More elements towards the XMDS full catalogue

The 2008 status report

L.Chiappetti[1]

INAF, IASF Milano, via Bassini 15, I-20133 Milano, Italy

Abstract. I supplement previous reports, describing the work I have done in 2008 to improve the catalogue of identified sources beyond the original XMDS/VVDS 4σ catalogue and the subsequent version used for the XMDS/VVDS 3σ hard sample. In this work I used all database tables available to me in the Milan database up to April 2008, using a procedure as automatic as possible and only objective criteria for identification ranking. We intend to use this catalogue to extract subsets for further work since now on.

Key words: LSS – XMDS –

1. Introduction

We originally published in Chiappetti et al. (2005) (hereafter Paper I) a catalogue of 286 tentative identifications for the X-ray sources detected in the XMDS fields at a significance above 4σ, and falling in the area covered by the VVDS survey (Le Fèvre et al., 2004),based on an highly manual procedure, described, together with its input tables, in section 6 of Paper I.

A further working catalogue using all (GO) XMDS sources (1147), inside and outside the VVDS area, using additional data tables which became available in the meantime (e.g. CFHTLS and SWIRE *pro tempore*), and automatizing the procedure as much as possible, was produced as described in Chiappetti (2006), hereafter Report I, and has been used as a basis for the study of the AGN VVDS 3σ hard sample (136 sources) reported by Tajer et al. (2007) and Polletta et al. (2007).

A working release for 1168 XMDS sources, which is the ultimate number for the XMDS area was presented in Chiappetti (2007), hereafter Report III (and used in M.Giorgetti's master thesis).

We present here a further working release with the same number of sources in which: *(a)* we added further optical tables for identification as they became available in the meanwhile; and *(b)* fully reprocessed the identification in an automated and objective way.

This report gives a short account of such procedure. For more details one can consult the web page `http://sax.iasf-milano.inaf.it/~lucio/LSS/TriIdent/procedure.html` which also contains a reference to similar pages for the previous 3 versions.

In section 2 I list the input database tables used as starting point for the identification, The procedure is described step-by-step in the various subsections of section 3. In particular the ranking (see 3.3) is now generated almost automatically from positional probabilities (see 3.2), although after data verification and eventual visual inspection. Other sections give some simple statistics and the coverage of the different surveys (see 3.4) and an updated comparison with the XMM-LSS catalogue version 1 published in Pierre et al. (2007) and with the `INTERIM` versions described in Chiappetti (2008) (see 3.5). Finally section 4 describes the implementation in the Milan database.

We intend to use the updated complete catalogue as the source from which to extract subsamples for further work (e.g. Trinchieri, Giorgetti et al. in preparation), possibly up to a full photometric redshift catalogue.

2. Data sources

Our starting point has been the **same** *glorified correlation table* (GCT; table of pointers into all possible combinations of database tables, each one correlated with the `xmdsepic` table with a "standard" correlation radius or criterion) used in Report III, of which we made a copy.

This table has been then modified to include new references to tables in other (optical, IR) bands which became available after the end of Report III (see 2.2). There were **no** change to the X-ray data.

2.1. X-ray data

No new X-ray data were added with respect to what described in section 2.1 of Report III. The only, absolutely marginal, update to the `xmdsepic` table was the correction of the flux errors on the total (ABCD; 0.3-10 keV) band, which were previously incorrect.

Table	Update	Content	History	(5)	(6)
xmdsepic	**May 08**	reference table with XMDS X-ray sources resulting from the Milan pipeline	**new:** fixed flux error bug; for the rest as used since Paper I and Report I	n/a	
xmdsdup	n/a	clone of xmdsepic	n/a	6″	a
nov06	**Aug 08**	X-ray sources from the Saclay pipeline, band merged within 6″	Used in XLSS catalogue (Pierre et al., 2007); **new:** fixed off-axis and position error bug ; fixed coordinate bug and *added* revised astrometry	10″	b
virphot	Jul 08	VVDS "good" UBVRI(JK) photometry	used since Paper I and Report I	6″	c
bad	Jul 08	VVDS photometry flagged "bad"	as above	6″	c
loiano	Jan 07	VVDS Loiano U filter photometry	as above	6″	c
vimos	Jul 08	VVDS spectroscopic information	as above	6″	c
sacphot	Oct 03	CFH12K observations made by Saclay	used since Paper I and Report I	6″	d
virradio	Aug 03	entire VIRMOS1.4GHz catalogue	used since Paper I and Report I	40″	e
radio	Nov 03	entire XMM-LSS own ("Leiden") VLA radio catalogue	used since Paper I and Report I	40″	f
specfup	Sep 03	spectroscopy campaign of Oct 2002	used since Report I	6″	
xlssc	**Dec 07**	list of XLSSC clusters (**Lyon LS3DB**)	used since Report I	2′	g
loto	Nov 07	Lotoweb (Lyon LS3DB)	present since PaperI and Report I	10″	
d1	Jan 07	CFHTLS D1 field $u^*g'r'i'z$ photometry supplied via IPAC in Feb 05	used since Report I	6″	h
d1t3	Sep 07	CFHTLS D1 field release T003 supplied by Saclay	used since Report III	6″	
d1t4	**Aug 08**	CFHTLS D1 field release T004 supplied by Saclay	**new!!** added Jan 08	6″	
w1	Jan 08	CFHTLS W1 fields $u^*g'r'i'z$ photometry supplied via IPAC in Jul 05	used since Report I	6″	h
w1t3	Dec 07	CFHTLS W1 fields release T003 supplied by Saclay	used since Report III	6″	
w1t4	**Aug 08**	CFHTLS W1 fields release T004 supplied by Saclay	**new!!** added Jan 08	6″	
swire	Jan 07	SWIRE data from above IPAC releases	used since Report I	6″	i
swires05	Aug 08	SWIRE Spring 05 public release (IRSA Gator)	used since Report III	6″	
swiredr6	**Aug 08**	SWIRE DR6 supplied by IPAC	**new!!** added Jan 08	6″	
ukidss	**Apr 08**	UKIDSS DR3plus public release	**new!!** old version entirely overwritten	6″	
galex	Jun 04	unofficial FITS file with GALEX data	used since Report I	6″	j
simbad	Aug 08	SIMBAD sources	used since Paper I and Report I	20′	k
ned	Aug 08	NED sources	used since Paper I and Report I	20′	k
usno	Oct 07	USNO A2 catalog as kept at ST-ECF.	used since Report I	6″	
tajer07	Apr 07	Tables A.1, B.1, B.2 from Tajer et al. paper	used since Report III	n/a	l
polletta07	Apr 07	Table I from Polletta et al. paper	used since Report III	n/a	m
garcet07	**Nov 07**	Table 4 from Garcet et al. paper	**new!!**	n/a	n

Table 1. Database tables used as input to the present XMDS catalogue

(5) column (5) is the correlation radius used to populate the GCT with the object around the X-ray sources

(6) column (6) refers to the notes indicated below

a xmdsdup used to tag sources detected in overlapping fields (see Report I)

b correlation within 10″ in old corrected coordinates, i.e. those published in Pierre et al. (2007)

c all VVDS "authorized subsets" share a system of common identifiers, in particular bad and virphot are disjoinct sets, while the other two are subsets of the union of the latter.

d CFH12K observations made by Saclay outside of the VVDS area. Now considered obsolete.

e Bondi et al. (2003)

f Cohen et al. (2003)

g reverse correlation table

h see detailed notes in Report I. Now considered obsolete.

i see detailed notes in Report I. Most data come from the July 2005 IPAC release (flagged as dataset=3). Obsolete.

j table does not contain any scientific information, but is used only to assess whether it is worth asking to proceed for a more formal collaboration.

k SIMBAD and NED may also include data from some of our catalogues (radio and XLSSC).

l Tajer et al. (2007)

m Polletta et al. (2007)

n Garcet et al. (2007) .

z the update date indicates the latest maintenance change (minor if not explicitly stated otherwise)
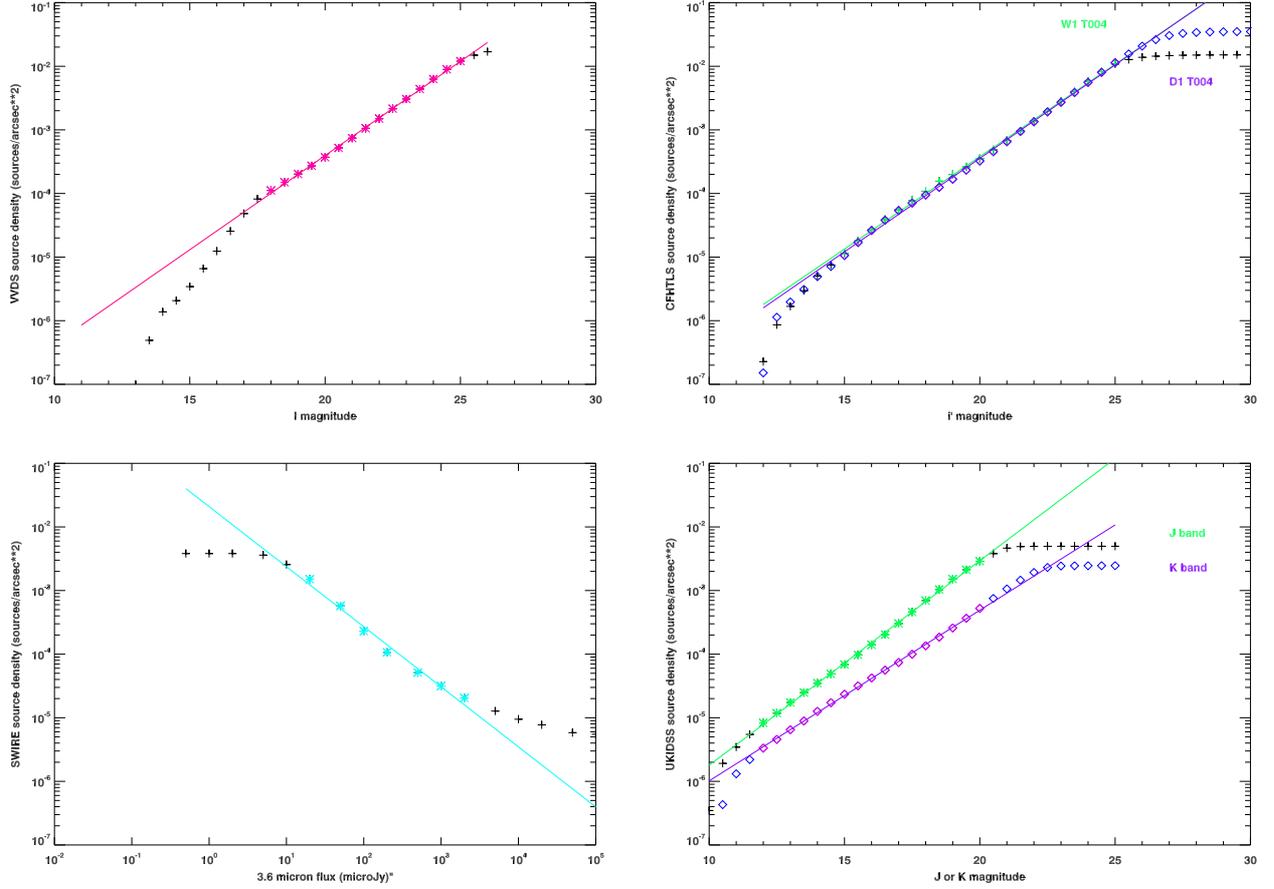
Fig. 1. Source count density for the VVDS (top left panel), the CFHTLS D1 (crosses) and W1 (diamonds) fields (top right panel), for SWIRE DR6 at $3.6\mu m$(aperture 2) fluxes (bottom left panel) and for UKIDSS J band (crosses) and K band (diamonds) (bottom right panel) The ranges used to produce the fits shown, whose parameters are given in Table 2 are shown in (lighter) colour.

Probability	m	density $n(brighter\ than\ m)$	a	b	tables
probvvds	I	$n(<I) = 10^{a+bI}$	-9.32636	0.29614	in order `virphot bad sacphot`
probd1	i'	$n(<i') = 10^{a+bi'}$	-9.32415	0.293833	for `d1t4`
			-9.23183	0.290519	for `w1t4` excluding ABC fields
probswire	F_λ	$n(>F_\lambda) = 10^{a+b*log(F_\lambda)}$			for `swiredr6`
	$\lambda = 3.6\mu m$		-1.68062	-0.944191	then in order of λ
	$\lambda = 4.5\mu m$		-1.73693	-0.976644	
	$\lambda = 5.8\mu m$		-2.04933	-0.829700	
	$\lambda = 8.0\mu m$		-1.49944	-1.07201	
	$\lambda = 24\mu m$		0.102480	-1.53410	
probukidss	J	$n(<J) = 10^{a+bJ}$	-8.67503	0.268272	taken best if both bands present
	K	$n(<K) = 10^{a+bK}$	-8.96264	0.321560	

Table 2. Parameters used for probability computation

There have been changes instead to the X-ray table referring to the Saclay pipeline (`nov06`), for the details of which we make reference to the `INTERIM` Report (Chiappetti , 2008). Essentially the source coordinates in `nov06` have changed because of the fix of an incorrect half-pixel offset and the repetition of the astrometric correction. However since the data had already been published in the first XLSS catalogue (Pierre et al., 2007), the new coordinates have been *added* in new database columns leaving the original ones alone. The new ones will be used in section 3.5.
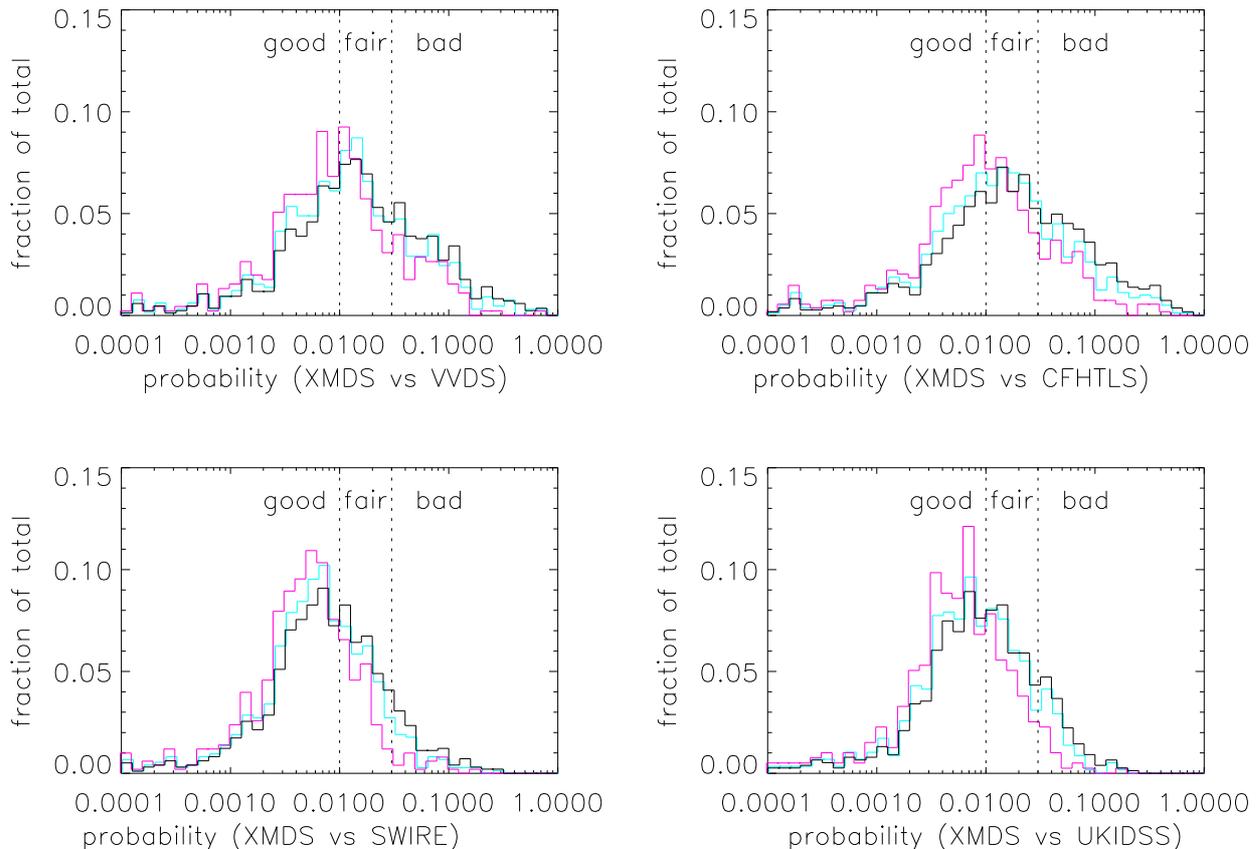
Fig. 2. Histograms of the four probabilities (*probvvds*, *probd*1, *probswire* and *probukidss* normalized to the total number of rank 0-1 objects with not undefined probability in the total sample (black), with a significance of at least 3σ in any band (cyan), or of at least 4σ (magenta). The dashed fiducial lines identify the boundaries of the loci with good, fair, or bad probability.

2.2. Input database tables

Table 1 gives a synoptic view of the database tables pointed from the GCT used for the present working catalogue.

Most tables where already present in the GCT used in Report I or Report III, and for them the counterpart association was not repeated (but the identification and ranking was repeated for all sources).

The tables which have been added ex-novo are those concerning CFHTLS release T004, the SWIRE DR6 IPAC release, and the UKIDSS DR3plus public release (which *replaces and overwrites* the previous table which was never used), plus tables referring to published papers.

For CFHTLS release T004, we used as input two tables elaborated by M.Polletta, one for the D1 field, and a comprehensive one for the W1 fields and "our" northern (ABC) fields (therefore superseding database table `cfhtnorth`), where duplicated sources in adjacent files had been natively removed. They have been ingested in temporary tables, and only the objects within $9''$ from an X-ray source are kept online (the correlation was done however within $6''$).

For SWIRE the latest release ("DR6") data have been supplied by IPAC in Jan 2008, with an update in Mar 2008 to remove some duplicated sources incorrectly left in. The files have been pre-processed by M.Polletta for simplification in the number of columns, classification of extended objects, and flagging of poor fluxes. With respect to the Spring 05 release, DR6 is less conservative and does not exclude sources below significance thresholds. Also DR6 natively includes MIPS data in all its bands (24, 70 and 160 μm. Data have been ingested in temporary tables, and only the objects within $10''$ from any X-ray source are kept online (the correlation was done however within $6''$). Technically there is an hidden table `swiredr6_ext` which contains both "aperture 2" and Kron fluxes (for IRAC, only PRF fluxes for MIPS), while table `swiredr6` is a *view* which selects "aperture 2" or Kron according to the fact the source is pointlike or extended following a recipe defined by M.Polletta.

For UKIDSS the latest release available when the ingestion was done was DR3plus. Data have been extracted

CFHTLS D1 (T004 and older) counterparts
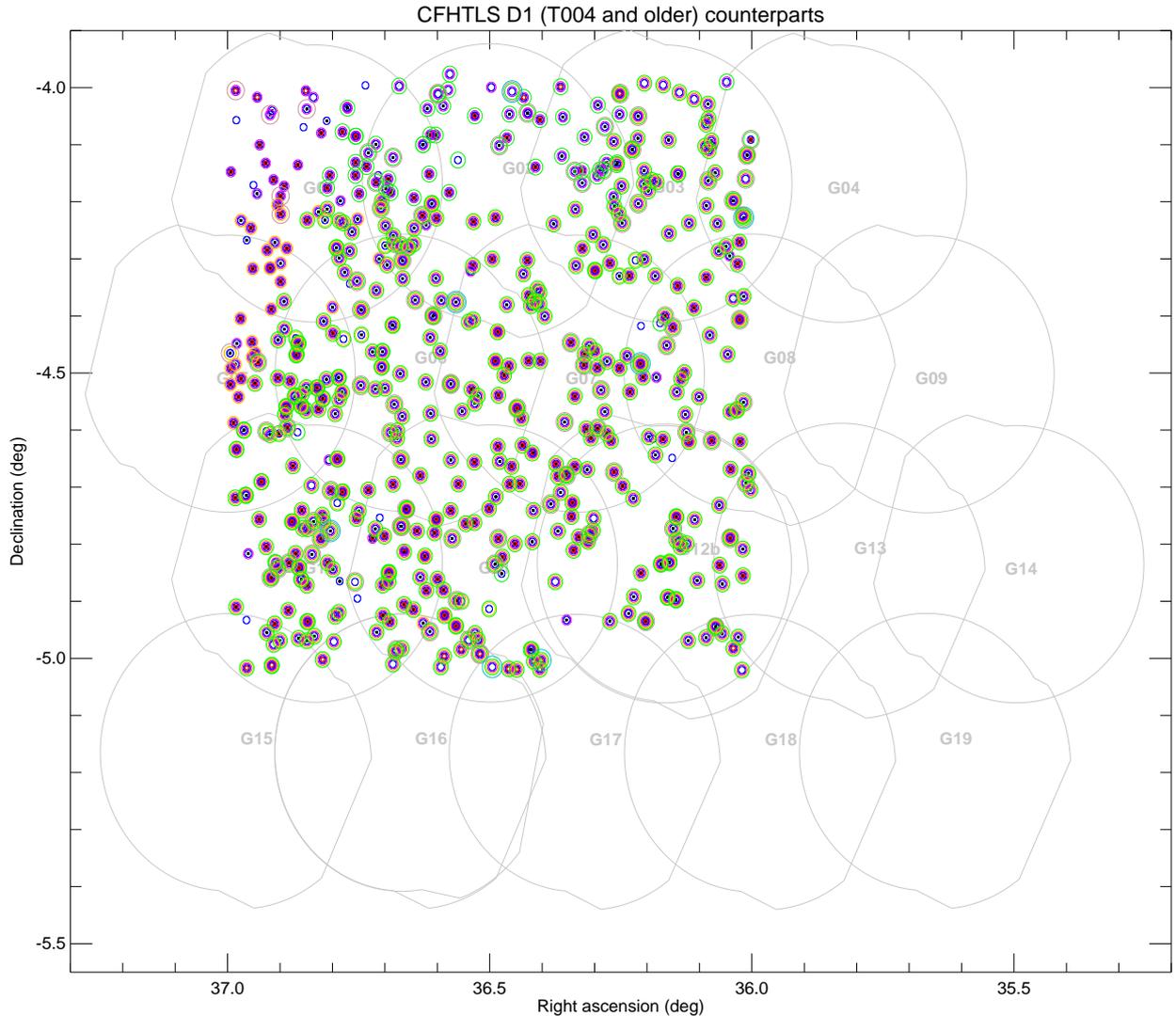
Fig. 3. Positions of the X-ray sources with a CFHTLS D1 counterpart. For symbols see 3.4 in text. The CFHTLS D1 covers the central part of the top three rows of fields.

at WSA within a radius of 10″ from a supplied list of positions, and then ingested *overwriting* any previous data (the DR1 UKIDSS data had not been used by anybody yet, so it was simpler this way). They have been taken from the DXS and UDS surveys, cover a larger area than before, and more bands (about half of the sources have both J and K, and one third has also H).

Concerning results of the Saclay X-ray analysis, the pointers provided are to the full band merged physical table nov06, inclusive of spurious sources, and fields not considered for the XLSS catalogue in Pierre et al. (2007). We however provide a flag xlsscat telling whether the nov06 entry is or not in the XLSS catalogue.

2.3. Astrometry

No new astrometric correction has been done in this release. The astrometric correction used remains the Paper I astrometry for the "VVDS fields" and the one described in Report I for fields G09 and G14 to G19, and in Report III for field G12bis. We refer therefore to Fig.4 of Report I, and Fig.1 of Report III.

3. The procedure

As said above, the starting point for the present version of the catalogue is the *same* GCT generated by the procedure described in Report III, i.e. the counterpart associations for all tables, but the ones marked as **new** in Table 1 (i.e. CFHTLS T004, SWIRE DR6 and UKIDSS) has *not* been repeated, but simply *preserved from the previous version*.

The details of our procedure are given in http://sax.iasf-milano.inaf.it/~lucio/LSS/ TriIdent/procedure.html or in internal notes.

Some of the steps listed there, and summarized below, have been in practice applied more than once at different
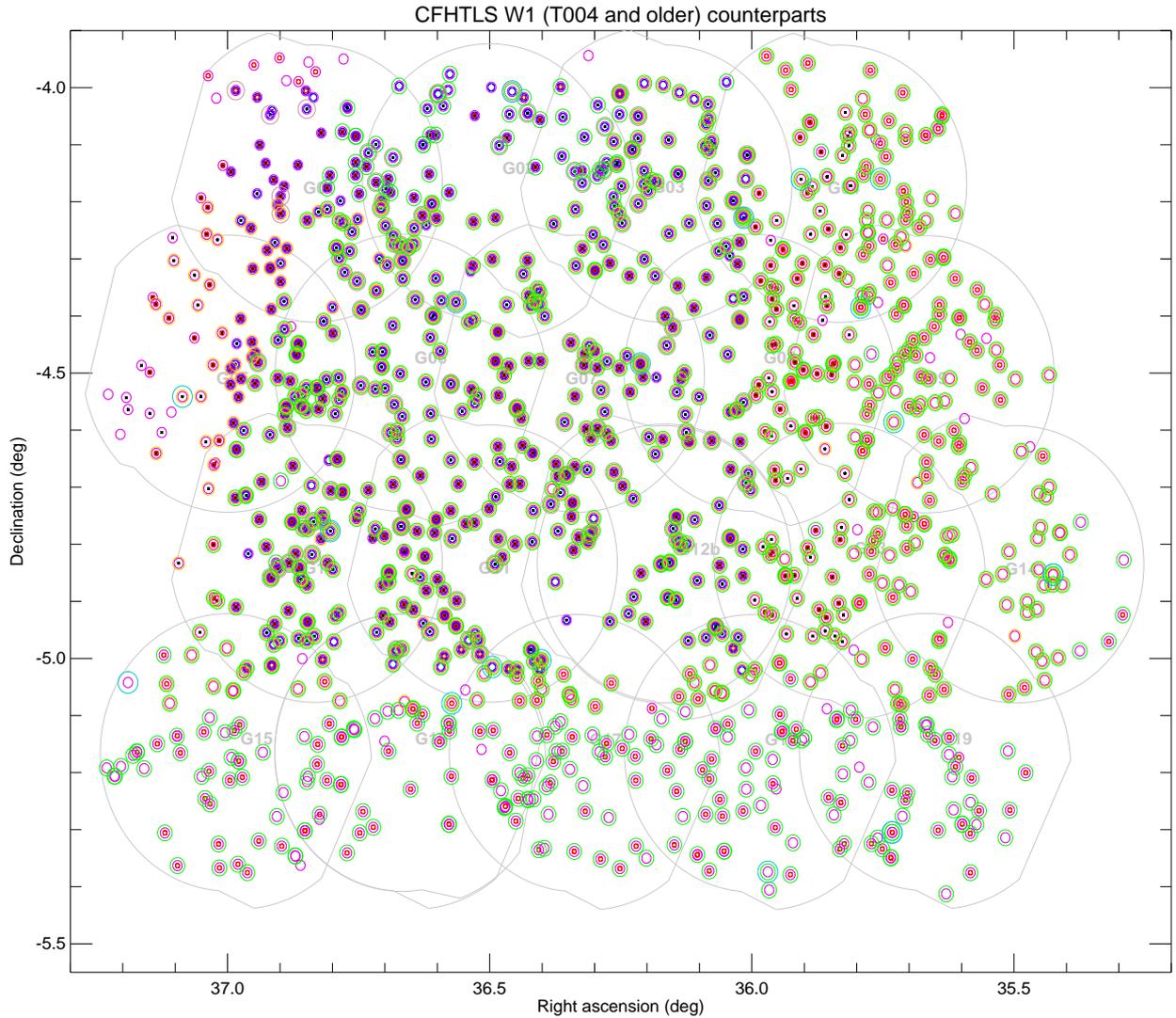
Fig. 4. Positions of the X-ray sources with a CFHTLS W1 counterpart. For symbols see 3.4 in text. The CFHTLS W1 covers all fields also outside the D1 area.

stages (sometimes altogether repeated), and/or not necessarily in the order in which they are listed.

3.1. Insertion of new tables

As in Report III the addition of counterparts from the new tables (in order `d1t4`, `w1t4`, `swiredr6`, and `ukidss`) has been managed in an *incremental* way, exploiting the associations already made.

– a preliminary step is to create a pointer column for the new table in the GCT.
– then one *inserts a pointer* to the new table entry into existing counterpart sets when the object in the new table is closer to one of the existing counterparts in other table within a predefined radius. E.g. in the case of `d1t4` objects they were first compared with CFHTLS objects of *earlier releases of the same table*

(here `d1t3` then `d1`), then with the other tables of the same origin (here for CFHTLS `w1t3` and `w1`), then with other optical (VVDS, `sacphot`), UKIDSS or SWIRE positions, and finally radio or external catalogues.
– the correlation radius used has been $0.5''$ when comparing positions of the same origin (e.g. CFHTLS to CFHTLS or SWIRE to SWIRE), $1''$ when comparing to other optical or SWIRE catalogues, $1.5''$ in case of external catalogues, and taking into account the radio position error box for radio catalogues.
– then one *inserts a pointer* to the new table entry into previous GCT entries flagged as blank fields.
– finally *new counterpart sets* are inserted when an object in the new table is close to an X-ray source within the appropriate correlation table, but has no correspondence in the other photometric catalogues.
– next a verification of the pre-flagging is made for all modified entries (a case flagged "unique counterpart"
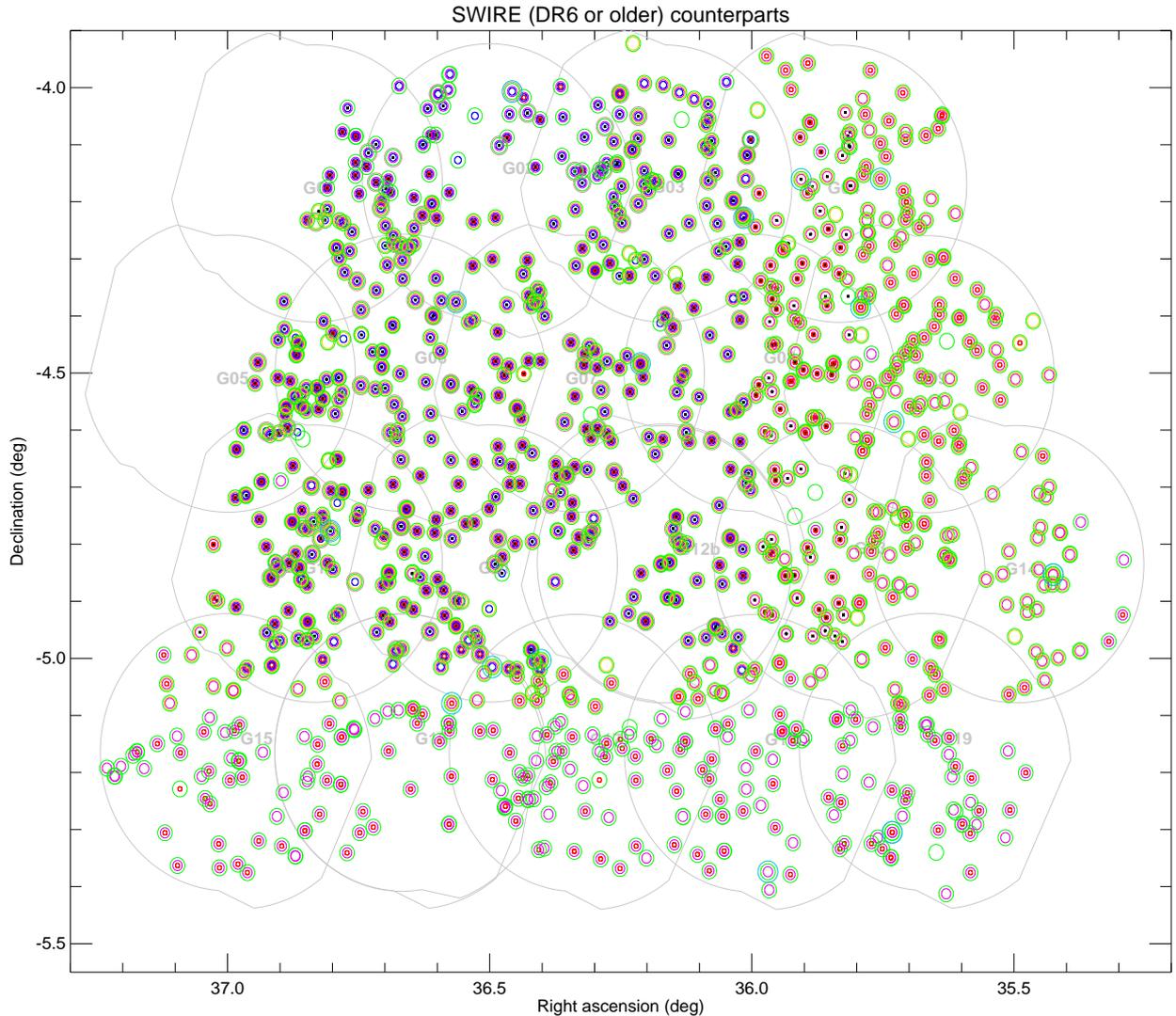
Fig. 5. Positions of the X-ray sources with a SWIRE counterpart. For symbols see 3.4 in text. SWIRE covers almost all the fields, but excludes most of G05 and G01.

can no longer be unique, or no longer a blank field, no longer the brightest, etc.). Also some of the flags and technical ranks can and have to be propagated to new entries.

– In particular in case of CFHTLS photometry one can reassign the star, galaxy or faint source and the saturated source flags. Since these are independent, a source can have multiple or contradictory flags (in particular can be saturated in VVDS but not in CFHTLS, and still be flagged saturated; or can be flagged as galaxy in one case and faint in another, considering that the magnitude threshold where the star/galaxy classification becomes undefined (the "faint limit") is I=21.5 for VVDS, i'=21.0 for CFHTLS).

We refer to Report I and III for the definition of *technical autoranks*, not intended for scientific usage, and beyond the scope of the present report.

3.2. Computing probabilities

Similarly to what was done in Report III, we computed the probability of chance coincidence between the X-ray source and its counterparts, based on the X-ray to optical (or IR) distance, the optical or IR intensity, and the density of sources brigther than a given intensity.

We computed `four` probabilities : *probvvds*, *probd*1, *probswire* and *probukidss*. The latter however, as in Report III, has not been used directly in the pre-ranking procedure.

All the probabilities have been *recomputed afresh*. They are based on a formula like

$$probability = 1 - exp(-\pi\ n(brighter\ than\ m)\ r^2)$$

where r is the largest between the X-ray to counterpart distance and 2″ (i.e. the probability is *capped* to a distance corresponding to an intrinsic position uncertainty), and
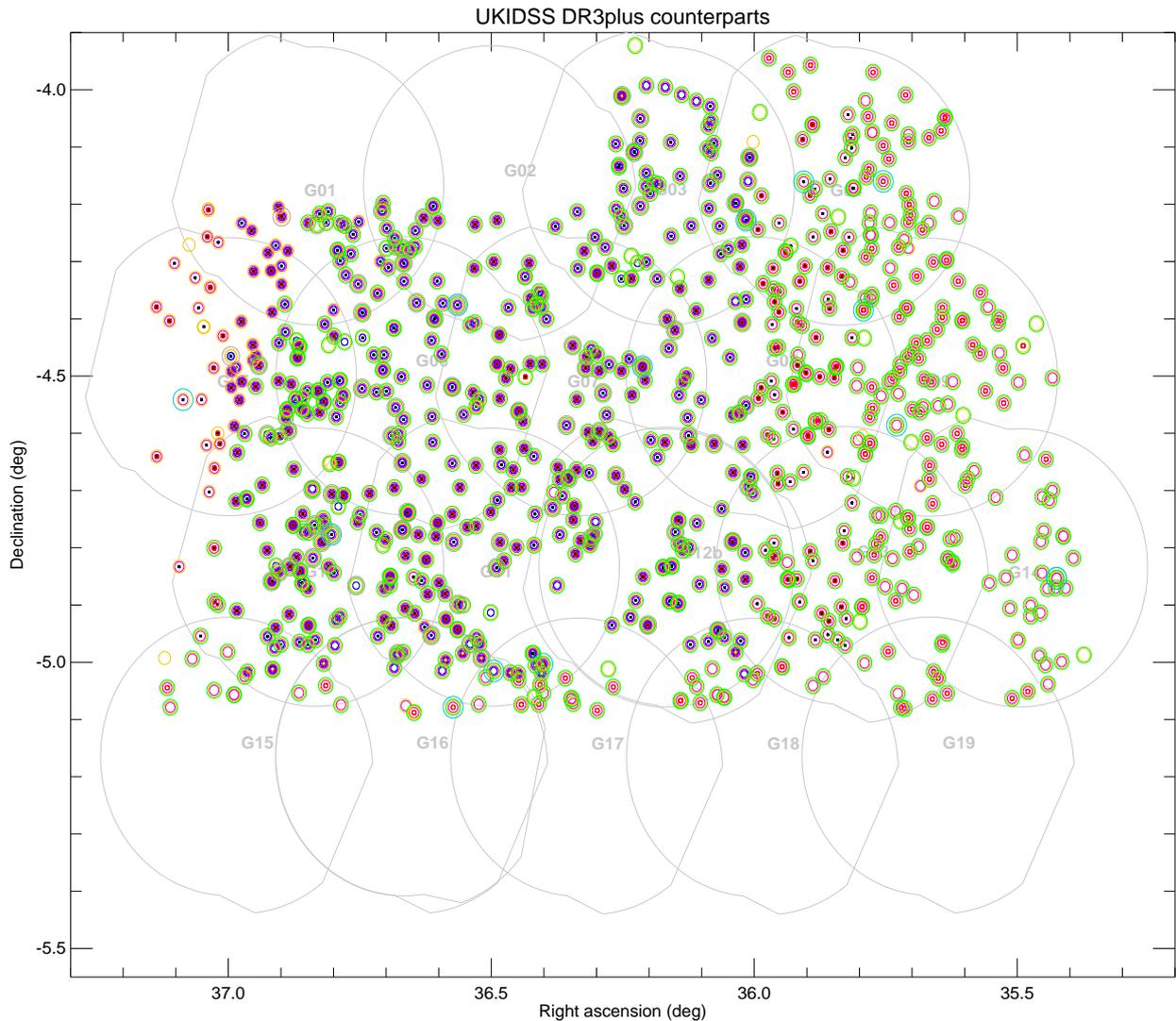
Fig. 6. Positions of the X-ray sources with a UKIDSS counterpart. For symbols see 3.4 in text. UKIDSS DR3plus covers only part of the XMDS area, although more than DR1plus (compare Fig. 11 in Report III).

the density $n(brighter\ than\ m)$ is computed from simple linear fits as reported in Table 2. The same table indicates also the magnitudes or fluxes used to look up the density for the appropriate band.

The usage of *capped probabilities* (which essentially privilege intensity with respect to distance in choosing between multiple counterparts closer than $2''$ to the X-ray source) was introduced in Report III.

Probability *probvvds* computed for sources with a VVDS counterpart (with preference to `virphot` then `bad`) has not been recomputed and is the same as in Report III.

CFHTLS probability, traditionally called *probd1*, has been recomputed for sources with a CFHTLS counterpart in the newer tables (in order `d1t4`, `w1t4`). For sources only in the older tables, it has not been recomputed, thus re-use the coefficients in Table 2 of Report III. For CFHTLS undefined magnitudes, the limiting magnitude of the field

(supplied by M.Polletta) has been used. CFHTLS sources in the northern ABC field have no i' magnitude. A different set of coefficients for the density fit using g' is available, but such case is never applicable to XMDS sources.

Probability *probswire* is computed in wavelength order for all sources in `swiredr6`. For sources only in the older SWIRE tables, it has not been recomputed, thus re-use the coefficients in Table 2 of Report III.

Probability *probukidss*, in the case both (J and K) magnitudes are present is the best (smallest) of the two.

A probability of 99 ("undefined") is assigned whenever it cannot be computed.

The density of VVDS sources is interpolated from the same data and with the same formula used for Paper I (see Fig. 1 top left panel).

The density of CFHTLS sources has been derived separately from the *totality* of the sources in the D1 T004

Objects	Total	$> 4\sigma$	$> 3\sigma$	$> 2\sigma$		
detections[a]	1358					
independent sources	1168	576	851	1165		
Condition	Total	unique	B & C[b]	blank fields	other	
rank 0 unambiguous	525	47	395	5[c]	78	
rank 1 unambiguous	450	17	166	1[d]	267	
rank 0/2 ambiguous	84	1[e]	56	0	27	
rank 1/2 ambiguous	109	0	39	0	70	
Condition	Total	score=3.0	2 to 2.5	1.5	1.0	0 to 0.5
All unambiguous	975	359	239[f]	151	83	143
All ambiguous (best candidate)	193	37	100	43	5	8
Condition vs "Brera rule"	Total	1 good	1 fair	2 `swires05`	2 (any)	3 prob.ratio
All unambiguous	975	509	775	751	824	510
All ambiguous (best candidate)	193	145	187	167	182	44

Rank 0 and 1 identifications						
with VVDS counterpart	595		+			
with CFHTLS D1 counterpart[g]	486 (526)	+				
with CFHTLS W1 counterpart[g]	1060 (1074)	=				
with either D1 or W1[g]		1088 (1111)	=			
with either VVDS or CFHTLS[g]			1102 (1122)			
with VVDS, CFHTLS or `sacphot`				1112 (1128)	+	
with SWIRE DR6 counterpart	988					
with any SWIRE counterpart	1006				=	
with optical[g] or SWIRE counterpart					(1161)	+
with UKIDSS	779					
with UKIDSS only	2					
with optical[g], UKIDSS or SWIRE						(1163)

Rank 0 and 1 identifications with counterpart in	Condition	Total	$p < 0.01$ good	$p < 0.03$ or fair	$p > 0.03$	
W1 T004		1060				
W1 or D1 T004		1088				
W1 or D1 T004 *and* SWIRE DR6		937				
W1 or D1 T004 *or* SWIRE DR6		1139	643	951	188	

Table 3. Basic statistics ot the present XMDS catalogue

a at $p < 2 \times 10^{-4}$

b brightest and closest

c blank fields flagged as autorank=4 or flag 01 set, see text

d affected by bright uncatalogued star (#624, see text)

e source #628 (unique catalogued optical ambiguous vs UKIDSS only, but uncatalogued optical visible in image)

f including 38 solitary counterparts for which max score is 2

g first value for T004, value in parenthesis also for older releases

and W1 T004 data (ingested in a temporary table), with a coarse fit to the data (see Fig. 1 top right panel)

The density of SWIRE sources has been derived in each waveband from the *totality* of sources in the DR6 catalogue (using IRSA Gator in count-only mode, which was not possible for data retrieval for the lack of the so-called "xpf" files) using aperture 2 fluxes; see Fig. 1 bottom left panel for 3.6μm (other bands not shown).

The density of UKIDSS sources has been derived separately for J and K bands from the *totality* of DXS data, using WSA in count-only mode: see Fig. 1 bottom right panel.

The computation of density is based on source counts, but requires the knowledge of a sky area, which I computed as in Report I, using a grid of cells 0.01×0.01 degrees and counting how many cells contain at least one object. I obtained for D1 an area of 1.02 deg^2, for W1 12.91 deg^2, for SWIRE 9.70 deg^2 and for UKIDSS DXS *now* 17.53 deg^2.

3.3. Ranking on probabilities

Once the probabilities have been computed, they can be used to assign *a priori* a preliminary rank (the *autorank*) to a particular association. For this I consider an individ-
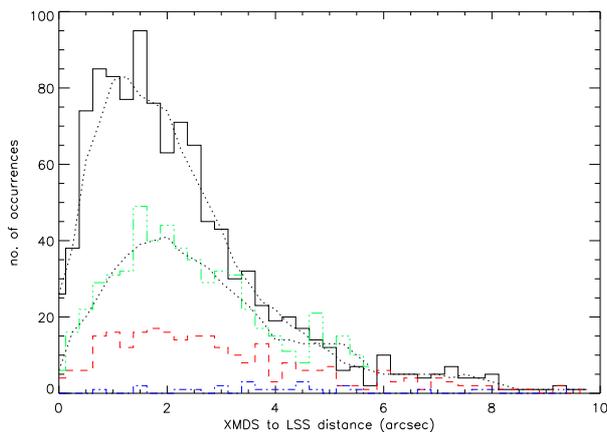
Fig. 7. Histogram of the angular distance between XMDS and XLSS X-ray positions in astrometrically corrected coordinates. The solid (black) line refers to all sources. The long dash (red) line to the source having a detection likelihood below 40 in the best band in which they are detected. The dash-dot (blue) line to extended sources. For comparison we report also the distribution of the inter-band distance `maxdist` between XLSS positions in bands B and CD, for sources detected in both bands (dash-dotted line, green). The dotted lines are smoothing of the first and last histogram.

ual probability p according to the following classification as in Report I and III :

– good if $p < 0.01$
– fair if $0.01 < p < 0.03$
– bad if $p > 0.03$
– undefined if $p = 99$

An *autorank=4* as well as flag 01 have been used to flag the blank fields (X-ray source is unidentified, has no counterpart).

For the rest *autorank*s have been re-assigned as described in Report III (including UKIDSS only sources). No further details are given here.

Ranks are also identical to what described in Report III, and defined according to the codes:

– rank -1 means *entry to be later totally removed*, i.e. technical autoranks, tiling overlaps, other duplicated or spurious cases
– rank 3 means *entry rejected* as counterpart
– rank 0, 1 and 2 indicate originally best, good or possible counterparts

The distinction between rank 0 and 1 is hardly relevant for sources with a single counterpart (or the few blank fields).

As in Report III, In the case of X-ray sources with more than one possible counterparts, a tie-break on probability is made so that one preferred counterpart receives a rank 0 or 1, and all other secondary counterparts receive a rank 2. Therefore single rank 2 entries never occur alone. To allow easy spotting of X-ray sources with more than one counterpart, I systematically set flag 09 "ambiguous" for all entries corresponding to an X-ray source with more than one (rank 0-2) counterparts.

For the ambiguous sources rank 0 and 1 has a precise definition, namely a rank 0 with one or more rank 2 companions means the association is only "possibly ambiguous" (the rank 0 should be definitely preferred), while associations of a rank 1 with one or more rank 2 companions are "definitely ambiguous" and intrinsically unresolvable.

As for Report III we use the so-called *three Brera rules*, and associated *scores*:

– rule 1: at least one of the three main probabilities is good (score=1.0) or fair (score 0.5)
– rule 2: The X-ray source has a SWIRE `swires05` counterpart (score 1.0) or a `swiredr6` or `swire` one (score 0.5)
– rule 3: The counterpart set with the best (smallest) probability (considered as the smallest of the three main ones) has a probability ratio of at least 10 with respect to all other possible counterparts (for verification purposes this is done considering rank 2 and 3 counterparts). If rule is met, a score of 1.0 is added.

As in Report III a data inspection was done to verify the score and solve possible problems. Any peculiarity was always noted in the "comments" associated to the source entry.

I recall that in the database one can combine rank, (autorank) and flag to select counterparts by quality. E.g. an expression like `rank=0 and not find_in_set('09',flags)` locates the single rank 0 sources (the best), while `rank=1 and find_in_set('09',flags)` locates the best counterpart of the "really ambiguous" cases.

3.4. General properties

At the end of the identification procedure, the catalogue contains 1403 valid entries (i.e. those with rank 0-2) corresponding to 1168 distinct X-ray sources.

Namely there are 609 rank 0 and 559 rank 1 identifications (each one corresponding by construction to a single X-ray source). In addition there are 235 entries with rank 2 corresponding to 193 distinct X-ray sources (84 of them have a rank 0 entry, i.e. are "possibly (0/2) ambiguous" and 109 have a rank 1 entry, i.e. are "definitely (1/2) ambiguous").

5 of the rank 0 entries have `autorank=4`, i.e. are blank fields (X-ray sources with no counterpart). There is an entry currently flagged with flag 01 (blank field) which is not a blank field, but spoiled by a bright uncatalogued star (X-ray source #624).

I summarize some more statistical information in Table 3, giving a breakdown by significance, identification reason, or kind of counterpart(s).

Sample	Survey Prob.	VVDS	CFHTLS	SWIRE	UKIDSS
Total	good	38.1	35.7	59.9	54.2
	or fair	69.1	64.8	88.6	84.5
3σ	good	43.2	42.7	68.0	60.4
	or fair	75.4	73.5	93.0	89.2
4σ	good	53.0	53.2	77.4	71.5
	or fair	82.4	81.6	96.2	95.1

Table 4. Percentages of proposed identifications with given probability range

Of the 11 objects without an optical or SWIRE identification 6 are the already discussed "pretended" blank fields inclusive of the contaminated #624, 1 is UKIDSS only, and the other are bad or faint photometry not confirmed in T004.

I note also that, not obvious from the table, there are also 33 objects which have a SWIRE counterpart without an optical counterpart in our catalogues, which reduce to just 14 without even an UKIDSS counterpart. Some of them have flag 02 ("weak sources") set. Possibly also others, specially when with an old `swire` counterpart only, could be classified as "weak source" or "blank" fields, but others again are confirmed in DR6 and sometimes correspond with an uncatalogue enhancement in the optical finding chart.

Fig. 2 represents the distribution of the probabilities defined in section 3.2. One can see that the choice of the good, bad and fair probability ranges made in 3.3 is sound. In particular we report in Table 4 the percentage of objects having a good or good or fair probability for the total, 3 and 4σ samples.

It is very useful to evaluate whether in a given region we do not find counterparts in a given table because either they do not exist or the region has not been observed. For these one can use the plots presented in Fig. 3 to 6 (the figures for VVDS and radio are not repeated here, see Fig. 6, 7 and 12 of Report III). They give the sky areas covered by the various surveys overplotted on the footprint of the FoV of our fields. Each figure lists *only* the (autorank 0-1) sources with a counterpart in a *given* table (i.e. a non null entry in the GCT). The symbols used indicate in which other tables there is *also* a counterpart. Such symbols are concentric circles of different colours, corresponding, as in Report III, from the inner to the outer to:

- black dot: VVDS counterpart
- red circle: `sacphot` (CFH12K) source
- blue: CFHTLS D1 source (T004 or older)
- magenta: CFHTLS W1 source (T004 or older)
- orange: UKIDSS DR3plus source
- green: SWIRE source (DR6 or older)
- pink: `virradio` source
- largest cyan circle: `radio` source

3.5. XMDS vs XMM-LSS

We remind here the main differences between the XMDS and the Xamin (Saclay) pipeline. *(a)* XMDS uses the SAS to detect candidates in 5 energy bands simultaneously operating on event files merged from all 3 XMM cameras and from the entire XMM field of view, and *(b)* then applies the Baldi et al. (2002) characterization; *(c)* the event pattern selection; *(d)* the removal of redundant sources; and *(e)* the astrometric correction are also different.

I update here the comparison between our XMDS sources and the sources in the `nov06` table and in the published XMM-LSS (`XLSS`) catalogue (Pierre et al., 2007). In particular I supplement it with the `INTERIM` catalogue (Chiappetti , 2008) and with a comparison with the most recent family ("corr2") of astrometrically corrected coordinates after the fixing of the Xamin half-pixel bug.

Of our 1168 XMDS sources, 1076 have a `nov06` counterpart, and in 951 cases they are in `XLSS`. Of the remaining 125, 55 are flagged as spurious on the basis of the Saclay detection likelihood (so are not eligible for `XLSS`). The other 70 are not spurious but some (20) fall in XMM fields considered not eligible for `XLSS` (mostly G12, while G12bis was not analysed in `XLSS` but is included in `INTERIM`, plus also B04 and B32).

Of 92 XMDS sources which are not in `nov06`, 51 are at high ($> 10'$) off-axis angles. About half (19) of the other are ultrasoft (band A) sources (with such band not considered in the Saclay pipeline), otherwise they are usually very weak sources anyhow.

As said above the XMDS catalogue includes 1168 (non redundant) X-ray sources above its probability threshold, while the present `XLSS` catalogue has 1574 X-ray sources in the G good fields. The `INTERIM` catalogue has 1618 sources in its good fields (adding G12bis) We correlate the catalogues within a distance of $10''$.

For the low significance or extended `XLSS` sources not appearing in XMDS, refer to Report III.

The number of common detections between ours and the XLSS catalogue is nominally 875 detected in the same field (871 for `INTERIM`), but this value can be raised to 1006 (or 995) if we consider the field of the primary detection of our stacked entries, or to 1064 (or 1054) taking into account the different handling of redundant detections (which each catalogue can, according to its own criteria, remove in different pointings). We concentrate now on the comparison of such common objects.

The distance between the (astrometrically corrected) X-ray positions is shown in Fig. 7. The newest astrometric correction of the `INTERIM` catalogue is used. However there is no great difference between Fig. 7 and Fig. 14 of Report III, so probably both astrometric corrections do an equivalent job. 57% of the sources are closer than $2''$, 88% closer than $4''$ and only 4% are more distant than $6''$, in general concentrated among the sources with lesser significance, and the few extended ones. The agreement between the

XMDS and `INTERIM` positions, peaking around $1''$, is better than the typical inter-band distance between `INTERIM` detections in the two energy bands, which peaks around $2''$.

Fig. 8 attempts to cross calibrate the detection likelihood of `Xamin` with the chance probability of the XMDS (for definition see Baldi et al. (2002)).

Alternatively one can use Fig. 9 to cross calibrate the detection likelihood of `Xamin` with the significance in terms of number of σ of the XMDS (see Paper I and references therein). Both figures update and supersede Fig. 15 and 16 of Report III.

The common subset of sources includes 24 potentially extended sources (14 classified C1, of which 3 detected in both bands and 1 detected only in the hard band but satisfying the extension (C1) criterion in such band); 10 classified C2 of which none detected in both bands, and 4 detected only in the hard band but satisfying the C2 criterion in such band).

To compare the count rates in the two catalogues, one has to note that XMDS operates on camera-merged event files, and therefore computes automatically a MOS1+MOS2+pn rate, while `Xamin` works on MOS1+MOS2 and pn separately. Therefore we compare the XMDS rate with an `INTERIM` camera merged rate computed as

$$rate = \frac{rate_{MOS}(exp_{MOS1} + exp_{MOS2}) + rate_{pn}exp_{pn}}{exp_{MOS1} + exp_{MOS2} + exp_{pn}}$$

(As in Report III this rate is not the same as the plain summed rate used in Pierre et al. (2007))

The count rates are compared in Fig. 10, updating Fig. 17 of Report III (see further considerations therein). The average ratio is closer to unity but different, namely the `INTERIM` rate is 1.103 times higher than the XMDS one in the B band, and 1.184 times higher in the CD band.

The fluxes, computed for XMDS according to the prescriptions of Baldi et al. (2002) and for `INTERIM` as explained in Pierre et al. (2007), are compared in Fig. 11, updating Fig. 18 of Report III (see further considerations therein). The systematic difference in flux is : the `XLSS` fluxes are 0.849 lower than the XMDS fluxes in the B band, while they are only 0.95 lower in the CD band.

4. Database tables and tools

The catalogue described here is still considered for internal use only. I describe here the ways and tools used to access it.

4.1. Tables and views

The underlying arrangement in our database is analogous to what used up to Report III included, i.e. there is a GCT (Glorified Correlation Table) linking all the 30 member tables (the physical tables listed in Table 1). However the GCT is not accessed directly. In the past it was accessed via a provisional virtual table, but usually one then specified columns in the member tables directly if they weren't defined as virtual columns.

Presently the best way of accessing the data (publicized in workspace `xmdsm`) is via a *view*, which is less clumsy than virtual tables. The view corresponding to the present catalogue is named `WORK4`. This gives access to a sensible selection of virtual columns. Access to columns in member tables remains possible via older (and clumsier) virtual table (not publicized) or, with some efficiency penalty, ticking on the appropriate tick box.

There are also *views within views* which limit access to a subset of rows (records, entries) in `WORK4`.

4.2. Data products

Currently the situation about data products associated to the present catalogue is not final. The following data products exist, but not all of them are formally associated to `WORK4` yet.

The formal association is indicated in italics in the following list, and means the product is accessible via the normal database interface. Other products are accessible so far only via a tool described in 4.3

- *comments* are the optional textual comments associated to each X-ray source when necessary, as already mentioned
- *SIMBAD pointers* and
- *NED pointers* are external links to the relevant web pages, when a source has a non-null counterpart there
- the remaining data products are FITS images with the WCS (and orientation) as supplied by whoever produced them
- *VVDS thumbnails* are $40 \times 40''$ I band images centered on X-ray sources falling inside the VVDS area
- CFHTLS thumbnails are $40 \times 40''$ i' band images centered on X-ray sources with a W1 T004 counterpart (from the T003 public image archive at CADC)
- SWIRE thumbnails are a family of up to 7 images (in the IRAC and MIPS bands) centered on X-ray sources with a SWIRE counterpart (in any release). Size is $30''$ for IRAC and $60''$ for MIPS.

It should be possible to retrieve also UKIDSS thumbnails from WSA, but since they use an unusual WCS (`RA---ZPN DEC--ZPN` currently unsupported by the tool described in 4.3), their retrieval has been defererred sine die.

4.3. Special tools

Besides the normal database interfaces, there are two additional tools which are useful to support the identification work.
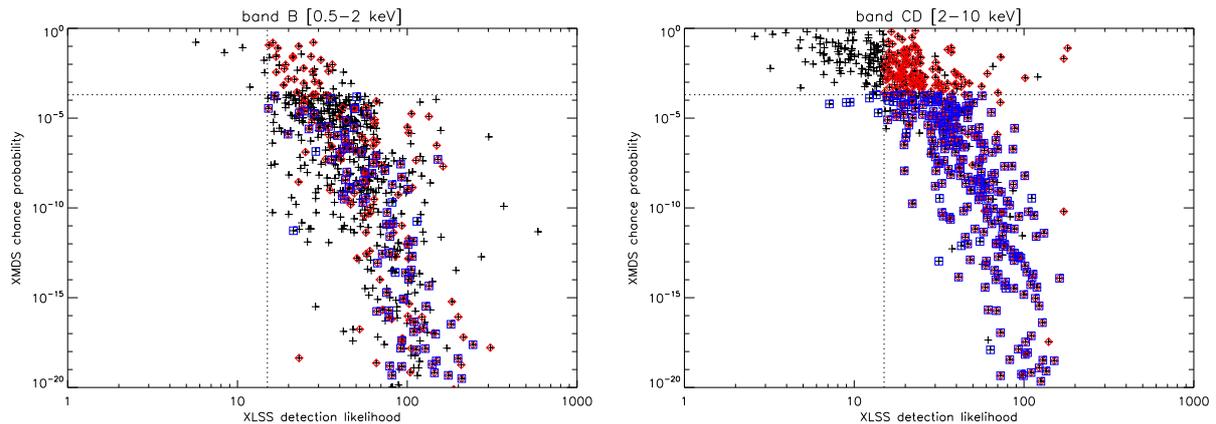
Fig. 8. Cross calibration between the `XLSS` (Xamin) detection likelihood and the XMDS chance detection probability. Left panel for the soft band, right panel for the hard band. The dashed lines indicate the two acceptance thresholds of $ML > 15$ and $p < 2 \times 10^{-4}$ (remember that a source can be accepted if it is above the threshold *in at least one band* but not necessarily in all). Crosses indicate all objects detected in the given band. A (red) diamond surrounds the sources detected above threshold *in both bands* in the `XLSS` or `INTERIM`, while a (blue) square surrounds those detected above threshold *in both bands* in the XMDS.
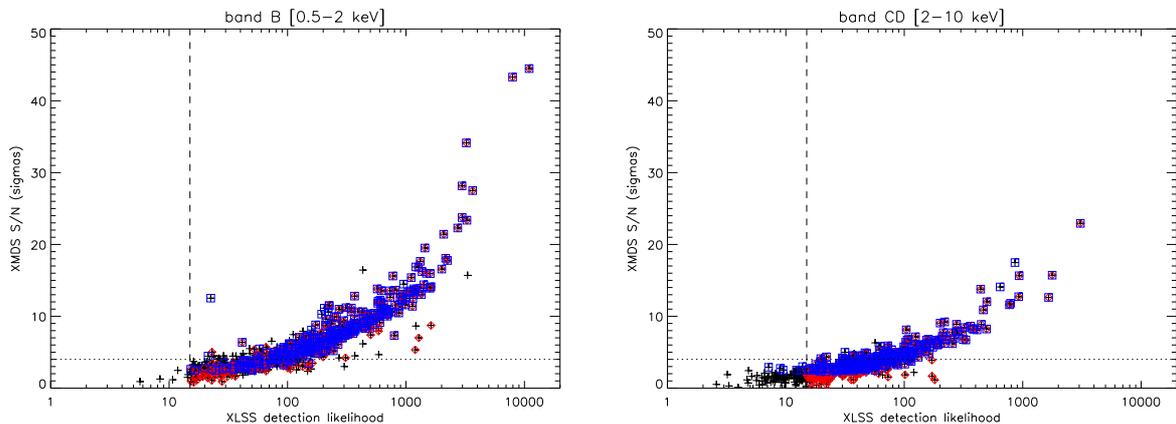
Fig. 9. Cross calibration between the `XLSS` (Xamin) detection likelihood and the XMDS signal to noise ratio. Left panel for the soft band, right panel for the hard band. The dashed vertical line indicates the `XLSS` acceptance thresholds of $ML > 15$ (remember that a source can be accepted if it is above the threshold *in at least one band* but not necessarily in all), while the dotted horizontal line shows the conventional level of 4σ. Crosses indicate all objects detected in the given band. A (red) diamond surrounds the sources detected above ML threshold *in both bands* in the `XLSS` or `INTERIM`, while a (blue) square surrounds those detected above chance proability threshold *in both bands* in the XMDS.

The former is the so-called *validation interface* (see Fig. 12), which was in use since Paper I as a maintenance tool for the current GCT. It is a web interface located at URL `http://cosmos.iasf-milano.inaf.it/~lssadmin/Website/LSS/Query/Validate/` which allows to view all entries of the GCT relevant to one X-ray source, and to edit them (entries can be deleted or added, counterparts can be removed or added, flags and ranks can be edited, comments may be entered).

The second tool is a graphical interface which could in the future be integrated in the main database interface, but currently is separated from it, pending the definition of the final protocol.

The interface itself, allowing to view thumbnail images, and overlaying onto them "regions" corresponding to the sources in the relevant counterpart sets, is based on a Java applet, and can therefore be used on any browser.

Access occurs presently through URL `http://sax.iasf-milano.inaf.it/~lucio/temp/Java/TestWeb/combo24.html`, which currently displays two horizontally adjacent frames. The left one is an interface to a servlet which provisionally emulates the main database interface (currently allows to select the X-ray source of interest), while the right one contains the applet itself.

The applet can display an image (and control its look and zoom), onto which regions are overlaid. The regions (corresponding to counterparts in all counterpart sets) are also listed in a table, and one can interactively hilight one
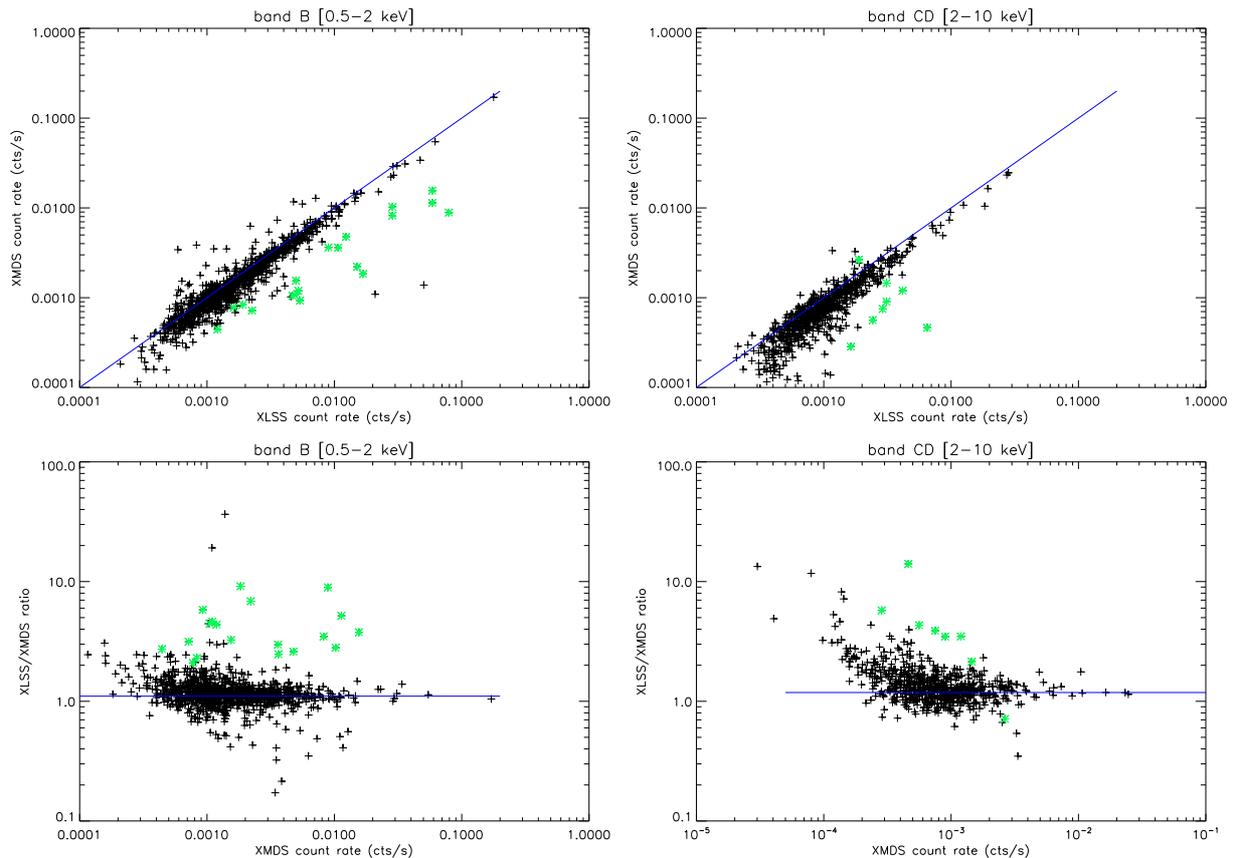
Fig. 10. Two alternate ways of comparing count rates. Top : The XMDS count rate vs the `INTERIM` camera-merged count rate for band B (left panel) and band CD (right panel). The diagonal solid line is a fiducial line corresponding to equal XMDS and `INTERIM` rates. Bottom : The ratio of the `INTERIM` camera-merged and XMDS count rates as function of the XMDS count rate for band B (left panel) and band CD (right panel). The horizontal solid line is a fiducial line corresponding to the actual average ratio in the band (see text). In both panels (black) crosses mark pointike sources detected in the band, while (green) asterisks correspond to extended sources.

or more of them selecting them with the mouse in the image or in the table (see left panel in Fig. 13).

The servlet computes the URLs of all existing images in the various wavebands. The user of the applet can then retrieve directly one of them (with no further servlet interaction, see right panel in Fig. 13), and also define a default waveband. Currently when no VVDS thumbnail exists, a thumbnail is created from DSS-II via a CGI accessing a repository at ESO (note that this may take time, while all other data products are local). The regions are retrieved from the servlet and, once stored in the applet, can be overlaid on any image.

References

Baldi, A., Molendi, S., Comastri, A., et al. 2002, Ap.J, 564, 190

Bondi, M., et al. 2003, A&A, 403, 857

Chiappetti, L., et al. 2005, A&A, 439, 413 (Paper I)

Chiappetti, L., 2006a, XMM-LSS Internal Report N. 1-Mi (Report I)

Chiappetti, L., 2007, XMM-LSS Internal Report N. 3-Mi (Report III)

Chiappetti, L., 2008, XMM-LSS Internal Report N. 5-Mi (`INTERIM` Report)

Cohen, A. S., et al. 2003, Ap.J 591, 640

Garcet, O., et al. 2007, A&A, 474, 473

Le Fèvre, O., et al. 2004, A&A, 417, 839

Pierre, M., et al. 2007, MNRAS, 382, 279

Polletta, M., et al. 2007, ApJ, 663, 81

Tajer, M., et al. 2007, A&A, 467, 73
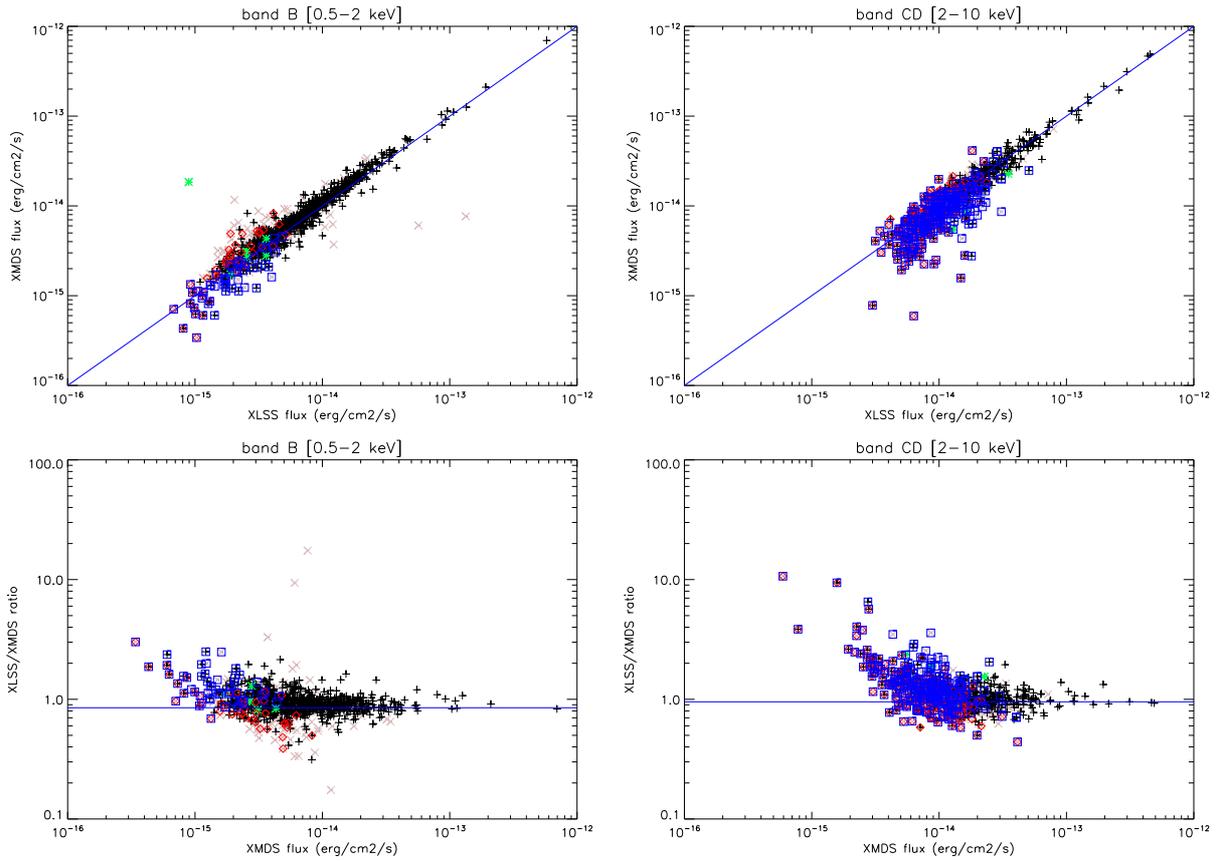
Fig. 11. Two alternate way of comparing fluxes. Top: The XMDS flux vs the `INTERIM` flux for band B (left panel) and band CD (right panel). The diagonal solid line is a fiducial line corresponding to equal XMDS and `INTERIM` fluxes. Bottom: The ratio of the `INTERIM` and XMDS fluxes as function of the XMDS flux for band B (left panel) and band CD (right panel). The horizontal solid line is a fiducial line corresponding to the actual average ratio in the band (see text). Symbols common for both: The (black) crosses indicate pointlike sources which have a `fluxflag` of 0 or 1, the (pink) X those with a `fluxflag` of 2, i.e. where the MOS and pn fluxes differ by more than 50%. The (green) asterisk correspond to extended C2 sources for which the flux is computed from the pointlike rates (C1 sources have flux set to undefined and are not plotted). A (red) diamond surrounds the points with a poor `XLSS` likelihood $15 < ML < 20$ in the band. A (blue) square surrounds the points with a poor XMDS probability $p > 2 \times 10^{-4}$ in the band. Note this symbology is different from the one used in Fig. 9.

Fig. 12. Appearance of the *validation interface* with a counterpart set loaded
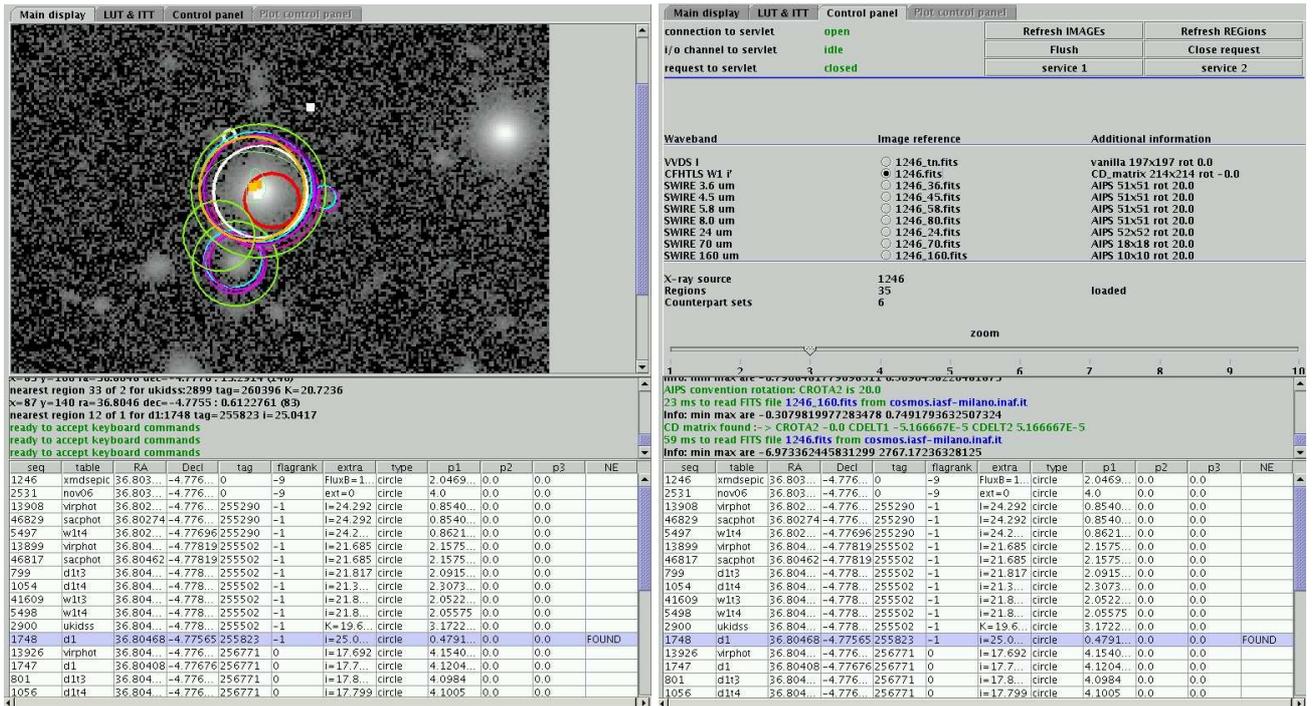
Fig. 13. Appearance of the *applet interface* for images and regions. In the left panel the top *tabbed pane* shows a panel with a CFTHLS image and all elements in counterpart sets of any rank, including rejected ones, shown as overlaid regions. One of the rejected counterparts is hilighted in the display and in the bottom table. In the right panel the top tabbed pane shows the control panel, listing, among other information, all the available image data products.