechter profile as line. The two different gray-scale regions indicate two extreme mass completeness limits for an old, passively evolving galaxy at the upper end of the redshift range in light grey and a starbursting galaxy at the lower end of the redshift range in darker grey. For both models we do not assume any dust extinction; for the old galaxy with little to no ongoing SF we do not expect significant amounts of dust, in agreement with spectroscopic observations (Kriek et al. 2006, 2009). For the starburst on the other hand the case without dust represent the extremely blue case, as in a more general case the addition of dust would shift the mass completeness limit to higher masses.

Comparing the mass-functions we generally find good agreement between the individual fields for masses $M \lesssim 5 \times 10^{11} \mathrm{M}_\odot$.

Larger scatter at higher masses: field/cluster variations?

what's wrong with fireworks?

plausible explanation for increase of mstar towards z=1?

$z > 1$ outliers at high masses are photo-scattered lower-mass low-z galaxies

6.3 Evolution of number density, characteristic mass and low-mass slope with redshift

6.4 Comparison to semi-analytical galaxy formation models

6.5 A combined MF across all fields

Somehow ???? combine all fields into one MF that covers all galaxies in all fields, then derive MF, fit parameters and hope the resulting values agree with the average of the individual.

with present and explain, show how well models at given redshifts are recovered, what range of uncertainties is.

7 DISCUSSION

GALEV models are for M_B^* galaxies of every type. Use observed L-Z relations for SFing and passive galaxies to estimate effects for super/sub-luminous galaxies: how much higher/lower is their metallicity ? L-Z relation available for local galaxies and for z=1 (Tremonti+)

Estimate from models changes expected for L-Z and TF (and FJ) relations: luminosity changes/ metallicity changes, compare to models with $Z_\odot$ only.
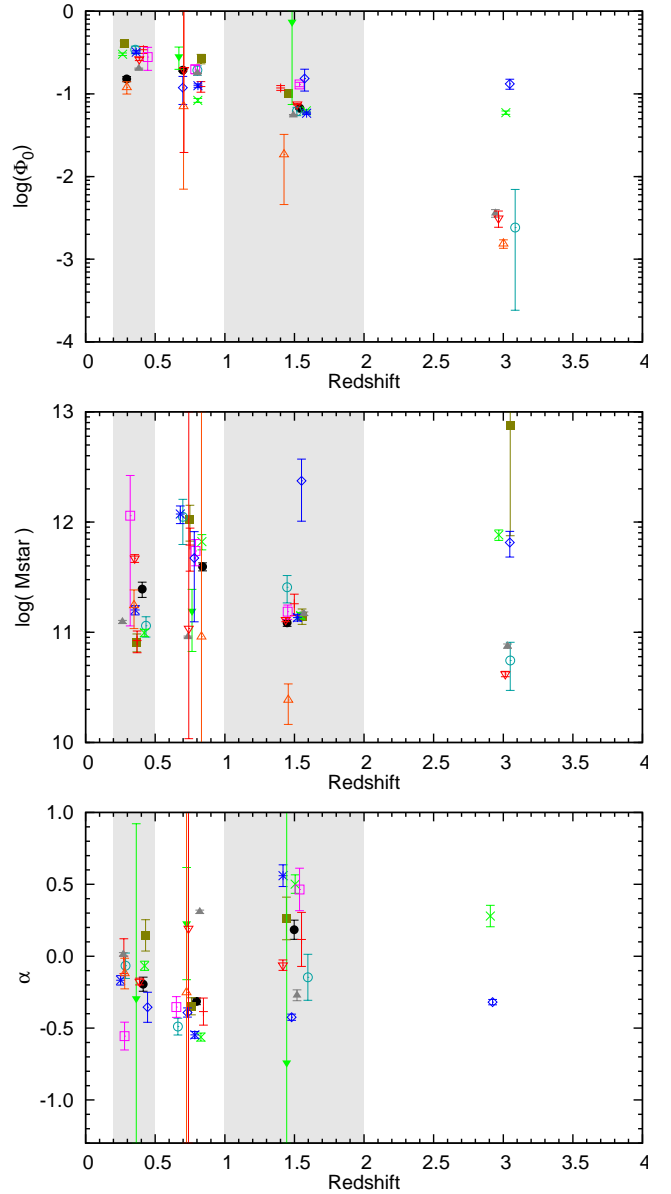
Extrapolate to $z = 2.5$????

Figure 12. Evolution of the number density parameter Φ_0 (top panel), characteristic mass $M_\star$ (middle panel) and low-mass slope α (bottom panel) with redshift. The mass functions were derived in four bins covering the range (0.2-0.5), (0.5-1.0), (1.0-2.0), and (2.0-4.0); these ranges are outlined by the gray-shaded regions. A small random offset in x-direction has been added to make individual data points more easily recognisable. Symbol definition are identical with the previous plots.
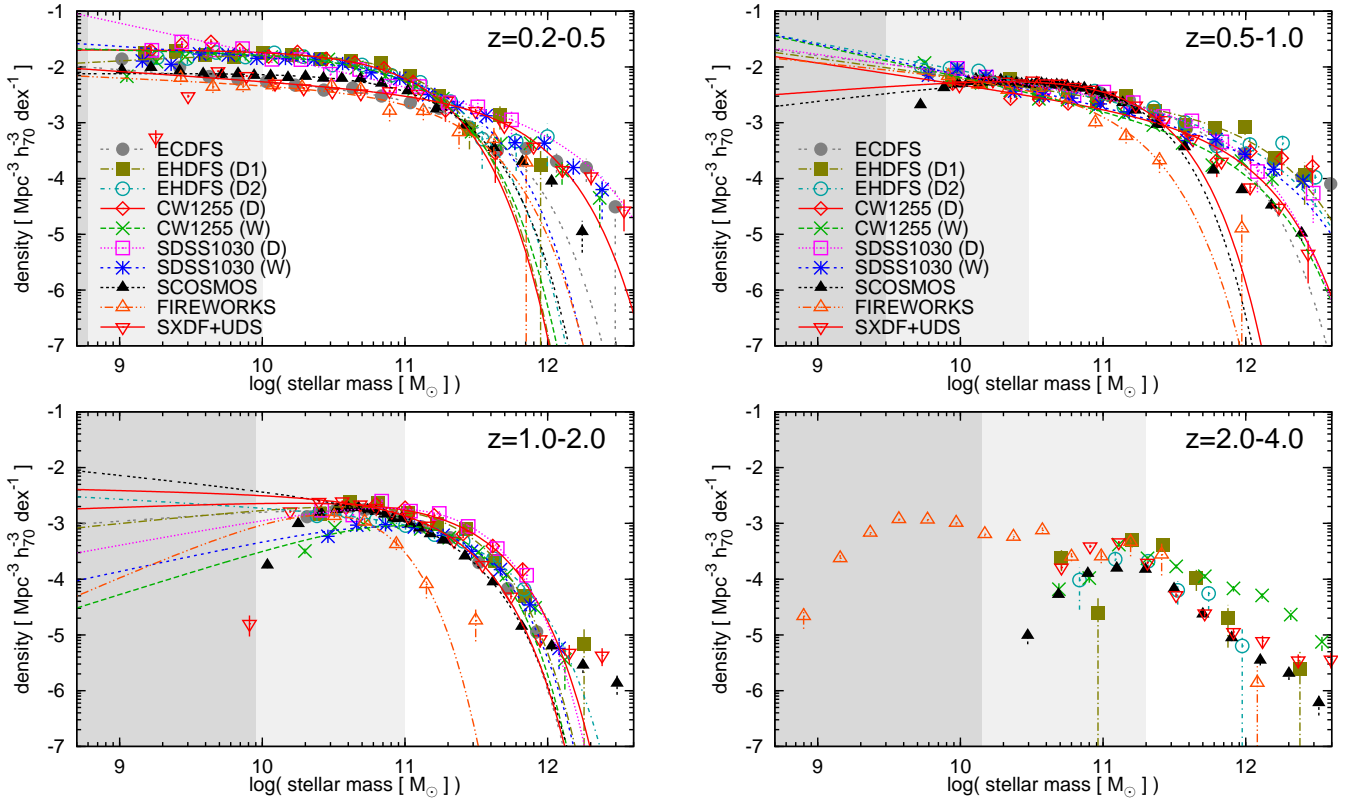
8 CONCLUSIONS

ACKNOWLEDGMENTS

Figure 11. Mass functions for each of the fields in four different redshift bins (0.2, 0.5), (0.5, 1.0), (1.0, 2.0), and (2.0, 4.0). Points represent the actual data, while lines represent the best-fit Schechter functions. The grey-shaded regions mark the regions where incompleteness affects a) passive and predominantly old galaxies at the upper end of the redshift bin (light grey) and b) actively star-forming galaxies at the lower end of the redshift range (darker grey), both assuming no dust extinctions. Symbols in the lower panels are identical with those in the upper panels.

partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The DEEP2 team and Keck Observatory acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community and appreciate the opportunity to conduct observations from this mountain.

Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.

REFERENCES

Anders, P., Bissantz, N., Fritze, U., & de Grijs, R. 2004, MNRAS, 347, 196

Anders, P. & Fritze, U. 2003, A&A, 401, 1063

Benítez, N. 2000, ApJ, 536, 571

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

Bicker, J., Fritze, U., Möller, C. S., & Fricke, K. J. 2004, A&A, 413, 37

Blanc, G. A., Lira, P., Barrientos, L. F., Aguirre, P., Francke, H., Taylor, E. N., Quadri, R., Marchesini, D., et al. 2008, astro-ph/0803.0763, 803

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476

Bondi, M., Ciliegi, P., Schinnerer, E., Smolčić, V., Jahnke, K., Carilli, C., Zamorani, G. 2008, ApJ, 681, 1129

Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S., Lacey, C. G. 2006, MNRAS, 370, 645

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503

Brusa, M., Zamorani, G., Comastri, A., Hasinger, G., Cappelluti, N., Civano, F., Finoguenov, A., Mainieri, V., et al. 2007, ApJS, 172, 353

Bundy, K., Ellis, R. S., Conselice, C. J., Taylor, J. E., Cooper, M. C., Willmer, C. N. A., Weiner, B. J., Coil, A. L., et al. 2006, ApJ, 651, 120

Capak, P., Aussel, H., Ajiki, M., McCracken, H. J., Mobasher, B., Scoville, N., Shopbell, P., Taniguchi, Y., et al. 2007, ApJS, 172, 99

Cappelluti, N., Hasinger, G., Brusa, M., Comastri, A., Zamorani, G., Böhringer, H., Brunner, H., Civano, F., et al. 2007, ApJS, 172, 341
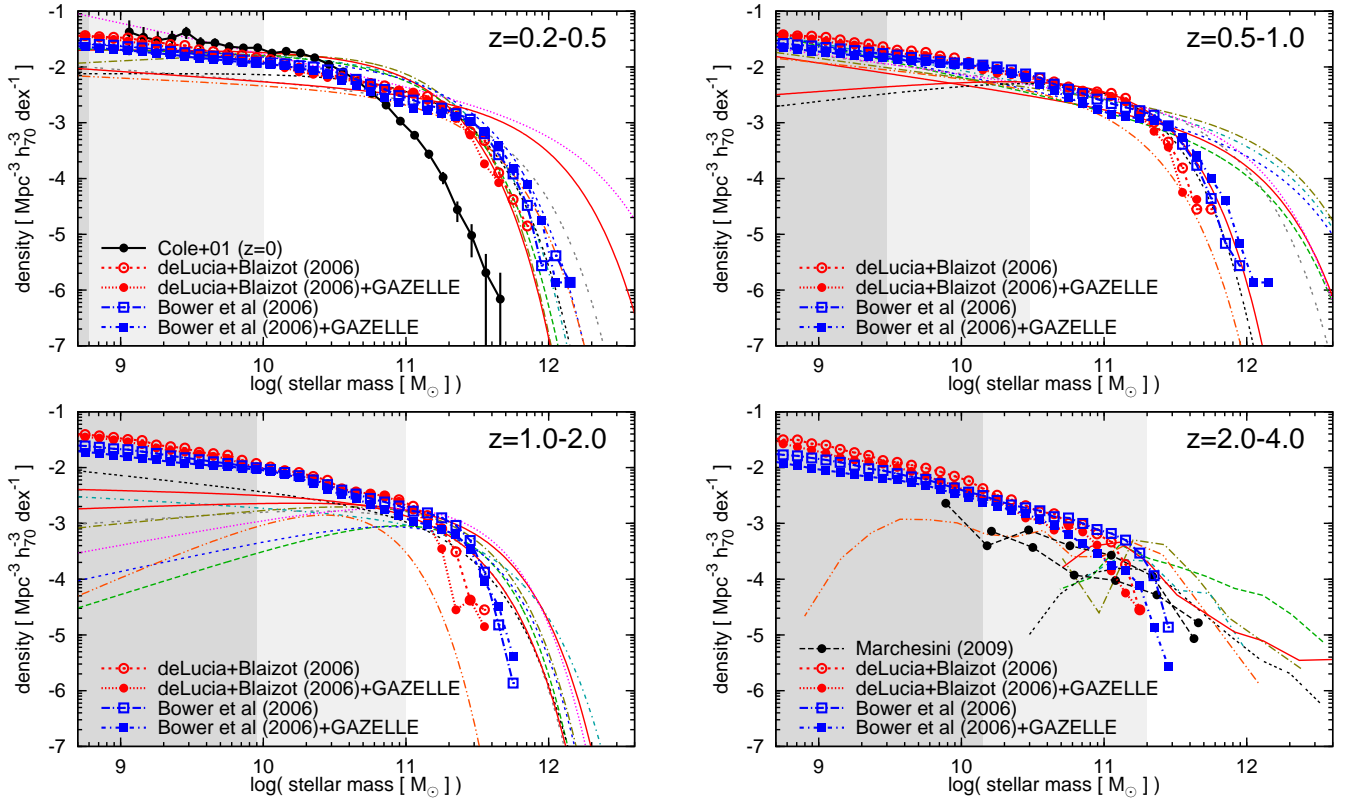
Figure 13. Mass functions for each of the fields in four different redshift ...

Coe, D., Benítez, N., Sánchez, S. F., Jee, M., Bouwens, R., Ford, H. 2006, AJ, 132, 926

Conroy, C., Gunn, J. E., & White, M. 2008, ArXiv e-prints

Conroy, C., White, M., & Gunn, J. E. 2009, ArXiv e-prints

Davis, M., Guhathakurta, P., Konidaris, N. P., Newman, J. A., Ashby, M. L. N., Biggs, A. D., Barmby, P., Bundy, K., et al. 2007, ApJ, 660, L1

De Lucia, G. & Blaizot, J. 2007, MNRAS, 375, 2

Doi, M., Tanaka, M., Fukugita, M., Gunn, J. E., Yasuda, N., Ivezić, Ž., Brinkmann, J., de Haars, E., et al. 2010, AJ, 139, 1628

Dressler, A., Smail, I., Poggianti, B. M., Butcher, H., Couch, W. J., Ellis, R. S., Oemler, A. J. 1999, ApJS, 122, 51

Falkenberg, M. A., Kotulla, R., & Fritze, U. 2009a, MNRAS, 397, 1940

Falkenberg, M. A., Kotulla, R., & Fritze, U. 2009b, MNRAS, 397, 1954

Feldmann, R., Carollo, C. M., Porciani, C., Lilly, S. J., Capak, P., Taniguchi, Y., Le Fèvre, O., Renzini, A., et al. 2006, MNRAS, 372, 565

Fernández-Soto, A., Lanzetta, K. M., & Yahil, A. 1999, ApJ, 513, 34

Finoguenov, A., Guzzo, L., Hasinger, G., Scoville, N. Z., Aussel, H., Böhringer, H., Brusa, M., Capak, P., et al. 2007, ApJS, 172, 182

Förster-Schreiber, N. M., Franx, M., Labbé, I., Rudnick, G., van Dokkum, P. G., Illingworth, G. D., Kuijken, K., Moorwood, A. F. M., et al. 2006, AJ, 131, 1891

Franx, M., Moorwood, A., Rix, H., Kuijken, K., Röttgering, H., van der Werf, P., van Dokkum, P., Labbe, I., et al.

2000, The Messenger, 99, 20

Giacconi, R., Rosati, P., Tozzi, P., Borgani, S., Hasinger, G., Bergeron, J., Gilmozzi, R., Nonino, M., et al. 2000, in Bulletin of the American Astronomical Society, Vol. 32, Bulletin of the American Astronomical Society, 1562–+

Grazian, A., Fontana, A., de Santis, C., Nonino, M., Salimbeni, S., Giallongo, E., Cristiani, S., Gallozzi, S., et al. 2006, A&A, 449, 951

Hasinger, G., Cappelluti, N., Brunner, H., Brusa, M., Comastri, A., Elvis, M., Finoguenov, A., Fiore, F., et al. 2007, ApJS, 172, 29

Hildebrandt, H., Arnouts, S., Capak, P., Moustakas, L., Wolf, C., Abdalla, F., Assef, R., Banerji, M., et al. 2010, A&A, submitted

Ilbert, O., Capak, P., Salvato, M., Aussel, H., McCracken, H. J., Sanders, D. B., Scoville, N., Kartaltepe, J., et al. 2009, ApJ, 690, 1236

Kashikawa, N., Shimasaku, K., Yasuda, N., Ajiki, M., Akiyama, M., Ando, H., Aoki, K., Doi, M., et al. 2004, PASJ, 56, 1011

Kitzbichler, M. G. & White, S. D. M. 2007, MNRAS, 376, 2

Koekemoer, A. M., Aussel, H., Calzetti, D., Capak, P., Giavalisco, M., Kneib, J.-P., Leauthaud, A., Le Fèvre, O., et al. 2007, ApJS, 172, 196

Kotulla, R. & Fritze, U. 2009, MNRAS, 393, L55

Kotulla, R., Fritze, U., Weilbacher, P., & Anders, P. 2009, MNRAS, 396, 462

Kriek, M., van Dokkum, P. G., Franx, M., Förster Schreiber, N. M., Gawiser, E., Illingworth, G. D.,

Labbé, I., Marchesini, D., et al. 2006, ApJ, 645, 44

Kriek, M., van Dokkum, P. G., Labbé, I., Franx, M., Illingworth, G. D., Marchesini, D., Quadri, R. F. 2009, ApJ, 700, 221

Labbé, I., Franx, M., Rudnick, G., Schreiber, N. M. F., Rix, H.-W., Moorwood, A., van Dokkum, P. G., van der Werf, P., et al. 2003, AJ, 125, 1107

Lawrence, A., Warren, S. J., Almaini, O., Edge, A. C., Hambly, N. C., Jameson, R. F., Lucas, P., Casali, M., et al. 2007, MNRAS, 379, 1599

Le Borgne, D. & Rocca-Volmerange, B. 2002, A&A, 386, 446

Lejeune, T., Cuisinier, F., & Buser, R. 1997, A&AS, 125, 229

Lejeune, T., Cuisinier, F., & Buser, R. 1998, A&AS, 130, 65

Lilly, S. J., Le Fèvre, O., Renzini, A., Zamorani, G., Scodeggio, M., Contini, T., Carollo, C. M., Hasinger, G., et al. 2007, ApJS, 172, 70

Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., Surace, J., Shupe, D., Xu, C., Oliver, S., Padgett, D., et al. 2003, PASP, 115, 897

Quadri, R., Marchesini, D., van Dokkum, P., Gawiser, E., Franx, M., Lira, P., Rudnick, G., Urry, C. M., et al. 2007, AJ, 134, 1103

Rudnick, G., Franx, M., Rix, H.-W., Moorwood, A., Kuijken, K., van Starkenburg, L., van der Werf, P., Röttgering, H., et al. 2001, AJ, 122, 2205

Schechter, P. 1976, ApJ, 203, 297

Schinnerer, E., Carilli, C. L., Scoville, N. Z., Bondi, M., Ciliegi, P., Vettolani, P., Le Fèvre, O., Koekemoer, A. M., et al. 2004, AJ, 128, 1974

Schinnerer, E., Smolčić, V., Carilli, C. L., Bondi, M., Ciliegi, P., Jahnke, K., Scoville, N. Z., Aussel, H., et al. 2007, ApJS, 172, 46

Schmidt, M. 1968, ApJ, 151, 393

Scoville, N., Abraham, R. G., Aussel, H., Barnes, J. E., Benson, A., Blain, A. W., Calzetti, D., Comastri, A., et al. 2007a, ApJS, 172, 38

Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C. M., Elvis, M., Giavalisco, M., Guzzo, L., et al. 2007b, ApJS, 172, 1

Sekiguchi, K. & et al. 2004, in Astrophysics and Space Science Library, Vol. 301, Astrophysics and Space Science Library, ed. M. Plionis, 169–+

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., et al. 2005, Nature, 435, 629

Taniguchi, Y., Scoville, N., Murayama, T., Sanders, D. B., Mobasher, B., Aussel, H., Capak, P., Ajiki, M., et al. 2007, ApJS, 172, 9

Trager, S. C., Worthey, G., Faber, S. M., Burstein, D., & Gonzalez, J. J. 1998, ApJS, 116, 1

Williams, R. E., Blacker, B., Dickinson, M., Dixon, W. V. D., Ferguson, H. C., Fruchter, A. S., Giavalisco, M., Gilliland, R. L., et al. 1996, AJ, 112, 1335

Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, ApJ, 691, 1879

Willmer, C. N. A., Faber, S. M., Koo, D. C., Weiner, B. J., Newman, J. A., Coil, A. L., Connolly, A. J., Conroy, C., et al. 2006, ApJ, 647, 853

Wuyts, S., Labbé, I., Schreiber, N. M. F., Franx, M., Rudnick, G., Brammer, G. B., van Dokkum, P. G. 2008, ApJ, 682, 985

XXX missing reference XXX. 2035, A&AS, 123, 305

Zamojski, M. A., Schiminovich, D., Rich, R. M., Mobasher, B., Koekemoer, A. M., Capak, P., Taniguchi, Y., Sasaki, S. S., et al. 2007, ApJS, 172, 468