Physical properties of galaxies and their evolution in the VIMOS VLT Deep Survey. I. The evolution of the mass-metallicity relation up to z 0.9

F. Lamareille, J. Brinchmann, T. Contini, C.J. Walcher, S. Charlot, E. Pérez-Montero, G. Zamorani, L. Pozzetti, M. Bolzonella, B. Garilli, et al.

Astronomy & Astrophysics manuscript no. mzpaper.hyper22868
13th November 2008

Physical properties of galaxies and their evolution in the VIMOS VLT Deep Survey⋆

I. The evolution of the mass-metallicity relation up to $z \sim 0.9$

F. Lamareille[1,3], J. Brinchmann[19], T. Contini[1], C.J. Walcher[7], S. Charlot[10,8], E. Pérez-Montero[1], G. Zamorani[3], L. Pozzetti[3], M. Bolzonella[3], B. Garilli[2], S. Paltani[15,16], A. Bongiorno[22], O. Le Fèvre[7], D. Bottini[2], V. Le Brun[7], D. Maccagni[2], R. Scaramella[4,13], M. Scodeggio[2], L. Tresse[7], G. Vettolani[4], A. Zanichelli[4], C. Adami, S. Arnouts[23,7], S. Bardelli[3], A. Cappi[3], P. Ciliegi[3], S. Foucaud[21], P. Franzetti[2], I. Gavignaud[12], L. Guzzo[9], O. Ilbert[20], A. Iovino[9], H.J. McCracken[10,11], B. Marano[6], C. Marinoni[18], A. Mazure[7], B. Meneux[22,24], R. Merighi[3], R. Pellò[1], A. Pollo[7,17], M. Radovich[5], D. Vergani[2], E. Zucca[3], A. Romano[5], A. Grado[5], and L. Limatola[5]

(Affiliations can be found after the references)

Preprint online version: 13th November 2008

Abstract

Aims. We want to derive the mass-metallicity relation of star-forming galaxies up to $z \sim 0.9$, using data from the VIMOS VLT Deep Survey. The mass-metallicity relation is commonly understood as the relation between the stellar mass and the gas-phase oxygen abundance.

Methods. Automatic measurement of emission-line fluxes and equivalent widths have been performed on the full spectroscopic sample of the VIMOS VLT Deep Survey. This sample is divided into two sub-samples depending on the apparent magnitude selection: wide ($I_{\rm AB} < 22.5$) and deep ($I_{\rm AB} < 24$). These two samples span two different ranges of stellar masses. Emission-line galaxies have been separated into star-forming galaxies and active galactic nuclei using emission line ratios. For the star-forming galaxies the emission line ratios have also been used to estimate gas-phase oxygen abundance, using empirical calibrations renormalized in order to give consistent results at low and high redshifts. The stellar masses have been estimated by fitting the whole spectral energy distributions with a set of stellar population synthesis models.

Results. We assume at first order that the shape of the mass-metallicity relation remains constant with redshift. Then we find a stronger metallicity evolution in the wide sample as compared to the deep sample. We thus conclude that the mass-metallicity relation is flatter at higher redshift. At $z \sim 0.77$, galaxies at $10^{9.4}$ solar masses have -0.18 dex lower metallicities than galaxies of similar masses in the local universe, while galaxies at $10^{10.2}$ solar masses have -0.28 dex lower metallicities. By comparing the mass-metallicity and luminosity-metallicity relations, we also find an evolution in mass-to-light ratios: galaxies at higher redshifts being more active. The observed flattening of the mass-metallicity relation at high redshift is analyzed as an evidence in favor of the open-closed model.

Key words. galaxies: evolution – galaxies: fundamental parameters – galaxies: abundances – galaxies: starburst

1. Introduction

The stellar mass and the gas-phase metallicity of a galaxy are two of the main parameters involved in the study of galaxy formation and evolution. As cosmological time progresses, theory predicts that both the mean metallicity and stellar mass of galaxies increase with age as galaxies undergo chemical enrichment and grow through merging processes. At any given epoch, the accumulated history of star formation, gas inflows and outflows, affects a galaxy mass and its metallicity. Hence one expects theses quantities to show some correlation and this will provide crucial information about the physical processes that govern galaxy formation.

First discovered for irregular galaxies (Lequeux et al. 1979), the mass-metallicity relation has been intensively studied (Skillman et al. 1989; Brodie & Huchra 1991; Zaritsky et al. 1994; Richer & McCall 1995; Garnett et al. 1997; Pilyugin & Ferrini 2000, among others) and is now well established in the local universe by the work of Tremonti et al. (2004) with SDSS data and Lamareille et al. (2004) with 2dFGRS data, the latter done on the luminosity-metallicity relation which is easier to derive when small number of photometric bands are available. These two studies have shown in two different ways that the mass-metallicity relation is mainly driven by the decrease of metal loss when stellar mass increases. Tremonti et al. (2004) have indeed observed an increase of the effective yield with stellar mass, while Lamareille et al. (2004) have shown the increase of the slope of the luminosity-metallicity relation. This last trend has also been observed down to much lower galaxy masses by Lee et al. (2006). We nevertheless note that they also observe a large scatter in the effective yield, which they find difficult to analyse in the context of more efficient mass loss among low mass galaxies.

Hierarchical galaxy formation models, that take into account the chemical evolution and feedback processes, are able to reproduce the observed mass-metallicity relation in the local uni-

⋆ based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, program 070.A-9007(A), and on data obtained at the Canada-France-Hawaii Telescope, operated by the CNRS in France, CNRC in Canada and the University of Hawaii.

verse (e.g. De Lucia et al. 2004; de Rossi et al. 2007; Finlator & Davé 2008). However these models rely on free parameters, such as feedback efficiency, which are not yet well constrained by observations. Alternative scenarios have been proposed to explain the mass-metallicity relation including low star formation efficiency in low-mass galaxies caused by supernova feedback (Brooks et al. 2007) and a variable stellar initial mass function being more top-heavy in galaxies with higher star formation rates, thereby producing higher metal yields (Köppen et al. 2007).

The evolution of the mass-metallicity relation on cosmological timescales is now predicted by semi-analytic models of galaxy formation, that include chemical hydrodynamic simulations within the standard Λ-CDM framework (De Lucia et al. 2004; Davé & Oppenheimer 2007). Reliable observational estimates of the mass-metallicity relation of galaxies at different epochs (and hence different redshifts) may thus provide important constraints on galaxy evolution scenarios. Estimates of the mass-metallicity - or luminosity-metallicity - relation of galaxies up to $z \sim 1$ are limited so far to small samples (Hammer et al. 2005; Liang et al. 2004; Maier et al. 2004; Kobulnicky et al. 2003; Maier et al. 2005, among others). Recent studies have been performed on larger samples (> 100 galaxies) but with contradictory results. Savaglio et al. (2005) concluded with a steeper slope in the distant universe, interpreting these results in the framework of the closed-box model. On the contrary, Lamareille et al. (2006a) did not find any significant evolution of the slope of the luminosity-metallicity relation, while the average metallicity at $z \approx 0.9$ is lowered by 0.55 dex at a given luminosity, and by 0.28 dex after correction for luminosity evolution. Shapley et al. (2005) and Liu et al. (2008) have also found $0.2 - 0.3$ dex lower metallicities at $z = 1$in the DEEP2 sample. At higher redshifts Erb et al. (2006) derived a mass-metallicity relation at $z \sim 2$ lowered by 0.3 dex in metallicity compared with the local estimate, a trend which could extend up to $z \sim 3.3$ (Maiolino et al. 2007). Please note that the above numbers are given after the metallicities have been renormalized to the same reference calibration in order to be comparable (Kewley & Ellison 2008).

In this work, we present the first attempt to derive the mass-metallicity relation at different epochs up to $z \sim 1$ using a unique, large (almost 20 000 galaxies) and homogeneous sample of galaxies selected in the VIMOS VLT Deep Survey (VVDS). The paper is organized as follows: the sample selection is described in Sect. 2, the estimation of stellar masses and magnitudes from SED fitting are described in Sect. 3, the estimation of metallicities from line ratios is described in Sec. 4, and we finally study the luminosity-metallicity (Sect. 5), and mass-metallicity (Sect. 6) relations. Throughout this paper we normalize the derived stellar masses, and absolute magnitudes, with the standard Λ-CDM cosmology, i.e. $h = 0.7$, $\Omega_{\rm m} = 0.3$ and $\Omega_\Lambda = 0.7$ (Spergel et al. 2003).

2. Description of the sample

The VIMOS VLT Deep Survey (VVDS, Le Fèvre et al. 2003b) is one of the widest and deepest spectrophotometric surveys of distant galaxies with a mean redshift of $z \approx 0.7$. The optical spectroscopic data, obtained with the VIsible Multi-Object Sectrograph (VIMOS, Le Fèvre et al. 2003a) installed at ESO/VLT (UT3), offers a great opportunity to study the evolution of the mass-metallicity relation on a statistically significant sample up to $z \approx 0.9$. This limitation is imposed, in the current study, by the bluest emission lines needed to compute a metallicity (i.e. [OII]λ3727, Hβ and [OIII]λ5007) being redshifted out of the wavelength range of the survey (approximately 5500Å $< \lambda <$ 9500Å). Metallicities measurements up to $z \approx 1.24$ using a different set of lines will be provided in a subsequent paper (Perez-Montero et al. 2008). The spectra have been taken in one arcsecond width slits, under a spectral resolution $R_{\rm s} \approx 230$**.**

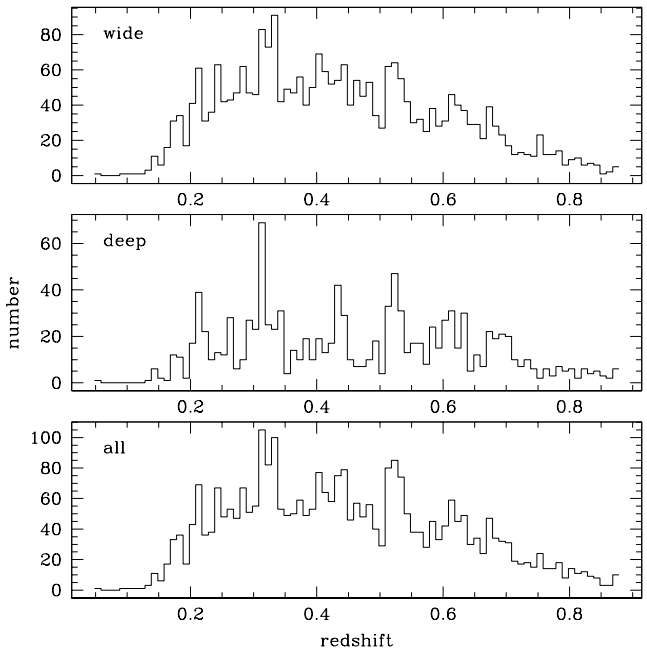


Figure 1. Redshift distributions of our wide and deep samples of star-forming galaxies, and of their union (see text for details).

The first epoch VVDS spectroscopic sample is purely apparent magnitude-selected and is divided into two deep fields ($17.5 \leq I_{\rm AB} \leq 24$): VVDS-02h (herefater F02, Le Fèvre et al. 2005) and CDFS (Le Fèvre et al. 2004); and three wide fields ($17.5 \leq I_{\rm AB} \leq 22.5$): VVDS-22h, VVDS-10h and VVDS-14h (hereafter F22, F10, and F14 respectively, Garilli et al. 2008). Spectroscopic data have been reduced using the *VIPGI* pipeline (Scodeggio et al. 2005), which performs automatic 1D spectra extraction, correction for telluric absorption lines, and flux calibration. We only kept objects with a redshift known at a confidence level greater than 75% (i.e. VVDS redshift flags 9, 2, 3 and 4). Duplicated observations were not used (some objects have been observed twice or more, either by chance or intentionally), we always kept the main observation. We also removed, for the purpose of our study (i.e. star-forming galaxies), all stars (zero redshift) and the sample of broad-line Active Galactic Nuclei (hereafter AGNs; Gavignaud et al. 2006).

The photometric coverage of the VVDS spectroscopic sample is as follows: CFH12k observations in *BVRI* bands for F02, F22, F10 and F14 fields (McCracken et al. 2003; Le Fèvre et al. 2004), completed by *U* band observations in the F02 field (Radovich et al. 2004), CFHTLS (Canada-France-Hawaï Telescope Legacy Survey) $u*g'r'i'z'$ bands for F02 and F22 fields, and *JKs* photometry available in F02 field (Iovino et al. 2005; Temporin et al. 2008). For the CDFS field, we use CFH12k *UBVRI* (Arnouts et al. 2001) and HST *bviz* (Giavalisco et al. 2004) observations. Finally, objects in F02 and F22 fields have been cross-matched with the UKIDSS public catalog (Warren et al. 2007), providing additional observations in *JK* bands.

Throughout this paper, we will use a *deep* and a *wide* samples. The deep sample is made of F02 and CDFS fields. Up to $z < 1.4$ (the limit for the [OII]λ3727 emission line to be in the observed wavelength range), the deep sample contains 7404 galaxies (non-broad-line AGN) with a redshift measured at a confidence level greater than 75%. The wide sample is made of F22, F10, F14, F02 and CDFS fields (the last two being limited out to $I_{AB} \leq 22.5$). Up to $z < 1.4$ the wide sample contains 13978 galaxies (non-broad-line AGN) with a redshift measured at a confidence level greater than 75%. We draw the reader's attention on the fact that the two samples overlap for galaxies observed in F02 or CDFS fields at $I_{AB} \leq 22.5$, giving a total number of 18648 galaxies. Fig. 1 shows the redshift distribution of the star-forming galaxies (see Sect. 2.2) in the wide and deep samples and in their union.

2.1. Automatic spectral measurements

The emission lines fluxes and equivalent widths, in all VVDS galaxy (non-broad-line AGN) spectra, have been measured with the *platefit_vimos* pipeline. Originally developed for the high spectral resolution SDSS spectra (Tremonti et al. 2004; Brinchmann et al. 2004), the *platefit* software has first been adapted to fit accurately all emission lines after removing the stellar continuum and absorption lines from lower resolution and lower signal-to-noise spectra (Lamareille et al. 2006b). Finally, other improvements have been made to the new *platefit_vimos* pipeline thanks to tests performed on the VVDS and zCOSMOS (Lilly et al. 2006) spectroscopic samples. A full discussion of *platefit_vimos* pipeline will be presented in Lamareille et al. (in preparation) but, for the benefit of the reader, we outline some of the main features in this section.

The stellar component of the spectra is fitted as a non-negative linear combination of 30 single stellar population templates with different ages (0.005, 0.025, 0.10, 0.29, 0.64, 0.90, 1.4, 2.5, 5 and 11Gyr) and metallicities (0.2, 1 and 2.5 $Z_{\odot}$). These templates have been derived using the Bruzual & Charlot (2003) library and resampled to the velocity dispersion of VVDS spectra. The dust attenuation in the stellar population model is left as a free parameter. Foreground dust attenuation from the Milky Way has been corrected using Schlegel et al. (1998) maps.

After removal of the stellar component, the emission lines are fitted together as a single *nebular spectrum* made of a sum of gaussians at specified wavelengths. All emission lines are tied to have the same width, with exception of the [OII]λ3727 line which is actually a doublet of two lines at 3726 and 3729 Å and appear broadened compared to the other single lines. The spectral resolution is also too low to clearly separate [NII]λ6584 and Hα emission lines. It has been shown however by Lamareille et al. (2006b) that the [NII]λ6584/Hα emission-line ratio, which is used as a metallicity calibrator, can be reliably measured above a sufficient signal-to-noise ratio even at the resolution of VVDS.

Note that because of the limited observed wavelength coverage of the spectra, we do not observe all well-known optical lines at all redshifts. Thanks to the stellar-part subtraction, no correction for underlying absorption has to be applied to the Balmer emission lines.

The error spectrum which is needed for both fits of the stellar or nebular components is calculated as follows: a first guess is obtained from photons statistics and sky subtraction and is calculated directly by the *VIPGI* pipeline. A fit of the stellar and nebular components is performed with *platefit_vimos* using this first error spectrum. The residuals of this first fit are then smoothed and added to the error spectrum quadratically, and a new fit is performed with *platefit_vimos*.

Finally, detected emission lines may also be removed from the original spectrum in order to get the observed *stellar spectrum* and measure indices from it, as well as emission-line equivalent widths. The underlying continuum is indeed obtained by smoothing the stellar spectrum. Then, equivalent widths are measured as direct integration over a 5 Gaussian-sigma bandpass of the emission-line Gaussian model divided by the underlying continuum.

For the absorption lines, one has to be aware that the resolution of VVDS spectra is too low in order to measure accurately Lick indices: the bandpasses of these indices are narrower than the width of the absorption lines at this resolution. In order to measure the absorption component in the Hδ line which is used to derive stellar masses (see below), we have defined a new broadened index for this specific line called Hδ_W. Table 1 gives the three bandpasses of this index. Figure 2 shows the comparison between this index and the standard Lick Hδ_A index measured on various model templates taken from Bruzual & Charlot (2003) library and covering a wide range of ages and metallicities. For future comparison purpose between VVDS measurements and other studies, we derive the following relation:

$$\mathrm{H}\delta_A = -4.69 + 1.691 \times \mathrm{H}\delta_W - 0.044711 \times \mathrm{H}\delta_W^2 \quad (1)$$

Table 1. Definition of the Hδ_W index

bandpass	λ (Å)
central	[4060.5; 4145]
blue	[4014; 4054]
red	[4151; 4191]

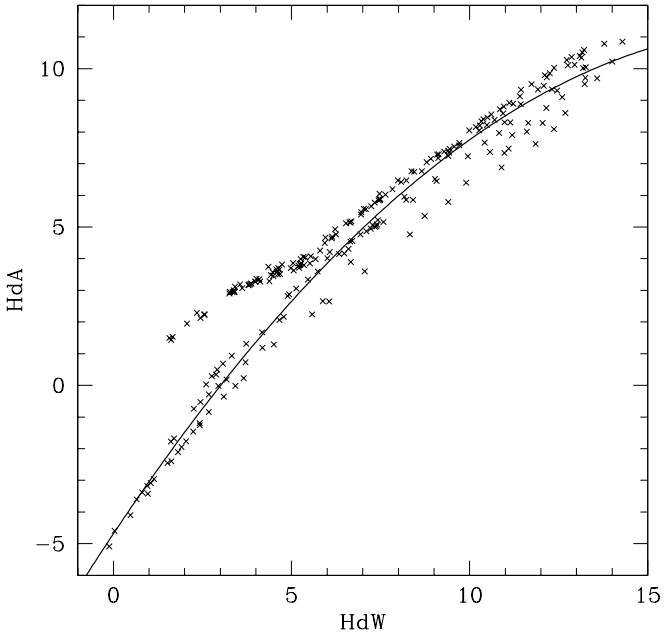


Figure 2. Comparison between the standard Lick Hδ_A index and the new broad index defined for low resolution VVDS spectra, namely Hδ_W (see Tab. 1), for a set of model templates taken from Bruzual & Charlot (2003) library and covering a wide range of ages and metallicities. The solid curve is the least square fit, which follows Eq. 1. Values are given in Å.

Table 2. Sets of emission lines which are required to have $S/N > 4$ in various redshift ranges associated to the various diagnostics used in this study.

Redshift	[OII]	Hβ	[OIII]	Hα	[NII]	[SII]
$0.0 < z < 0.2$				✓	✓	✓
$0.2 < z < 0.4$		✓	✓	✓	✓*	✓*
$0.4 < z < 0.5$		✓	✓			
$0.5 < z < 0.9$	✓	✓	✓			

* For the red diagnostic the [NII]λ6584 and [SII]$\lambda\lambda$6717+6731 emission lines may not be used at the same time.

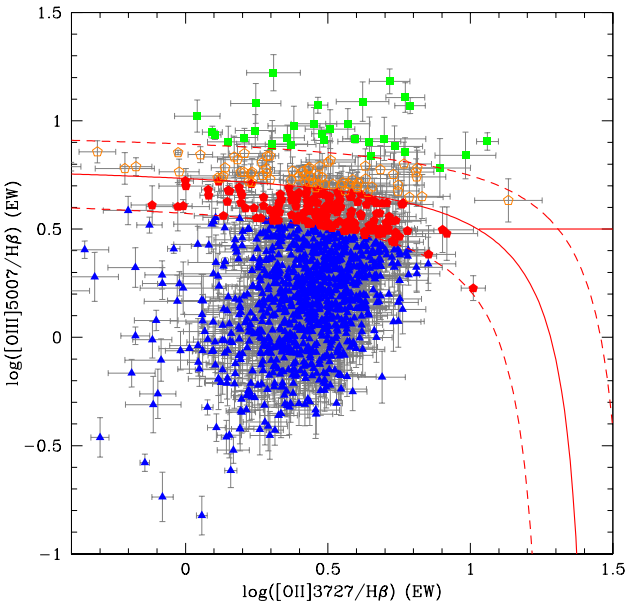


Figure 4. Blue spectral classification of 1060 narrow emission-line galaxies in the redshift bin $0.5 < z < 0.9$. The emission-line ratios – i.e. [OIII]λ5007/Hβ and [OII]λ3727/Hβ – are calculated using equivalent widths. The red solid curve is the empirical separation defined by Lamareille et al. (2004). The dashed curves delimits the error domain where both star-forming and AGNs galaxies show similar blue line ratios. The star-forming galaxies are plotted as blue triangles, the Seyfert 2 galaxies as green solid squares, the candidate star-forming galaxies as red solid pentagons, and the candidate AGNs as orange open pentagons. The error bars are shown in grey.

Note that the majority of points lying outside the fitted relation in Fig. 2 (for H$\delta_W < 5$Å) are models with a very young stellar population (< 10 Myr), hence this relation is only valid for stellar populations older than 10 Myr.

2.2. Selection of star-forming galaxies

2.2.1. Description of the various diagnostics

To ensure accurate abundance determinations, we restrict our attention to galaxies with emission lines detected at $S/N > 4$, where the set of lines considered for this S/N cut varies with redshift. Table 2 provides a summary of the emission lines required to have $S/N > 4$ in the various redshift ranges associated to the various diagnostics defined below.

Our work intends to understand the star-formation process in galaxies. Thus we have to remove from the sample of emission-line galaxies those for which the source of ionized gas, responsible for these lines, is not hot young stars. Emission-line galaxies can be classified in various spectral types, which depend on the nature of their nebular spectrum. The two main categories are the star-forming and AGN galaxies. Their source of ionizing photons are respectively hot young stars, or the accretion disk around a massive black-hole. As stated before, broad-line AGNs (also called Seyfert 1 galaxies) are already taken out of our sample. The narrow-line AGNs are divided in two sub-categories: Seyfert 2 galaxies , and LINERs .

Both Seyfert 2 galaxies and LINERs can be distinguished from star-forming galaxies using standard diagnostic diagrams (Baldwin et al. 1981; Veilleux & Osterbrock 1987), which are based on emission-line ratios. The most commonly used classification is provided by the [OIII]λ5007/Hβ vs. [NII]λ6584/Hα diagram (Fig. 3a), for which a semi-empirical limit between star-forming, Seyfert 2 galaxies, and LINERs has been derived by Kewley et al. (2001) from photo-ionization models. Due to our low spectral resolution, the [NII]λ6584 line is not always detected and its detection is less reliable than other lines, because of possible deblending problems with the bright Hα line. Consequently, most of the galaxies in the same redshift range are actually classified with the [OIII]λ5007/Hβ vs. [SII]$\lambda\lambda$6717+6731/Hα diagram (see Fig. 3b).

The two diagrams shown in Fig. 3, which we also refer to as the red diagnostics, can only be used with our data in the $0.2 < z < 0.4$ redshift range, as the desired emission lines are not visible outside this domain. At higher redshifts, one can use the alternative blue diagnostic which was empirically calibrated by Lamareille et al. (2004) from 2dFGRS data, in order to give similar results as the red diagnostics. The blue diagnostic is based on the [OIII]λ5007/Hβ vs. [OII]λ3727/Hβ diagram (line ratios calculated on rest-frame equivalent widths), and may be applied in the $0.5 < z < 0.9$ redshift range (see Fig. 4).

For the other redshift ranges, where neither the red nor the blue diagnostic applies, we use a minimal classification based on Hα, [NII]λ6584 and [SII]$\lambda\lambda$6717+6731 for $0.0 < z < 0.2$, and on Hβ and [OIII]λ5007 for $0.4 < z < 0.5$. We propose the following selection (called Hβ diagnostic) for star-forming galaxies in the $0.4 < z < 0.5$ redshift range: $\log([\mathrm{OIII}]\lambda5007/\mathrm{H}\beta) < 0.6$ (rest-frame equivalent widths, see Fig. 5a). We note that real star-forming galaxies which are lost with the Hβ diagnostic are mainly low metallicity ones. In any case a quick check on SDSS data has told us that no more than 40% of star-forming galaxies with $12 + \log(\mathrm{O/H}) < 8.1$ actually fall in the AGN region of the Hβ diagnostic, this proportion being negligible at higher metallicities.

Figure 5b shows the minimum classification for the $0.0 < z < 0.2$ redshift range (called Hα diagnostic). The proposed separation has been derived based on the 2dFGRS data (Lamareille 2006) and follows the equation (rest-frame equivalent widths):

$$\log([\mathrm{NII}]\lambda6584/\mathrm{H}\alpha) = \begin{cases} -0.4 & \text{if} \quad \log([\mathrm{SII}]\lambda\lambda6717+6731/\mathrm{H}\alpha) \geq -0.3 \\ -0.7 - 1.05\log([\mathrm{SII}]\lambda\lambda6717+6731/\mathrm{H}\alpha) & \text{otw.} \end{cases} \quad (2)$$

This equation have been derived to efficiently reduce both the contamination by real AGNs in the star-forming galaxies region ($< 1\%$), and the fraction of real star-forming galaxies which are lost ($< 4\%$).

The results of all diagnostics, in the wide, deep, and global samples, are shown in Table. 3.

Table 3. Statistics of star-forming galaxies and narrow-line AGNs among emission-line galaxies for various diagnostics depending on which emission lines are observed (see text for details). The results are presented for the two wide and deep samples used in this paper, and for their union (some objects are in common). We also mention for each sample the total number of objects which includes emission-line, faint and early-type galaxies.

Sample/Diagnostic	red	*%*	blue	*%*	Hα	*%*	Hβ	*%*	all	*%*
wide (13978)										
emission-line (total)	455	*100*	924	*100*	47	*100*	1402	*100*	2828	*100*
star-forming	415	*91*	768	*83*	30	*64*	978	*70*	2191	*77*
candidate s.-f.	0	*0*	103	*11*	0	*0*	306	*22*	409	*14*
candidate AGN	0	*0*	35	*4*	0	*0*	74	*5*	109	*4*
Seyfert 2	18	*4*	18	*2*	17	*36*	44	*3*	97	*3*
LINER	22	*5*	0	*0*	0	*0*	0	*0*	22	*1*
deep (7404)										
emission-line (total)	151	*100*	568	*100*	11	*100*	574	*100*	1304	*100*
star-forming	136	*90*	412	*73*	10	*91*	312	*54*	870	*67*
candidate s.-f.	0	*0*	92	*16*	0	*0*	166	*29*	258	*20*
candidate AGN	0	*0*	47	*8*	0	*0*	47	*8*	94	*7*
Seyfert 2	5	*3*	17	*3*	1	*9*	49	*9*	72	*6*
LINER	10	*7*	0	*0*	0	*0*	0	*0*	10	*1*
union of wide and deep (18648)										
emission-line (total)	469	*100*	1213	*100*	47	*100*	1671	*100*	3400	*100*
star-forming	427	*91*	945	*78*	30	*64*	1074	*64*	2476	*73*
candidate s.-f.	0	*0*	166	*14*	0	*0*	404	*24*	570	*17*
candidate AGN	0	*0*	70	*6*	0	*0*	111	*7*	181	*5*
Seyfert 2	20	*4*	32	*3*	17	*36*	82	*5*	151	*4*
LINER	22	*5*	0	*0*	0	*0*	0	*0*	22	*1*

2.2.2. Discussion of possible biases

It is very important to evaluate the possible sources of biases, which may affect the spectral classification, coming from the use of various diagnostics at different redshift ranges.

As the [OII]λ3727/Hβ line ratio is less accurate than [NII]λ6584/Hα, or [SII]$\lambda\lambda$6717+6731/Hα, to distinguish between star-forming and Seyfert 2 galaxies, the blue diagnostic is defined with an error domain (see the dashed curves in Fig. 4), inside which individual galaxies cannot be safely classified. We thus introduce two new categories of galaxies: candidate star-forming galaxies and candidate AGNs, which fall respectively in the lower-half or the upper-half part of the error domain.

We know however that AGNs are in minority in the universe. Thus we emphasize that *i)* still no reliable classification can be performed for any individual galaxy falling inside the error domain of the blue diagnostic, *ii)* the search for AGNs is highly contaminated in the candidate AGN region, *but iii)* any study involving statistically significant samples of star-forming galaxies should not be biased when candidate star-forming galaxies and candidate AGNs are included. As it will be shown in Sect. 5 and 6, this ends up to a negligible bias on the derived metallicities.

As for the blue diagnostic, we have defined an error domain for the Hβ diagnostic in the following range: $0.4 < \log([\mathrm{OIII}]\lambda5007/\mathrm{H}\beta) < 0.6$. Star-forming galaxies or AGNs falling inside this domain are classified as candidates.

One can see in Table. 3 that the red diagnostic is the only one able to find LINERs. Indeed, LINERs fall in the star-forming galaxies region with the blue or Hβ diagnostic, while they fall in the Seyfert 2 region with the Hα diagnostic. We know however that the contamination of star-forming galaxies by LINERs, in the blue or Hβ diagnostic, is less than 1%.

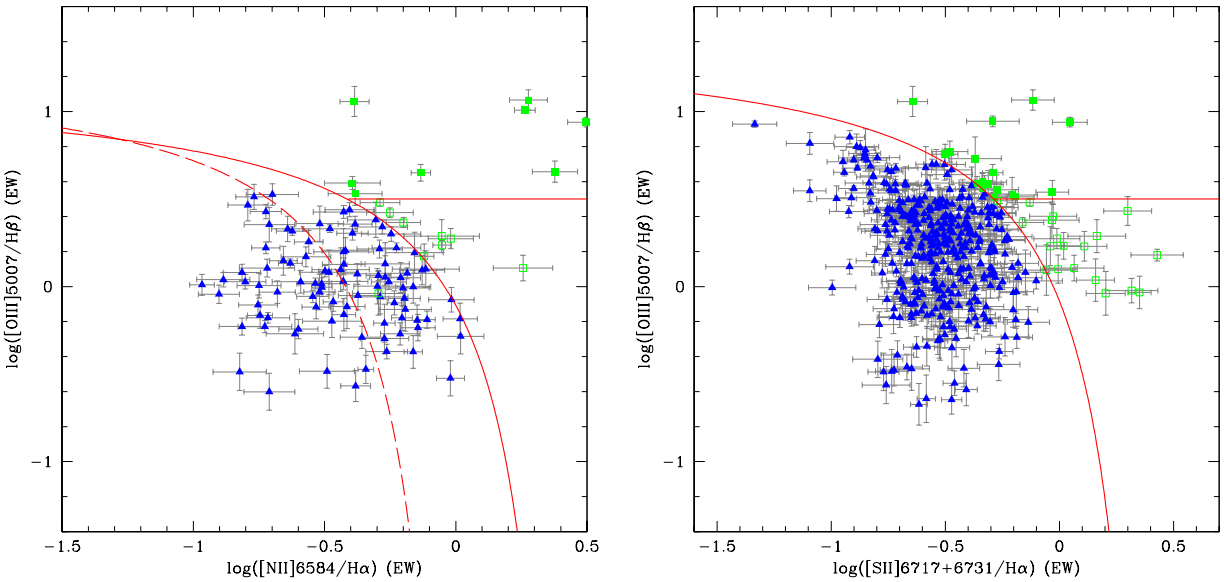
Another bias could come from the population of composites galaxies, i.e. for which the ionized gas is produced by both an AGN and some star-forming regions. When looking for AGNs in the SDSS data, Kauffmann et al. (2003a) have defined a new, less restrictive, empirical separation between star-forming galaxies and AGNs (see the dashed curve in Fig. 3). They have then classified as composites all galaxies between this new limit and the old one by Kewley et al. (2001). This result is confirmed by theoretical modeling: Stasińska et al. (2006) have found that composite galaxies are indeed falling in the region between the curves of Kauffmann et al. (2003a) and Kewley et al. (2001) in the [OIII]λ5007/Hβ vs. [NII]λ6584/Hα diagram. Moreover, they have found that composite galaxies fall in the star-forming galaxy region in the other diagrams ([OIII]λ5007/Hβ vs. [SII]$\lambda\lambda$6717+6731/Hα or [OII]λ3727/Hβ).

It is difficult to evaluate the actual bias due to contamination by composite galaxies, in the red or blue diagnostics that we use in this study. This difficulty comes mainly from the fact that, if composite galaxies actually fall in the composite region as defined by Kauffmann et al. (2003a), not all galaxies inside this region are necessarily composites. A large majority of them might be normal star-forming galaxies. One way to evaluate the contamination by composite galaxies is to look for star-forming galaxies with an X-ray detection. Such work has been performed with zCOSMOS data with similar selection criteria than the wide sample: Bongiorno et al. (in preparation) have found that the contamination of star-forming galaxies by composites is approximately 10%.

3. Estimation of the stellar masses and absolute magnitudes

3.1. The Bayesian approach

The stellar masses are estimated by comparing the observed Spectral Energy Distribution (hereafter SED), and two spectral



(a) (b)

Figure 3. Red spectral classification of 412 narrow emission-line galaxies in the redshift bin $0.2 < z < 0.4$. The emission-line ratios – i.e. [OIII]λ5007/Hβ and [NII]λ6584/Hα (a) or [OIII]λ5007/Hβ and [SII]$\lambda\lambda$6717+6731/Hα (b) – are calculated using equivalent widths. The red solid curves are the semi-empirical separations defined by Kewley et al. (2001). The dashed line is the separation proposed by Kauffmann et al. (2003a). The star-forming galaxies are plotted as blue triangles, the Seyfert 2 galaxies as green solid squares, the LINERs as green open squares. The error bars are shown in grey.

features (Hδ_W absorption line and $D_n(4000)$ break), to a library of stellar population models. The use of two spectral features reduces the well-known age-dust-metallicity degeneracy in determining the mass-to-light ratio of a galaxy. Compared to pure photometry, Balmer absorption lines are indeed less sensitive to dust, while the $D_n(4000)$ break is less sensitive to metallicity, both being sensitive to the age.

One observation is defined by a set F_i of observed fluxes in all photometric bands, a set σF_i of associated errors, a set $I_{i'}$ of observed indices, and a set $\sigma I_{i'}$ of associated errors. The χ^2 of each model, described by a set F_i^0 of theoretical photometric points, and a set $I_{i'}^0$ of theoretical indices, is calculated as:

$$\chi^2 = \sum_i \frac{\left(F_i - A \cdot F_i^0\right)^2}{\sigma F_i^2} + \sum_{i'} \frac{\left(I_{i'} - I_{i'}^0\right)}{\sigma I_{i'}^2} \quad (3)$$

where A is the normalization constant that minimizes the χ^2. Note that the normalization constant has not to be applied to the spectral indices as they are already absolute.

Still for one observation, the set of χ_j^2 calculated on all models are summarized in a PDF (Probability Distribution Function) which gives the probability of each stellar mass, given the underlying library of models (the prior). Each stellar mass $M_\star$ is assigned a probability described by the following normalized sum:

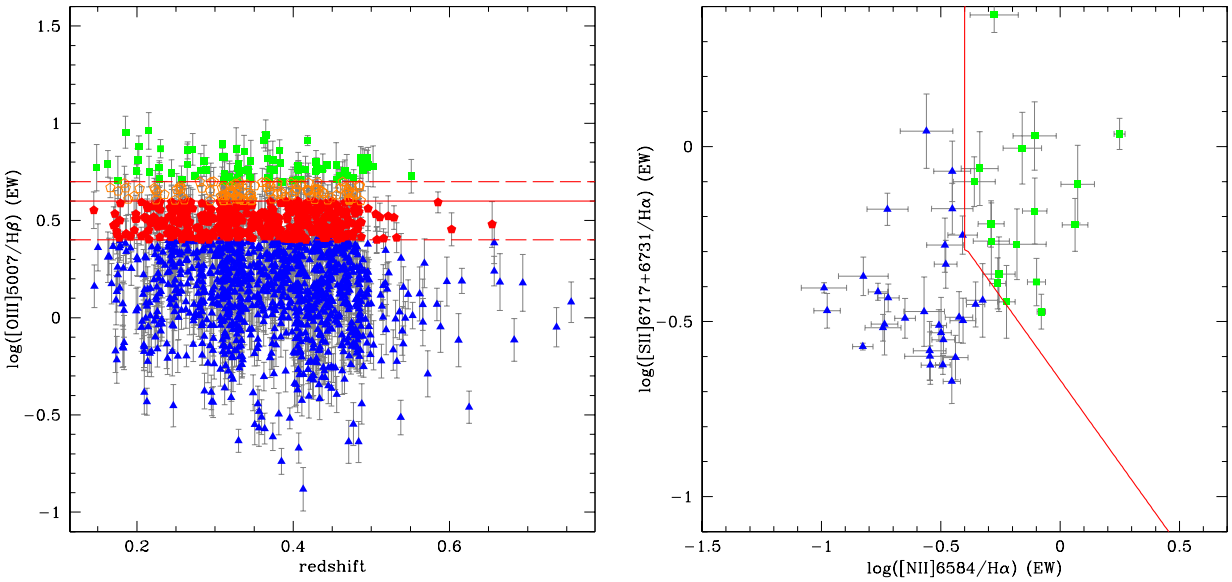
$$P(M_\star|\{F_i\}) = \frac{\sum_j \delta(M_\star - A \cdot M_\star^j) \cdot \exp\left(-\chi_j^2/2\right)}{\sum_j \exp\left(-\chi_j^2/2\right)} \quad (4)$$

where $M_\star^j$ is the stellar mass of the model associated to χ_j^2. In practice, the PDF is discretized into bins of stellar masses. Our stellar mass estimate is given by the median of the PDF .

This method, based on the Bayesian approach, has been introduced by the SDSS collaboration (Kauffmann et al. 2003b; Tremonti et al. 2004; Brinchmann et al. 2004) in order to carry SED fitting estimates of the physical properties of galaxies, and starts now to be widely used on this subject. Its main advantage is that it allows us, for each parameter, to get a reliable estimate, *independently* for each parameter, which takes all possible solutions into account, not only the best-fit. Thus, this method takes into account degeneracies between observed properties in a self-consistent way. It also provides an error estimate of the derived parameter from the half-width of the PDF .

3.2. Description of the models

We use a library of theoretical spectra based on (Bruzual & Charlot 2003, hereafter BC03) stellar population synthesis models, calculated for various star formation histories. Unless many other studies which only use a standard declining exponential star formation history, we have used an improved grid including also secondary bursts (stochastic library, Salim et al. 2005; Gallazzi et al. 2005). The result of the secondary bursts, compared to previous methods, is mainly to get higher masses as we are now able to better reproduce the colors of galaxies containing both old and recent stellar populations. When one uses a prior with a smooth star formation history, such objects are better fitted by young models which correctly reproduce the recent star formation history and the colors in the bluer bands. But such models fail to reproduce the underlying old stellar population which affects the colors of the redder bands, therefore they lead to an underestimate of the final stellar masses by a factor



(a) (b)

Figure 5. Intermediate spectral classifications of 1479 (a) and 43 (b) narrow emission-line galaxies. The two plots show the intermediate diagnostic diagrams, used in the redshift bins $0.4 < z < 0.5$ (a) and $0.0 < z < 0.2$ (b), or in other redshifts bins for galaxies with insufficient number of detected emission lines (therefore not classified with the red or blue diagnostic). The emission-line ratios – i.e. [OIII]λ5007/Hβ, [SII]$\lambda\lambda$6717+6731/Hα, and [NII]λ6584/Hα – are calculated using equivalent widths. The red solid lines are the empirical separations defined by Lamareille (2006). In the left panel, the dashed lines delimits the error domain where both star-forming and AGNs galaxies show similar blue line ratios. The star-forming galaxies are plotted as blue triangles, the Seyfert 2 galaxies as green solid squares, the candidate star-forming galaxies as red solid pentagons, and the candidate AGNs as orange open pentagons. The error bars are shown in grey.

≈ 1.4 (see Pozzetti et al. 2007, for a discussion of the different methods).

Masses derived with this method are also relative to the chosen IMF (Chabrier 2003) . The models include self-consistent two-component dust corrections (Charlot & Fall 2000).

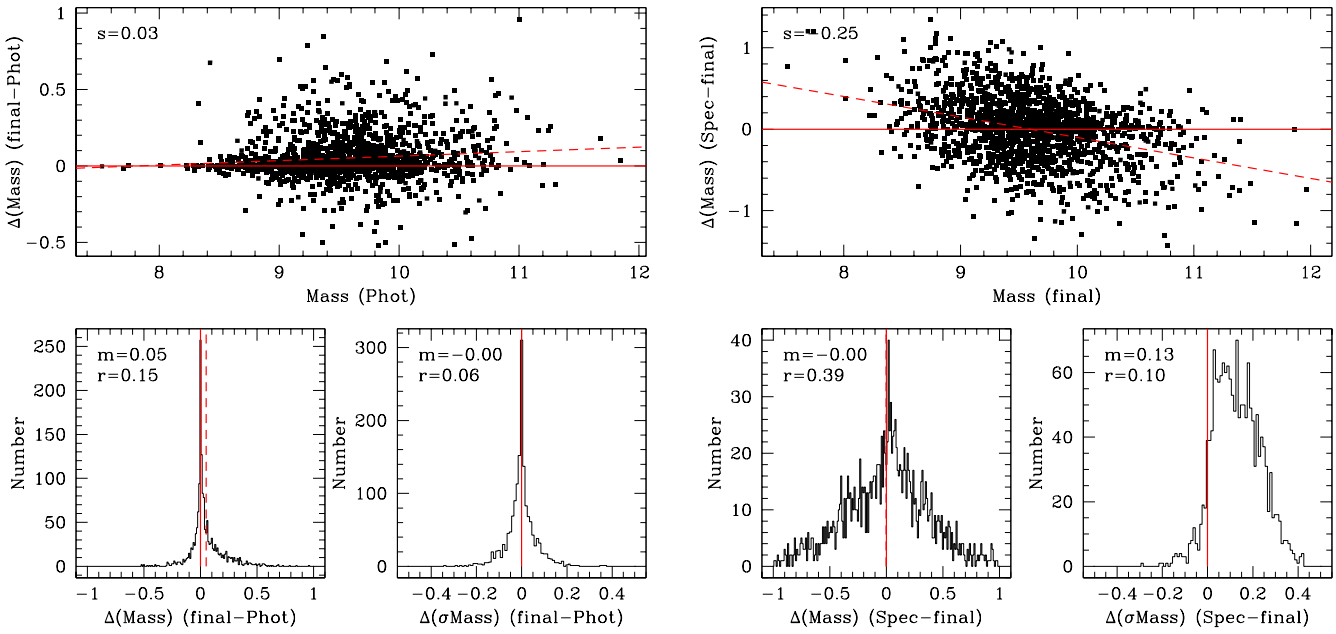
In addition to the secondary bursts, the models used in this study also include a new treatment of the TP-AGB stars (Marigo & Girardi 2007) (Charlot & Bruzual, in preparation, hereafter CB07), in order to provide a better fit of the near-infrared photometric bands (see also Maraston 2005; Maraston et al. 2006). We have derived the stellar masses of our whole sample of galaxies with BC03 and CB07 models. Figure 6 shows the results of this comparison. As expected, the masses derived with CB07 models are smaller, since the near-infrared flux of TP-AGB stars was underestimated in BC03 models. The mean shift is -0.07 dex, with a dispersion of 0.11 dex, which corresponds to an overestimate of the stellar masses by a factor ≈ 1.2 with BC03 models. Their is also a small trend for more massive galaxies to be more affected by the difference between the two models. This comparison has been studied in more details by Eminian et al. (2008).

The same comparison have been performed on masses derived only using photometry: the mean shift is -0.09 dex, with a dispersion of 0.19 dex. The spectral indices are not expected to vary much with the new treatment of TP-AGB stars introduced by CB07 models. It is thus expected that masses computed only with photometry vary more from BC03 to CB07 models than those computed also with spectral indices.

We note that the stellar masses derived with BC03 models and a smooth star formation history are very similar to the ones derived with CB07 models and secondary bursts. Stronger effects have to be expected on derived star formation rates or mean stellar ages.

Figure 7 now compares the stellar masses obtained with CB07 models when using only photometry, only spectroscopy, and both photometry and spectroscopy. Figure 7(a) shows the comparison between the mass obtained using only photometry, and the one obtained by adding the two spectral indices mentioned above. The derived mass and associated error are not significantly biased towards the two different methods (photometric masses are -0.05 dex lower). Nevertheless, we have compared the χ^2 values in both cases and found that it never increases when we add the spectral indices in the analysis. Moreover, the χ^2 obtained in this latter case is lowered by a factor whose mean value is 2.3. The dispersion of 0.15 dex in the comparison between the two masses has thus to be taken into account as the minimum uncertainty of our stellar masses.

Figure 7(b) shows what would happen if only spectral indices were used. We see that the derived mass does not suffer from any bias, but is subject to a large dispersion of 0.39 dex (a factor 2.45). A slope of -0.25 dex/decade is found, but is not significant given the dispersion. The comparison of the derived errors shows that this additional uncertainty is mostly taken into account: both measurements have to account for a minimum uncertainty of 0.15 dex and the measurement obtained using only spectroscopy show a 0.13 dex additional uncertainty.



(a) (b)

Figure 7. Same legend as in Fig. 6, except that we now compare the stellar masses that we obtain with CB07 models with photometry and spectroscopy (final) to the ones obtain: (a) using only photometry, (b) using only spectroscopy. Only objects with a sufficient number of photometric bands *and* spectral indices (i.e. more than two of each) are used in these plots.

The reason of this comparison is that Tremonti et al. (2004) have not used full SEDs to derive stellar mass-to-light but only two spectral indices: $D_n(4000)$ and $\mathrm{H}\delta_\mathrm{A}$, the mass being then scaled from z-band luminosity. We note that their spectral indices have been measured at a much better signal-to-noise and resolution on SDSS spectra as compared to VVDS spectra. We conclude from Fig. 7(b) that Tremonti et al. (2004) masses computed only with spectroscopy are directly comparable to our final masses computed also with photometry, provided that we take scale them by -0.07 dex and -0.056 dex to account respectively for the difference between BC03 and CB07 models, and between Kroupa (2001) and Chabrier (2003) IMF.

3.3. Defining volume-limited samples

Figure 8 shows the stellar masses obtained, as a function of redshift, for the galaxies of our various samples (wide or deep, all galaxies or star-forming galaxies only). As expected in a magnitude-selected sample, the minimum stellar mass which is detected depends on the redshift, because of the decrease of apparent magnitude when the distance increases. This effect is well-known as the Malmquist bias. The upper-limit of the mass distribution is in part due to the bright cut-off in the magnitude selection, while in addition to that a larger volume is observed at higher redshift which allows a better sampling of the upper mass region. The VVDS is a constant solid-angle survey: the observed volume increases with distance and rare objects, like very massive galaxies, get more chances to be detected.

There are two ways to account for these two biases. The first one is to use the Vmax technique: each object is weighted by the inverse of the maximum volume in which it could be observed, given the selection function of the survey. The result is a density which is statistically corrected for the selection effects, but may suffer from evolution effects. Another, simpler, technique is to build volume-limited samples: a redshift range is defined and only objects which may be observed in the entire volume between these two redshifts (given the selection function) are kept. We use this last technique in our study.

In practice, for magnitude-selected surveys, one has to calculate the minimum stellar mass which is detected at the upper limit of the redshift range, and then to throw out the objects below this limit since they are not observed in the entire volume. Nevertheless the minimum mass which may be detected at a given redshift, for a given limiting apparent magnitude, depends on the stellar mass-to-light ratio. One way to solve this issue is to calculate, for each galaxy, the mass $M^\star_\mathrm{lim}$ it would have if its apparent magnitude I was equal to the selection magnitude I_sel (i.e. 24 for the deep sample, and 22.5 for the wide sample):

$$M^\star_\mathrm{lim} = M^\star + 0.4 \cdot (I - I_\mathrm{sel}) \quad (5)$$

The result is a distribution of minimum stellar masses which reflects the distribution of stellar mass-to-light ratios in the sample. The red, green, and blue curves in Fig. 8 show respectively the 50%, 80% and 95% levels for the cumulative sum of this distribution. The limiting masses derived for the 4 different samples mentioned above are given in Table 4. As expected, we see that the limiting mass increases with redshift and is lower for the deep sample. The derived completeness levels on the deep sample are similar but higher to the ones found by Meneux et al. (2008) on the same data. The difference is due to the different cosmology used in this paper, i.e. $h = 1$.

We see also that the limiting mass is lower when we consider only star-forming galaxies, compared to the whole samples. This

Table 4. Limiting masses (in logarithm of solar masses) of our various samples in order to define volume-limited samples, as a function of the redshift ranges and for various mass-to-light completeness levels.

	$z < 0.5$			$z < 0.6$			$z < 0.7$			$z < 0.8$			$z < 0.9$		
	50%	80%	95%	50%	80%	95%	50%	80%	95%	50%	80%	95%	50%	80%	95%
wide (all)	9.6	10.0	10.2	9.7	10.2	10.4	9.9	10.3	10.6	10.0	10.5	10.7	10.2	10.6	10.9
wide (star-forming)	9.3	9.6	9.9	9.4	9.7	10.1	9.6	9.9	10.2	9.7	10.0	10.3	9.8	10.1	10.4
deep (all)	8.7	9.2	9.5	8.9	9.3	9.7	9.0	9.5	9.9	9.1	9.6	10.0	9.2	9.7	10.1
deep (star-forming)	8.6	8.9	9.2	8.7	9.0	9.3	8.8	9.0	9.3	8.8	9.1	9.4	8.9	9.2	9.5

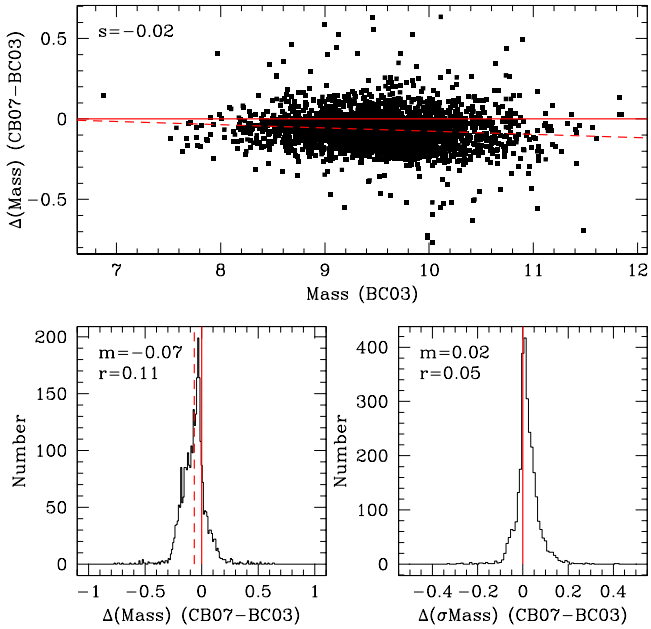


Figure 6. Comparison of the stellar masses of star-forming galaxies in our sample, obtained with BC03 and CB07 models, using both photometry and spectroscopy. The latter provides a better treatment of the TP-AGB stars in the near-infrared. The top panel shows the difference between the two masses, the bottom panels show the associated histogram (left) and the histogram of the differences between the two error estimates (right). The slope of the difference (s) is shown in the plot. The mean value (m) and the rms (r) of the two histograms are given in the plots.

is due to star-forming galaxies having lower mass-to-light ratios. This shows that if a sub-sample has a smaller range in mass-to-light ratios, then a less conservative mass limit can be adopted than if a single fixed maximum mass-to-light ratio is taken.

Finally, we note that the wide sample contains more massive objects since it covers a larger volume of universe.

3.4. Absolute magnitudes

As a by-product of the estimation of stellar masses, we have also calculated the *k*-corrected absolute magnitudes, in the rest-frame *B*-band, using the Bayesian approach and the CB07 models. The results for star-forming galaxies are shown in Fig. 9 as a function of redshift. Because of the *k*-correction, the maximum absolute magnitude which may be detected, as a function of redshift, is also affected by the stellar mass-to-light ratio. The absolute magnitude in rest-frame *B*-band is affected by the ratio between the flux in this band, and the flux in the observed *I*-band, which ratio depends on the type of the galaxy and on redshift We also note that as observed *I*-band is fairly close to rest-frame *B*-band, the variation in *k*-correction is small compared to the variation in mass-to-light ratio as can be seen by the reduced spread between the different completeness level curves in Fig. 8 as compared to Fig. 9.

The limiting *B*-band absolute magnitudes, associated to various redshift ranges and *k*-correction completeness levels, are given in Table 5. We note that the absolute magnitudes used in this study may differ from previous works performed on the same dataset (e.g. Ilbert et al. 2005). For self-consistency, we have chosen to calculate the absolute magnitudes with the same models than the ones used for the stellar masses, namely CB07 models with secondary bursts.

4. Estimation of the metallicities

In this study, the metallicities are estimated as the gas-phase oxygen abundances $12 + \log(\mathrm{O/H})$.

4.1. The empirical approach

Strong-line empirical calibrators may be used to compute the gas-phase metallicities from the available measured emission lines, e.g. Hα, Hβ, [OIII]λ5007, [NII]λ6584, [OII]λ3727. As shown by Lamareille et al. (2006b), the [NII]λ6584 and Hα emission lines are reliably deblended for sufficiently high signal-to-noise ratio even at the resolution of the VVDS.

We use the *N*2 calibrator up to $z \sim 0.2$ (Denicoló et al. 2002; van Zee et al. 1998), the *O*3*N*2 calibrator for the redshift range $0.2 < z < 0.5$ (Pettini & Pagel 2004), and the R_{23} calibrator for the redshift range $0.5 < z < 0.9$ (McGaugh 1991).

At high redshifts, where nitrogen or sulfur lines are not observed, the metallicity is degenerate because of radiative cooling at high metallicities which lowers the intensities of oxygen lines. Thus, low oxygen-to-hydrogen line ratio can be interpreted as either low oxygen abundance or high oxygen cooling, i.e. high oxygen abundance. The degeneracy of the R_{23} calibrator is broken using the *B*-band absolute magnitude and the reference luminosity-metallicity relation (see Lamareille et al. 2006a, for a detailed discussion of this method).

As clearly shown by Kewley & Ellison (2008), the difficulty of using different metallicity calibrators at different redshifts comes from the huge differences, up to ± 0.7 dex, which exist between them. Therefore, we have decided to renormalize the different calibrators to the Charlot & Longhetti (2001, hereafter CL01) calibrator, which has been used to estimate the mass-metallicity relation from SDSS data (Tremonti et al. 2004). The correction formulas are found in Table B3 of Kewley & Ellison (2008).

Finally, a consequence of the small spectral coverage is also that Hα and Hβ lines are not observed together for most redshifts, making difficult the estimation of the dust attenuation

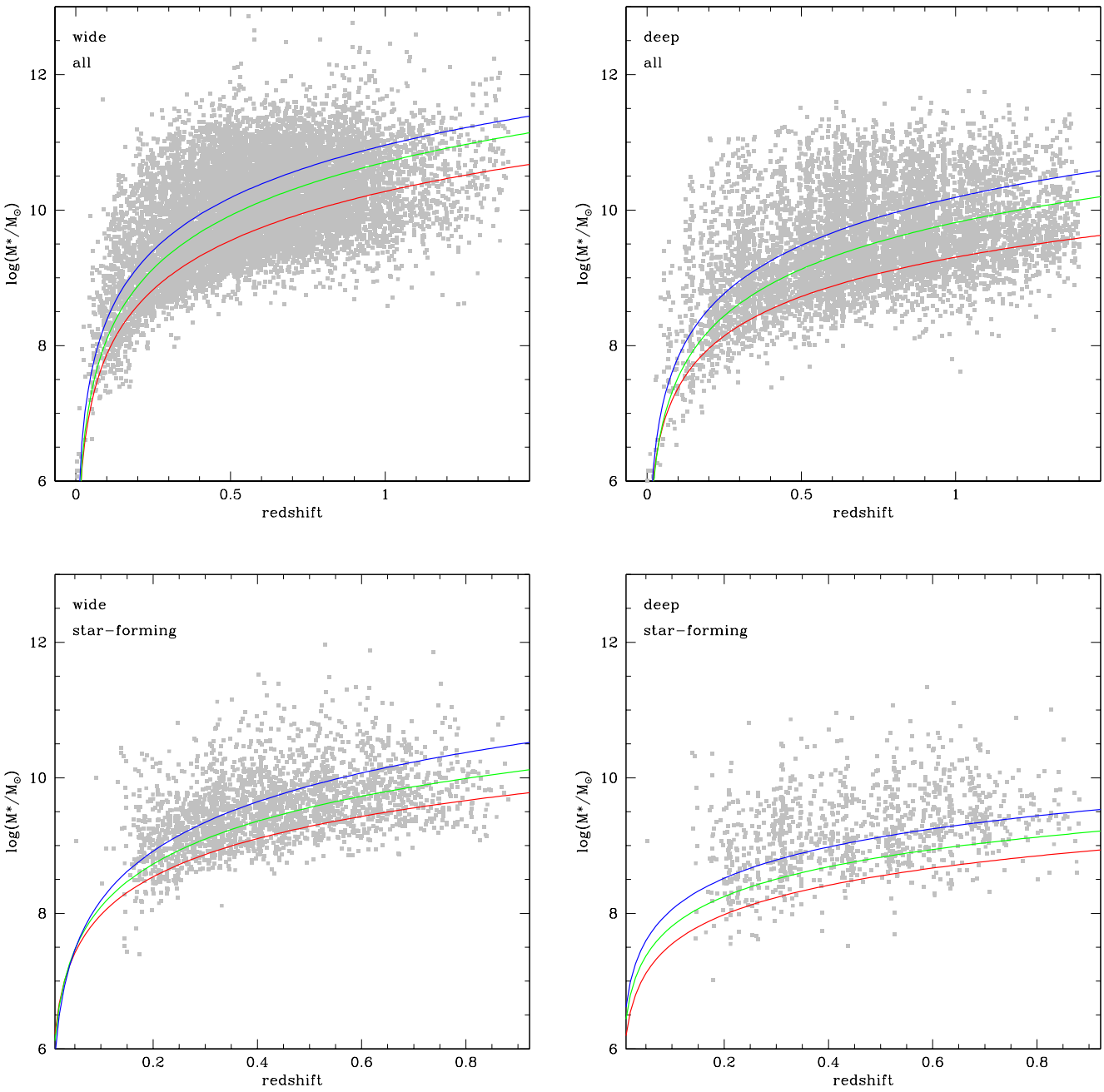


Figure 8. Stellar mass estimates (in logarithm of solar masses), as a function of redshift, of the galaxies in our samples: left: the wide sample, right: the deep sample, top: all galaxies (except broad-line AGNs), bottom: only star-forming galaxies (see Sect. 2.2). The red, green and blue curves illustrate respectively the 50%, 80% and 95% mass-to-light completeness levels (see text for details).

from the standard Balmer decrement method. As already shown by Kobulnicky & Phillips (2003), and Lamareille et al. (2006a), metallicities can be reliably derived using equivalent width measurements instead of dust-corrected fluxes.

Nevertheless, Liang et al. (2007) have shown that the [OIII]λ5007/[OII]λ3727 emission line ratio, which is used in R_{23} calibrator, does not give the same value if it is calculated using equivalent widths or dust-corrected fluxes: the ratio between the two results depends on the differential dust attenuation between stars and gas, and on the slope of the stellar continuum. They have found that the derived metallicities are affected by a factor ranging from -0.2 to 0.1 dex, but they also showed that this factor has a mean value of only -0.041 dex. It is thus not significant for the estimation of the mass-metallicity relation on statistical samples. Nevertheless, in order to avoid any possible bias on a different sample, we decided to use their corrective formula based on the $D_n(4000)$ index.

4.2. The Bayesian approach

The metallicities may also be estimated using again the Bayesian approach. The relative fluxes of all measured emission lines are compared to a set of photoionization models, which predicts the theoretical flux ratios given four parameters: the gas-phase

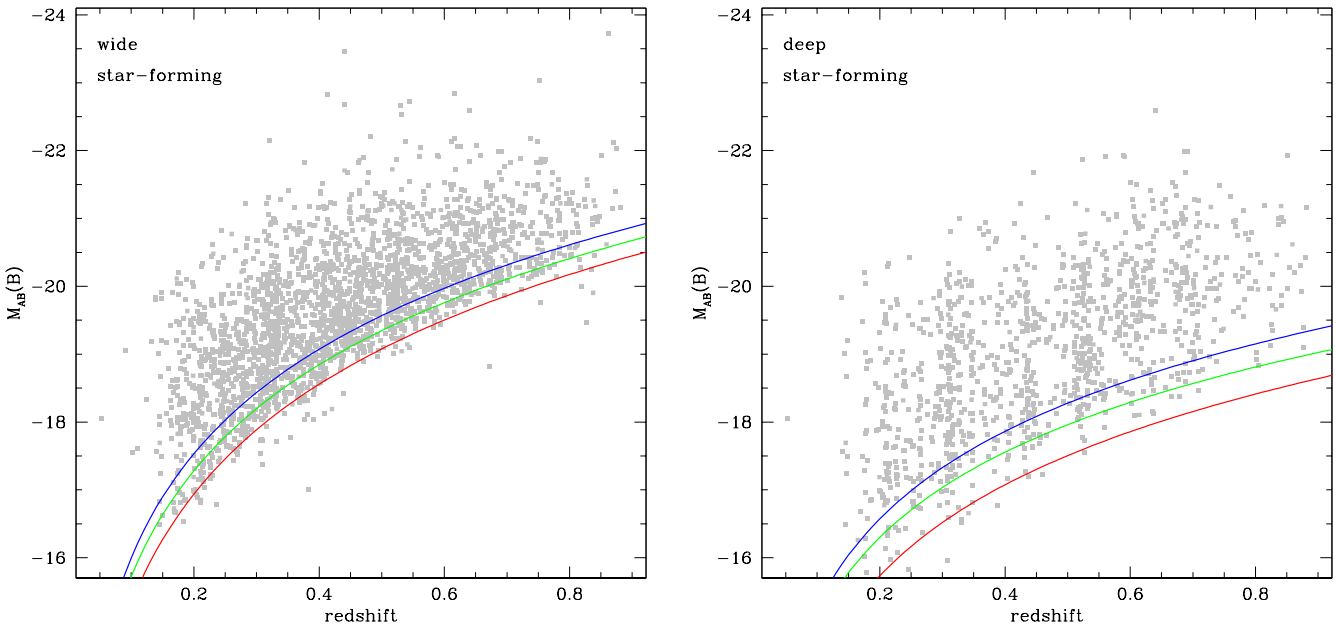


Figure 9. Absolute magnitudes estimates in the rest-frame *B*-band (AB system), as a function of redshift, of the star-forming galaxies of the wide (left) and deep (right) samples. The red, green and blue lines illustrate respectively the 50%, 80% and 95% *k*-correction completeness levels (see text for details).

Table 5. Limiting *B*-band absolute magnitudes (AB system) of our wide and deep samples of star-forming galaxies, in order to define volume-limited samples, as a function of the redshift ranges and for various *k*-correction completeness levels.

	$z < 0.5$		$z < 0.7$		$z < 0.9$	
	50%	95%	50%	95%	50%	95%
wide	−19.1	−19.6	−19.9	−20.3	−20.5	−20.9
deep	−17.5	−18.3	−18.2	−18.9	−18.6	−19.3

metallicity, the ionization level, the dust-to-metal ratio and the reddening (Charlot & Longhetti 2001). The CL01 models are based on population synthesis for the ionizing flux (Bruzual & Charlot 2003), emission line modeling (Cloudy, Ferland 2001) and a two-component dust attenuation law (Charlot & Fall 2000).

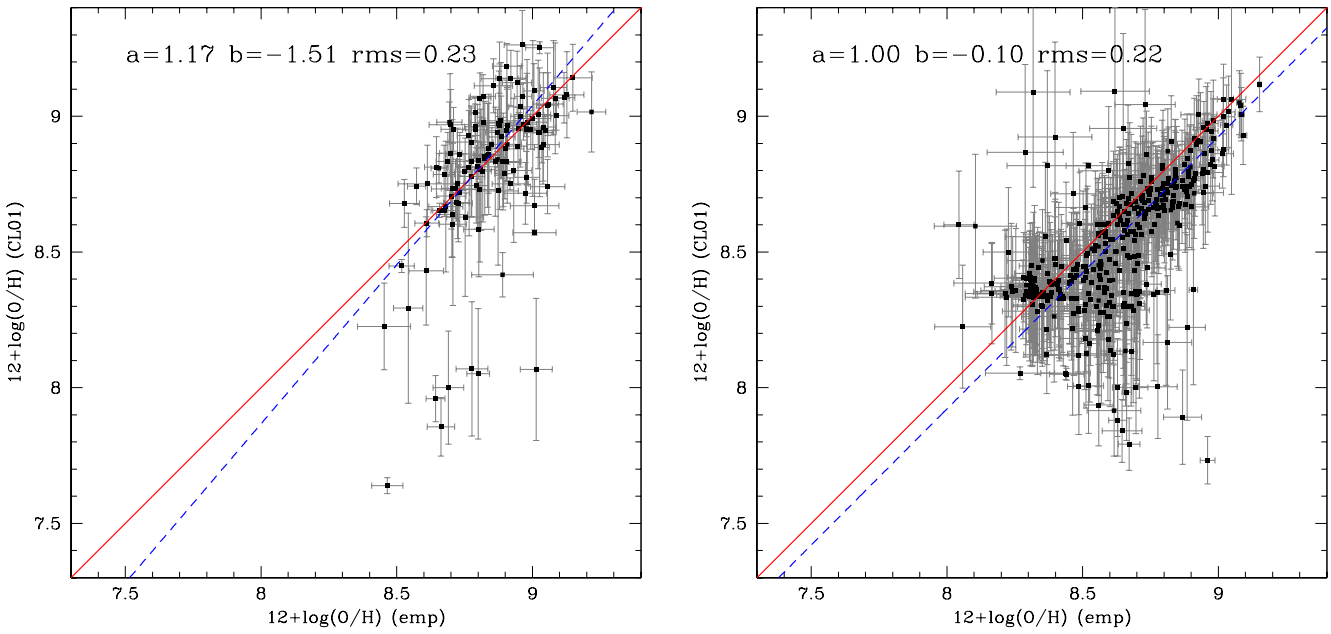
We calculate the χ^2 of each model and summarize them in the PDF of the metallicity using a similar method than described above in Eq. 3 and 4, but now applied to emission-line fluxes instead of photometric points. Only the emission lines with enough signal-to-noise (S/N¿4) are used in the fit. This method is applied directly on observed line fluxes: the correction for dust attenuation is included self-consistently in the models.

The O/H degeneracy at high redshift ends up with double-peaked PDFs: one peak for the low metallicity solution and another one at high metallicity. However, other information such as dust extinction, ionization level, or star formation rate help affecting to the two peaks two different probabilities. Thus, we solve the degeneracy by fitting two peaks in the PDFs, and by keeping the one with the highest probability. As already discussed in Lamareille et al. (2006a), this method cannot be used to chose the metallicity of one single galaxy, but it can be used statistically to derive a mean metallicity, e.g. as a function of mass.

In this study, given the relatively low spectral resolution and spectral coverage of our spectra, we use the CL01 method only as a check of the quality of empirically-calibrated metallicities. The three main advantages of checking our results with the CL01 method are: *(i)* the use of one unique calibrator for all redshift ranges; *(ii)* the different method to deal with dust extinction, i.e. the self-consistent correction instead of the use of equivalent widths; *(iii)* the different method to break the O/H degeneracy, i.e. the fit of the double-peaked PDF instead of the use of the luminosity.

Fig. 10 shows the comparison between the empirically-calibrated and the CL01 metallicities. Apart from a small number of outliers (7%), the two results are in good agreement within a dispersion of approximately 0.22 dex. This dispersion can be compared with the intrinsic dispersion of the mass-metallicity relation, which is also of ~ 0.22 dex. We have also calculated the mean of the error bars for the two methods: 0.06 dex for the empirical method, and 0.18 dex for the CL01 method.

Let us summarize all the contributions to the global dispersion. *(i)* The empirically-calibrated errors have been estimated from basic propagation of the line measurement errors. We know from tests performed on duplicated observations that *platefit_vimos* errors are not underestimated. They reflect the negligible contribution of noise to the global dispersion. *(ii)* Conversely, the CL01 errors have been estimated from the width of the PDF. They also reflect the additional dispersion due to the degeneracies in the models, which are taken into account thanks to the Bayesian approach. *(iii)* Finally the difference between the mean CL01 error and the global dispersion reflects real variations in



(a) $z < 0.4$ (b) $z > 0.4$

Figure 10. Comparison between empirically-calibrated (*x*-axis), and CL01 metallicities (*y*-axis) for our whole sample of star-forming galaxies. The red solid line shows the $y = x$ relation, while the blue dashed line is a least-square fit to the data points. The parameters of the fit are given in the two plots (*a*: slope, *b*: zero-point).

the physical parameters of galaxies, which are not taken into account in both models.

The comparison between the two methods have shown that: *(i)* the various empirical calibrators used at different redshift ranges give consistent results with the CL01 metallicities, thanks to the correction formulas provided by Kewley & Ellison (2008); *(ii)* the empirical correction for dust attenuation using equivalent widths, and the correction formula provided by Liang et al. (2007), is consistent with CL01's results, obtained directly from observed line fluxes and a self-consistent reddening correction; *(iii)* apart from a relatively small number of outliers, the O/H degeneracy is correctly broken using the luminosity diagnostic with the empirical method.

4.3. Possible biases

Fig. 10 also indirectly shows that the metallicities obtained with $N2$ and R_{23}calibrators at different redshifts are directly comparable, even if there is no galaxy in our sample where these two calibrators can be applied at the same time. Indeed we see that both $N2$ and R_{23}metallicities are in agreement with CL01 metallicities, apart from the small number of outliers. Moreover we have investigated the possible biases introduced by the use of different sets of lines, at different redshifts, with the CL01 method: we have confirmed that this method gives, as expected, comparable results at different redshifts.

Another possible bias comes from the fact that the slit used for the spectroscopy not necessarily covers all the light of the observed galaxy. We have taken the spatial extension of the galaxies calculated from photometry, and compared it to the width of the slit, i.e. one arcsecond. This gives the aperture fraction of the galaxy, expressed in percent. Fig. 11(bottom) shows the distribution of the aperture fractions for the star-forming galaxies in our

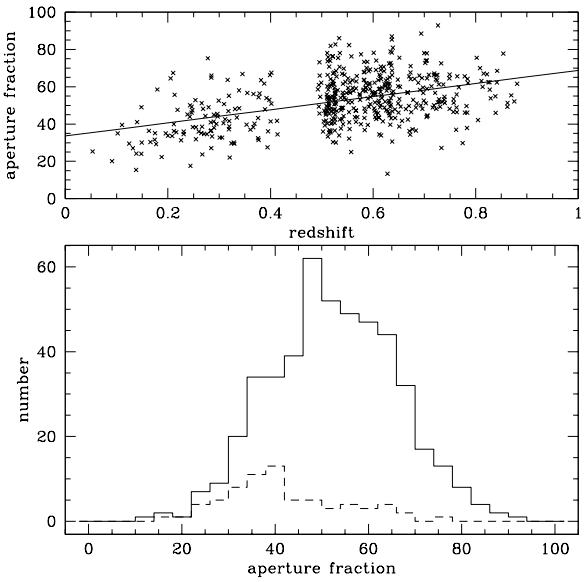


Figure 11. Bottom: distribution of the aperture fractions (in percent) for the star-forming galaxies in our sample. The dashed histogram shows the distribution for the galaxies at low redshift ($z < 0.5$) with a high metallicity ($12 + \log(\mathrm{O/H}) > 8.8$). Top: evolution of the apertures fraction as a function of redshift. The solid line is the fit to the data points.

sample (not that this information was not available for the CDFS sample). The mean aperture fraction is 52%, and we see that the large majority of objects has an aperture fraction greater than 20%, which is the minimum aperture fraction given by Kewley, Jansen, & Geller (2005) in order to minimize aperture effects on derived parameters like the metallicity or the star-formation rate.

In Fig. 11, we also observe the expected correlation between redshift and the aperture fraction, both increasing at the same time. The aperture fractions show an increase of 35% per unit-redshift. Since the aperture fractions are already high enough at low redshift not to affect the derived metallicities, and since they increase with redshift, this slope is likely to have a marginal effect on the derived evolution of metallicity as a function of redshift.

5. The luminosity-metallicity relation

5.1. Study of the derived fits

We study in this section the relation between the rest-frame *B*-band luminosity and gas-phase metallicity, for the star-forming galaxies of our wide and deep samples. Here we use the metallicities estimated with the empirical approach, and renormalized to the CL01 method (see Sect. 4.1). Figure 12 shows the luminosity-metallicity relation of the wide and deep samples in three redshift ranges: $0.0 < z < 0.5$, $0.5 < z < 0.7$ and $0.7 < z < 0.9$. The results are compared to the luminosity-metallicity relation in the local universe derived by Lamareille et al. (2004) with 2dFGRS data, and renormalized to the CL01 method. In order to do this comparison, we have done a linear fit to the data points.

As already shown before, the luminosity-metallicity relation is characterized by an non-negligible dispersion of the order of ≈ 0.25 dex (higher than the one of the mass-metallicity relation). Therefore, the method used to perform the fit has a huge impact on the results, and it is mandatory to use the same method before doing comparisons between two studies. Thus, we have used the same method than Lamareille et al. (2004), i.e. the ols bisector fit (Isobe et al. 1990), starting from the 50% *k*-correction completeness level (see Sect. 3.4). The results are shown in Fig. 12 and in Table 6.

The slope is by $\sim 2\sigma$ steeper than the one of the reference relation in the local universe, this slope being similar in all redshift ranges (≈ -0.75), except $0.5 < z < 0.7$ in the deep sample. As already shown by Lamareille et al. (2004) who have compared the slopes of the luminosity-metallicity relation for low and high metallicity objects, this slope is steeper for high metallicity objects. This effect has been confirmed by Lee et al. (2006) who have extended the mass-metallicity relation of Tremonti et al. (2004) to lower masses and metallicities and found a flatter slope. We emphasize that we are not discussing here the saturation at $12 + \log(\mathrm{O/H}) \approx 9.2$ as observed by Tremonti et al. (2004), and which might be confused with a flatter slope at very high metallicity.

Thus, given that we expect a steeper slope at high metallicity, we analyze our results as a lack of data points in the low metallicity region: the wide sample does not go deep enough to detect low-metallicity and low-luminosity objects above the completeness limit. Moreover, in the $0.0 < z < 0.5$ redshift range of wide and deep samples, the S/N¿4 cut on emission lines introduces a bias towards high metallicity objects: the faint and blended [NII]λ6584 line, used to compute metallicities in this redshift range, becomes rapidly undetectable at low metallicities.

The apparently high number of galaxies with low redshift, low luminosity and high metallicity seen in Fig. 12(left) may be explained by two effects. First, as said in previous paragraph, the [NII]λ6584 line is more likely to be observed for high metallicity objects. There is consequently a lack of objects with low luminosity and low metallicity which makes the distribution apparently wrong (as compared to previous studies). Second, these objects at low redshift may have lower aperture fractions, which may introduce a bias towards higher metallicities given that we preferentially observe their central part. Indeed, we have plot the distribution of the aperture fractions for objects with a low redshift and a high metallicity in Fig. 11(bottom, dashed histogram): they are clearly lower, with a mean of 40%.

We recognize that the results obtained by doing a direct linear fit to data points suffer from a drawback: the correlations between metallicity and luminosity shown in Fig. 12 are weak. This weakness is characterized by the non-negligible error bars for the slopes and the zero-points of the relations provided in Table 6, and also by low Spearman correlation rank coefficients: of the order of -0.3 for the bottom-center panel, and-0.1 for the other panels.

5.2. Global metallicity evolution

Despite the weakness of the correlations found in our data, we know from previous studies that the luminosity-metallicity exists. Thus we can use the existence of this relation as an assumption and find new results.

In order to quantify the evolution of the luminosity-relation, we can also derive the mean evolution of metallicity. This is done with the additional assumption that the slope of the luminosity-relation remains constant slope at zero-order. The results are shown in Fig. 12 and in Table 6. As expected, the evolution is barely significant in the $0.0 < z < 0.5$ redshift range. In the $0.5 < z < 0.7$ and $0.7 < z < 0.9$ redshift ranges, the evolution is stronger. It is similar in the wide sample to what has been found by Lamareille et al. (2006a). In the $0.5 < z < 0.7$ redshift range of the deep sample, the results are similar with both methods, which confirms the hypothesis that a steeper slope is found only when low-metallicity points are not included in the fit.

Finally, the comparison of the wide and deep samples show a stronger evolution of the metallicity with redshift in the wide sample. This result tells us that the slope *does not actually remain constant*, and that the evolution of the metallicity is stronger in more luminous objects. At $z \sim 0.76$, galaxies with an absolute *B*-band magnitude of ~ -20.1 have -0.32 dex lower metallicities than galaxies of similar luminosities in the local universe, while galaxies with an absolute *B*-band magnitude of ~ -21.2 have -0.65 dex lower metallicities.

We have checked that the results are stable when breaking the O/H degeneracy at high redshift with a different reference relation (e.g. lowered in metallicity). As stated before by Lamareille et al. (2006a), the use of a lower reference relation to break the degeneracy only changes the metallicities of the galaxies in the intermediate region ($12 + \log(\mathrm{O/H}) \approx 8.3$), thus non-significantly changing the whole luminosity-metallicity relation.

We have also checked the effect of not including the candidate star-forming galaxies in the fit. As stated in Sect. 2.2, the contamination of candidate star-forming galaxies by AGNs is less than 1%, and is thus not expected to affect significantly the luminosity-metallicity relation. Therefore, the results of not including the candidate star-forming galaxies are only given as information in Table 6. We have also checked the effect of including candidate AGNs. In both cases, the effect on the luminosity-metallicity relation is very small.

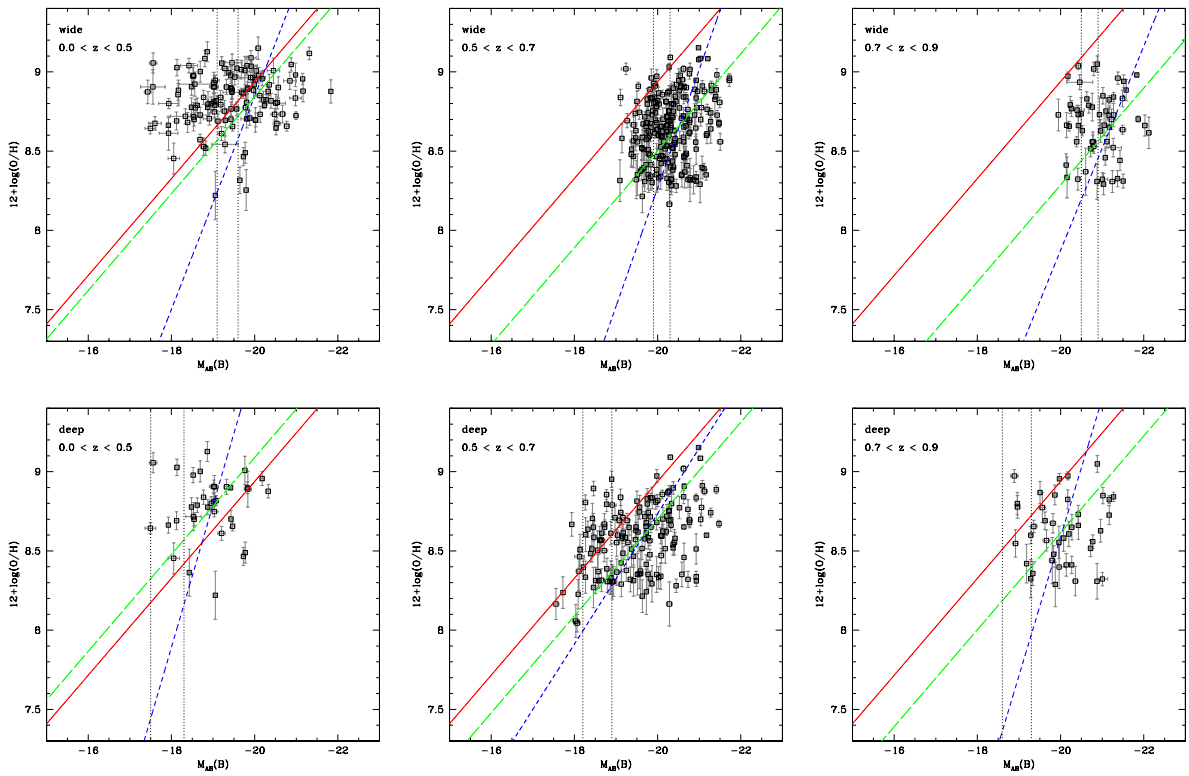


Figure 12. Rest-frame *B*-band luminosity-metallicity relation for the wide (top) and deep (bottom) samples, for three redshift ranges: from left to right $0.0 < z < 0.5$, $0.5 < z < 0.7$ and $0.7 < z < 0.9$. The metallicities have been estimated using the empirical approach (see Sect. 4.1). The solid line shows the luminosity-metallicity relation at low redshift derived by Lamareille et al. (2004), and renormalized to the CL01 method. The short-dashed line shows the fit to the data points using the ols bisector method (see text). The long-dashed lines shows the fit to the data points assuming a constant slope.

Table 6. Evolution of the luminosity-metallicity relation for the wide and deep samples. The reference relation is the one obtained by Lamareille et al. (2004) with 2dFGRS data and renormalized to the CL01 method. In each redshift ranges, we give the results of the ols bisector fit (slope and zero-point) and of the constant-slope fit (mean shift, see Fig. 12), together with the mean redshift and magnitude, and the dispersion of the relation. The metallicity shift is given in three cases: *a)* using only star-forming galaxies, *b)* adding candidate star-forming galaxies, and *c)* adding also candidate AGNs.

sample	slope	zero-point	z	$M_{\rm AB}(B)$	$\Delta\log({\rm O/H})^L$			rms
					a	*b*	*c*	
reference	-0.31	2.83						
wide								
$0.0 < z < 0.5$	-0.67 ± 0.17	-4.60 ± 3.4	0.31	-19.99	*−0.08*	-0.10	*−0.10*	0.23
$0.5 < z < 0.7$	-0.75 ± 0.08	-6.67 ± 1.6	0.59	-20.53	*−0.39*	-0.44	*−0.45*	0.23
$0.7 < z < 0.9$	-0.64 ± 0.21	-5.01 ± 4.4	0.77	-21.22	*−0.58*	-0.65	*−0.70*	0.24
deep								
$0.0 < z < 0.5$	-0.90 ± 0.45	-8.28 ± 8.61	0.30	-18.99	*0.17*	0.15	*0.15*	0.28
$0.5 < z < 0.7$	-0.41 ± 0.06	0.50 ± 1.29	0.59	-19.67	*−0.20*	-0.24	*−0.25*	0.27
$0.7 < z < 0.9$	-0.87 ± 0.41	-8.89 ± 8.20	0.76	-20.05	*−0.24*	-0.32	*−0.39*	0.29

We need to disentangle between the two effects which make these objects more luminous, i.e. a higher mass or a lower mass-to-light ratio, which one is responsible of the stronger evolution of the metallicity. Thus, we now study the mass-metallicity relation in the next section.

6. The mass-metallicity relation

6.1. Global metallicity evolution

We now derive the mass-metallicity relation for the star-forming galaxies of the wide and deep samples, using the empirical approach for computing the metallicities, and in three redshift ranges: $0.0 < z < 0.5$, $0.5 < z < 0.7$ and $0.7 < z < 0.9$. The results are shown in Fig. 13 and Table 7, and are compared to the reference mass-metallicity relation in the local universe, derived by Tremonti et al. (2004) with SDSS data. The reference relation has been shifted in mass, in order to take into account the effect of using different models (see Sect. 3.2).

As in Fig. 12, the data points shown in Fig. 13 do not show strong Spearman rank correlation coefficients. We thus skip the step of doing a fit to these data points. We directly probe the global evolution in metallicity of star-forming galaxies compared to the reference relation, doing the likely assumption that

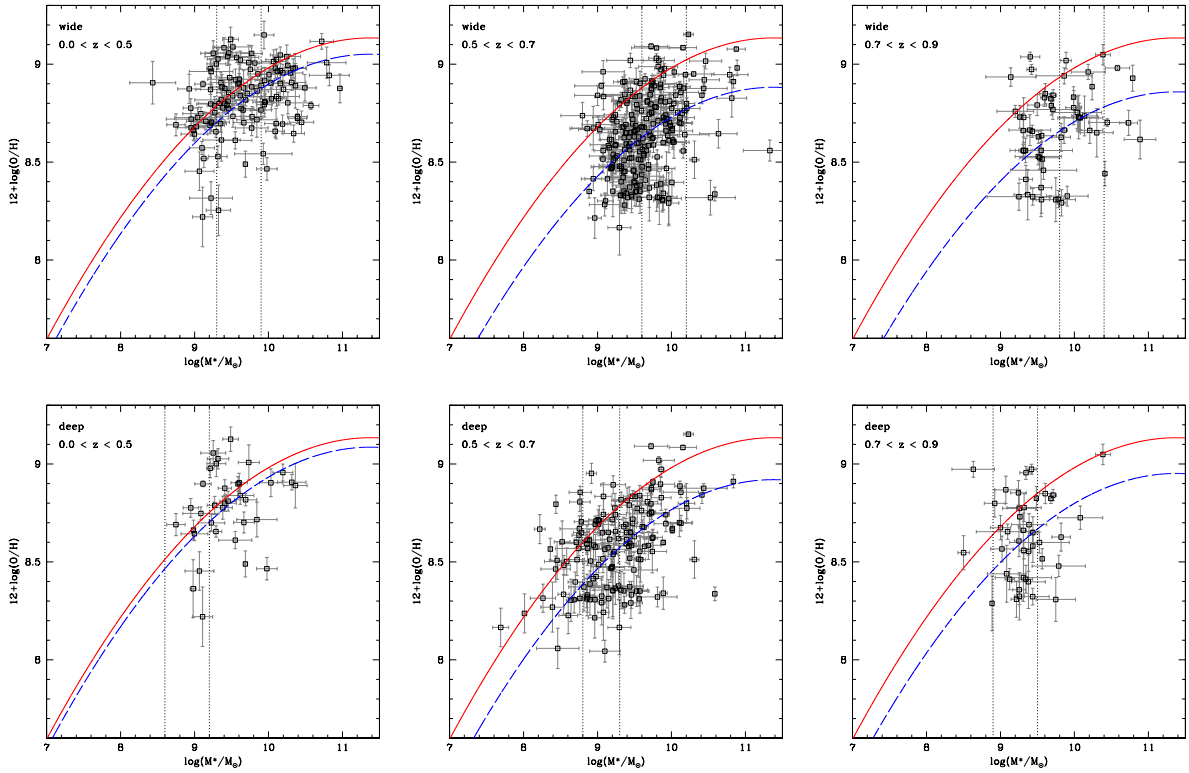


Figure 13. The mass-metallicity relation of star-forming galaxies for the wide (top) and deep (bottom) samples, for three redshift ranges: from left to right $0.0 < z < 0.5$, $0.5 < z < 0.7$ and $0.7 < z < 0.9$. The metallicities have been estimated using the empirical approach (see Sect. 4.1). The solid curve shows the mass-metallicity relation at low redshift derived by Tremonti et al. (2004). The long-dashed curves shows the fit to the data points assuming that the SDSS curve is only shifted down in metallicity.

Table 7. Evolution of the mass-metallicity relation for the wide and deep samples. The reference relation is the one obtained by Tremonti et al. (2004) with SDSS data and the CL01 method, renormalized in mass. In each redshift ranges, we give the mean redshift, stellar mass, and metallicity shift (assuming that the shape of the relation remains constant, see Fig. 13), and the dispersion of the relation. The metallicity shift is given in three cases: *a)* using only star-forming galaxies, *b)* adding candidate star-forming galaxies, and *c)* adding also candidate AGNs.

sample	z	$\lg(M_\star)$	$\Delta\log(\mathrm{O/H})^M$			rms
			a	*b*	*c*	
wide						
$0.0 < z < 0.5$	0.30	9.87	*−0.08*	−0.08	*−0.09*	0.17
$0.5 < z < 0.7$	0.59	9.97	*−0.22*	−0.25	*−0.26*	0.20
$0.7 < z < 0.9$	0.78	10.19	*−0.23*	−0.28	*−0.36*	0.19
deep						
$0.0 < z < 0.5$	0.29	9.45	*−0.04*	−0.05	*−0.05*	0.20
$0.5 < z < 0.7$	0.59	9.45	*−0.17*	−0.21	*−0.23*	0.19
$0.7 < z < 0.9$	0.76	9.40	*−0.12*	−0.18	*−0.23*	0.20

the mass-metallicity relation exists also at high redshift. To do so, we calculate the mean shift in metallicity by fitting to the data points the same curve than Eq. 3 of Tremonti et al. (2004), allowing only a different zero-point. The zero-order assumption is indeed that the shape of the mass-metallicity relation does not vary with redshift. The results are shown in Fig. 13 and in Table 7. The fit is performed above the 50% mass-to-light completeness level.

There are some caveats in the interpretation of the mass and metallicity evolution of the deep sample, because of selection and statistical effects. The mean observed stellar seams to decrease with redshift, which is in contradiction of what one would expect from the Malmquist bias. First, we find higher stellar masses than expected in the lowest redshift bin. The sample is actually not complete down to the limiting mass: lower mass galaxies would have lower metallicities, and low metallicities are indeed difficult to measure because of the blending of [NII]λ6584 and Hα lines. Conversely, the highest redshift bin has a lower mean mass than expected, which is more probably due to a statistical effect because of the small solid angle of the deep sample. This later effect also explains why the metallicity evolution seems smaller in the highest redshift bin.

As observed on the luminosity-metallicity relation in previous section, we clearly see a stronger metallicity evolution in the wide sample than in the deep sample. The wide and deep samples span interestingly different ranges in masses, but are otherwise identical. This effect thus shows that the most massive galaxies have experienced the most significant evolution in metallicity. At $z \sim 0.77$, galaxies at $10^{9.4}$ solar masses have -0.18 dex lower metallicities than galaxies of similar masses in the local universe, while galaxies at $10^{10.2}$ solar masses have -0.28 dex lower metallicities. We therefore conclude that the shape of the mass-metallicity relation varies with redshift, so that it was flatter in an earlier universe.

We also remark that like for the luminosity-metallicity relation, the potential influence of candidate AGNs or candidate star-forming galaxies is negligible (see Table 7).

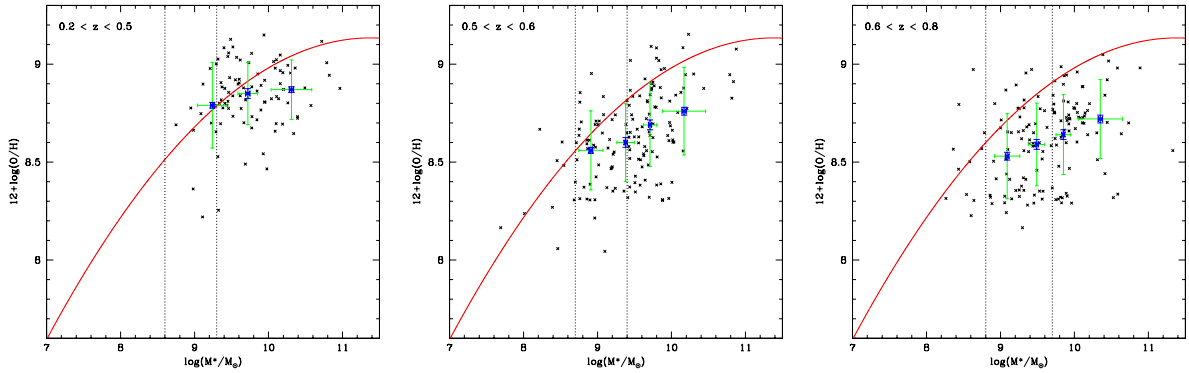


Figure 14. The mass-metallicity relation of star-forming galaxies for three redshift ranges: from left to right $0.2 < z < 0.5$, $0.5 < z < 0.6$ and $0.6 < z < 0.8$. The metallicities have been estimated using the empirical approach (see Sect. 4.1). The solid curve shows the mass-metallicity relation at low redshift derived by Tremonti et al. (2004). We show the mean metallicities by bins of stellar masses. The blue error bars represent the uncertainty on the mean, while the green error bars represent the dispersion of the data points. The 50% mass-to-light completeness levels for the deep and wide samples (see Table 4) are shown as vertical dotted lines.

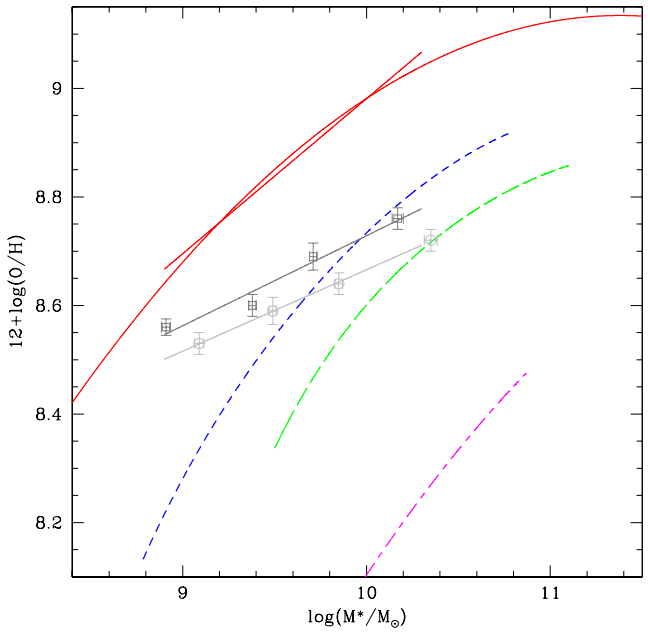


Figure 15. The mass-metallicity relation of star-forming galaxies for two redshift ranges: $0.5 < z < 0.6$ (dark grey) and $0.6 < z < 0.8$ (light grey). The metallicities have been estimated using the empirical approach (see Sect. 4.1). The solid curve shows the mass-metallicity relation at low redshift derived by Tremonti et al. (2004). This relation has also been linearized in the range $8.9 < \log(M_\star/M_\odot) < 10.3$. The short-dashed curve is the relation derived by Savaglio et al. (2005) at $z \sim 0.7$, the long-dashed line is the relation derived by Erb et al. (2006) at $z \sim 2.0$, and the long-dashed–short-dashed line is the relation derived by Maiolino et al. (2008) at $z \sim 3.5$.

6.2. Evolution of the shape of the mass-metallicity relation

We now evaluate the evolution of the shape of the mass-metallicity relation as a function of redshift, by co-adding a number of data points together in bins of mass, thus increasing the signal-to-noise ratio. Figure 14 shows the results in three redshift ranges: $0.2 < z < 0.5$, $0.5 < z < 0.6$ and $0.6 < z < 0.8$. The mean redshifts are respectively $z \sim 0.33$, $z \sim 0.54$ and $z \sim 0.67$. The mean masses are respectively $10^{9.75}$, $10^{9.54}$ and $10^{9.70}$ solar masses. We have joined for this study the wide and deep samples: only the deep sample is used from its 50% completeness limit to the 50% completeness of the wide sample, above which both samples are used.

The mean masses and metallicities have been estimated in respectively 3, 4, and 4 stellar mass bins. This approach is similar to the one used at low redshift by Tremonti et al. (2004). The bins are not equally spaced in mass,but always contain a similar number of data points, i.e. ~ 45 in our case. We have set two consecutive mass bins to have 25% of their data points in common. We evaluate the errors on the means and the dispersions of the data points, which are both shown in Fig. 14. The errors on the means are very low thanks to the number of available data points. The dispersion in metallicity stays almost constant and equal to ≈ 0.22 dex. All results are calculated as the median of 1000 bootstrap estimates.

Fig. 15 shows the evolution of the slope in the two last redshift bins. We confirm that the mass-metallicity relation tends to become flatter at higher redshifts: the slope is 0.166 ± 0.02 dex/decade at $z \sim 0.54$ and 0.150 ± 0.01 dex/decade at $z \sim 0.67$, which has to be compared to the slope 0.29 dex/decade of the Tremonti et al. (2004) curve linearized in the same mass range. The slope evolves by -0.25 dex/decade by unit of redshift in the range $0 < z < 0.7$. These results also confirm the assumption, used in previous section, that the mass-metallicity relation exists at high redshift.

6.3. Comparison with previous works

Fig. 15 shows also the comparison between our results, and other studies performed at high redshifts (Savaglio et al. 2005; Erb et al. 2006; Maiolino et al. 2008). All curves have been renormalized in stellar masses using the shifts summarized in Table. 8. Metallicities derived by Savaglio et al. (2005), Erb et al. (2006) and Maiolino et al. (2008) have been converted respectively from Kobulnicky & Kewley (2004), Pettini & Pagel (2004) $N2$, and Kewley & Dopita (2002) methods to the CL01 method.

The mass-metallicity relation at redshift $z \sim 0.7$ is given by Eq. 8 of Savaglio et al. (2005), shifted by 0.47 dex in stellar mass. The mass-metallicity relation at redshift at redshift $z \sim 2.0$ is given by Eq. 3 of Tremonti et al. (2004), shifted by -0.56 dex

Table 8. Different setups for different mass-metallicity relations found in the literature, and associated shifts to be applied to their stellar masses in order for them to be comparable with our results. This table shows the type of data, the models, the presence or not of secondary bursts, and the IMF used to compute the stellar masses. The last column gives the global shift that has to be applied to the logarithm of the stellar mass.

References†	data⋆	model	bursts	IMF†	sum
our study	P+S	CB07	yes	C03	
T04	S	BC03	yes	K01	
	+0.00	−0.07	-	−0.056	−0.126
S05	P	Pégase	yes	BG03	
	+0.05	−0.09	-	+0.024	−0.016
E06	P	BC03	no	C03	
	+0.05	−0.09	+0.14	-	+0.1
M08	P	BC03	no	S55	
	+0.05	−0.09	+0.14	−0.232	−0.132

⋆S stands for spectroscopy, P stands for photometry.
†T04: Tremonti et al. (2004); S05: Savaglio et al. (2005); E06: Erb et al. (2006); M08: Maiolino et al. (2008); C03: Chabrier (2003); K01: Kroupa (2001); BG03: Baldry & Glazebrook (2003); S55: Salpeter (1955).

in metallicity, as found by Erb et al. (2006). The mass-metallicity relation at redshift $z \sim 3.5$ is given by Eq. 2 of Maiolino et al. (2008), with the parameters given in their Table 5.

Contrary to Savaglio et al. (2005), we find a flatter slope of the mass-metallicity relation at $z \sim 1$. Nevertheless, the comparison of data points shows that their results and ours are in good agreement for the highest mass bins. The larger difference comes from the lowest mass bins, in which they are probably not complete. We know indeed that lower mass-to-light ratio galaxies are preferentially observed in the lowest incomplete mass bins, and that such galaxies show smaller mean metallicities (Ellison et al. 2008).

The comparison with the data of Erb et al. (2006), which are taken at $z \sim 2$, is less straightforward: there is a fairly good agreement in metallicity with our high-mass end data, but actually at a rather different redshift. This could mean that there have been very little metallicity evolution from $z \sim 2$ to $z \sim 1$, but this would be hard to understand when looking also at the other results. The metallicity evolves indeed strongly from $z \sim 3.5$ (Maiolino et al. 2008), and between $z \sim 1$ and the local universe (our data and Savaglio et al. 2005). Nevertheless we note that this later discrepancy with Erb et al. (2006) results may be understood: according to the downsizing scenario, the evolution of the most massive galaxies plotted here should be actually smaller between $z = 2$ and $z = 1$ than between $z = 1$ and $z = 0$. Among others, Pérez-González et al. (2008) have quantified that galaxies below $\log(M_\star/M_\odot) = 11.5$ have formed half of their stars at $z < 1$. The main reason of a possible overestimate of Erb et al. (2006) metallicities is probably statistical variation effects. They have indeed based their results on stacked spectra of very few galaxies. The effect of the selection function is therefore difficult to analyze.

The type of galaxies observed by Erb et al. (2006) at $z \sim 2$, which are active galaxies, may also have later evolved to "red and dead" passive galaxies, and got higher metallicities than the ones actually observed by us or by Tremonti et al. (2004) at lower redshifts. Such dead galaxies would unfortunately not satisfy any more the selection function of any work based on emission-line measurements, and would not be observed.

6.4. Evolution of the mass-to-light ratio

The metallicity evolution on the wide sample is very similar to what has been found by Lamareille et al. (2006a) on the luminosity-metallicity relation, after having applied a correction for luminosity evolution (Ilbert et al. 2005). This tells us that the stronger evolution of the luminosity-metallicity relation, compared to the mass-metallicity relation, is effectively due to an additional luminosity evolution. This luminosity evolution has

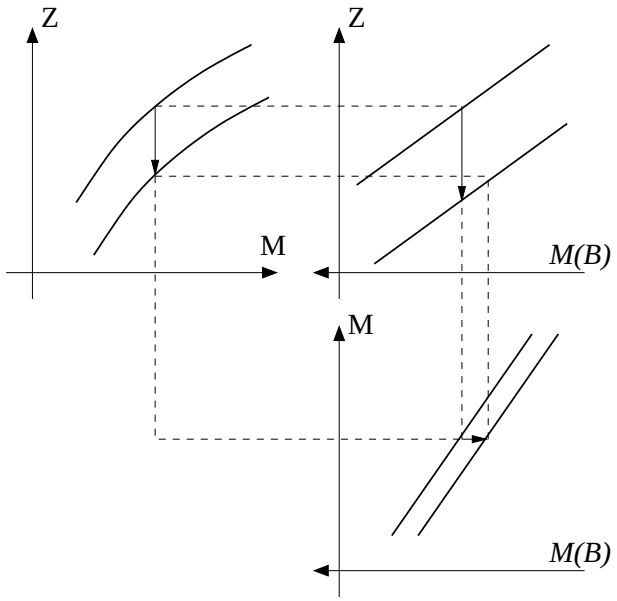


Figure 16. This plot shows, for galaxies with similar stellar masses, their evolution when redshift increases: for metallicity in the mass-metallicity plane (top-left), for metallicity in the luminosity-metallicity plane (top-right), and for luminosity in the mass-luminosity plane (bottom-right).

Table 9. Evolution of the mass-to-light ratio for the wide and deep samples. This evolution is computed as an absolute rest-frame *B*-band magnitude evolution at constant stellar mass, from the comparison of the luminosity-metallicity (see Fig. 12 and Table 6) and mass-metallicity (see Fig. 13 and Table 7) relations.

	$\Delta M_{\rm AB}(B)$	
redshift range	*wide*	*deep*
$0.0 < z < 0.5$	−0.06	0.65
$0.5 < z < 0.7$	−0.61	−0.10
$0.7 < z < 0.9$	−1.19	−0.45

to be understood for galaxies with similar masses, which means we can measure the evolution of the mass-to-light ratio in galaxies using metallicity as a pivot.

Fig. 16 shows schematically the global evolution of metallicity in the luminosity-metallicity plane (top-right). When there is both an evolution of the metallicity in the mass-metallicity plane (top-left), and an evolution of the luminosity in the mass-luminosity plane (bottom-right), galaxies being more luminous at a given mass, the result is a stronger evolution of the metallicity in the luminosity-metallicity plane.

The luminosity evolution at constant mass of our galaxies can be calculated using the following formula:

$$\Delta M_{\rm AB}(B) = a^{-1} \times \left(\Delta \log({\rm O/H})^M - \Delta \log({\rm O/H})^L\right) \quad (6)$$

where a is the slope of the luminosity-metallicity relation, i.e. -0.31 dex/mag. The results are given in Table 9. The positive evolution for the lowest redshift bin in the deep sample is due to the incompleteness of the $N2$ calibration: at a given metallicity, our sample is biased towards higher luminosity. In the other redshift ranges and the whole wide sample, we see that galaxies at higher redshifts have lower mass-to-light ratios than today, and that this evolution is more significant for massive galaxies (i.e. the wide sample compared to the deep). These two results are in good agreement with the general scenario of downsizing (Cowie et al. 1996) for the evolution of the star formation rates in galaxies. Less massive galaxies show a smaller evolution in their star formation rate at $z < 1$ since they are still actively forming stars.

It will be further analyzed in a subsequent paper of this series.

6.5. Derived star formation rates

We now discuss the evolution of metallicity at constant stellar mass, in terms of star formation rates.

Assuming constant star formation rates (SFR) and the closed-box model, one can calculate a stellar mass evolution and relates it to a metallicity evolution using the following equations:

$$M_\star(t) = M_\star(0) + SFR \times t \quad (7)$$

$$M_\star(t) + M_g(t) = M_{\rm tot} \quad (8)$$

$$Z(t) = y \times \ln\left(1 + \frac{M_\star(t)}{M_g(t)}\right) \quad (9)$$

where $M_\star$, M_g and $M_{\rm tot}$ are respectively the stellar mass, the gas mass, and the total baryonic mass of the galaxy (which remain constant in the closed-box model); and where Z and y are respectively the metallicity and the true yield (which depends only on the stellar initial mass function).

We know the mass-metallicity relation at redshift $z = 0$ (Tremonti et al. 2004), and the Table 7 gives the metallicity evolution for various stellar masses and redshifts. Thus, we can revert Eq. 7, 8 and 9 to derive the star formation rate which explains the observed values. The results are quoted in Table 10 (case *a*). They are calculated for a given cosmology, and for a true yield $y = 0.0104$ (Tremonti et al. 2004).

A better modeling can be performed by taking into account the time evolution of the star formation rates in galaxies. We may for example assume an exponentially decreasing star formation rate, i.e. $SFR(t) = SFR(0) \times \exp(-t/\tau)$, where τ gives the characteristic *e*-folding time of the galaxy. The exponentially decreasing law is a good choice when analyzing a population of galaxies, with respect to the global cosmic evolution of star formation rate. Eq. 7 is then replaced by the following formula:

$$M_\star(t) = M_\star(0) + SFR(0) \times \tau \times \left(1 - e^{-t/\tau}\right) \quad (10)$$

Table 10 (case *b*) gives the results obtained when assuming the relation given by Eq. 12 of Savaglio et al. (2005) between the *e*-folding time and the total baryonic mass of the galaxy. Comparing cases *a* and *b*, we see that the results are not dramatically affected by the assumption of a decreasing SFR. The star formations activities in case *b* are systematically higher than in case *a*, which is expected as with a decreasing SFR a higher initial value is needed to explain the same metallicity evolution, as compared to a constant SFR.

We note also that the derived star formation activities in the wide and deep samples are very close, despite their different ranges in stellar masses. In case *a*, star formation activities in the deep sample are nevertheless higher than in the wide sample, which is expected as the wide sample spans higher stellar masses. Conversely, in case *b*, we find higher star formation activities for more massive objects. This comes from the assumption, made by Savaglio et al. (2005) and used in our equations, that less massive galaxies have a longer *e*-folding times. Consequently the less massive galaxies of the deep sample, which are assigned long *e*-folding times ans thus can be approximated as constant SFR, do not vary much from case *a* to case *b* while the more massive galaxies of the wide sample, which are assigned short *e*-folding times, are more affected by the decreasing SFR hypothesis.

Preliminary results on the evolution of star formation activities of VVDS galaxies, which will be presented in a subsequent paper **(see also Walcher et al. 2008)**, give at redshift $z \approx 0.6$ (our best sampled redshift bin) the following numbers: $\approx 10^{-9.8}$ yr^{-1} for galaxies at $\approx 10^{9.5}$ solar masses (mean mass of the deep sample), and $\approx 10^{-10.3}$ yr^{-1} for galaxies at $\approx 10^{10}$ solar masses (mean mass the wide sample).

Assuming the closed-box hypothesis, our results in the wide sample are in fairly good agreement with the observed star formation activities at similar masses, this agreement being better in case *b*. But in the deep sample, the inferred star formation activities which explain our observed metallicity evolution seem significantly underestimated by a factor ≈ 5. Although the varying *e*-folding times hypothesis is in agreement with the slower evolution of the mass-to-light ratio found in the deep sample in previous section, it fails in providing a good agreement between our inferred star formation activities in the deep sample and observed values at similar masses. Moreover this hypothesis make more massive galaxies forming stars more actively, which is in contradiction with observations.

We therefore conclude that the closed-box hypothesis is not valid to explain our observed metallicity evolutions (assuming that the preliminary results mentioned above are confirmed).

7. Conclusion

We have calculated rest-frame luminosities, stellar masses and gas-phase oxygen abundances of a statistically significant sample of star-forming galaxies, selected from the VIMOS VLT Deep Survey. This has allowed us to derive luminosity-metallicity and mass-metallicity relations in various redshift ranges up to $z \sim 0.9$. These relations have also been derived in two sub-samples: the deep sample which extends to lower observed luminosities thanks to a deeper magnitude selection, and the wide sample which extends to higher luminosities thanks to

Table 10. Derived star formation activities needed to explain the observed values quoted in Table 7, assuming the closed-box model, and $h = 0.7$, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$ cosmology. The values are given in the wide and deep samples, and for two different assumptions: *a)* constant star formation rate, *b)* exponentially declining star formation rate (see text). In case *b*, each value is the instantaneous star formation rate at the specified redshift (see Eq. 10).

mean redshift	$\log(\mathrm{SFR}/M_\star)$			
	wide		*deep*	
	a	*b*	*a*	*b*
0.3	-11.06 ± 0.7	-10.93 ± 0.7	-11.05 ± 0.9	-10.99 ± 0.9
0.6	-10.64 ± 0.5	-10.36 ± 0.5	-10.51 ± 0.4	-10.40 ± 0.5
0.77	-10.73 ± 0.4	-10.27 ± 0.5	-10.65 ± 0.5	-10.53 ± 0.5

a larger solid angle. The selection function have been taken into account in order to define volume- and mass-limited samples.

For both the wide and the deep samples, we have studied as a function of redshift: the evolution in metallicity at constant luminosity, the evolution in metallicity at constant stellar mass, the associated evolution in luminosity at constant stellar mass, and the evolution of the slope of the mass-metallicity relation.

Additionally, we have also:

- Measured the fraction of star-forming galaxies and AGNs as a function of redshift.
- Calculated the difference in terms of stellar masses between the old BC03 and new CB07 stellar population models.
- Studied the effect on the computation of stellar mass of the use of spectral indices rather than only photometric points.
- Checked that gas-phase oxygen abundance measurements in low and high redshift ranges are comparable even if calculated with different methods.
- Studied the marginal effect of candidate star-forming galaxies and candidate AGNs on the derived luminosity-metallicity or mass-metallicity relations.

In the wide sample, the mean evolution of metallicity at constant luminosity, and of luminosity at constant stellar mass, are in good agreement with previous studies by e.g. Lamareille et al. (2006a); Ilbert et al. (2005); Hammer et al. (2005) among others.

By doing the assumption that the slope of the luminosity-metallicity relation, or the shape of the mass-metallicity relation, remain constant with redshift, we finally found different metallicity evolutions for the wide and deep samples. The wide and deep samples span different ranges in masses, which shows that the assumption is wrong and that this slope or shape actually evolves with redshift. We have found that the most massive galaxies show the strongest metallicity evolution. At $z \sim 0.77$, galaxies at $10^{9.4}$ solar masses have -0.18 dex lower metallicities than galaxies of similar masses in the local universe, while galaxies at $10^{10.2}$ solar masses have -0.28 dex lower metallicities.

We have then studied the mass-metallicity relation on co-added data points by bins of stellar masses. The inferred slopes show an evolution to a flatter mass-metallicity relation at $z \sim 1$ as compared to the local universe, which confirms that higher mass galaxies show a stronger metallicity evolution during this period. The slope evolves by -0.25 dex/decade by unit of redshift in the range $0 < z < 0.7$.

Moreover, we have inferred from our observations the star formation activities which explain this evolution in metallicity. To do so, we have assumed the closed-box hypothesis. We have found that the observed metallicity evolution would be explained by similar star formation activities at any mass, which is in contradiction with the well-known correlation between stellar mass and star formation activity. Assuming preliminary observations of the star formation activities in VVDS galaxies, we conclude that our results are underestimated by a factor ≈ 5 in the deep sample, while their are in fair agreement in the wide sample.

These results make a strong evidence against the closed-box model. Indeed, we know from previous studies that the star formation activity of galaxies decreases while stellar mass increases. The only way to explain the smaller metallicity evolution of the less massive galaxies is thus to replace in Eq. 9 the true yield by a smaller effective yield. Assuming a smaller effective yield allows galaxies to show smaller metallicity evolution, even with a stronger star formation activity.

The smaller effective yield of less massive galaxies can be understood in the "open-closed" model. In this model, galaxies with small stellar masses evolve like open-boxes: most of the metals produced during star formation are ejected in the intergalactic medium by stellar winds and supernovae feedback, the effective yield is thus very small. Conversely, galaxies with high stellar masses evolve in the open-closed model like closed-boxes: the metals are retained in the galaxy thanks to a high gravitational potential, the effective yield is thus close to the true yield.

The dependence of the effective yield with the gravitational potential of the galaxies, itself related to their total baryonic mass, has been already shown by Tremonti et al. (2004). It is naturally explained in the hierarchical galaxy formation scenario: in contrary to the closed-box model, the total baryonic mass does not remain constant and increases with time, together with the stellar mass, thanks to galaxy merging and accretion. This relation between the stellar mass and the total baryonic mass is at the origin of the mass-metallicity relation in the open-closed model, and is clearly supported by our data.

In the open-closed model, the smaller metallicity evolution of less massive is naturally explained despite their stronger star formation activities.

We emphasize that Eq. 7, 8 and 10 are not valid any more in the open-closed model. It is not enough to replace, in Eq. 9, the true yield by the effective yield in order to infer the right star formation rates which explain the observed metallicity evolution. These equations must be modified to take also the difference between the true and the effective yield into account. The way to modify these equations depends on the model used to explain smaller effective yields (e.g. gas loss). Replacement equations can be found in other studies (e.g. Erb et al. 2006; Erb 2008; Dalcanton 2007; Finlator & Davé 2008). We also refer the reader to subsequent papers of this series, for a more detailed discussion of the relation between the present results and the evolution of the star formation rates of galaxies, and for comparisons with simulations (e.g. De Lucia et al. 2004).

Acknowledgements. We thank C. Tremonti for having made the original *platefit* code available to the VVDS collaboration. F. Lamareille thanks the Osservatorio di Bologna for the receipt of a post-doctoral fellowship.

This work has been partially supported by the CNRS-INSU and its Programme Nationaux de Galaxies et de Cosmologie (France), and by Italian Ministry (MIUR) grants COFIN2000 (MM02037133) and COFIN2003 (num.2003020150) and by INAF grants (PRIN-INAF 2005).
The VLT-VIMOS observations have been carried out on guaranteed time (GTO) allocated by the European Southern Observatory (ESO) to the VIRMOS consortium, under a contractual agreement between the Centre National de la Recherche Scientifique of France, heading a consortium of French and Italian institutes, and ESO, to design, manufacture and test the VIMOS instrument.
This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.

References

Arnouts, S., Vandame, B., Benoist, C., et al. 2001, A&A, 379, 740
Baldry, I. K. & Glazebrook, K. 2003, ApJ, 593, 258
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Brodie, J. P. & Huchra, J. P. 1991, ApJ, 379, 157
Brooks, A. M., Governato, F., Booth, C. M., et al. 2007, ApJ, 655, L17
Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000
Chabrier, G. 2003, PASP, 115, 763
Charlot, S. & Fall, S. M. 2000, ApJ, 539, 718
Charlot, S. & Longhetti, M. 2001, MNRAS, 323, 887
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
Dalcanton, J. J. 2007, ApJ, 658, 941
Davé, R. & Oppenheimer, B. D. 2007, MNRAS, 374, 427
De Lucia, G., Kauffmann, G., & White, S. D. M. 2004, MNRAS, 349, 1101
de Rossi, M. E., Tissera, P. B., & Scannapieco, C. 2007, MNRAS, 374, 323
Denicoló, G., Terlevich, R., & Terlevich, E. 2002, MNRAS, 330, 69
Ellison, S. L., Patton, D. R., Simard, L., & McConnachie, A. W. 2008, ApJ, 672, L107
Eminian, C., Kauffmann, G., Charlot, S., et al. 2008, MNRAS, 384, 930
Erb, D. K. 2008, ApJ, 674, 151
Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, ApJ, 644, 813
Ferland, G. J. 2001, PASP, 113, 41
Finlator, K. & Davé, R. 2008, MNRAS, 385, 2181
Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41
Garilli, B., Le Fèvre, O., Guzzo, L., et al. 2008, A&A, 804, accepted (astro-ph/08044568)
Garnett, D. R., Shields, G. A., Skillman, E. D., Sagan, S. P., & Dufour, R. J. 1997, ApJ, 489, 63
Gavignaud, I., Bongiorno, A., Paltani, S., et al. 2006, A&A, 457, 79
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ, 600, L93
Hammer, F., Flores, H., Elbaz, D., et al. 2005, A&A, 430, 115
Ilbert, O., Tresse, L., Zucca, E., et al. 2005, A&A, 439, 863
Iovino, A., McCracken, H. J., Garilli, B., et al. 2005, A&A, 442, 423
Isobe, T., Feigelson, E. D., Akritas, M. G., & Babu, G. J. 1990, ApJ, 364, 104
Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003a, MNRAS, 346, 1055
Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003b, MNRAS, 341, 33
Kewley, L. J. & Dopita, M. A. 2002, ApJS, 142, 35
Kewley, L. J. & Ellison, S. L. 2008, AJ, 801, accepted (astro-ph/08011849)
Kewley, L. J., Heisler, C. A., Dopita, M. A., & Lumsden, S. 2001, ApJS, 132, 37
Kewley, L. J., Jansen, R. A., & Geller, M. J. 2005, PASP, 117, 227
Kobulnicky, H. A. & Kewley, L. J. 2004, ApJ, 617, 240
Kobulnicky, H. A. & Phillips, A. C. 2003, ApJ, 599, 1031
Kobulnicky, H. A., Willmer, C. N. A., Phillips, A. C., et al. 2003, ApJ, 599, 1006
Köppen, J., Weidner, C., & Kroupa, P. 2007, MNRAS, 375, 673
Kroupa, P. 2001, MNRAS, 322, 231
Lamareille, F. 2006, Ph.D. Thesis, http://tel.archives-ouvertes.fr/tel-00090655/fr/
Lamareille, F., Contini, T., Brinchmann, J., et al. 2006a, A&A, 448, 907
Lamareille, F., Contini, T., Le Borgne, J.-F., et al. 2006b, A&A, 448, 893
Lamareille, F., Mouhcine, M., Contini, T., Lewis, I., & Maddox, S. 2004, MNRAS, 350, 396
Le Fèvre, O., Mellier, Y., McCracken, H. J., et al. 2004, A&A, 417, 839
Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003a, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. Edited by Iye, Masanori; Moorwood, Alan F. M. Proceedings of the SPIE, Volume 4841, pp. 1670-1681 (2003)., ed. M. Iye & A. F. M. Moorwood, 1670–1681
Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A, 439, 845
Le Fèvre, O., Vettolani, G., Maccagni, D., et al. 2003b, in Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II. Edited by Guhathakurta, Puragra. Proceedings of the SPIE, Volume 4834, pp. 173-182 (2003)., 173–182
Le Fèvre, O., Vettolani, G., Paltani, S., et al. 2004, A&A, 428, 1043
Lee, H., Skillman, E. D., Cannon, J. M., et al. 2006, ApJ, 647, 970
Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A., & Torres-Peimbert, S. 1979, A&A, 80, 155
Liang, Y. C., Hammer, F., Flores, H., et al. 2004, A&A, 423, 867
Liang, Y. C., Hammer, F., & Yin, S. Y. 2007, A&A, 474, 807
Lilly, S. J., Le Fevre, O., Renzini, A., et al. 2006, ArXiv Astrophysics e-prints
Liu, X., Shapley, A. E., Coil, A. L., Brinchmann, J., & Ma, C.-P. 2008, ApJ, 678, 758
Maier, C., Lilly, S. J., Carollo, M., Stockton, A., & Brodwin, M. 2005, ApJ, in press (astro-ph/0508239)
Maier, C., Meisenheimer, K., & Hippelein, H. 2004, A&A, 418, 475
Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
Maiolino, R., Nagao, T., Grazian, A., et al. 2007, ArXiv e-prints, 712
Maraston, C. 2005, MNRAS, 362, 799
Maraston, C., Daddi, E., Renzini, A., et al. 2006, ApJ, 652, 85
Marigo, P. & Girardi, L. 2007, A&A, 469, 239
McCracken, H. J., Radovich, M., Bertin, E., et al. 2003, A&A, 410, 17
McGaugh, S. S. 1991, ApJ, 380, 140
Meneux, B., Guzzo, L., Garilli, B., et al. 2008, A&A, 478, 299
Pérez-González, P. G., Trujillo, I., Barro, G., et al. 2008, ApJ, 687, 50
Pettini, M. & Pagel, B. E. J. 2004, MNRAS, 348, L59
Pilyugin, L. S. & Ferrini, F. 2000, A&A, 358, 72
Pozzetti, L., Bolzonella, M., Lamareille, F., et al. 2007, A&A, 474, 443
Radovich, M., Arnaboldi, M., Ripepi, V., et al. 2004, A&A, 417, 51
Richer, M. G. & McCall, M. L. 1995, ApJ, 445, 642
Salim, S., Charlot, S., Rich, R. M., et al. 2005, ApJ, 619, L39
Salpeter, E. E. 1955, ApJ, 121, 161
Savaglio, S., Glazebrook, K., Le Borgne, D., et al. 2005, ApJ, 635, 260
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Scodeggio, M., Franzetti, P., Garilli, B., et al. 2005, PASP, 117, 1284
Shapley, A. E., Coil, A. L., Ma, C.-P., & Bundy, K. 2005, ApJ, 635, 1006
Skillman, E. D., Kennicutt, R. C., & Hodge, P. W. 1989, ApJ, 347, 875
Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175
Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., & Asari, N. V. 2006, MNRAS, 371, 972
Temporin, S., Iovino, A., Bolzonella, M., et al. 2008, A&A, 482, 81
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
van Zee, L., Salzer, J. J., Haynes, M. P., O'Donoghue, A. A., & Balonek, T. J. 1998, AJ, 116, 2805
Veilleux, S. & Osterbrock, D. E. 1987, ApJS, 63, 295
Walcher, C. J., Lamareille, F., Vergani, D., et al. 2008, A&A, accepted (astro-ph/08074636)
Warren, S. J., Cross, N. J. G., Dye, S., et al. 2007, ArXiv Astrophysics e-prints
Zaritsky, D., Kennicutt, R. C., & Huchra, J. P. 1994, ApJ, 420, 87

[1] Laboratoire d'Astrophysique de Toulouse-Tarbes, Université de Toulouse, CNRS, 14 av. E. Belin, F-31400 France
[2] IASF-INAF, Via Bassini 15, I-20133, Milano, Italy
[3] INAF-Osservatorio Astronomico di Bologna, Via Ranzani 1, I-40127, Bologna, Italy
[4] IRA-INAF, Via Gobetti 101, I-40129, Bologna, Italy
[5] INAF-Osservatorio Astronomico di Capodimonte, Via Moiariello 16, I-80131, Napoli, Italy
[6] Università di Bologna, Dipartimento di Astronomia, Via Ranzani 1, I-40127, Bologna, Italy
[7] Laboratoire d'Astrophysique de Marseille, UMR 6110 CNRS-Université de Provence, BP8, F-13376 Marseille Cedex 12, France
[8] Max Planck Institut für Astrophysik, D-85741, Garching, Germany
[9] INAF-Osservatorio Astronomico di Brera, Via Brera 28, I-20021, Milan, Italy
[10] Institut d'Astrophysique de Paris, UMR 7095, 98 bis Bvd Arago, F-75014, Paris, France
[11] Observatoire de Paris, LERMA, 61 Avenue de l'Observatoire, F-75014, Paris, France
[12] Astrophysical Institute Potsdam, An der Sternwarte 16, D-14482, Potsdam, Germany
[13] INAF-Osservatorio Astronomico di Roma, Via di Frascati 33, I-00040, Monte Porzio Catone, Italy

[14] Universitá di Milano-Bicocca, Dipartimento di Fisica, Piazza delle Scienze 3, I-20126, Milano, Italy
[15] Integral Science Data Centre, ch. d'Écogia 16, CH-1290, Versoix, Switzerland
[16] Geneva Observatory, ch. des Maillettes 51, CH-1290, Sauverny, Switzerland
[17] Astronomical Observatory of the Jagiellonian University, ul Orla 171, PL-30-244, Kraków, Poland
[18] Centre de Physique Théorique, UMR 6207 CNRS-Université de Provence, F-13288, Marseille, France
[19] Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, P-4150-762, Porto, Portugal
[20] Institute for Astronomy, 2680 Woodlawn Dr., University of Hawaii, Honolulu, Hawaii, 96822, USA
[21] School of Physics & Astronomy, University of Nottingham, University Park, Nottingham, NG72RD, UK
[22] Max Planck Institut für Extraterrestrische Physik (MPE), Giessenbachstrasse 1, D-85748 Garching bei München,Germany
[23] Canada France Hawaii Telescope corporation, Mamalahoa Hwy, Kamuela, HI-96743, USA
[24] Universitätssternwarte München, Scheinerstrasse 1, D-81679 München, Germany