XMM-LSS
INTERNAL
REPORT n°6-Mi

October 27, 2009

The 2XLSS catalogue

A status report in advance of publication

L.Chiappetti[1]

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

Abstract. I report on the recently released 2XLSS catalogue, giving reference information similar to the published XLSS catalogue or to what provided in internal reports for the XMDS and INTERIM catalogues. The catalogue tables contain X-ray results deriving from the reprocessing (jun09) with the latest (Py3.2) Xamin pipeline of all our observations up to AO7 included, and associated optical, IR and UV information.

Key words: LSS

1. Introduction

The *published* catalogues deriving from the XMM-LSS so far have been only the XMDS/VVDS 4σ catalogue (Chiappetti et al., 2005) with just 286 entries, and the XMM-LSS catalogue version 1 hereafter XLSS (Pierre et al., 2007) with 3385 entries. They are supplemented by a number of papers describing other samples of AGN (Tajer et al., 2007; Polletta et al., 2007; Garcet et al., 2007) or clusters (Pierre et al., 2006; Pacaud et al., 2007) extracted sometimes from (larger) working catalogues available internally in the Milan database (or elsewhere).

Various versions of such working catalogues for the XMDS, or more precisely, for the results of the analysis of XMDS fields (so called G fields) done with the Milan pipeline (Baldi et al., 2002), for a total of 1358 detections and 1168 independent X-ray sources, have been documented in internal reports (Chiappetti , 2006a,b, 2007, 2008a).

The database tables and the working INTERIM catalogues derived from the results of the analysis of fields up to AO5 (B fields *and* G fields), and including also the SXDS fields, with the *pro-tempore* version of the Xamin Saclay pipeline (Pacaud et al., 2006) were documented in Chiappetti (2008b), hereafter Report V. They were not published, were little used internally, and did not include AO7 data obtained in early 2009 (so called feb09).

All X-ray data from GTO to AO7 were recently reprocessed systematically at Saclay with the latest (Py3.2) Xamin pipeline. Such result constitute the so-called jun09 release of database tables, and were the starting point for the current 2XLSS catalogue.

I therefore present here the 2XLSS catalogue with 6282 entries. The SXDS fields (Ueda et al., 2008) which are located in a "hole" of the XMM-LSS, are presently excluded. The 2XLSS catalogue has just been released for internal use, and it is intended to publish it after a period of usage to assess its properties and quality.

In section 2 I list the input database tables used as starting point, namely X-ray data (2.1) and optical-IR-UV data (2.2), while other ancillary tables contained in the Milan database are briefly mentioned in section 2.3, and the astrometric correction in section 2.4. The procedure used to create the 2XLSS catalogue is described step-by-step in the various subsections of section 3, with particular regard to the X-ray tables (3.3), the X-ray/optical catalogue (3.5) and the data products (3.7). A comparison with earlier releases is presented in 3.4, namely the raw database tables are compared in 3.4.1 and the catalogues in 3.4.2. Section 4 discusses perspective for future identification work, in particular the pre-ranking (4.1) based on the probabilities described in 3.6, and possible aid tools (4.2). Section 5 gives some summary statistics on the catalogues, the X-ray one (5.1) and the X-ray/optical one (5.2).

2. Data sources

The starting point for the 2XLSS X-ray catalogue proper have been the latest release (jun09) of X-ray tables (see 2.1). For the 2XLSSOPT virtual table (and the astrometric correction, see 2.4 !) some other recently ingested or pre-existing optical,IR and UV tables have been used (described in 2.2). All used physical tables are listed in Table 1.

The ending point, analogous in this to what done for the XLSS catalogue version I (Pierre et al. (2007), hereafter the XLSS paper), are a number of *glorified correlation tables* (GCTs; tables of pointers into a predefined combination of database tables, each one correlated with the main X-ray table with a "standard" correlation radius

Table	Update	Content	History	(5)	(6)
jun09*	**Jul 09**	X-ray sources from latest Saclay pipeline, band merged within 6″	**new!! added Jul 09**	6″	a
subaru*	Aug 08	X-ray sources from the Saclay pipeline, band merged within 6″	SXDS observations analysed by us, originally ingested in Oct 07; In Aug 08 fixed zero-exposure bug and coordinate bug and *replaced* revised astrometry	6″	a
d1t4	Feb 09	CFHTLS D1 field release T004 supplied by Saclay	in use since Jan 08	6″	
w1t4	Jun 09	CFHTLS W1 fields release T004 supplied by Saclay	In use since Jan 08; **added objects**	6″	
swiredr6	Aug 09	SWIRE DR6 supplied by IPAC	in use since Jan 08; **added objects**	6″	
ukidssdr5	**Aug 09**	UKIDSS DR5plus public release	**new!!** replaces completely old DR3 ukidss table; first time used with XMM-LSS	6″	
galex	**Aug 09**	GALEX GR4/5 public release	populated since Nov 08 for XMDS; **new!!** first time used with XMM-LSS	6″	
simbad	Aug 09	SIMBAD sources	present since 2003 and regularly updated	20′	b
ned	Aug 09	NED sources	present since 2003 and regularly updated	20′	b
usno	Oct 07	USNO A2 catalog as kept at ST-ECF.	present since 2005 and regularly updated	6″	
stalin09	**Sep 09**	Table 2(,3 and 4) from Stalin et al. paper		n/a	c

Table 1. Database tables used as input to the present 2XLSS 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 the radius in column (5) is used for band merging and overlap removal (see 3.2) in the case of X-ray tables
b SIMBAD and NED may also include objects from some of our catalogues (e.g. radio and XLSSC).
c Stalin et al. (2009)

or criterion), above which the catalogue *virtual tables* are based.

2.1. X-ray data

The starting point for the 2XLSS catalogue proper are the family of physical tables jun09 (constituted, as usual, by the two single-band tables jun09b and jun09cd, and by the band-merged table jun09, see 2.3). Such tables were ingested from FITS catalogues supplied by Saclay and produced by the latest Xamin pipeline version (Py 3.2) used to reanalyse afresh all fields (GTO, AO1 and AO2 whose previous processing was in nov06; AO5 previously in jul07; and AO7 previously in feb09).

It has to be noted that the SAS steps preliminary to Xamin introduce little random variations which may affect fainter sources. In particular, although feb09 and jun09 were processed with the same Py3.2 release, they are not identical. Also for such reason, the jun09 tables are intended to *supersede* all previous versions.

SXDS full exposures used in INTERIM should be compliant with the Py3.2 release, while "chunk" 10ks exposures were analysed with an earlier release. Some tests using subaru data were done around Oct 2009 and included in an unofficial draft of the present report but have now been removed.

The ingestion and in particular band merging for jun09 was done as described in section 2.3.5 of the XLSS paper, and it is outside of the scope of the present report. Similarly the computation of fluxes, and the extended source classification was also done at ingestion time, as described in sections 2.3.2 and 2.3.4 of the XLSS paper.

During the ingestion, caution was used so that the sequence numbering of sources in jun09 started where feb09 ended. This means the source numbering is unique and distinct from all previous tables (nov06, jul07, subaru and feb09), which prevents confusion and potentially allows jun09 and subaru to be concatenated.

I remind here the (pointing) field numbering and naming conventions. In particular I note that the field numbering (column field in physical tables and Xfield in catalogues) has remained the same as in the past, while the field *naming* convention has changed, and is now consistent with the one used in Saclay (field names are only relevant for filenames like those of data products, see 3.7).

- the original observation of a B field in any AO (up to AO5 included) is numbered n (e.g. field B01, observed only once, is 1, and field B04a, reobserved later, is also 4).
- some AO1-2 B fields were bad and were reobserved in AO5. They are numbered 500+n (e.g. field B04b, in the past called B04bis, is 504). Note that a field

observed for the first time in AO5 is numbered n (B33 is 33, B35a is 35).
- some AO5 B fields were also bad, and were reobserved in AO7. All AO7 fields are repeats, and are numbered 700+n (e.g. B04c, in the past called B04ter, is 704 and B35b is 735).
- the original observation of a G field is numbered 1000+n (e.g. field G07 is 1007)
- however field G16 was observed in two chunks (G16a and G16b) which are numbered 1116 and 1216
- additionally field G12a was bad, and was reobserved in AO5 as G12b, which is numbered 1112
- the 7 SXDS fields (full exposures) are numbered 2000+n (e.g. S01 is 2001)
- the first 10 ks of the SXDS field exposures (Saclay S0nc) are numbered 2100+n (e.g. S01c is 2101)
- the next 10 ks of the SXDS field exposures (Saclay S0nd) are numbered 2200+n (e.g. S01d is 2201)

Fields flagged as bad (typically those with the pn exposure under 7ks) are marked by a boolean flag column `badfield=1`. Such column name is for the physical `jun09` tables. The `2XLSS` catalogues use instead `Xbadfield=1`.

2.2. Optical, IR and other data

For CFHTLS release T004, we used (already at the time of the `INTERIM` catalogue) as input two files elaborated by M.Polletta, one for the D1 field, and a comprehensive one for the W1 fields and "our" northern (ABC) fields (therefore superseding previous database table `cfhtnorth`), where duplicated sources in adjacent files had been natively removed. They were 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″). This population procedure was updated using the `jun09` coordinates. It shall be noted that the `d1t4` table uses the standard CFHTLS undefined magnitude marker (99), while the `w1t4` follows the convention by M.Polletta, and replaces the undefined magnitude with the *negative* value of the limiting magnitude in the band for the specific W1 field. For the three northern field, where only $g'r'z'$ photometry is available u^* and i' are set to zero.

Earlier release tables existing in the database (`d1`, `w1`, `d1t3`, `w1t3`), used for the XMDS catalogues (see e.g. Chiappetti (2008a), hereafter Report IV), were **not** considered for the `INTERIM` catalogue, nor for the `2XLSS` one.

For SWIRE the latest release ("DR6") data were supplied by IPAC in Jan 2008, with an update in Mar 2008 to remove some duplicated sources incorrectly left in. The files were pre-processed by M.Polletta for simplification in the number of columns, classification of extended objects, and flagging of poor fluxes (already at the time of the `INTERIM` catalogue). With respect to the public 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 were 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″). This population procedure was updated using the `jun09` coordinates. 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.

Earlier release tables existing in the database (`swire` and `swires05`), used for the XMDS catalogues (see Report IV), were **not** considered for the `INTERIM` catalogue, nor for the `2XLSS` one.

For UKIDSS the latest release ("DR5plus"), containing data from the two surveys which overlap with us, DXS and UDS (the latter particularly covers the SXDS or `subaru` area) has become available in Aug 09 while we were processing `2XLSS`. For this reason the earlier release (table `ukidss`, used only with the XMDS; see Report IV) was abandoned, and a new table `ukidssdr5` was ingested retrieving from the WSA public archives all objects within 10″ from any X-ray source (in `jun09`, `subaru` and XMDS), using the *crossId form*. Such data could then be ingested directly.

For GALEX the public data available on the NASA MAST (GR4) were originally retrieved in the surrounding of XMDS sources and ingested in a database table. Such procedure was recently repeated, always using a radius of 10″, from the latest release called GR4/GR5 and the list of `jun09` positions. A tool called CasJobs available at MAST was used to do the correlation. The material ingested in our database includes all GALEX objects within 10″ of XMDS, `nov06`, `jul07`, `subaru`, `feb09` and `jun09` sources. Since it well known that the MAST GALEX catalogue contains redundant sources where GALEX pointings overlap (so called *tiling artifacts*), we have run a procedure to flag GALEX objects within 1.5″ from any other observed in a different tile, and to prefer one (observed in two bands, or with smallest inter-band separation, or with smallest off-axis angle).

The tables referring to external catalogues (SIMBAD, NED, USNO) have been recently updated with pointers to objects in the surrounding of `jun09` X-ray sources, and can be accessed in correlation with the `2XLSS` catalogue, although not members of it.

The tables referring to published papers, although sometimes correlated with one or more of the three original X-ray tables (`nov06`, `jul07`, `subaru`) are presently not correlated with `2XLSS`, with the exception of the latest table `stalin09` (for which however the author consulted our public `XLSS` catalogue), and in a limited way `garcet07`.

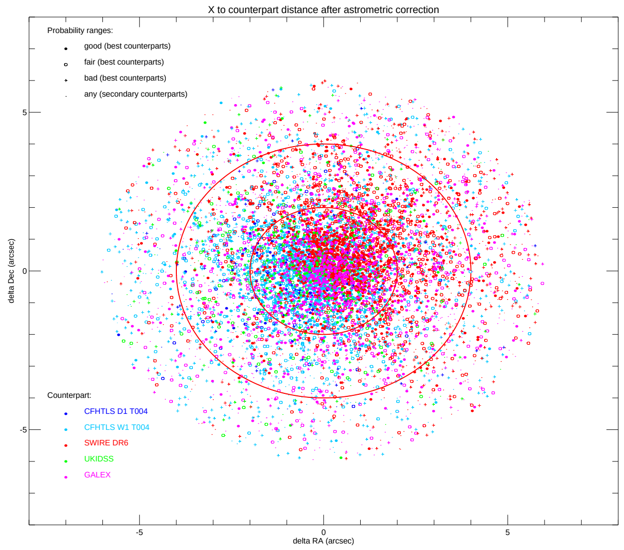

Fig. 1. Distances in RA and Dec between the X-ray corrected position and the counterpart position. Different symbols indicate the identification quality. A circle is plotted when the counterpart is the best one, and the chance probability is good or fair (filled in case of good probability). A cross is plotted for the best counterpart when the probability is bad. A dot is plotted for secondary (ambiguous) counterparts, irrespective of probability, but only if it is good or fair. Different colours (as shown on figure) indicate the origin of the counterpart position for the distance calculation. Two fiducial radii of 2 and 4″ are also shown.

2.3. Database technicalities

Each physical X-ray table is actually a family of X-ray tables (that's why I use an indication like e.g.`jun09*`). There are two *individual band* tables (e.g. `jun09b` and `jun09cd` which contain detail data coming from the original XAMIN FITS catalogue for the separate detections in the B (0.5-2 keV) and CD (2-10 keV) bands), and one *band merged* table (e.g. `jun09`) with the most relevant information. Band merging is described in section 2.3.5 of the `XLSS` paper.

The optical, IR and UV tables are usually single physical tables, unless otherwise stated in 2.2.

The database contains also *correlation tables* which link one X-ray table to a single other table. They have just two columns, with the *sequence pointers* in the two tables (e.g. a correlation table may say that X-ray object 8 is associated with optical object 5968, that X-ray object 2 is associated with optical objects 834 and 835, and that X-ray object 11 is associated with none). The association is precomputed using a predefined criterion (usually a distance within a given radius, but not necessarily). Correlation tables allow to speed up two-table queries.

The database contains also *views* which are a way to see the result of a query on a subset of a table (rows or columns), or on more than one table, as if it were a real table.

In particular there may be views like the *unions* which may concatenate `jun09` and `subaru` tables in a "combo" (see 3.1), and the four *virtual tables* `2XLSS`, `2XLSSB`, `2XLSSCD`, `2XLSSOPT` which are the preferred and recommended way for the user to access the catalogue.

Virtual tables are based on a GCT (which extend the concept of correlation tables to associations of more than two tables).

The database tables pointed from the GCTs used for the present working catalogue (i.e. *member tables*) are those above the dividing line in Table 1.

The tables below the line are accessed only as a result of a two-table query between a virtual table and one of them at a time.

2.4. Astrometry

Astrometric correction offsets were generated afresh using SAS task EPOSCORR in a manner analogous to what described in section 2.3.3 of the `XLSS` paper, but using a different (and homogenous) optical reference catalogue.

The optical reference files were generated taking all objects in `w1t4` within 6″ from the X-ray source position, brighter than $i' = 25$ (or $r' = 25$ for the ABC fields), and having a chance probability (as defined in 3.6) $p < 0.03$. In case of more possible counterparts the one with the smallest probability was taken.

The new astrometric offsets are reported with their numeric values in `http://cosmos.iasf-milano.inaf.it/~lssadmin/Website/LSS/List/.newastroreport.html` Appropriate colour coding in such page shows which XMM fields have been corrected using W1 or ABC optical fields, or a mixture. Fields B68a and B68b (bad) had no CFHTLS counterparts and were corrected using stars in USNO A2.0. Field G12a (bad) had no counterparts at all and was not corrected.

The `astrocorr` (or `Xastrocorr` in `2XLSS`) flag, used at some time to cope with different optical references used in the astrometric correction, is now mostly irrelevant for the newer corrections (with the exception of B68a/b, `astrocorr=5` i.e. USNO and G12a `astrocorr=0` i.e. not corrected). For all other fields which derive from the same W1 T004 (with ABC extension) reference, it is identically `astrocorr=4`.

We have discontinued the production of a plot with the astrometric correction offsets for the individual pointings, and refer to the URL quoted above for the values of the offsets.

Fig. 1, comparable with Fig. 9 of Chiappetti et al. (2005) or Fig. 1 of Chiappetti (2007), gives instead the distances in RA and Dec between the X-ray corrected position and the counterpart position. The best or secondary counterpart is selected based on probability, as described in 5.2. The catalogue (colour-coded in figure) from which to extract the counterpart position (if a given counterpart is present in more than one) is the one giving the smallest distance.

The results in term of positional accuracy are as follows. 88% of the sources have both RA and Dec offsets lower than 4″, and 56% have both within 2″. If one restricts to the best counterparts with good probability, as defined in 4.1, one has more than 96% within 4″, and 78% within 2″ (93% and 65% respectively including those with good or fair probability).

In terms of true distance 83% of the total is within 4″, which makes 90% of the good-or-fair associations (the circles in Fig. 1) and 94% of the good ones (the filled circles in Fig. 1).

There is some evidence from Fig. 1 of a systematics of the deviations between X-ray positions and positions in the various catalogues. The average deviation for the optical and UKIDSS catalogues clusters around a point in the third quadrant (e.g. -0.35″,-0.06″for W1), while the one for SWIRE clusters around a point in the first quadrant (0.77″,0.55″).

3. The procedure

The final procedure leading to the `2XLSS` catalogue can be applied to our (`jun09`) data alone or potentially to the concatenation of `jun09` and `subaru` tables. The procedure was run a few times in test modes before the Escorial consortium meeting, and has been finalized as described below.

3.1. Table concatenation

What described here was used only for a temporary test, not released and deferred to the future.

The first step of the procedure would be to generate a "combo", concatenating `jun09` and `subaru` (technically this applies not just to the band-merged table, but also to individual band tables and dependent correlation tables), defining a *view* named `newcombo` which allows to access, as if it were a single table, the concatenation of: (a) all sources in `jun09`, (b) all sources in TBD full exposures (`field < 2100`) of `subaru`.

By "all sources" I mean all fields (good and bad), and all detection likelihoods (including the "spurious" ones $ML < 15$). Only 438 duplicated detections in the `subaru` 10ks chunks were excluded (they remain available via the `subaru` table, 45% of them are spurious in one band, and only 11% are above 4σ).

The `newcombo*` tables were not released to the users, but were used as starting point for the remainder of the procedure, which is the same for `jun09` or for `newcombo`. However its result were not released and are not reported any further (will be reprocessed later in a different way).

3.2. Overlap removal

The procedure for removal of redundant sources detected in the regions where pointings overlap is similar to the one described in section 2.3.6 of `XLSS` paper. Namely :

- only *merged* sources which are *non-spurious* ($ML > 15$) are considered
- the search radius is 6″
- for each couple of nearby sources, the one with the smallest off-axis angle is preferred *except that if one source is detected in a good field and the other in a bad field, the source in the good field prevails unconditionally*, i.e. the off-axis angle is used only when both fields are good, or both fields are bad
- overlaps between 3 or more fields were manually arbitrated

Note that pointings which were *later* repeated (the first of a sequence of two like B22a and B22b, or the first two of a sequence of three like B04a, B04b and B04c) are by definition always bad, while the last repeat is usually good. However four AO7 fields which conclude such a sequence (B17c, B45b, B47b and B68b) are also bad. Note also that multiple detection of sources can occur between *adjacent* fields which overlap at their edges, but also over the *entire Field of View* of "repeated" fields. In all cases it is possible that a source in an overlapping *region* is detected in a single pointing. Such source will not be subject to overlap removal and will be preserved in the final catalogue. To allow discrimination of such detections deriving from bad fields, one can use the condition `Xbadfield=1` to take them, or `Xbadfield=0` to exclude them.

`2XLSS` includes 117 pointings, of which 30 are flagged as bad fields.

The removal procedure removes 1148 entries, leaving 6282 sources in the GCT for the `2XLSS` catalogue.

Note that in some cases this implies that a source published in the `XLSS` catalogue is now superseded by a different choice. The implication of this on source naming are discussed in 3.3.1 below, while a comparison between `2XLSS` and `XLSS` is presented in 3.4.2.

3.3. The `2XLSS` X-ray catalogues

For analogy with the published `XLSS` catalogue (see Table 11 of the `XLSS` paper) I provide three virtual tables for the X-ray data: a merged catalogue `2XLSS` (analogous of `XLSS`), and two single-band ones `2XLSSB` and `2XLSSCD`, analogous of `XLSSB` and `XLSSCD`.

The naming and meaning of the columns in such catalogues are as far as possible identical to the ones listed in

Tables 4 and 5 of the XLSS paper. A detailed explanation is available on line at http://cosmos.iasf-milano.inaf.it/~lssadmin/Website/LSS/List/2XLSS.html. This is a summary of the differences :

- all *non-raw* sky coordinates refer to the astrometrical correction described in 2.4
- the Xastrocorr flag is set to 4, 5 or 0 as described in 2.4
- the catalogue names are as described in 3.3.1
- there is an additional column Xlsspointer to provide a match with the XLSS catalogue, as explained in 3.3.1 and 3.4.2
- there is an additional column Xbadfield to flag bad fields, as explained in 2.1 and 3.2.

The number of sources in the merged catalogue is 6282 (5377 in 2XLSSB and 2324 in 2XLSSCD).

3.3.1. Source naming

There is an IAU requirement that once a source in a catalogue has been assigned a name (even if this is a "coordinate name"), the name cannot change even if the actual coordinates are improved (modified), unless a completely new catalogue is issued.

Considering that the raw input coordinates in jun09 are different, the astrometric correction is different, *the actual detections by* XAMIN *are different* and the effect of overlap removal may select different sources, it is justified to consider this a new issue of the XMM-LSS catalogue. Therefore:

- the official catalogue name Xcatname is now of the form 2XLSS Jhhmmss.s-ddmmss, where coordinates are based on the corr set
- Pending registration with IAU of the 2XLSS prefix and publication of the catalogue, it is advised to publish an *unofficial*, provisional catalogue name of the form XLSSU Jhhmmss.s-ddmmss. Note that the prefix XLSSU is registered with the IAU
- the single-band catalogue names Bcatname and CDcatname are neither official, nor registered with the IAU. So they use the prefixes 2XLSSB or 2XLSSCD in all cases.
- the reference to the XLSS source replaced by a 2XLSS source is possible using column Xlsspointer which contains the value of Xseq in table XLSS (an explicit lookup in such table is necessary to find its name or other characteristics). There is no explicit way to locate XLSS sources not confirmed in 2XLSS. For details consult section 3.4.2.

	jun09	old	notes
non-spurious	7430	7349	1
spurious	5584	11935	1
without counterpart	2733	8976	2
spurious	2218	8287	
non-spurious	515	689	
jun09 associated with old	10409		3
multiple counterparts	128	101	3
spurious in both	2965		4
demoted	426		5
promoted	704		6
non-spurious in both	6314		good !
both band detection	2228	2372	
same "P/E" class	5589		good !
different class	725		7
change P to E	48		
now both band detection	257		
now single band detection	401		
within 2″	5045		good !

Table 2. Statistics of the comparison between jun09 and earlier releases, using a correlation radius of 6″

1 ratio non-spurious to spurious improved
2 note most non-associated are spurious in both
3 note most associated are single unambiguous
4 spurious in both can be ignored
5 demoted means non-spurious to spurious
6 promoted means spurious to non-spurious
7 include also e.g. P– to PP

As for the XLSS catalogue, there is a limited number of cases where the band merging is ambiguous, and a source in a band happens to be associated with two different objects in the other band. This is discussed at the end of section 2.3.7 of the XLSS paper (column Xlink and eventual addition of an a|b suffix to the catalogue name to disambiguate it). There are only 2 new cases, and none of them requires disambiguation. The 8 old ambiguous cases in XLSS are now unambiguous (7 cases) or no longer present (1 case).

3.4. Comparison with earlier releases

3.4.1. Comparison with physical tables

We compare here the "raw" data in jun09 with the combination of the earlier releases (nov06, jul07 and feb09). By "raw" we mean here spurious and non-spurious sources, and before overlap removal and astrometric correction. See the next section 3.4.2 for a comparison with the published XLSS catalogue.

One noteworthy difference between jun09 and earlier releases is that the ratio non-spurious:spurious (where for pointlike sources spurious means $ML < 15$ – extended sources are non-spurious by construction) definitely improved (this actually began to occur already in jul07, with the transition from the IDL to the Python pipeline,

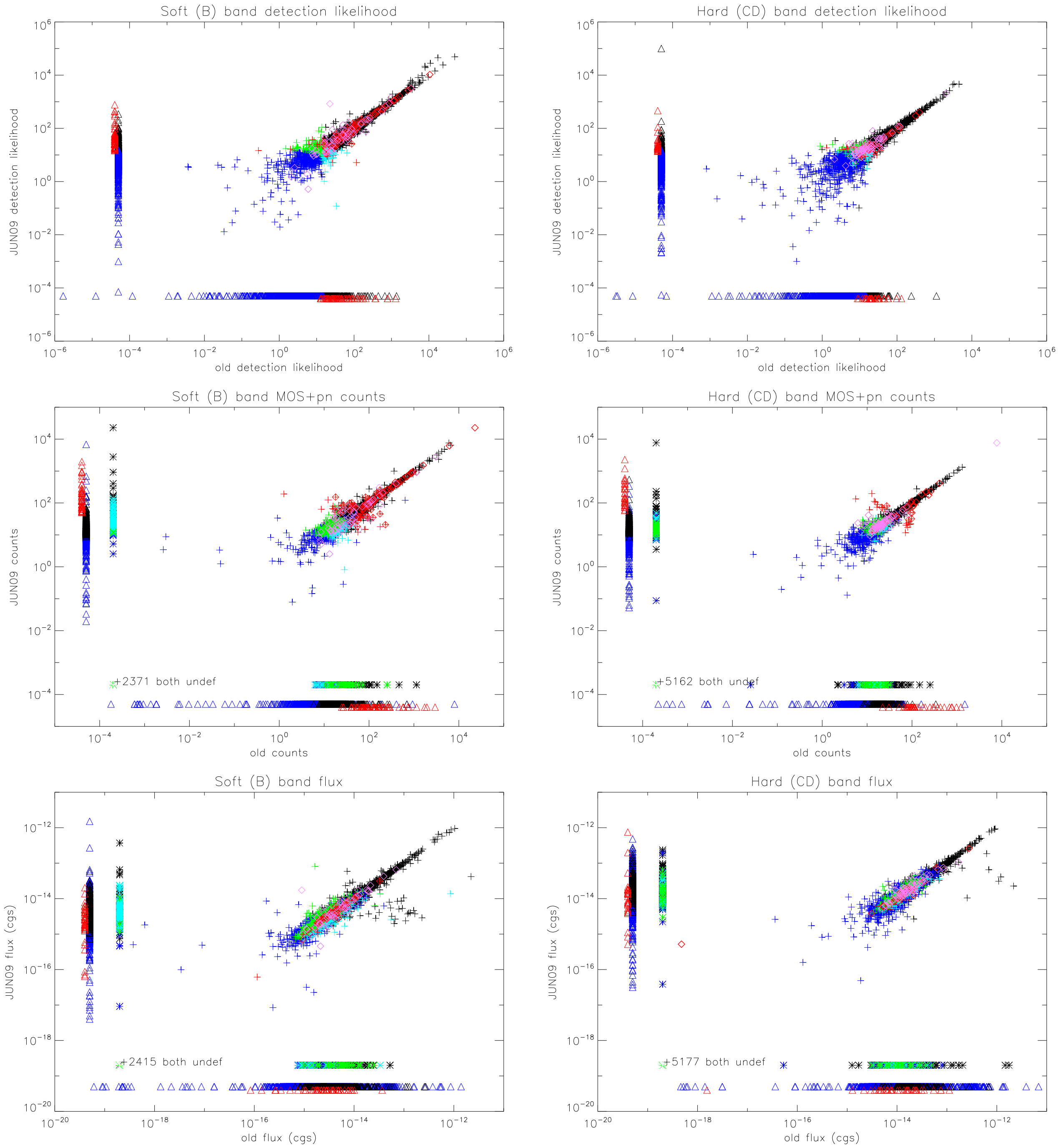

Fig. 2. Comparison of the detection likelihood (top row), of the number of MOS+pn counts (middle row) and of the flux (bottom row) in the soft (left column) and hard (right column) energy bands. between `jun09` and earlier releases. Crosses indicate an *unambiguous association* (single or best) and diamonds indicate the second choices for an ambiguous association for pointlike detections. Asterisks indicate counts or flux are *undefined* in one table, while triangles indicate sources detected only in one release (both are placed at a conventional out-of-range X or Y position). The number of objects with undefined counts or flux in *both* releases in a given band, but nevertheless associated, is indicated near the bottom left corner of each panel. Colour coding for pointlike sources is as follows: blue indicates detections spurious in both `jun09` and old; cyan objects demoted from non-spurious to spurious in `jun09`; green objects promoted to non-spurious; and black objects *confirmed* as non-spurious. Red triangles (at a conventional out-of-range position) indicate extended sources detected only in one release, thick red diamonds are *confirmed extended sources*, while thin red diamonds are old or `jun09` extended sources classified pointlike non-spurious in the other dataset; thin red crosses are old or `jun09` extended sources classified pointlike spurious, and pink diamonds are ambiguous associations to extended objects.

and is mainly due to different handling of numeric precision).

Other differences are possible in the source detection, which affects positioning but also the likelihoods and the classification depending on them. As a result some sources could change from spurious to non-spurious and v.v., or from extended to pointlike and v.v.. In addition there can be new detections in `2XLSS` or lost detections with respect to `XLSS`. Table 2 gives a summary of such occurrences, using an association done field-by-field using the raw position (before astrometric correction). Two detections are associated if they occur in the same pointing and the raw positions are closer than $6''$.

Correlations tables (both direct and reverse, i.e. `jun09` to old and old to `jun09`) created for this purpose are available if one wants to repeat the comparison (it has to be done separately for `nov06`, `jul07` and `feb09`).

Concerning extended sources, 80 of the `XLSS` ones are confirmed in `2XLSS`, of which 74 with the same C1/C2 classification, while 6 have been demoted from C1 to C2. There are 109 `XLSS` extended sources not confirmed at all, and 90 completely new extended sources in `2XLSS`. In addition 44 `XLSS` extended sources are now classified pointlike, and 42 old pointlike sources are now extended in `2XLSS`. The similarity of the numbers makes one think that it is perhaps incorrect to say *not confirmed* and one should say instead *not associated within 6″*. It is possible that a larger correlation radius shall be used for extended objects.

Comparison of detection likelihoods, number of counts and fluxes for sources associated within $6''$ as said above are plotted in Fig. 2.

In conclusion the results in `jun09` look sensible and compatible with earlier ones. The limited number of discrepancies should be due either to marginal cases (objects near the ML=15 threshold going up and down) or to somewhat better resolution of blends of nearby sources. A very limited number of cases deserve further investigation (particulary for extended sources where a match radius of $6''$is possibly not enough).

The `jun09` to old distance peaks rather sharply at 1-1.5″, while the inter-band distance (column `maxdist` in database) peaks at 2.5″ (and is less peaked) for both `jun09` and old, indicating reproducibility in position between pipelines is better than reproducibility of position between detections in the 2 bands with the same pipeline.

3.4.2. Comparison with `XLSS`

Of the 6282 sources, 3431 derive from the 52 GTO, AO1 or AO2 fields (formerly in `nov06`), 1869 from the 46 AO5 fields (formerly `jul07`) and 982 from the 19 AO7 fields (formerly `feb09`). 5871 of the total refer to detections in good fields, and 411 to bad fields (109 of which in the four AO7 bad fields mentioned in 3.2). More specifically, of the 3431 from `nov06`, 3356 come from good fields, and 75 from 7 bad fields; of the 1869 from `jul07`, 1642 come from good fields, and 227 from 19 bad fields; of the 982 from `feb09`, 873 come from good fields, and 109 from 4 bad fields. It shall be noted that only 79 (17%) of the sources in bad fields have a likelihood above 40 (i.e. the 3σ level), and 37 of them are located in the AO7 fields.

However `XLSS` contained 3385 sources taken exclusively from `nov06` good fields. (Sources in bad fields, and sources in AO5 and AO7 fields should not be part of `XLSS` by construction).

The differences are due to the effect of the different positions and characteristics (as likelihood and hence "spuriosity") in the results of the most recent pipeline (i.e. `jun09` , see 3.4.1 above), and to the overlap removal procedure when the overlaps between `XLSS` fields and AO5 or AO7 fields (and the latest astrometrically corrected coordinates) are taken into account.

More specifically 2933 `2XLSS` sources are associated to 2895 distinct `XLSS` objects. These are in `2XLSS` those with non-zero `Xlsspointer`. 2866 of them are detected in `2XLSS` in `nov06` fields: more specifically 2821 in the same field, and 45 in other `XLSS` fields, inclusive of 3 now detected in a bad field (2 in G16a and 1 in B17a). 63 `2XLSS` objects are instead now detected in good fields not included in `XLSS` (G12, but also several B fields adjacent to the `nov06` fields), and 4 in bad fields.

It is worth reminding how the `Xlsspointer` was generated. It was done correlating `jun09` with `nov06` (and later restricting the check on `nov06` sources included in `XLSS`) in three different ways using a radius of $6''$. First comparing the `2XLSS` astrometrically corrected coordinates (`corr`in `jun09`) with the latest set (`cor2`) of coordinates in `nov06` (those used in the `INTERIM` catalogue, see Report V), then comparing them with the published (`corr`) `XLSS` coordinates, and finally comparing uncorrected coordinates. Each of these associations could give a single counterpart, or more than one where the closest is obviously preferred. The `Xlsspointer` was inserted in the GCT entry referring to an `2XLSS` source "filling the slot" in the following priority order

Case		occurrences
single	`corr=cor2`	2443
closest	`corr=cor2`	394
single	`corr=corr`	7
closest	`corr=corr`	none
single	uncorrected	5
closest	uncorrected	none
other	`corr=cor2`	79
other	`corr=corr`	4
other	uncorrected	1

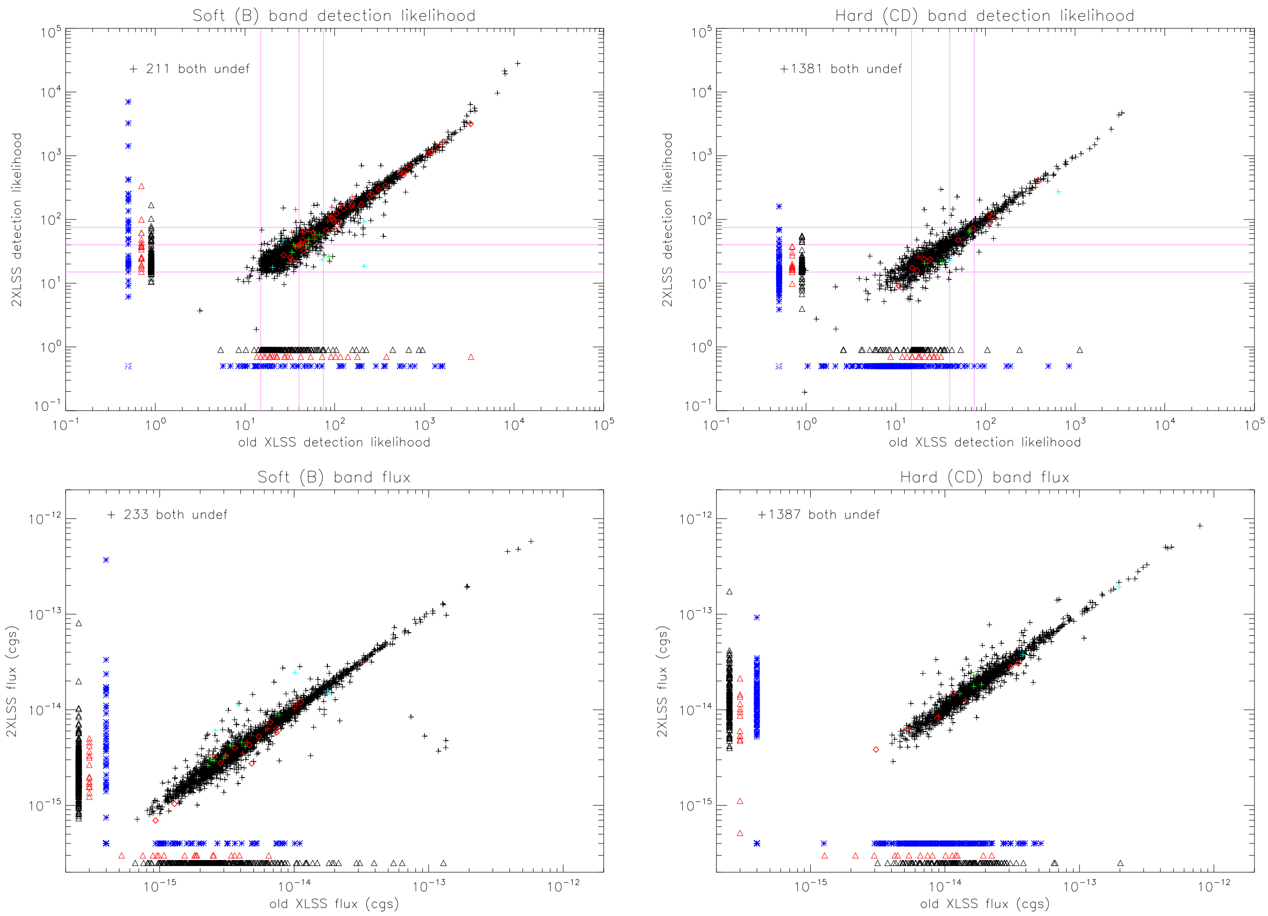

Fig. 3. Comparison of the detection likelihood (top row) and of the flux (bottom row) in the soft (left column) and hard (right column) energy bands. between `2XLSS` and `XLSS` catalogues. Crosses and diamonds indicate pointlike or extended objects associated in the two catalogues as described in text. Blue asterisks indicate likelihood or flux are present but *undefined* in one catalogue, while triangles indicate sources present only in one catalogue (both are placed at a conventional out-of-range X or Y position). The number of objects with undefined values in *both* catalogues in a given band, but nevertheless associated, is indicated near the top left corner of each panel. Colour coding is as follows: black cross for pointlike common sources in `2XLSS` good fields; cyan cross idem for bad fields; green cross for `XLSS` extended object pointlike in `2XLSS`; viceversa for red cross; red diamond for extended sources in both `2XLSS` and `XLSS`. Triangles are black or red for pointlike or extended sources which are either new in `2XLSS` or were present in `XLSS` and have been lost in the new catalogue. In the likelihood plots, the thin pink lines are fiducial marks corresponding to the spurious/non-spurious threshold (15) and to the conventional 3σ (40) and 4σ (75) levels.

The information on the above classification is available in the GCT as the hidden column `glorlss09.reliable`.

The reason why sometimes the closest `nov06` counterpart is not chosen, is that maybe the closest one was not a `XLSS` source because of its overlap removal.

It shall be noted that there are 493 `2XLSS` sources detected in fields potentially part of `XLSS`, as well as 490 `XLSS` sources which aren't associated to an `2XLSS` source. Of the 493, 94% have a somewhat marginal likelihood ($ML < 40$ i.e. below 3σ) Similarly of the 490, 90% are below 3σ. These figures may give an indication where to put a tradeoff between depth and reliability.

Of the 2933 `XLSS` objects confirmed in `2XLSS`, 99% are classified extended or pointlike in both (only 24 are different); 87% have exactly the same "PE" classification, and more precisely 97% have the same classification in the soft band, and 90% in the hard band (detection or non-detection can also be a reason for different classification, e.g. a PP vs a P–).

Changes in likelihood and flux are shown in Fig. 3 and show a good compatibility between the two catalogues. The few outliers are not associated to any peculiar condition (not to large distances nor to bad likelihoods).

The distribution of the distance between `2XLSS` and `XLSS` positions peaks sharply at $1''$. 75% of the sources have distances below $2''$, and 95% below $4''$.

The `2XLSS` and `XLSS` fluxes are rather similar, although it is difficult to quantify this in lack of error bars. For the 2623 non-undefined soft-band fluxes the difference is contained within 10%, 20% and 50% respectively for 69%, 89% and 98% of the cases. For the 1263 non-undefined hard-band fluxes the equivalent percentages are 60%, 86% and 99%.

3.5. The X-ray/optical catalogue

The 2XLSSOPT virtual table provides a synoptic view of the X-ray sources from 2XLSS, together with the nearby optical, IR and UB candidates. It is mimicked on the XLSSOPT table described in the XLSS paper, but provides information on the *latest* (T004) CFHTLS D1 and W1 fields (and on "our" ABC fields), on SWIRE, UKIDSS and GALEX, using the tables described in 2.2.

3.5.1. Optical pre-identification

Unlike the brute force approach used originally for the XMDS (Chiappetti (2006a) aka Report I, i.e. considering all possible combinations of counterparts given by the individual correlation tables with X-ray sources, and then doing a radical cleanup of spurious combinations), I elaborated a variant of the incremental addition used in the latest XMDS versions (Chiappetti (2008a) aka Report IV) described below. This procedure was already tested for the INTERIM catalogue (Chiappetti (2008b) aka Report V), although with CFHTLS, SWIRE and UKIDSS only.

- a preliminary step is to create a GCT and initialize it. The member tables of such GCT are the three X-ray tables (jun09, jun09b, jun09cd) used for 2XLSS a clone of the main X-ray table used to keep track of X-ray duplications, and d1t4, w1t4, swiredr6, ukidssdr5 and galex. The GCT is initialized copying into it the content of the GCT underlying 2XLSS (i.e. the list of all X-ray sources in the band-merged catalogue together with the pointers to the single-band catalogues).
- immediately afterwards a correlation of the main X-ray table with itself within 30″ is used to insert a "clone pointer". This is not used for the optical identification work, but could be useful in the future to study how many X-ray sources are there surrounding another X-ray source, and perhaps to assist in the comparison with XLSS (see 3.4.2). Note that if one X-ray source has more than one nearby objects, additional *placeholder records* are inserted in the GCT (with all other table pointers set to -1). These placeholder records are **not** visibile in the 2XLSSOPT catalogue.
- then one *inserts a pointer* to the first optical table (d1t4) using the existing correlation table, and limiting to the objects within 6″. If the X-ray source has one optical counterpart only, the pointer is *inserted* in the existing primary record (placeholders are ignored). If it has more, the pointer of the closest candidate is inserted, while *additional records are added* copying from the primary one and replacing the pointer. A record in the GCT is also called a *counterpart set.*
- then one *inserts a pointer* to the next table entry into existing counterpart sets when the object in such table is closer to one of the existing counterparts in other optical tables within a predefined radius. E.g. in the case of w1t4 objects are compared with d1t4, while swiredr6 objects are compared first with w1t4, then d1t4, UKIDSS objects are compared with preceding tables (in order W1, D1, SWIRE), and GALEX objects are compared with all other tables (in order W1, D1, SWIRE, UKIDSS). The objects within 6″ from each X-ray source are considered, while a correlation radius of 0.5″ is used when comparing positions of the same origin (i.e. D1 and W1), and 1″ when comparing to other optical, SWIRE or UKIDSS catalogues, and 1.5″ when comparing to GALEX.
- In all cases the pointer is *inserted* in an existing record when there is a single match with the X-ray position and all the positions in the pre-existing catalogues. *Additional records are added* in all other cases (typically an independent counterpart of the X-ray source with no counterpart in previous catalogues, but could also be an ambiguous association of more sources in the current catalogue with a previously defined counterpart set)
- Finally the chance probability of the association of a counterpart with the X-ray source are computed as described in 3.6

3.5.2. The 2XLSSOPT table

2XLSSOPT loosely mimics XLSSOPT as described in Table 10 of the XLSS paper, but provides a number of additional columns (see http://cosmos.iasf-milano.inaf.it/~lssadmin/Website/LSS/List/2XLSS.html or the main database interface for details). It provides essential information on the X-ray sources, the position and $u^*g'r'i'z$ magnitudes of the optical candidates (as for XLSSOPT), the position and fluxes of the SWIRE candidates, the position and magnitudes of the UKIDSS candidates, the position and fluxes of the GALEX candidates, together with all distances from the X-ray position and chance probabilities (see 3.6).

2XLSSOPT contains 19168 counterpart sets, which on average means that an X-ray source has 3 possible optical or IR *not validated* associations within 6″. De facto 45% of the X-ray sources have from none to two possible counterparts, and only 19% more than 4.

2XLSSOPT provides also a flag comparing our optical-SWIRE association with the one provided by IPAC in early 2008 (see 3.5.3).

Probability	m	density $n(brighter\ than\ m)$	a	b	tables
probXO	i'	$n(< i') = 10^{a+bi'}$	-9.32415	0.293833	for d1t4
			-9.23183	0.290519	for w1t4 excluding ABC fields
	r'	$n(< r') = 10^{a+br'}$	-9.18619	0.279706	for w1t4 ABC fields
probXS	F_λ	$n(> F_\lambda) = 10^{a+b*log(F_\lambda)}$			in order swires05 swire
	$\lambda = 3.6\mu m$		-1.68062	-0.944191	for swires05 then swire
	$\lambda = 4.5\mu m$		-1.73693	-0.976644	then in order of λ for swire
	$\lambda = 5.8\mu m$		-2.04933	-0.829700	
	$\lambda = 8.0\mu m$		-1.49944	-1.07201	
	$\lambda = 24\mu m$		0.102480	-1.53410	
probXU	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	
probXG	NUV	$n(< NUV) = 10^{a+bJ}$	-11.0875	0.326965	taken best if both bands present
	FUV	$n(< FUV) = 10^{a+bK}$	-13.9827	0.433838	

Table 4. Parameters used for probability computation

ipacflag	meaning	count
0	records not in the IPAC files, i.e. no SWIRE counterpart	12282
1	SWIRE-W1 association confirmed by IPAC	5393
2	SWIRE-D1 association confirmed by IPAC	48
3	SWIRE with no optical association confirmed by IPAC	1046
11	our SWIRE-W1 association and IPAC's are to different objects	46
12	we associate SWIRE-D1 while IPAC associates SWIRE-W1	10
13	IPAC associates a W1 which we do not associate	225
14	IPAC associates a D1 which we do not associate	7
15	IPAC associates an optical object we ignore since it is farther than 6″ from X-ray	7
21	The W1 object is associated to SWIRE by IPAC, but we prefer another W1	none
22	The W1 object is associated to SWIRE by IPAC, but we prefer a D1	2
23	The W1 object is associated to SWIRE by IPAC, but ignored by us	97
24	The D1 object is associated to SWIRE by IPAC, but ignored by us	5

Table 3. Values, explanation and statistics of the ipacflag The flags 21 to 24 indicate entries with no SWIRE counterpart but whose optical counterpart was associated to a SWIRE object by IPAC. An X-ray source with a counterpart set with such flag has always also a corresponding entry with a flag 11 to 14, which allows to reconstruct the IPAC choice. The reverse is not true, however when an X-ray source has a counterpart set with a flag 11 to 14, usually it has more than one, indicating an ambiguity in the association of *one* SWIRE object with more optical ones.

3.5.3. Comparison with the CFTHLS-SWIRE correlation made at IPAC

Files with the correlation between SWIRE DR6 and CFHTLS T004 release full area (including ABC fields) were made available in early 2008 by IPAC to the XMM-LSS Consortium for usage under the MoU rules (so called swireXt004 datasets). Such files contained about half-million sources and were not ingested in the database. It shall be noted that they are: (a) *SWIRE oriented* : the correlation is done starting from the SWIRE objects; (b) *single source* : only one (the closest ?) CFHTLS counterpart is returned (or none); (c) *W1 preferred* : such counterpart is looked for in W1, and only if none found in D1; (d) presumably they used a larger correlation radius than us.

Our 2XLSSOPT catalogue is instead: X-ray oriented, multiple source and D1 preferred. Therefore if an X-ray source has one or more D1 counterparts, they are associated. If they then correspond to W1 objects these are associated, and finally if they correspond to SWIRE objects, these are associated. So we can come out with different CFHTLS-SWIRE associations than in the IPAC file.

A flag column in 2XLSSOPT, named ipacflag, provides information whether our association and IPAC's matches. It can assume the values listed in Table 3. More than 98% of the X-ray sources with at least one SWIRE counterpart show a match between our and IPAC's choices (the cases with ipacflag between 1 and 3).

3.6. Computing probabilities

I computed the probability of chance coincidence between the X-ray source and its counterparts, based on the X-ray to optical (or IR or UV) distance, the optical, IR or UV intensity, and the density of sources brigther than a given intensity.

I computed four probabilities : *probXO*, *probXS*, *probXU* and *probXG*. They are based on a formula like

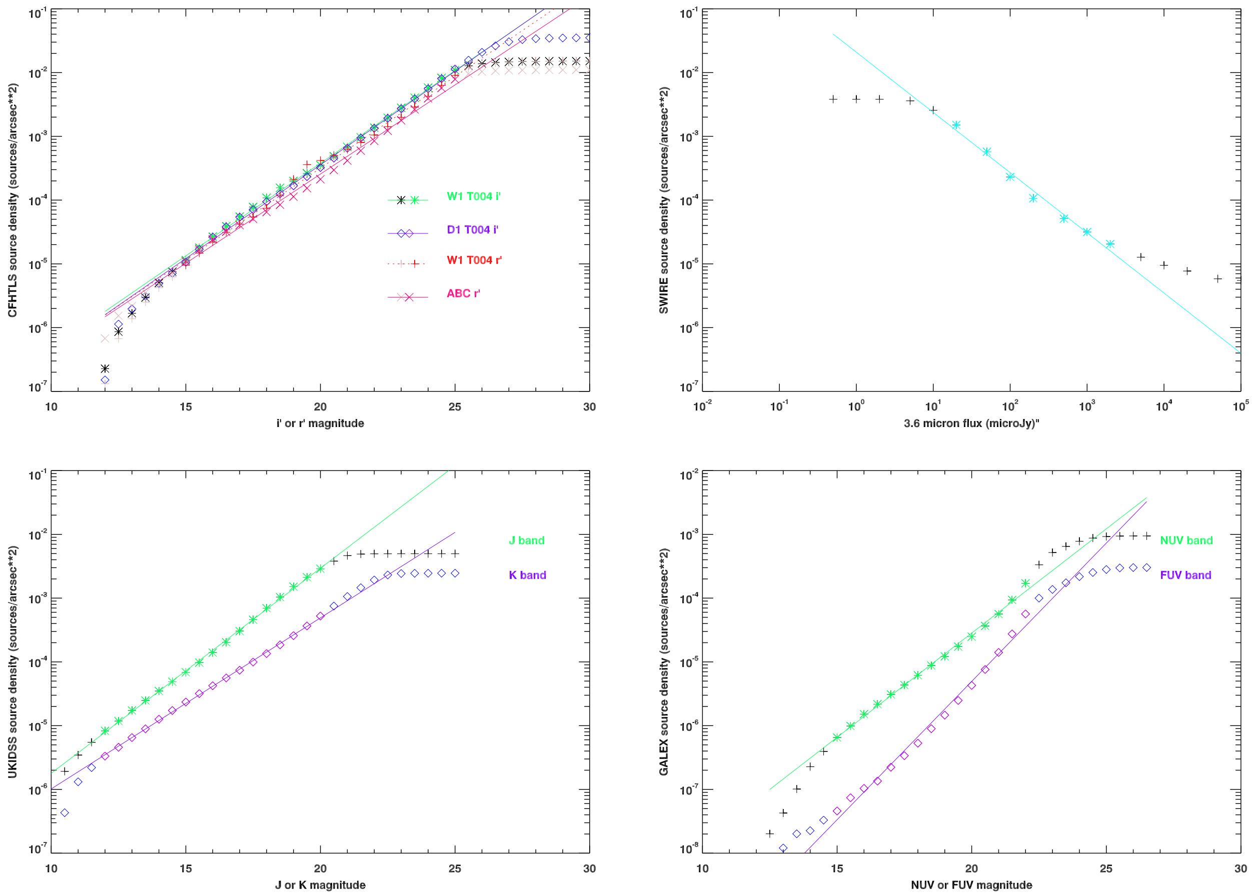

Fig. 4. Source count density for the CFHTLS D1 (asterisks) and W1 (diamonds) fields i' band, as well as for the W1 (crosses) and ABC (X) fields r' band (top left panel) ; for SWIRE DR6 at $3.6\mu m$(aperture 2) fluxes (top right panel); for UKIDSS J (crosses) and K band (diamonds) (bottom left panel); and for GALEX NUV (crosses) and FUV band (diamonds) (bottom right panel). The ranges used to produce the fits shown, whose parameters are given in Table 4 are shown in (lighter) colour. Note the GALEX Y-axis is displaced by one decade.

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

where r is the X-ray to counterpart distance (unlike what done for the XMDS since Chiappetti (2007) and in Report IV the distance has not been *capped* to $2''$), and the density $n(brighter \; than \; m)$ is computed from simple linear fits as reported in Table 4. The same table indicates also the magnitudes or fluxes used to look up the density for the appropriate band.

The coefficients are the same used for XMDS in Report IV (with the exception of those for the r' magnitude in the northern ABC fields, which are used only here) and for the INTERIM catalogue in Report V (with the exception of the GALEX ones which are new).

X-ray to CFHTLS probability, called $probXO$, is computed for sources with a CFHTLS counterpart in order d1t4, then w1t4. In the case of undefined CFHTLS magnitudes, the field limiting magnitude was used (read directly from w1t4, or fixed to $i' = 25$ for D1).

X-ray to SWIRE probability $probXS$ is computed in wavelength order.

X-ray to UKIDSS probability $probXU$, in the case both (J and K) magnitudes are present, is the best (smallest) of the two.

X-ray to GALEX probability $probXG$, in the case both (NUV and FUV) magnitudes are present, is the best (smallest) of the two. Note that such (AB) magnitudes are available in database table galex but are not present as virtual columns in 2XLSS, where only the corresponding fluxes are reported.

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

The density of CFHTLS sources has been derived separately from the *totality* of the sources in the D1 T004 and W1 T004 data (ingested in a temporary table), with a coarse fit to the data (see Fig. 4 top left panel). For the r' magnitudes two fits have been done separately, one for the W1 area proper, and one for the ABC fields alone. Both are shown in Fig. 4 top left panel, however only the fit for the ABC fields is reported in Table 4 and has been used for probability computation.

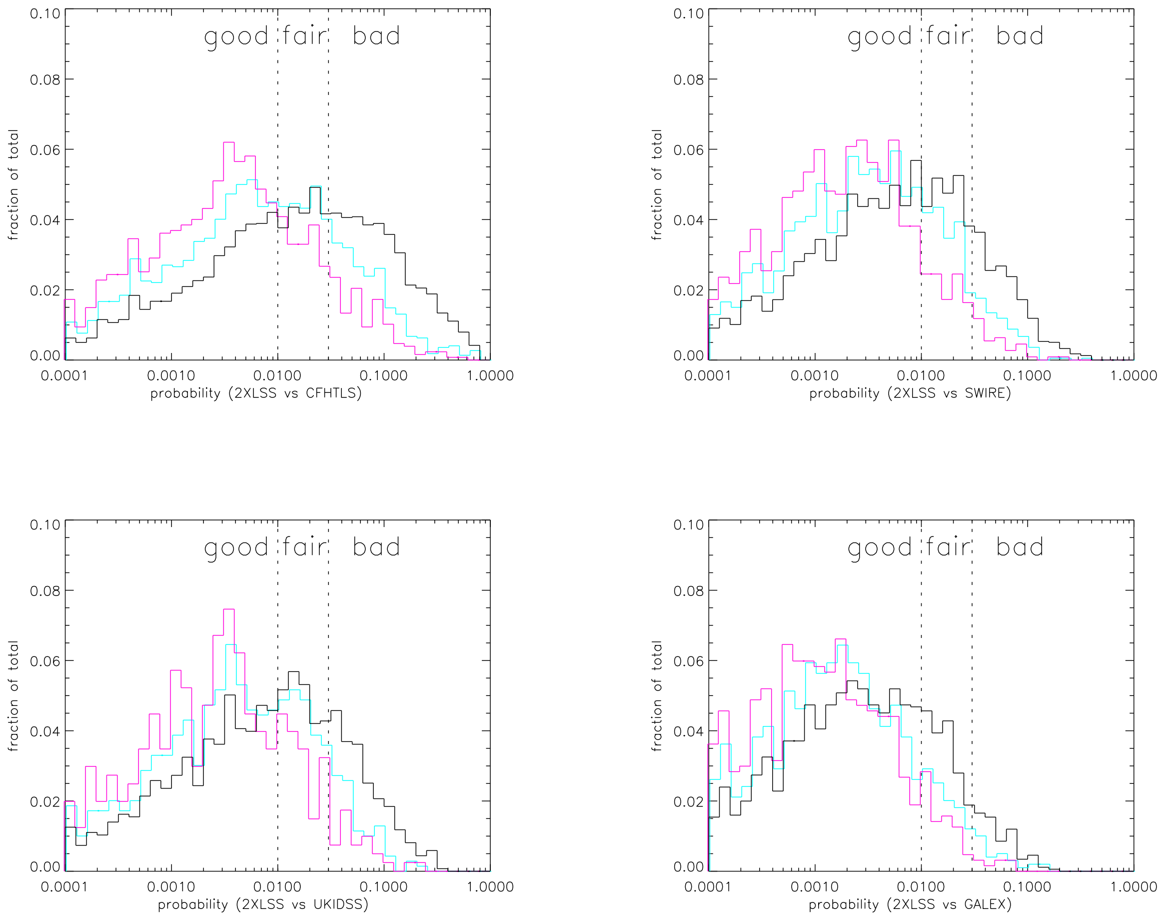

Fig. 5. Histograms of the four *uncapped* probabilities (*probXO*, *probXS*, *probXU* and *probXG*) normalized to the total number of best counterparts with not undefined probability in the total sample (black), with a detection likelihood of at least 40 (3σ) in the best band (cyan), or of at least 75 (4σ, magenta). The dashed fiducial lines identify the loci with good, fair, or bad probability.

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. 4 top right panel for 3.6μm (other bands not shown).

The density of UKIDSS sources was derived using the DR3 release (*sic*!), separately for J and K bands from the *totality* of DXS data, using WSA in count-only mode: see Fig. 4 bottom left panel.

The density of GALEX sources was derived from the GR4 release using MAST CasJobs in the sky region $30° \leq \alpha \leq 40°$ and $-10° \leq \delta \leq 0°$: mode: see Fig. 4 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 IV, 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 proper 12.91 deg^2, for the ABC fields (using r') 2.97 deg^2, for SWIRE 9.70 deg^2 and for UKIDSS DXS 17.53 deg^2. For UKIDSS such calculation was done for the DR3 release. For GALEX an area of 95.87 deg^2 was found for GR4 in the sky region quoted above.

3.7. Data Products

Currently the data products associated to the `2XLSS` tables, are the same associated to the `jun09` tables, i.e. the X-ray field-related data products (images, exposure maps, wavelet images and `ds9` contours) supplied by Saclay.

The `jun09` tables alone are associated also with the original XAMIN FITS catalogues.

`2XLSS` provide additionally as *object-related* data products the SIMBAD and NED pointers associated to the X-ray sources.

Fig. 6. Positions of the X-ray sources with a CFHTLS D1 counterpart. For symbols see 5 in text. The CFHTLS D1 covers the central part of the XMDS (G) fields. In this and in the next 3 figures the EPIC FoV footprint appears in light pink-gray for good fields, and in azure-gray for bad fields.

However it will be possible (as it was been done for demo purposes for a dozen of X-ray sources in `INTERIM`) to load thumbnail images from the CFHTLS, SWIRE and potentially UKIDSS as described in Reports IV and V, namely :

- CFHTLS thumbnails i.e. $40 \times 40''$ i' band images centered on X-ray sources with a W1 T004 counterpart (from the T004 public image archive at CADC). *Note that now T004 images are public, not only T003 as at the time of Reports IV and V,*
- Also the ABC fields are now in the public archive at CADC (although the stacking procedure may be slightly different). In this case the g' band images were chosen because that's the only band present for all 3 fields.
- SWIRE thumbnails i.e. 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.
- UKIDSS thumbnails could in principle be retrieved from WSA, but they use an unusual WCS (`RA---ZPN DEC--ZPN` currently unsupported by the tool described in 4.2).

These products could be used to support a future identification and validation program using a tool like the one proposed in 4.2.

4. Perspectives for future work

I propose that the `2XLSS` catalogue constitute the test bed to define the procedure for the validation of optical identifications in a way similar to what I did for the XMDS (see

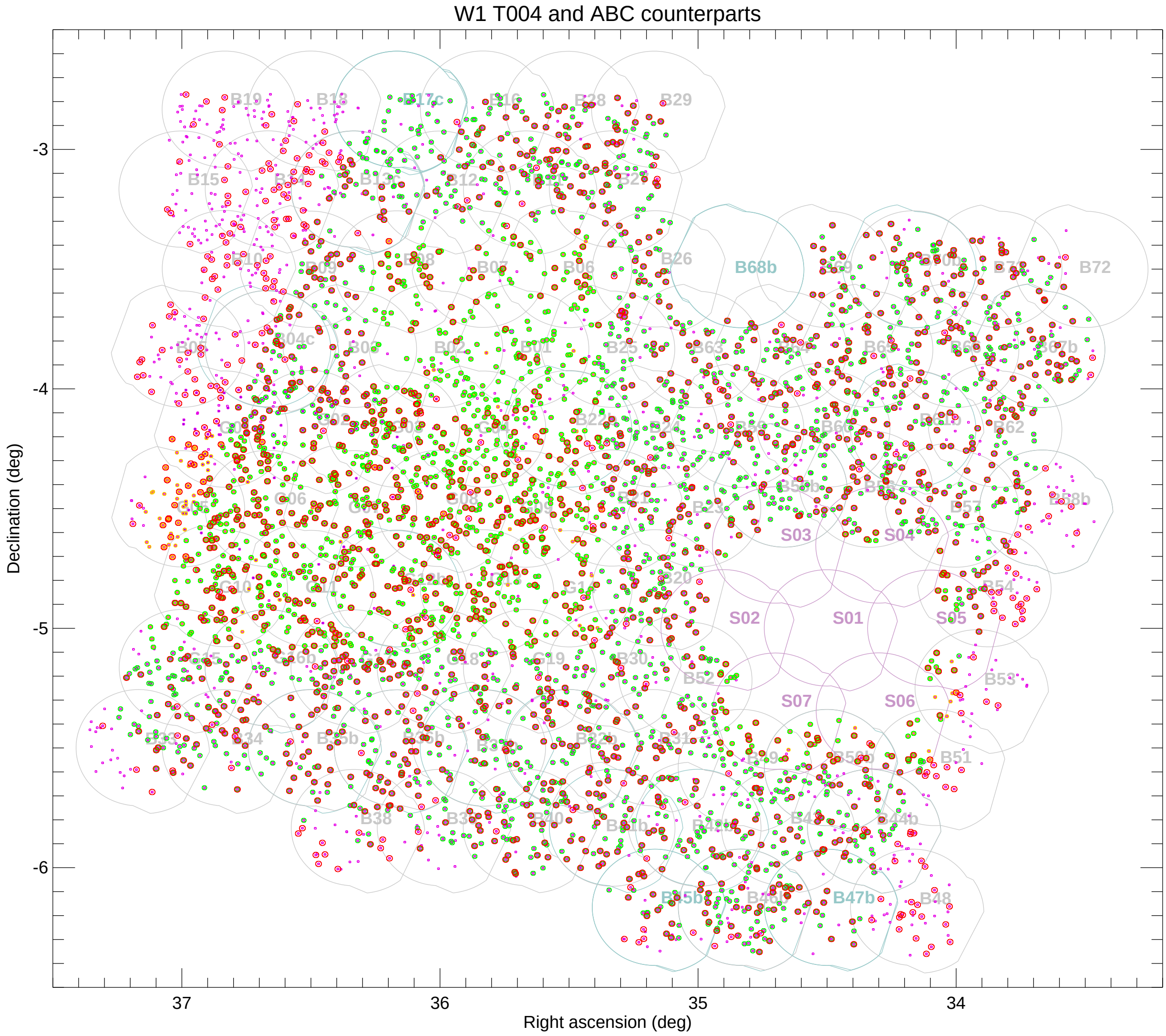

Fig. 7. Positions of the X-ray sources with a CFHTLS W1 counterpart. For symbols see 5 in text. The CFHTLS W1 extended with the northern ABC fields covers almost all of our fields.

Report IV), and which could be done in a collaborative way by a team of volunteers using methods and tools as described below.

Such validated identification could be the starting point for systematic computation of photometric redshift for all sources.

4.1. Ranking on probabilities

It should be possible to select the best or preferred counterpart of an X-ray source ranking the probabilities in 2XLSSOPT (see 3.6) in a way like this (used for XMDS, see Report IV).

- good if $p < 0.01$
- fair if $0.01 < p < 0.03$
- bad if $p > 0.03$ (however undefined if $p = 99$)

An overall rank could be assigned automatically using the above definition and some agreed criterion to combine the results in the different bands and for the different counterpart sets.

A statistics of the probability ranges is presented in 5.2 (see also Fig. 5).

4.2. Identification support tools

While the ranks could be tentatively assigned in an automatic way, there will necessarily be a number of ambiguous cases requiring visual inspection with appropriate tools.

I propose to adapt such tools from those existing for the XMDS. I refer to Report V and to my presentation at the Escorial consortium meeting for a visual impression of the tools and a summary of their capability.

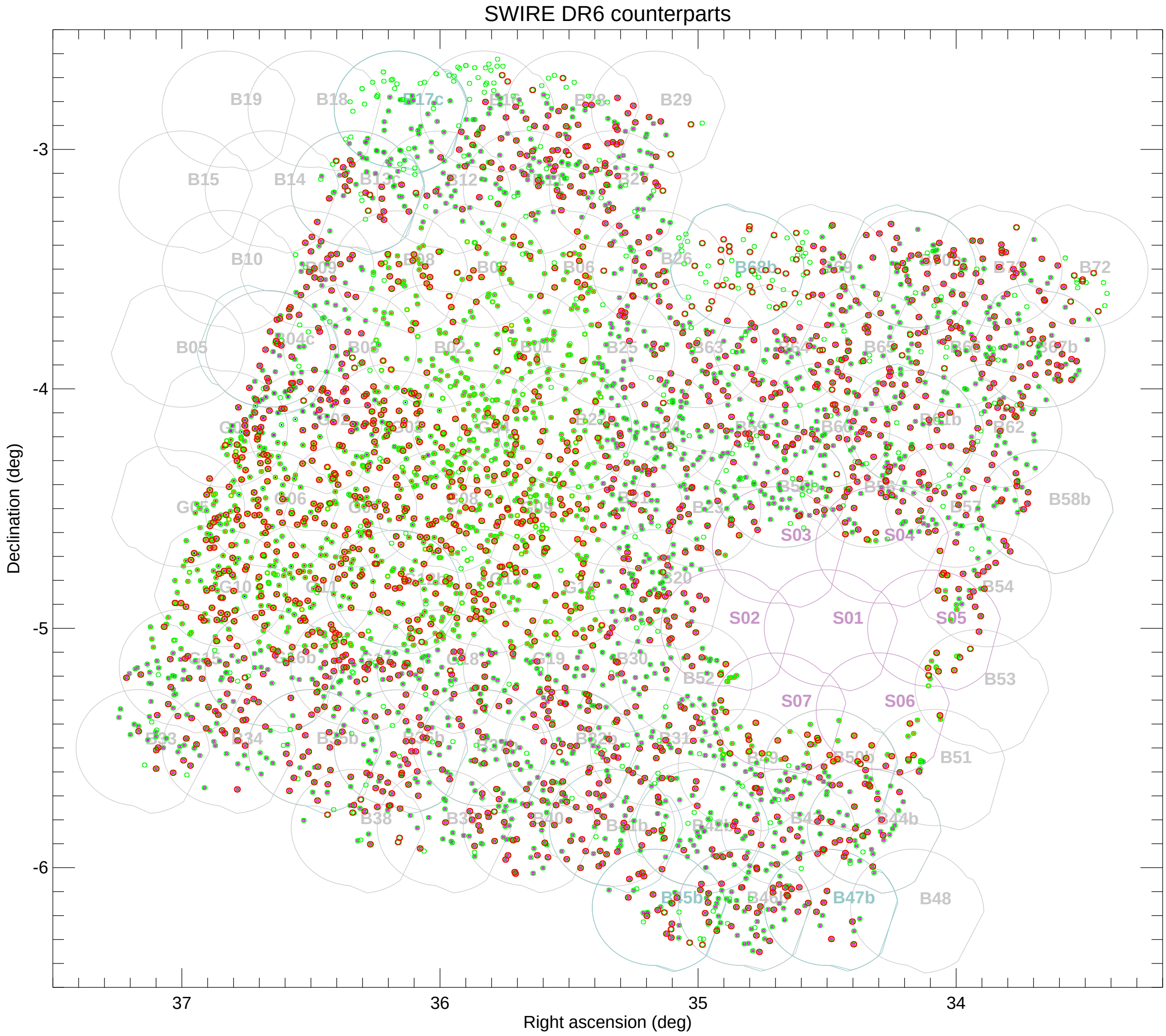

Fig. 8. Positions of the X-ray sources with a SWIRE counterpart. For symbols see 5 in text. SWIRE covers almost all the fields, except the E and W edges.

4.2.1. The validation interface

The *validation interface* will be a tool which will allow to confirm or alter the automatic ranking, and at the same time to edit the GCT underlying the 2XLSSOPT catalogue. Such interface could be used to assign ranks to preferred counterpart sets and to reject unwanted ones.

4.2.2. The graphical interface

A second tool exists already in prototypal form. See URL in Report V. I have adjusted the demo prototype (so called "combo26") to read regions from 2XLSSOPT and to support i' or g' images, as well as supporting gzipped FITS images. Such a tool should allow to display a thumbnail image (and control its look and zoom), onto which one can overlay the regions (corresponding to counterparts in all counterpart sets, or to objects in the external non-member tables) and interact with them.

This will assist in telling *which is which*, and ultimately in confirming which counterpart sets are to be preserved or rejected, in conjunction with their ranks and probabilities.

5. Catalogue statistics

5.1. The X-ray catalogue

The 2XLSS table contains a total of 6282 X-ray sources, of which 1879 are detected in both energy bands, 3576 only in the soft band, and 827 only in the hard band.

Of a total of 193 extended sources (53 C1 and 140 C2), there are 10 extended sources classified C1, and 7 classified C2 detected in both bands (of these only 3 C1 are detected as *extended in both bands*); there are 38 extended sources classified C1, and 89 classified C2 detected only in the soft

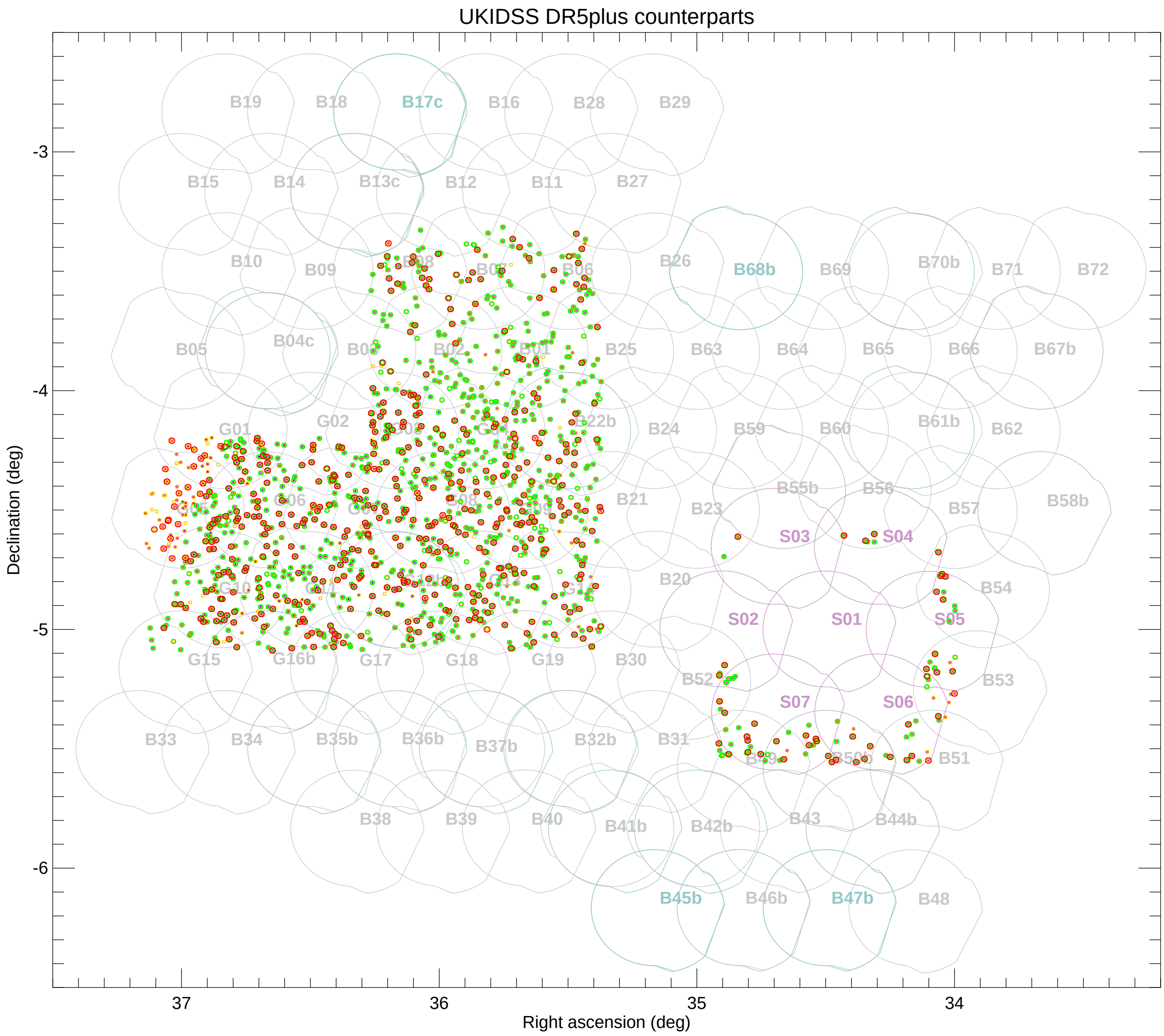

Fig. 9. Positions of the X-ray sources with a UKIDSS counterpart. For symbols see 5 in text. UKIDSS DR5plus still covers so far two disjoint areas (DXS and UDS), one of which covers the SXDS fields.

band; there are 5 extended sources nominally classified C1, and 44 classified C2 detected only in the hard band.

The number of pointlike sources (6089 total) is 1862 (99%) detected in both bands, 3449 (96%) in the soft band and 778 (94%) in the hard band.

Of the pointlike sources, 61% of those with a detection in both bands are detected, in the best band, with a likelihood above 75 (which, according to the calibration with the XMDS reported in Report IV, should correspond to the 4σ level), and 82% above likelihood 40 (3σ level). For 87% of the sources the best band (highest detection likelihood) is the soft band.

For the detections only in the soft band, only 13% are above 4σ, and 34% above 3σ. In the hard band the percentages are 3% above 4σ, and 11% above 3σ.

These results throw some doubt on the significance of detections in a single band.

5.2. The joint X-ray/optical catalogue

2XLSSOPT contains nominally 19168 counterpart sets.

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. I include 5 figures (from Fig. 6 to Fig. 10) which give the sky areas covered by the various surveys used by us overplotted on the footprint of the FoV of our fields. Each figure lists *only* the *best* (see below) 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 from the inner to the outer to :

- a small blue dot indicates a CFHTLS D1 source
- a small magenta circle a CFHTLS W1 (or ABC) source

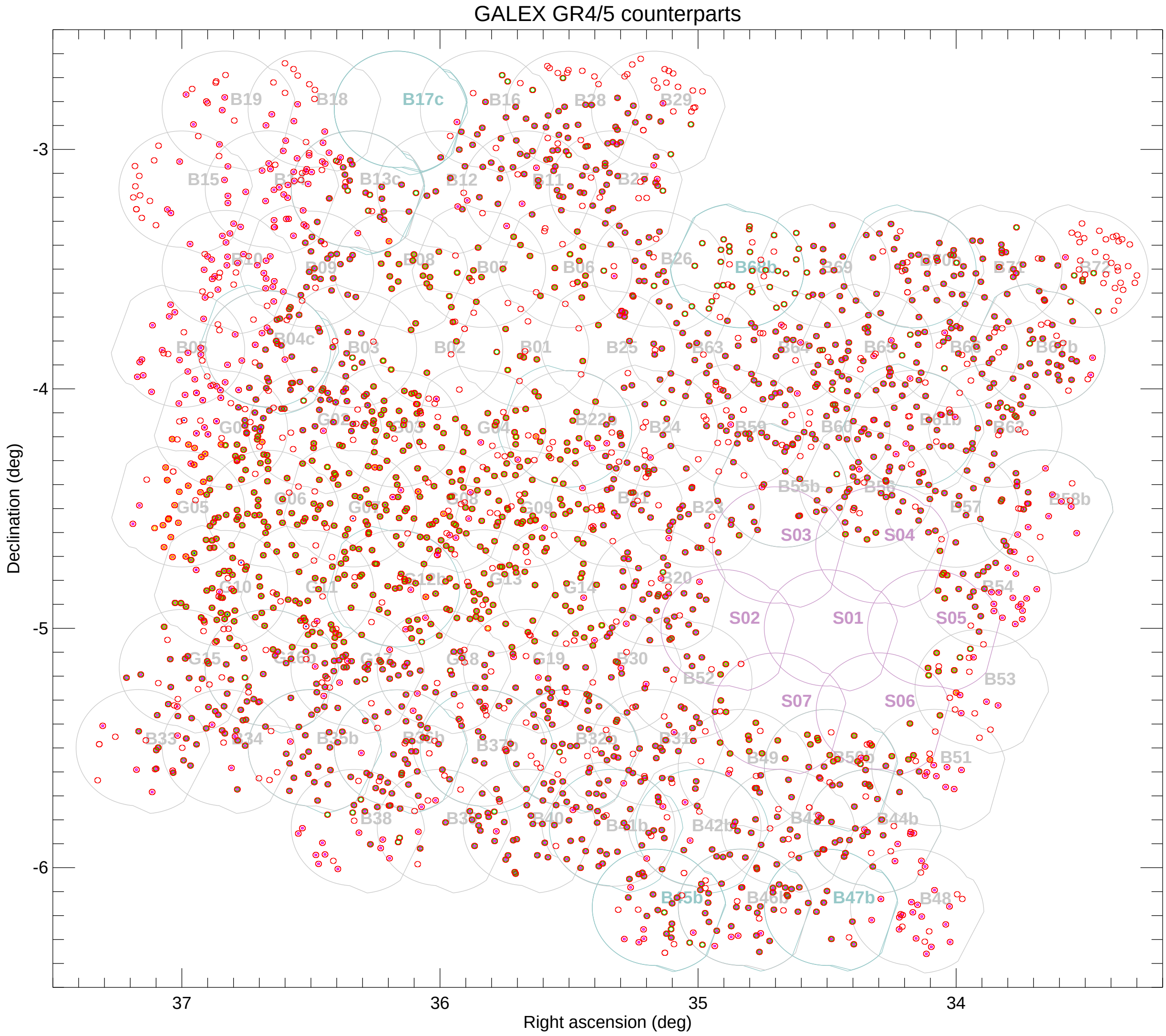

Fig. 10. Positions of the X-ray sources with a GALEX counterpart. For symbols see 5 in text. GALEX GR4/5 covers almost all of our fields.

– a larger orange circle an UKIDSS DR3plus source
– a larger green circle a SWIRE source
– a larger red circle a GALEX source

For each X-ray source we have taken as "*best*" counterpart the one with the smallest chance probability in *any* catalogue.

The outline of the FoV is drawn in light pink for good fields, and in light azure for bad fields. Bad fields are usually not labelled with the field name, unless they are the last repetition of a given pointing (this occurs for B17c, B45b, B47b and B68b). SXDS fields are currently drawn in darker tone, since the relevant sources have not yet been included. Figures for the test including them were available but not shown.

268 X-ray sources correspond to *blank fields* i.e. have no CFHTLS, SWIRE, UKIDSS or GALEX counterpart within 6″.

1097 X-ray sources have a *single counterpart*, while the rest has potentially more counterparts.

I have attempted a rough characterization using the criteria defined in 4.1. No source has a probability classified *good* in all four optical/IR/UV catalogues. The results are summarized in Table 5 which has to be interpreted as follows:

Looking at the row "best and single", 204 (19%) of the single counterparts are detected in 3 out of 4 catalogues with a good probability in all three. 132 (12%) are detected in 2 of the 4 catalogues (and not detected in the other), and with a good probability in all the catalogues where they are detected. 138 (13%) are detected in only one catalogue with a good probability. 66 (6%) are detected in two or three catalogues, and in one of them with a good probability (the other can be fair or bad). Similarly for the cases having all or at least fair probability (48, 55, 66 and 68). Finally 320, despite being the only possible

Probability class / in how many catalogues ? / Counterpart set	n/a	good $p < 0.01$ 3	2	1	some	fair $0.01 < p < 0.03$ 3	2	1	some	bad $p > 0.03$ all
Blank field	268									
Best and single		204	132	138	66	48	55	66	68	320
Best		707	595	616	603	355	337	279	553	872
Secondary		65	114	710	183	68	178	556	500	10512

Table 5. Basic statistics ot the 2XLSSOPT

counterpart, are detected in a number of catalogues from 1 to 3, but always with a bad probability.

Similarly 707 of the best *non-single* counterparts are detected in 3 out of 4 catalogues with a good probability in all three, etc. etc. up to 872 cases which, despite being the best counterpart, are detected always with a bad probability.

Considering the secondary counterparts, 10512 (about 90%) are always bad and could surely be rejected. To be precise, one of such secondaries has all undefined probabilities, because it is a single 160 μm SWIRE source. There are however e.g. 65 cases where the secondary counterpart has a good probability in three catalogues (although however worse than the best counterpart), which probably indicates intrinsically ambiguous cases. Similarly for at least those which have at least one good probability.

Summarizing, 48.7% of the sources have a best counterpart with a good probability, 28.0% a fair one, and 4.3% are blank fields.

One can also view things in a different way, and evaluate how many of the good, fair or bad best counterparts are detected below a given significance (using the Report IV calibration between likelihood and number of σ).

42% of the best good counterparts are detected above 4σ; 17% of the fair ones; 5% of the bad ones and 12% of the blank fields. Or conversely, of the 1656 X-ray sources above 4σ, 77% have a good counterpart, 18% a fair one, 3% a bad one and 2% are unidentified.

Similarly at 3σ 62% of the best good counterparts are detected above such level; 41% of the fair ones; 16% of the bad ones and 24of the blank fields. Or conversely, of the 2862 X-ray sources above 3σ, 66% have a good counterpart, 25% a fair one, 7% a bad one and 2% are unidentified.

Fig.5 gives the distribution of the probabilities in their three ranges. This figure shall be compared with Fig.2 of Report IV, bearing however in mind that Report IV uses *capped probabilities* (which are worse i.e. higher for objects closer than the capping distance of $2''$, which result in the histograms shown here to be less peaked and with a tail at low probabilities). While the two figures are similar, one can note that, in particular for the CFHTLS catalogue, there seems to be a worse tuning with 2XLSS then with the XMDS catalogue. The match is better for the 3σ and 4σ samples, strengthening the idea that 2XLSS extends to lower significances than the XMDS catalogue. The GALEX data are perhaps *overtuned* in the sense there is an excess of good probabilities. This may indicate that the probability computation has to be revised.

Acknowledgements. I acknowledge the work done by N.Clerc (who ran the XAMIN pipeline on all data) and by M.Polletta (who prepared the CFHTLS and SWIRE data used as input to the relevant database tables).

I acknowledge the help of J.J. Kavelaars of the CADC Helpdesk about thumbnail cutout for the ABC fields.

References

Baldi, A., Molendi, S., Comastri, A., et al. 2002, Ap.J, 564, 190
Chiappetti, L., et al. 2005, A&A, 439, 413 (XMDS/VVDS 4σ)
Chiappetti, L., 2006a, XMM-LSS Internal Report N. 1-Mi (Report I)
Chiappetti, L., 2006b, XMM-LSS Internal Report N. 2-Mi (Report II)
Chiappetti, L., 2007, XMM-LSS Internal Report N. 3-Mi (Report III)
Chiappetti, L., 2008a, XMM-LSS Internal Report N. 4-Mi (Report IV)
Chiappetti, L., 2008b, XMM-LSS Internal Report N. 5-Mi (Report V)
Garcet, O., et al. 2007, A&A, 474, 473
Pacaud, F., et al. 2006, MNRAS, 372, 578 (XAMIN paper)
Pacaud, F., et al. 2007, MNRAS, 382, 1289
Pierre, M., et al. 2006, MNRAS, 372, 591
Pierre, M., et al. 2007, MNRAS, 382, 279 (XLSS paper)
Polletta, M., et al. 2007, ApJ, 663, 81
Stalin, C. S., Petitjean, P., Srianand, R., Fox, A. J., Coppolani, F., & Schwope, A. 2009, arXiv:0909.0464
Tajer, M., et al. 2007, A&A, 467, 73
Ueda, Y., et al. 2008, ApJS, 179, 124