Photometric redshifts of the CFHTLS Wide - T0006
Release v1.0, May 2010

Arnouts S., Coupon J., Ilbert O., McCracken H.J., Mellier Y.
& the Terapix/CFHTLS/VVDS/VIPERS[1] teams

This documentation describes the photometric redshift catalogues for the CFHTLS-Wide T0006 release. We refer to Ilbert et al. (2006) and Coupon et al. (2009) for a detailed description of the methods (these papers should be cited if you use this release). This document describes what has changed compared to previous photo-zed releases.

1 Data

1.1 Photometric data

We used version T0006 of the CFHTLS-Wide catalogues delivered by Terapix. The CFHTLS-Wide data cover 155 deg^2 in four different fields (W1, W2, W3, W4) (see Figure.1). A full description of this input catalogue can be found in the terapix release documentation [2].

Filters
All fields are observed in u*, g, r, i, z filters. However, the megacam i band filter was broken during the survey and a new i band filter has been used denoted by "y". We considered these two filters separately.

Selection
Photometric redshifts are highly unreliable at $i > 24$ in the CFHTLS-Wide. Catalogues are selected at $i < 24$ (or $y < 24$) to decrease their size and minimise computation time. This cut has been applied in the original catalogues (before correcting for the Milky Way extinction and color terms).

Merged catalogues
One catalogue is provided by Terapix for each tile of 1 deg^2. Some sources located in the border of the images are included in two different catalogues (orange points in Figure.1). In order to facilitate the use of the CFHTLS-Wide photo-z catalogues, we merged all the catalogues in a single catalogue for each field. Therefore, we provide four merged catalogues (one for each field). We kept all the sources in the merged catalogues, even the ones observed twice. The flag *inout* (column 3) allows the users to remove the less reliable observation when an object is observed twice. We set the flag *inout* as follow:

- the sources which are observed only once have a flag *inout=1*.
- when a same source is included in two different catalogues A and B, we checked if the source is closer to the center of the tile A or B. If the source is closer to the center of tile A, we set the flag *inout=1* for the source in the catalogue A (and *inout=0* for the source in the catalogue B).

By selecting the sources with *inout=1*, the user keeps only one entry per source and avoid duplicated sources.

[1] http://vipers.inaf.it
[2] http://terapix.iap.fr/cplt/T0006-doc.pdf

We also generate a new ID for the merged catalogue. This ID is positive for the sources having *inout=1*. Table. 1 lists the number of objects in the four merged catalogues.

Zero-points
The accuracy of the zero-points is crucial to avoid biases in the photo-z estimate. We already calibrate the zero-points using the spectroscopic samples (see section 2). However, the spectroscopic samples do not provide complete coverage of the wide fields. Therefore, it is crucial to ensure homogeneous photometry over the full fields. Terapix provides colour correction terms which can be applied to the catalogues in order to match the star sequence of the SDSS in each tile. For W1, W3, W4, we applied these color corrections given in tables 17-20 of the T0006 catalogue documentation from Terapix. For W2, the corrections have been derived from the W1 and W3 SDSS star sequence (private communication from Terapix).

Magnitudes
Colours are measured using MAG_AUTO magnitudes. These magnitudes are in the AB system and have been corrected from the Milky Way extinction using the Schlegel et al. (1998) maps.

1.2 Spectroscopic data

Spectroscopic redshifts are crucial to calibrate and test the accuracy of the photometric redshifts. We use only the most secure redshifts of each survey. We list below the various spectroscopic samples which were used.

- **W1 field -** The VVDS Deep (Le Fèvre et al. 2005) and VVDS Ultra-Deep (Le Fèvre et al., in prep) are available in this field. The VVDS Deep and Ultra-deep are purely selected in magnitude at $I_{AB} < 24$ (blue points in Figure 1) and $23 < I_{AB} < 24.75$ (cyan points in Figure 1), respectively. After having matched the photometric and spectroscopic samples, we found 5926 and 246 secure redshifts available from the Deep and Ultra-deep, respectively.
- **W2 field -** no spectroscopy is available in this field
- **W3 field -** 1267 redshifts from the DEEP2 survey (Davis et al. 2003) selected at $R_{AB} < 24.1$ (blue points in Figure 1).
- **W4 field -** the VVDS Wide targets (Garilli et al. 2008) (pink points in Figure 1) are selected at $I_{AB} < 22.5$. 10149 secure redshifts are available (4514 galaxies). A preliminary sample from the VIPERS survey (Guzzo et al., in prep) is also available in this field. These 2413 VIPERS targets were selected at $I_{AB} < 22.5$ (orange points in Figure 1). A color-color selection criteria produces a sample with $z > 0.5$.

field	all sources $i < 24$	no duplicated sources $i < 24$	final redshift $0 < z < 6$ $i < 24$	final redshift $0 < z < 6$ $i < 22.5$
W1	5346671	4794688	4042097	1243192
W2	2100453	1911609	1321661	405004
W3	3858738	3343478	2704642	823912
W4	2040682	1853429	1370237	455285

Table 1: Number of sources

2 Method

We used the public code Le_Phare (Arnouts et al. 2002, Ilbert et al. 2006).[3] Le_Phare is based on a standard template fitting procedure. The templates are redshifted and integrated through the instrumental transmission curves. The opacity of the IGM is taken into account and internal extinction can be added as a free parameter to each galaxy. The photo-z are derived by comparing the modelled fluxes and the observed fluxes with a χ^2 merit function. A probability distribution function is associated to each photo-z.

We adopted a configuration similar to the one used in Ilbert et al. (2006) and Coupon et al. (2009):

- *templates:* our primary template set are the four Coleman Wu and Weedman (CWW) observed spectra (Ell, Sbc, Scd, Irr) from Coleman et al. (1980) complemented with two observed starburst SED from Kinney et al. (1996). We optimised this primary set of templates using the VVDS spectroscopic sample.
- *training:* we performed an automatic calibration of the zero-points using the spec-z sample. The calibration is obtained by comparing the observed and modeled fluxes. The calibration is done iteratively until the zero-points converge. This step helps in removing bias.
- *extinction:* for spectral types later than Sbc, we introduce a reddening $E(B-V) = 0 - 0.35$ using the Calzetti et al. (2000) extinction law.
- *prior:* we applied a prior on the redshift distribution following a Bayesian approach similar to Benitez (2000). Such prior is necessary when no NIR is available. No redshift solutions are allowed which will produce a galaxy brighter than $M(G) = -24$.
- *redshift:* we adopted a slightly different method in order to extract the redshift from the probability distribution function (PDF). Rather than using directly the redshift which minimizes the χ^2 distribution function, we used the median of the PDF. Using the median of the PDF limits the risk of preferred solutions in narrow redshift ranges.

A new configuration of the photo-zed code which includes emission lines and a new template set (Polletta et al. 2007) has been successfully applied to the COSMOS survey (Ilbert et al. 2009). However, we prefer to remain consistent with previous CFHTLS photo-zed releases and use the same configuration as before which includes templates optimised for the CFHTLS. We defer an application of the Ilbert et al. 2009 method to the final CFHTLS release expected later this year.

3 Clean the galaxy photo-z catalogue

3.1 Select a galaxy sample

Our aim is to provide a galaxy photo-z catalogue. For this reason we removed stars and QSOs.

We separated galaxies and stars by combining a morphological criteria (the half-light radius) and a color criteria (the χ^2). Figure 2 shows the half-light radius as a function of magnitude for one image of the W1 field. The star sequence is clearly visible for bright objects. The position of the star sequence depends on the considered tile (the seeing changes between various images). For each pointing, we derived a limit in radius to select the point-like sources. This limit is given in column 8 of the catalogue. In order to define this limit, we computed the median med and the dispersion σ of the half-light radius for bright point-like sources (classified by SExtractor). Sources are classified as star if one of this condition is full-filled:

- $r2 < med + 5\sigma$ and $i < 20$;

[3] http://www.cfht.hawaii.edu/~arnouts/LEPHARE/cfht_lephare

- the source is classified as stars by sextractor in the original catalogue (*flag=1*);
- the source is faint $20 < i < 23$ and is a potential point-like source with $r2 < med + 5\sigma$. If $\chi^2(star) < \chi^2(gal)$, this source is classified as a star.

When the source is considered as a star, we set the final redshift at 0 (column 10). Red points in Figure 2 represent the sources classified as stars.

Some $i < 23$ point-like sources have a large χ^2 and are still present in the catalogue (they are not removed by the star selection criteria). If the fit obtained using an AGN template (Salvato et al. 2009) provides a better χ^2, we set the final photo-z at z=9.99 (column 10). Such objects are potential QSOs. An attempt was also made to provide a redshift using AGN templates (Salvato et al. 2009) but these photo-z have not been checked (column 26).

3.2 Removed unreliable photo-zeds

Masked areas
Some areas have low-quality photometry (halos around bright stars, borders) and are masked. Sources which are in masked regions are kept in the merged catalogues and they are flagged (column 7, $flag > 1$). Since the quality of the photometry in these areas is uncertain, we advise users to not use photometric redshifts in the masked areas. Therefore, we set the final photo-z (column 10) at -99.9 in masked areas. If users really needs to use these areas, the photo-z can still be found in column 11 (no star/galaxy classification for this column).

Low quality fit
When the fit is done with less than three filters, we considered that the quality of the photo-z is poor and we set the final photo-z at 9.99 (column 10). If the fit produces a really high reduced χ^2 value, we also considered that the fit is poor and we flagged such objects. If reduced $\chi^2 > 100$ (Figure 3 shows the χ^2 distribution), we set the final photo-z at 9.99 (column 10). We checked that most of the objects with $\chi^2 > 100$ are located in really specific areas of the fields.

Less accurate photo-z
Fainter galaxies have in general less accurate photo-zeds. There is also a dependency of the accuracy on the redshift. A 68% confidence-level error is provided for each photo-z (column 12 and 13). These errors could be use to sub-select a sample of accurate photo-z. Galaxies with two peaks in the PDFz are also flagged (column 22, *zp_2*). When *zp_2* has a positive value, two peaks are identified in the PDF. The chance of catastrophic failures is higher in this sample. The redshift distribution of this population (Figure 4) reflects the risk of degeneracy between the Balmer and the Lyman breaks.

To summarize, the user should select sources with a final photo-z (column 10) greater than 0 and lower than 6. This sample will include sources classified as galaxies, outside the masked areas, with a reasonable fit and without duplicated sources. Figure 5 shows the magnitude distribution of the various sub-samples.

4 Photometric redshift accuracy

Accuracy versus redshift
These photo-z are computed using optical data only (near-infrared data are not available). Therefore, these photo-z are reliable at $z < 1.3 - 1.5$ only, since the Balmer break is redshifted outside

the optical coverage at $z > 1.3 - 1.5$.

Accuracy versus magnitude
Figures 6, 7, 8 show the comparison between photometric and spectroscopic redshifts in the four fields. The accuracy and the fraction of catastrophic redshifts are indicated in the top left of each panel. The redshift accuracy is noted $\sigma_{\Delta z/(1+z_s)}$ and estimated using the normalized median absolute deviation (NMAD: Hoaglin et al. 1983, $1.48 \times \mathrm{median}(|z_\mathrm{p} - z_\mathrm{s}|/(1 + z_\mathrm{s}))$). We defined as catastrophic redshift an object with $|z_\mathrm{p} - z_\mathrm{s}|/(1 + z_\mathrm{s}) > 0.15$ (dotted lines in the figures). The percentage of catastrophic redshifts is denoted by η. The dispersion and the fraction of catastrophic redshifts increases dramatically from a bright sample at $i < 21.5$ (see Figure 6) to a faint sample at $i > 22.5$ (see Figure 8). The accuracy is around $\sigma_{\Delta z/(1+z_s)} \sim 0.034 - 0.037$ at $i < 21.5$ and reaches $\sigma_{\Delta z/(1+z_s)} \sim 0.068 - 0.081$ at $22.5 < i < 23.5$. The failure fraction increasing from $\eta \sim 2\%$ at $i < 21.5$ to $\eta \sim 12 - 14\%$ at $22.5 < i < 23.5$. Significant differences are seen between the four fields. The differences are mainly explained by the selection function of the spectroscopic samples. For instance, the DEEP2 spectroscopic sample (which covers W3) has been preselected to be more efficient at $0.7 < z < 1.5$. Since this redshift range is less reliable for the photo-z, it explains why the statistics measured in the W3 field are poorer than in other fields. The VVDS sample has been purely magnitude selected in the W1. Therefore, this sample is the most representative to estimate the photo-z accuracy.

We advise the users to apply a cut in magnitude, in order to isolate a sample of accurate photo-z. The choice of this cut depends on the scientific application. Figure 9 shows the evolution of the median 1σ redshift error as a function of magnitude [4] (the errors are not divided by $(1 + z)$ in this figure). The user should use this figure in order to decide which cut to apply depending on the scientific application.

Finally, we performed two other checks. In Figure 10, we show the redshift distribution in the four fields for different magnitude cuts. The median redshift is clearly shifted at higher redshift when we apply a fainter cut. Figure 11 shows $U - R$ versus $R - I$ in the four fields. The position of the star sequence is well consistent between the four wide fields. A similar color-color plot has been used by the DEEP2 team to preselect their targets at $z > 0.75$. The red points indicate the galaxies at $z > 0.75$ which are well separated from the galaxies at $z < 0.75$, as expected.

[4] The 1σ errors are derived for every object using the χ^2 distribution. This error is derived statistically. Therefore, possible biases are not taken into account

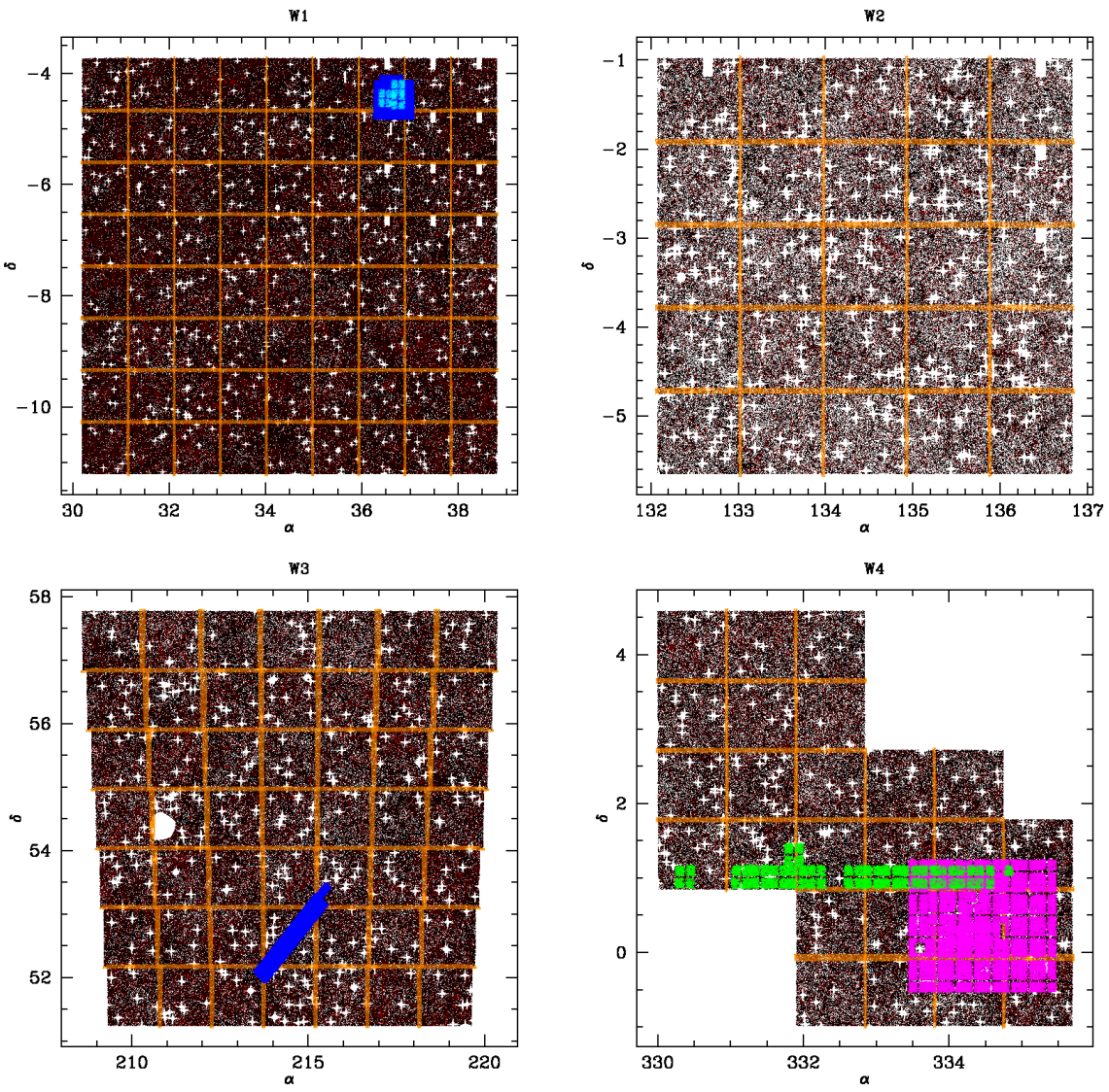

Figure 1: Area covered by the four wide fields and corresponding spectroscopic coverage. Orange points are duplicated objects observed in two different tiles.

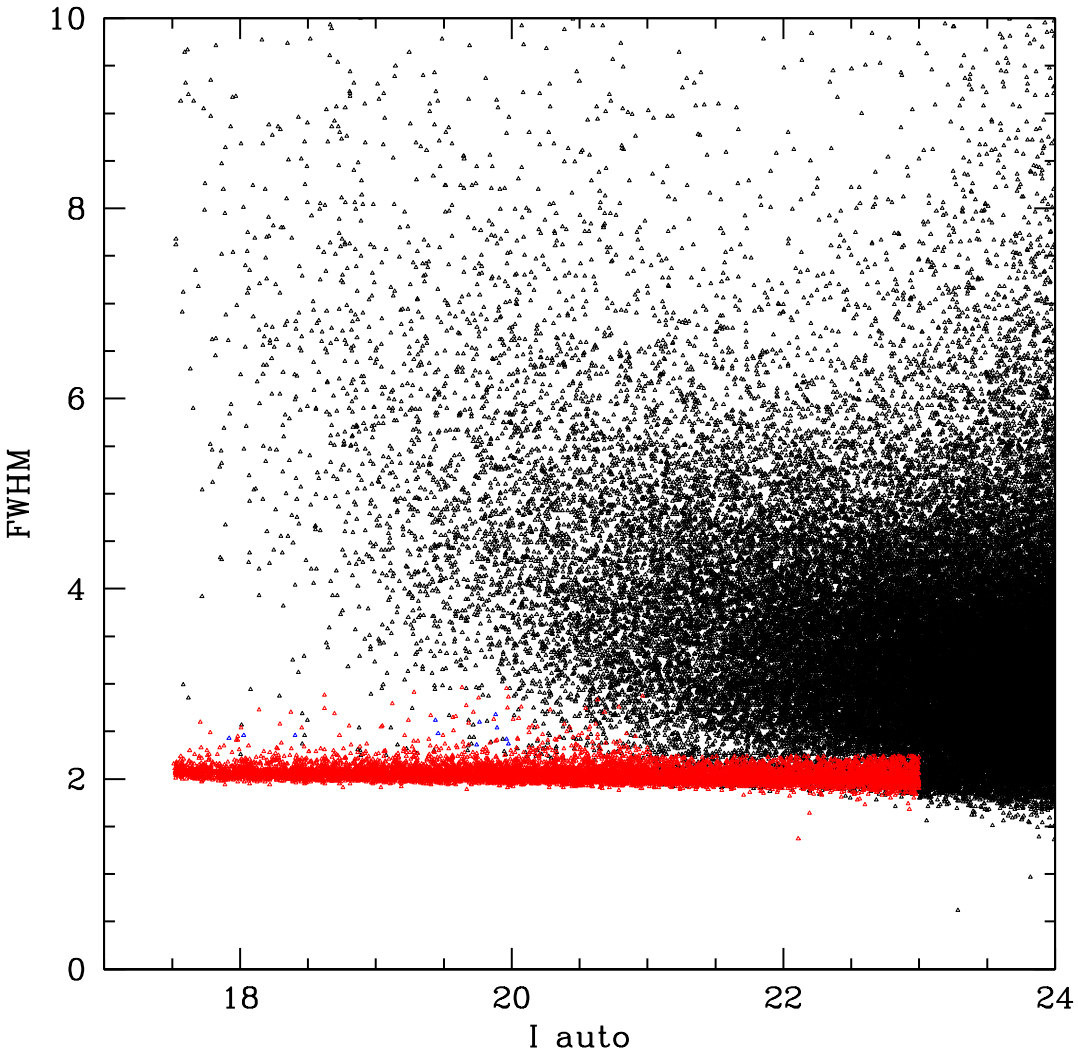

Figure 2: Half light radius as a function of magnitude in one of the tile of W1. The red points represents the sources classified as stars.

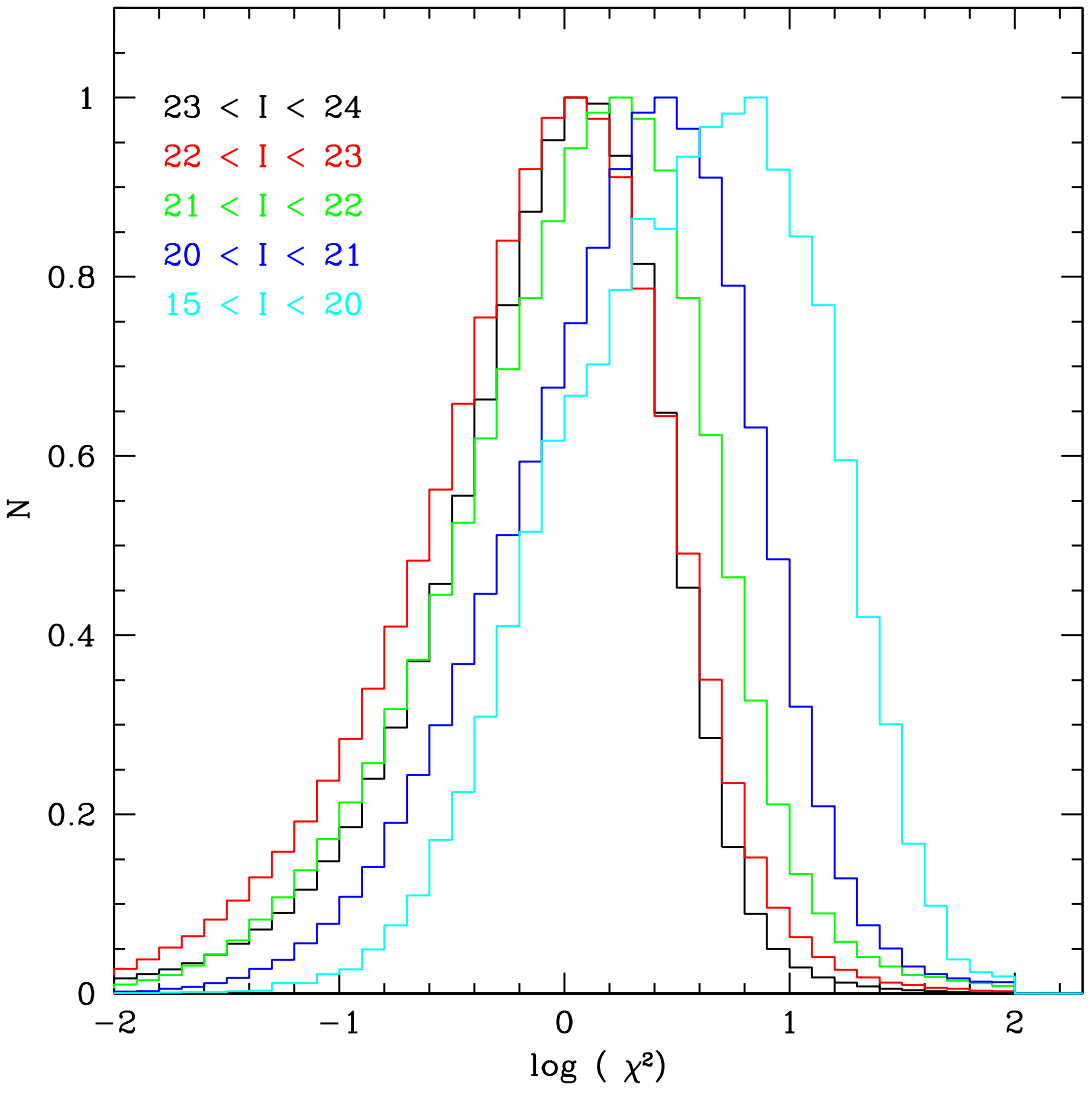

Figure 3: χ^2 distribution (W1 field).

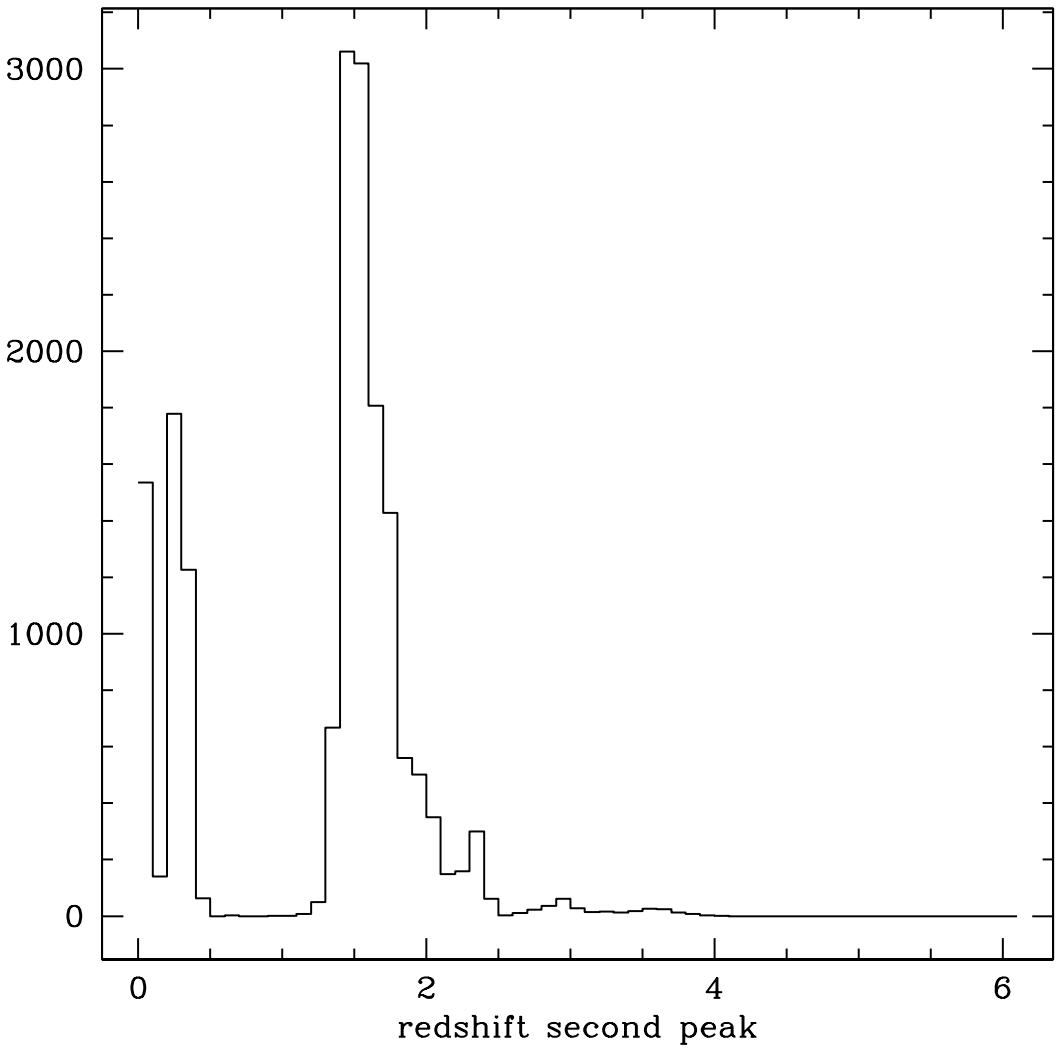

Figure 4: Redshift distribution of the second peak solution in the PDF (W1 field).

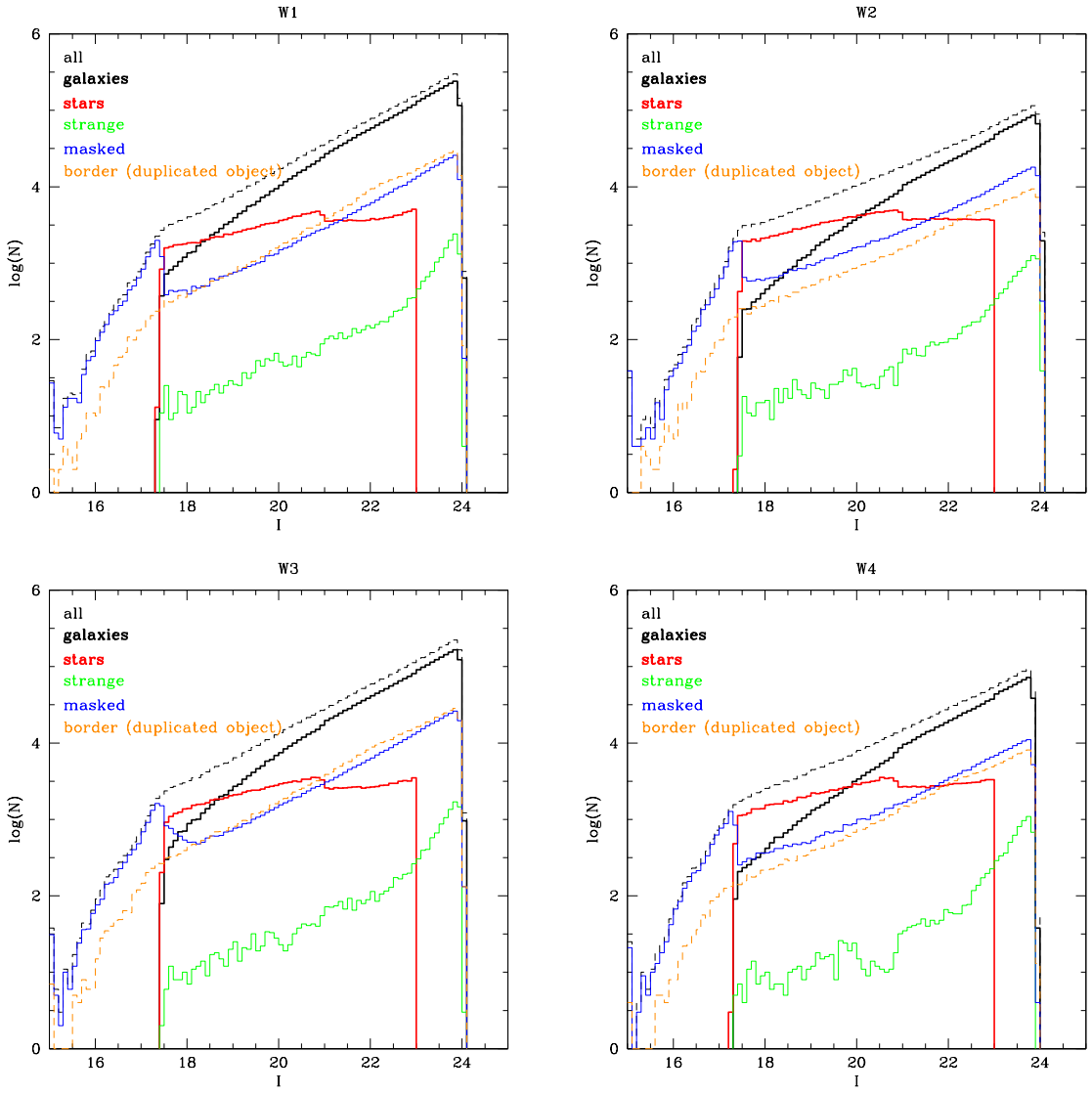

Figure 5: Number counts in i band. The merged catalogue are cut at $i < 24$. The number counts are sub-divided in galaxies, stars, masked objects. The objects flagged as "strange" are sources with a really high χ^2, possible QSO or sources with only 3 filters in the fit (final z set to 9.99).

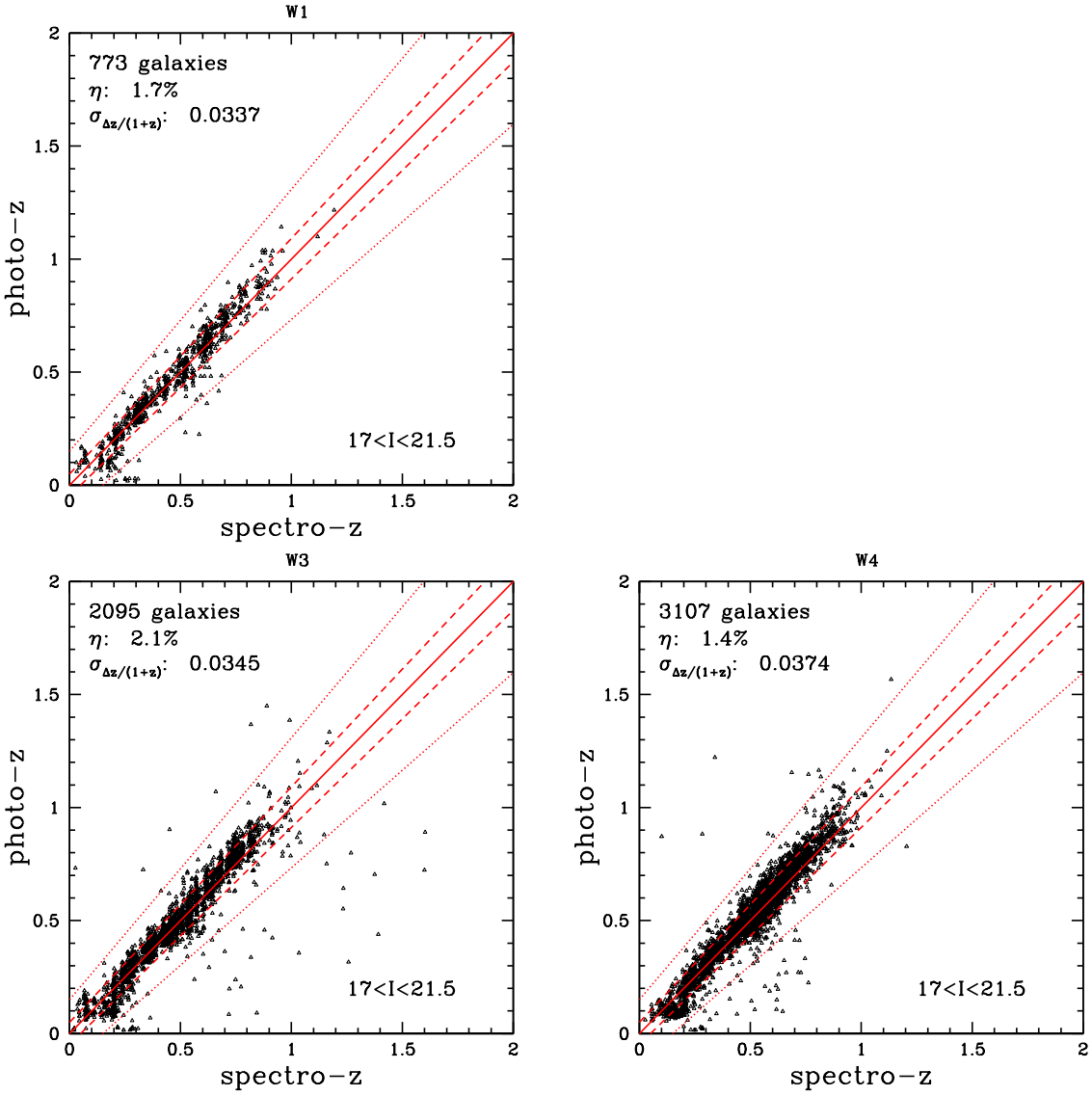

Figure 6: Photometric redshifts versus spectroscopic redshifts for a bright sample $i < 21.5$

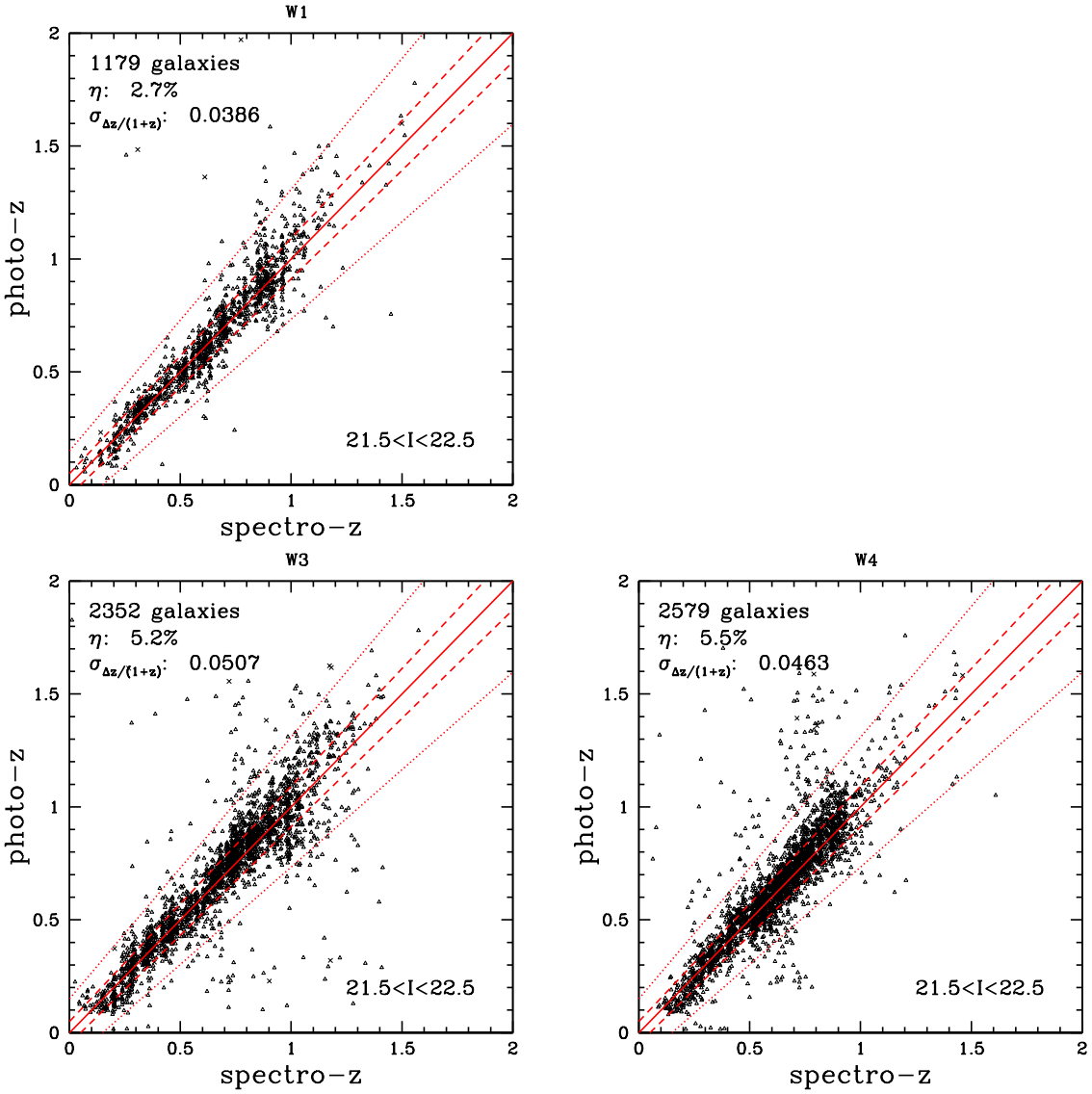

Figure 7: Photometric redshifts versus spectroscopic redshifts for the intermediate magnitude sample $21.5 < i < 22.5$

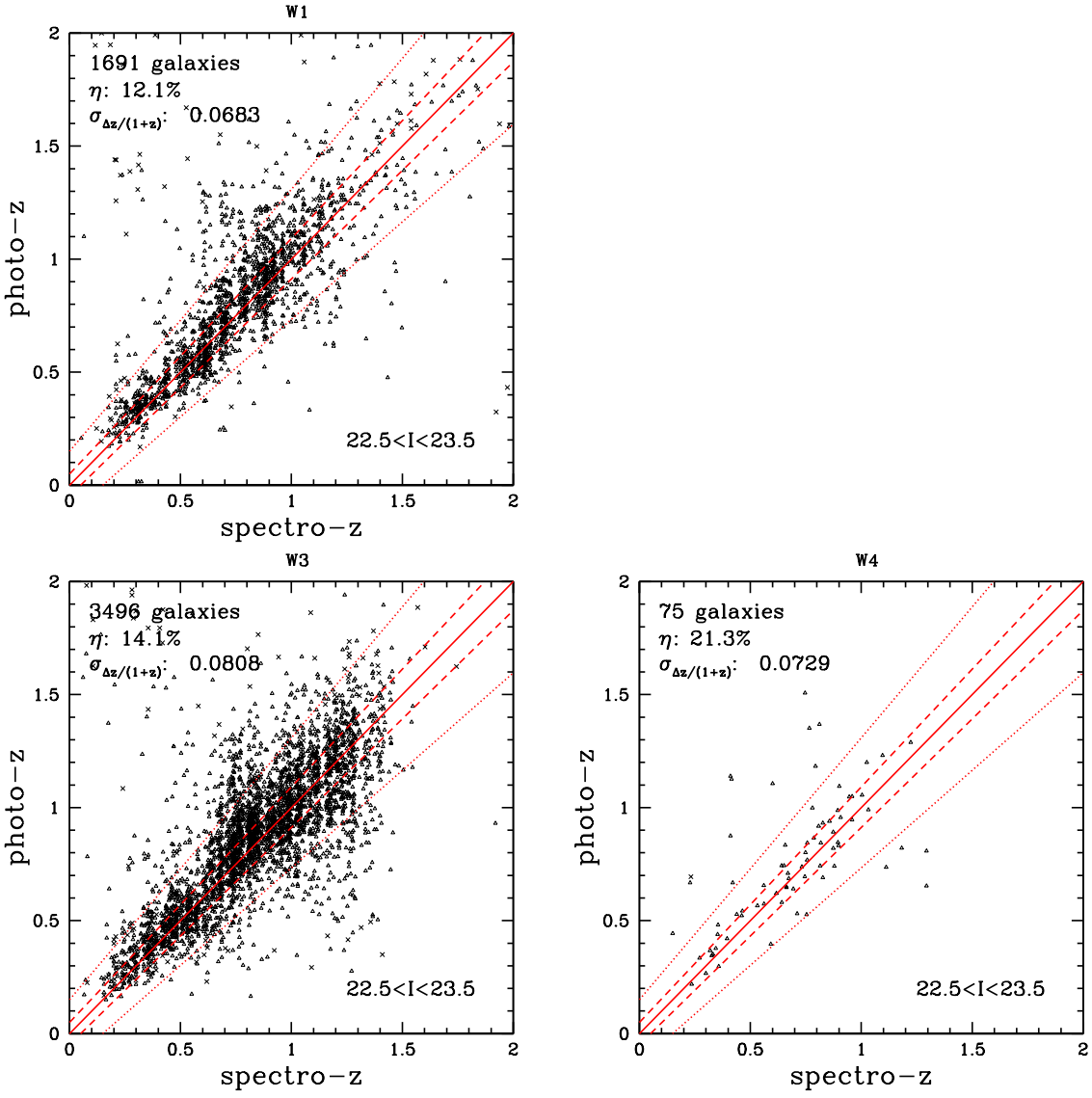

Figure 8: Photometric redshifts versus spectroscopic redshifts for a faint sample $i > 22.5$

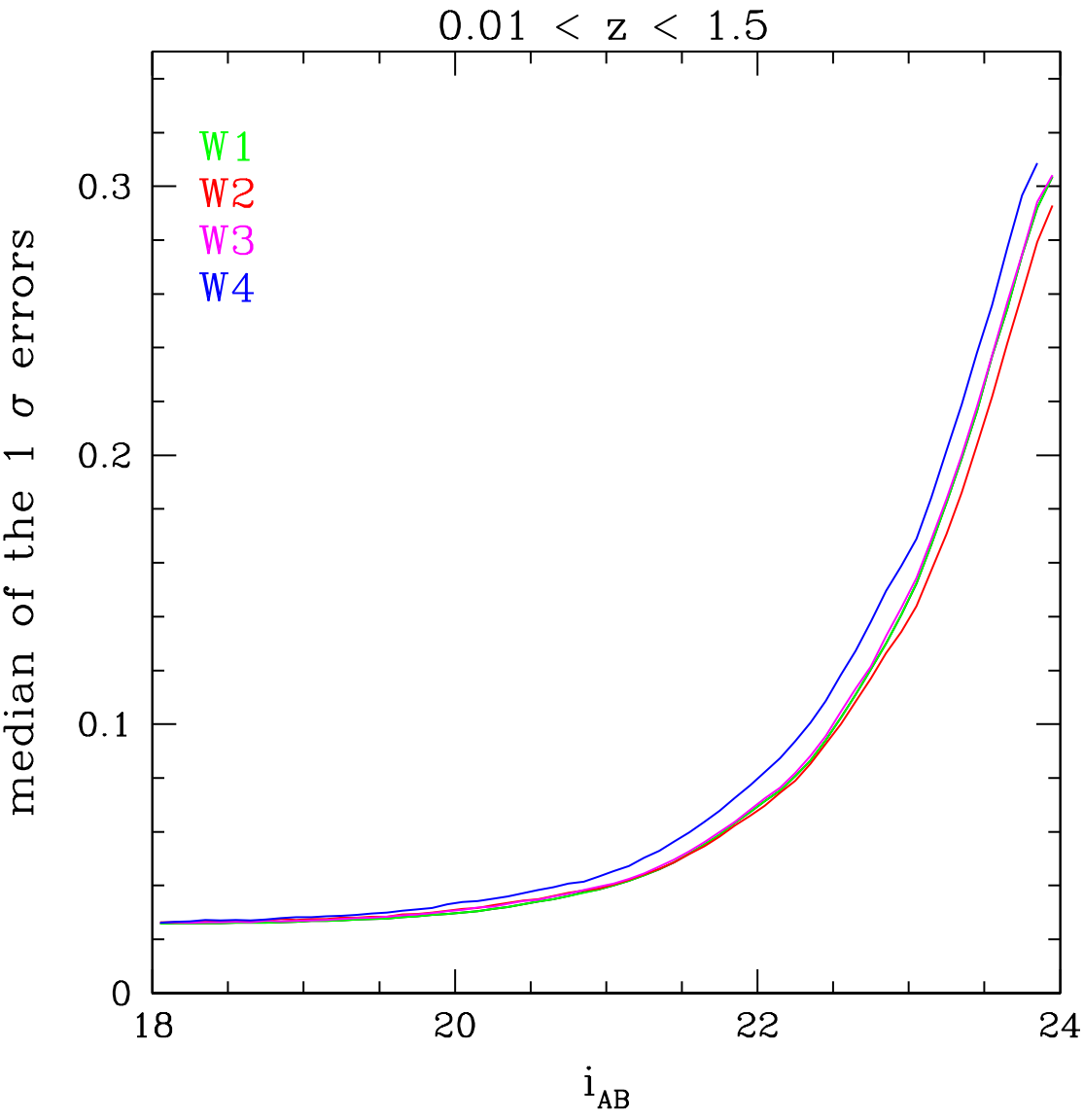

Figure 9: Median of the 1 σ errors (maximum value between the negative and positive 68% errors) as a function of i band magnitude. The errors are not divided by $(1 + z)$. The W4 field is slightly offseted in comparison to other fields. This offset is explained by the high Milky way extinction in this field (average E(B-V) of 0.06 in W4 while 0.02 in the 3 other fields).

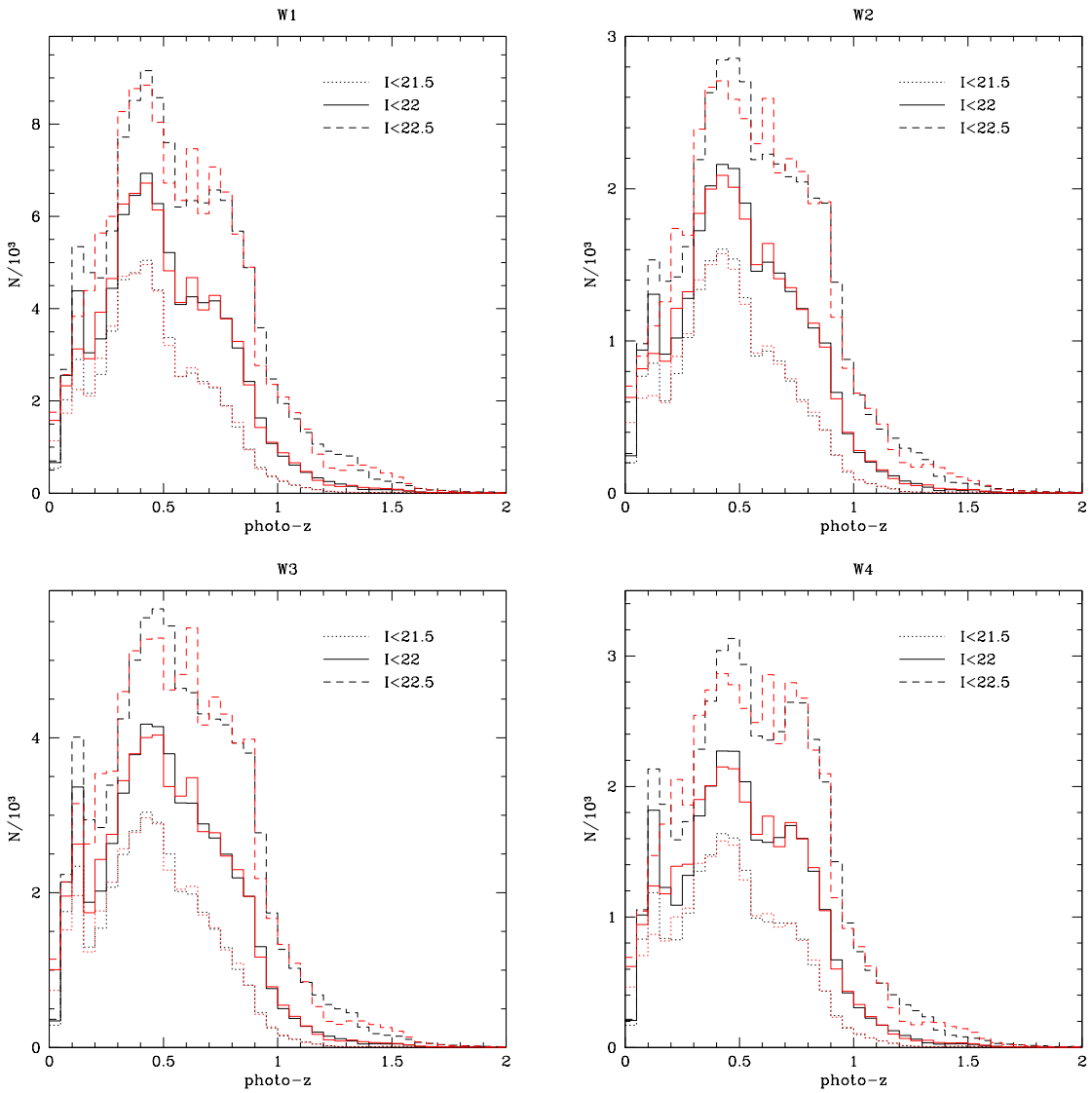

Figure 10: Redshift distributions in the four wide fields. The black lines are obtained using the median of the PDF. The red lines correspond to the minimum χ^2.

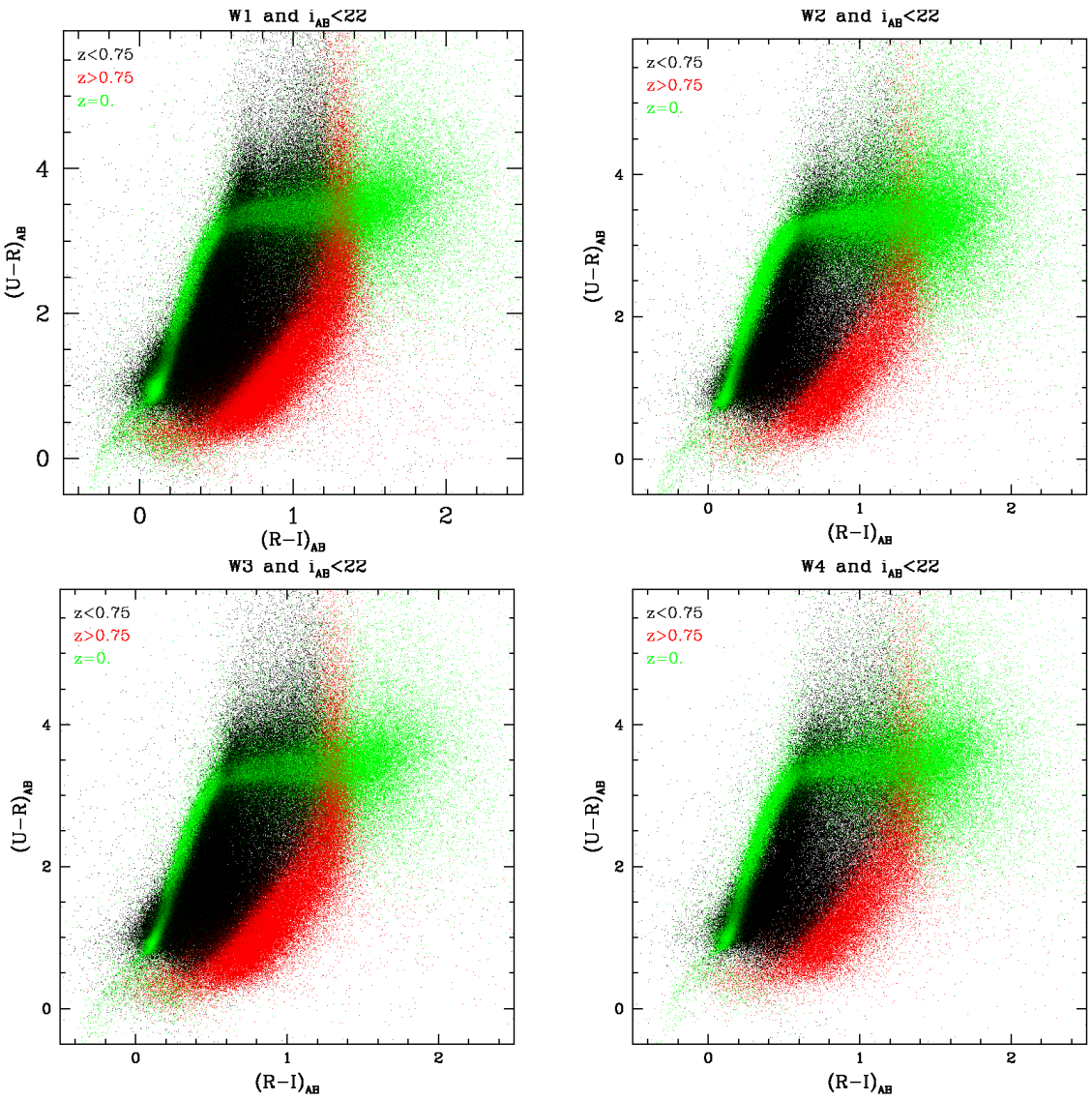

Figure 11: Color-color plot $U - R$ versus $R - I$, in each field. Black and red points are located at $z < 0.75$ and $z > 0.75$, respectively. The star sequence is shown in green.