XMM-LSS INTERNAL REPORT n°7-Mi

September 9, 2010

Extending 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 possibilities and choice for extension of the 2XLSS catalogue with data from the SXDS fields, giving reference information similar to the one provided in previous internal reports, and additional information about identifications, and possible other problems. The catalogue tables will 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 with the addition of the SXDS fields, and associated optical, IR and UV information.

Key words: LSS

1. Introduction

This document supplements the report issued in October 2009 (Chiappetti , 2009), describing the 2XLSS catalogue, hereafter Report VI, and presents possible extensions to 2XLSS (choosing one variant) which adds the X-ray data of the SXDS (Ueda et al., 2008) to our X-ray data from GTO to AO7, all reprocessed in an uniform way at Saclay with the latest (Py3.2) XAMIN pipeline. In addition it describes work done on 2XLSS and its extensions in *ranking* identifications in the optical and other wavebands.

For details on previous work on XMM-LSS catalogues we refer to the introduction of Report VI and references therein. We recall here only an essential list of references.

Published catalogues are represented by the XMDS/VVDS 4σ catalogue (Chiappetti et al., 2005) and the XMM-LSS catalogue version 1 hereafter XLSS (Pierre et al., 2007), supplemented by 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).

Internal reports instead document yet unpublished working catalogues for the complete XMDS (Chiappetti , 2006a,b, 2007, 2008a), produced with the Milan pipeline (Baldi et al., 2002), and for the XMM-LSS using the XAMIN Saclay pipeline (Pacaud et al., 2006), like a poorly used INTERIM version in Chiappetti (2008b), and the 2XLSS catalogue released for internal use in Oct 2009 in Chiappetti (2009) aka Report VI.

I present here the extended catalogue called 2XLSSe (pre-released in July and released to the consortium in August 2010) and a variant 2XLSSf used in earlier tests. Both versions combine the jun09 data used already for 2XLSS with the SXDS data reanalysed with our pipeline in subaru. The difference between 2XLSSe and 2XLSSf concerns only field S01. The current choice has been to use the 40 ks exposure of S01 (2XLSSe) instead of the longer full exposure (2XLSSf) .

The 2XLSSe catalogue includes 7083 entries (would be 7082 in 2XLSSf) , to be compared with 6282 in 2XLSS.

The 2XLSSe catalogue has been just released for internal use within the consortium in order not to hinder other people's work, but is likely not to be the final version. There are in fact two *caveats* about its usage.

- One concerns the choice of SXDS data, while confronting the fact that SXDS exposures are typically longer (our own good fields are in range 10 ks to 27 ks, while S02-07 are in the range 34-47 ks, S03 is 19 ks and S01 reaches 80 ks). Such inhomogeneity might give problems in the statistics and in the selection function.
- The other one concerns the 6″ radius used for band merging and overlap removal since XLSS, which might be too small. There is the possibility that some soft- and hard-band detections whose distance is between 6 and 10″, currently considered separate sources, are instead missed mergers.

The latter issue is discussed in some detail in section 6.1 of the present report. Possible solutions could be using a larger radius (10″, 8″?) or living with the side-effects of the current 6″ radius.

The solution to the former issue, if 2XLSSe is not confirmed, could be to use only one 10 ks chunk for each of the SXDS fields in a revised catalogue.

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

Table	Update	Content	History	(5)	(6)
jun09*	Jul 09	X-ray sources from latest Saclay pipeline, band merged within 6″	unchanged since Jul 09	6″	a
subaru*	**Apr 10**	X-ray sources from the Saclay pipeline, band merged within 6″	SXDS observations analysed by us, S01-07 full ingested in Aug 08; **new!!** 10ks chunks replaced in Mar 10 and S01 40ks in Apr 10	6″	a
d1t4	Feb 09	CFHTLS D1 field release T004	in use since Jan 08	6″	
w1t4	**May 10**	CFHTLS W1 fields release T004 both supplied by Saclay	In use since Jan 08; **added objects**	6″	
swiredr6	**May 10**	SWIRE DR6 supplied by IPAC	in use since Jan 08; **added objects**	6″	
ukidssdr5	**May 10**	UKIDSS DR5plus public release	in use since Aug 09; **added objects**	6″	
galex	**May 10**	GALEX GR4/5 public release	in use since Nov 08 for XMDS; **added objects**	6″	
simbad	**Aug 10**	SIMBAD sources	present since 2003 and regularly updated	20′	b
ned	**Aug 10**	NED sources	present since 2003 and regularly updated	20′	b
usno	Mar 09	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 extended 2XLSSe 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. (2010)

used to create the 2XLSSe catalogue is described step-by-step in the various subsections of section 3, with particular regard to the combination with subaru data (3.1), 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 3.6 discusses perspectives for identification work, in particular the pre-ranking (3.6.1) based on the probabilities, and a more refined ranking (3.6.2), using possible aid tools (4). 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 2XLSSe X-ray extended catalogue have been the latest release (jun09 and subaru) of X-ray tables (see 2.1). For the 2XLSSOPT* virtual tables (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 or criterion), above which the catalogue *virtual tables* are based.

2.1. X-ray data

The starting point for the 2XLSS catalogue proper were 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), ingested from FITS catalogues supplied by Saclay and produced by the Xamin (version Py3.2) reanalysis of all our fields (GTO, AO1 and AO2, AO5 and AO7), as described in Report VI.

The additional starting point for the *extended* 2XLSSe catalogues are the equivalent subaru family of tables with SXDS observations. The 7 SXDS full exposures (S01 "*full*" and S02 to S07) were compliant with the Py3.2 release of the pipeline already from the INTERIM catalogue. "Chunk" 10ks exposures were instead reanalysed recently with Py3.2 and fully supersede previous (incomplete) data. In addition a 40ks chunk of S01 has been recently reanalysed with Py3.2. Its usage instead of the longer full exposure is recommended for uniformity with the remainder of XLSS and to avoid source confusion problems. In fact one of the variants discussed later at the end of 3.1 for the *extended* 2XLSS catalogue is just the usage of S01 full vs S01 40ks.

The ingestion and in particular band merging was done as described in section 2.3.5 of the XLSS paper, and it is outside of the scope of the present report, but for 6-10″ issue discussed in 6.1. 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 (and subaru new data) be 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 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 changed for 2XLSS , and since Report VI is consistent with the one used in Saclay (field names are only relevant for filenames like those of data products, see 3.7). The convention for SXDS fields is *partially new* for this report.

- 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. The second pointings 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 *full* is 2001, and S02 is 2002)
- there is now a complete sequence of 10 ks chunks for each SXDS exposure. They vary in number from 7 for S01 to 1 for S03). The Saclay abbreviated name for chunk *m* of field S0*n* is S0*n*_c*m*, and the equivalent number is 2000+100*m+n (e.g. S05_c3 is 2305)
- the 40 ks chunk of S01 (Saclay short name S01_40) is numbered 2901

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 and subaru tables. The 2XLSS* catalogues use instead Xbadfield=1.

For subaru no pointings are actually bad, and the badfield=1 is used to *flag the chunks or exposures which are not used for the catalogues.* The default variant is to use the S01 40ks chunk plus full S02-207, so 2901 is flagged good and 2001 is flagged bad.

For jun09 usually bad fields were re-observed once or twice, and the most recent pointing is good. However *B17c, B45b, B47b, B68b* (717, 745, 747, 768) are nominally bad, but *should be used in the catalogues to avoid holes*, since they are the latest (and best though bad) pointings.

2.2. Optical, IR and other data

This section is virtually identical to what included in Report VI, but is included for self-completeness. The tables were however updated with the sources in the surrounding of latest X-ray ingestions (i.e. in the subaru fields).

For CFHTLS release T004, we use (since 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 where duplicated sources in adjacent files had been natively removed (with benefit of inventory). 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″). 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.

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. 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″). 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 ("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.

New: on request by O.Melnyk also the total (aka Hall) magnitudes were later loaded for all UKIDSS sources. Note that Hall magnitudes are not present in any of the 3 JHK bands for UDS. Conversely UDS has JHK aperture 3 magnitudes, while DXS has no magnitude in the H band. We also checked release DR7 which became recently available, but it does not provide additional coverage in sky nor in bands, and does not include UDS yet, so it won't be useful for us.

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 repeated, always using a radius of 10″, from the latest release called GR4/GR5 and the list of jun09 and subaru 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 and NED, this was unnecessary for USNO) have been *recently updated* with pointers to objects in the surrounding of jun09 and subaru X-ray sources, and can be accessed in correlation with the 2XLSS* catalogues, although not members of them. *Note that SIMBAD and NED provide indirectly also the correlation with some of our own catalogues or published subsets (XMDS,* XLSS *, XLSSC) and to other catalogues which we have also in the database (VVDS, VIRMOS 1.4GHz).*

New: A correlation with the 2XLSS* catalogues is also provided for some of the tables referring to published papers, namely the recent table stalin09 (for which however the author consulted our public XLSS catalogue), ueda08 (Ueda et al., 2008) and in a limited way garcet07 (Garcet et al., 2007). A correlation is provided also for the table vimos with VIMOS spectra. O.Melnyk has prepared a list of all sources with spectroscopic redshifts, which, when finalized, could be made available in the database.

2.3. Database technicalities

This section is virtually identical to what included in Report VI, but is included for self-completeness. Each physical X-ray table is actually a family of X-ray tables (that's why I use an indication like e.g.jun09* or subaru*). 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 are views like the *unions* which concatenate jun09 and subaru tables in a "combo" (see 3.1), and views like the groups of four *virtual tables* 2XLSSe, 2XLSSBe, 2XLSSCDe, 2XLSSOPTe 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.

Correlation tables between a virtual table and one of the non-Xray physical tables are technically emulated as "correlation views".

2.4. Astrometry

Astrometric correction offsets were generated afresh at the time of Report VI 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. We did not *compute* new astrometric corrections for the recently ingested subaru chunks, but merely *applied* those already generated using the full exposures. *We report here most of the text included in Report VI with due changes.*

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.

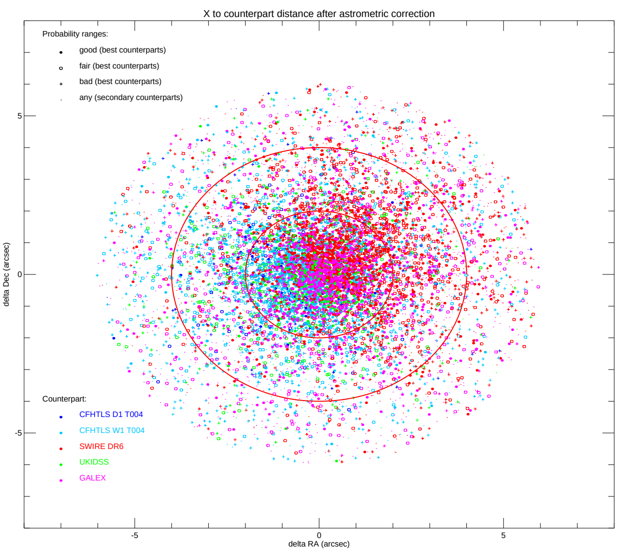

Fig. 1. Distances in RA and Dec between the X-ray corrected position and the counterpart position. Different symbols indicate the identification quality. A circle is plotted when the counterpart is the best one, and the chance probability is good or fair (filled in case of good probability). A cross 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.

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 3.6.1, one has more than 96% within 4″, and 79% within 2″ (93% and 66% 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. -0.41″,-0.05″for W1), while the one for SWIRE clusters around a point in the first quadrant (0.79″,0.53″).

3. The procedure

The final procedure leading to the `2XLSS*` catalogues was applied to our (`jun09`) data alone as described in Report VI, and to a concatenation of `jun09` and `subaru` tables for the extended catalogues described here.

3.1. Table concatenation

The first step of the procedure (*specific for this extension*) was to generate a "combo", concatenating `jun09` and `subaru` (technically this applied not just to the band-merged table, but also to individual band tables and dependent correlation tables), defining a *view* named `sdscombo` which allows to access, as if it were a single table, the concatenation of: (a) all sources in `jun09`, (b) all sources in full exposures (`field < 2100`) or in the S01 40ks exposure (`field=2901`) of `subaru`.

By "all sources" I mean all fields (good and bad), and all detection likelihoods (including the "spurious" ones $ML < 15$). Detections in the (new) `subaru` 10ks chunks are excluded (they remain available via the `subaru` table). *This choice was based on the assumption that longer exposures and going deeper to fainter fluxes is desirable. This assumption could be reconsidered in the future in favour of uniformity of exposures.*

The `sdscombo*` tables are presently not released to the users, but are used as starting point for the remainder of the procedure, which is the same for `jun09` (`2XLSS` described in Report VI) or for `sdscombo` (extended catalogues described here).

From a single `sdscombo` tables we generated *initially and provisionally* two variants of the extended catalogues, toggling the `badfield` flag for fields 2001 and 2901.

- The *default extended* catalogue (or 2XLSSe) considers 2901 (S01 40ks) as good and 2001 (S01 full exposure) as bad. *This is the catalogue currently released to the consortium.*
- The *full exposure* (or *full* for short or 2XLSSf) catalogue considers 2001 good and 2901 bad. *Material related to this provisional test appears in blue colour throughout this report.*

The variants were generated in the form of two GCTs glorsdscombo and glorsdscombofull, which were then subjected to the overlap removal procedure described in the next subsection.

3.2. Overlap removal

This section is adapted and amended from Report VI. 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″(see discussion in 6.1)
- 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. *For a conservative usage exclude bad fields, but maybe include the four AO7 "last repeats" mentioned above e.g. ANDing condition* find_in_set(Xfield,'717,745,747,768').

The same fictitious overlap occurs of course for fields 2001 and 2901, and its effects represent the only source of differences between the default and full extended catalogues. *Note that* 2XLSSe *includes 64 sources (flagged* Xbadfield=1 *in field 2001, i.e. detected only in the full exposure S01, and which one should prudentially exclude. Conversely* 2XLSSf *includes 8 sources from field 2901, i.e. detected only in the 40ks exposure of S01 . The rest of the sources in S01 (ascribed to 2901 in* 2XLSSe*) are detected in both cases.*

2XLSS proper includes 117 pointings, of which 30 are flagged as bad fields. The extension adds 7 SXDS pointings, of which one (S01) occurs in two incarnations (40ks and full), alternately toggled good or bad.

The removal procedure removes 1148 entries, leaving 6282 sources in the GCT for the 2XLSS catalogue. It leaves 7083 and removes 1419 for the default 2XLSSe, and respectively 7082 and 1420 for the 2XLSSf .

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 the variants of 2XLSS and XLSS is presented in 3.4.2. For 2XLSS sources detected in jun09 and replaced with an equivalent source detected in subaru in 2XLSSe, or differences between 2XLSSe and 2XLSSf see instead 3.4.3.

3.3. The 2XLSS *extended X-ray catalogues*

This section is adapted and amended from Report VI. For analogy with the published XLSS catalogue (see Table 11 of the XLSS paper) and 2XLSS (see Report VI) I provide three virtual tables for the X-ray data: a merged catalogue 2XLSSe (analogous of XLSS), and two single-band ones 2XLSSBe and 2XLSSCDe, analogous of XLSSB and XLSSCD. No virtual tables beyond the GCT have been generated for 2XLSSf. The X-ray/optical tables 2XLSSOPTe and 2XLSSOPTf are described in 3.5.

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 (versus 6282 for 2XLSS) 7083 or 7082 for 2XLSSe and 2XLSSf respectively (6041 in 2XLSSBe and 2702 in 2XLSSCDe).

Technically the 2XLSS* extended tables are realized as *union of joins*. This way they do not require the "combo" tables (although they were necessary to build the underlying GCTs), and a speed improvement of two orders of magnitude for the database queries can be achieved.

3.3.1. Source naming

This section is adapted and amended from Report VI. 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 2XLSS a new issue of the XMM-LSS catalogue. Therefore:

- the official catalogue name Xcatname is now generated in 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. This avoids problems with the few sources in 2XLSS proper which are not preserved in 2XLSSe or 2XLSSf or any difference between the two variants if one wish to exploit them.
- 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 in Report VI.

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 is only one new couple in 2XLSS (present also in 2XLSSe and 2XLSSf) and an additional couple in 2XLSSe only (in field S01 40ks). 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

The comparison of the "raw" data in jun09 with the combination of the earlier releases (nov06, jul07 and feb09), where by "raw" we mean here spurious and non-spurious sources, and before overlap removal and astrometric correction, *was reported in section 3.4.1 of Report VI and is not repeated here*.

See the next sections 3.4.2 and 3.4.3 for a comparison with the published XLSS catalogue. and between 2XLSS, 2XLSSe and 2XLSSf.

For subaru the recent reprocessing with the Py3.2 pipeline for full and chunk exposures *fully supersedes* any previous data. A comparison between the 10 ks chunks and full exposures has been done in April 2010 (material including figures discussed in correspondence with Saclay is not reported here but is available on request). If one compares the detection likelihood distribution of a full exposure with a typical 10ks chunk, one notes that there is a definite excess of spurious ($ML < 15$) sources in the full exposure, and some excess in the range $ML = 15 - 50$. Above $ML = 50$ the distributions are similar.

A comparison between the S01 full and 40 ks exposures (more relevant to compare 2XLSSe vs 2XLSSf) was reported in a message to Saclay on 9 Apr 2010 (material including figures also available on request). Some key points are reported here: Full field 2001 in subaru includes 285 detections, the 40 ks field 2901 only 222. Of these 167 are in common (within customary 6″) with just 1 ambiguity. 128 of the 167 have an identical classification. 127 are non-spurious (and therefore will be considered for the 2XLSSe catalogue) and 21 are spurious in both, the remaining 19 are spurious in 2901 and "promoted" in the full exposure (no full exposure detection is "demoted"). The detections (120 in 2001, 56 in 2901) not confirmed in the other pointing are usually detections in a single energy band, and either spurious or poor ($ML < 20$). For common objects, the detection likelihood is about a factor 2 lower in 2901 than in 2001, but the fluxes match at better than 16% (of course there is no indication on the flux error).

3.4.2. Comparison with XLSS

The comparison between XLSS and 2XLSS proper was reported in section 3.4.2 of Report VI (inclusive of a detailed comparison of flux and detection likelihood) and *is not repeated here*, since there is essentially no overlap between the XLSS and SXDS areas (with SXDS being what is added in the extension to 2XLSS).

The reliability flag tabulated in Report VI is available for 2XLSSe (with *identical* content) in the GCT as the hidden column glorsdscombo.reliable or glorsdscombofull.reliable.

3.4.3. Comparison of 2XLSS with the extensions

The generation of the two variants 2XLSSe and 2XLSSf has been described in detail in a report mailed to Saclay on 27 April 2010 (available on request with attached plots).

Before overlap removal one starts from 8502 (non-spurious) entries of which 7430 in jun09, 938 in full subaru exposures (of which 191 in S01), and 134 in the S01 40 ks exposure. The procedure considers overlaps between our and the SXDS fields, between different SXDS fields and also between the two S01 cases 2001 and 2901.

For what concerns a comparison with 2XLSS , both of our variants cause the *loss of the same 51 sources detected in B fields, in favour of those detected in SXDS fields*. Note that the 51 sources are the same in both variants since S01 (the only difference between the two variants) is at the centre of the other 6 SXDS fields and therefore has no overlap with our own (B) fields. The 51 sources are overlaps between B fields and S02-07 fields, for which *now* the SXDS source is preferred. Moreover the extended catalogue sees the *addition of new SXDS sources*, namely:

- for 2XLSSe: 654 in fields S02-S07, 134 in field 2901, and 64 in field 2001 (the latter *present but* flagged bad)
- for 2XLSSf: 653 in fields S02-S07, 190 in field 2001 and 8 in field 2901 (in this case flagged bad)

Of the 127 detections common to 2901 and 2001, one is disposed differently. While 2XLSSf prefers an S01 source, 2XLSSe prefers one in S06 (source 19027) to one in field 2901.

Note that 97 of the common sources have the same classification in 2XLSSe and 2XLSSf . Instead 23 sources detected as pointlike in both energy bands in 2XLSSf are detected in a single band in 2XLSSe (1 even as extended). Only 4 objects are "upgraded" from one to two-band detection going to 2XLSSe, while 3 sources change from extended to pointlike or vv.

The sources detected only in field 2001 (full) (preserved with bad field flagging in 2XLSSe) have low likelihoods but not necessarily so marginal. The few field 2901 ones (preserved with bad field flagging in 2XLSSf) are marginal or spurious. Sources detected only in one of 2001 and 2901 look only marginally fainter than those detected in both.

For common sources the detection likelihood in field 2901 is about 2 times lower than in field 2001, while flux matches within less than 6%.

Considering instead the 51 2XLSS sources replaced by SXDS sources in both extended catalogues (and whose list is available on request), 23 have the same classification, 25 are upgraded from single band to double band detection, 2 are detected in a single different band, and 1 is degraded from double to single band detection. The fluxes (when both are defined) are compatible though with some scatter, curiously larger in the soft band (see Fig. 2).

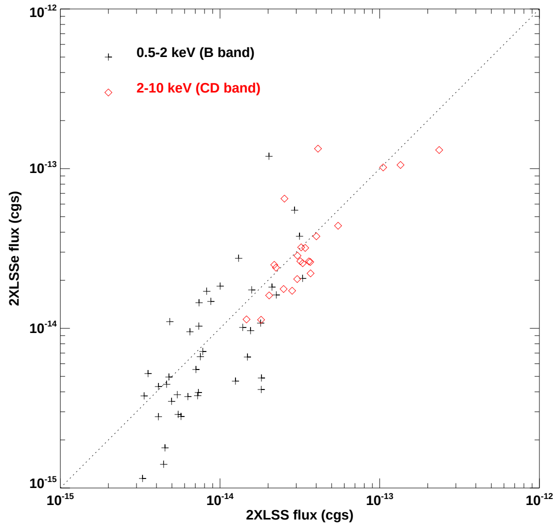

Fig. 2. Comparison between the fluxes in the 2XLSS and 2XLSSe catalogues for the 2XLSS sources replaced by a SXDS source in 2XLSSe. Black crosses are for the soft band. Red diamonds for the hard band. The dashed line is the locus of equal fluxes.

3.5. The X-ray/optical catalogue

The 2XLSSOPT* virtual tables provide a synoptic view of the X-ray sources from 2XLSS, together with the nearby optical, IR and UV candidates. The three variants (2XLSSOPT [see Report VI], 2XLSSOPTe, and 2XLSSOPTf not released for official use) are mimicked on the XLSSOPT table described in the XLSS paper, but provide 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

This section is virtually identical to what included in Report VI, but is included for self-completeness. 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*nd is the same described in Report VI.

- 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) or the three "combo" tables (sdscombo, sdscombob, sdscombocd) used respectively for 2XLSS or 2XLSSe, 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 the corresponding X-ray-only catalogue 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* catalogues.
– 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* tables

2XLSSOPT, 2XLSSOPTe and 2XLSSOPTf loosely mimic 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). They provide 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), the identification rank (see 3.6.2), and pointers to eventual comments.

2XLSSOPT, 2XLSSOPTe and 2XLSSOPTf contains respectively 19156, 21909 or 21902 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. *Note that the above figures are slightly inhomogeneous, in the sense that* 2XLSSOPTe *has been subject to some manual editing (as described at the end of 3.6.2) which was not applied to the other tables (since it is simpler and preferred to mantain only one).*

The 2XLSSOPT* tables provide also a flag comparing our optical-SWIRE association with the one provided by IPAC in early 2008. Such flag is described and analyzed in section 3.5.3 of Report VI for 2XLSSOPT. The results are very similar for the extended tables and are omitted from the present report.

3.6. Computing probabilities

This section is identical to the one in Report VI since the same computation was used for 2XLSS*extended*.

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

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

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 2. The same table indicates also the magnitudes or fluxes used to look up the density for the appropriate band.

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 2. Parameters used for probability computation

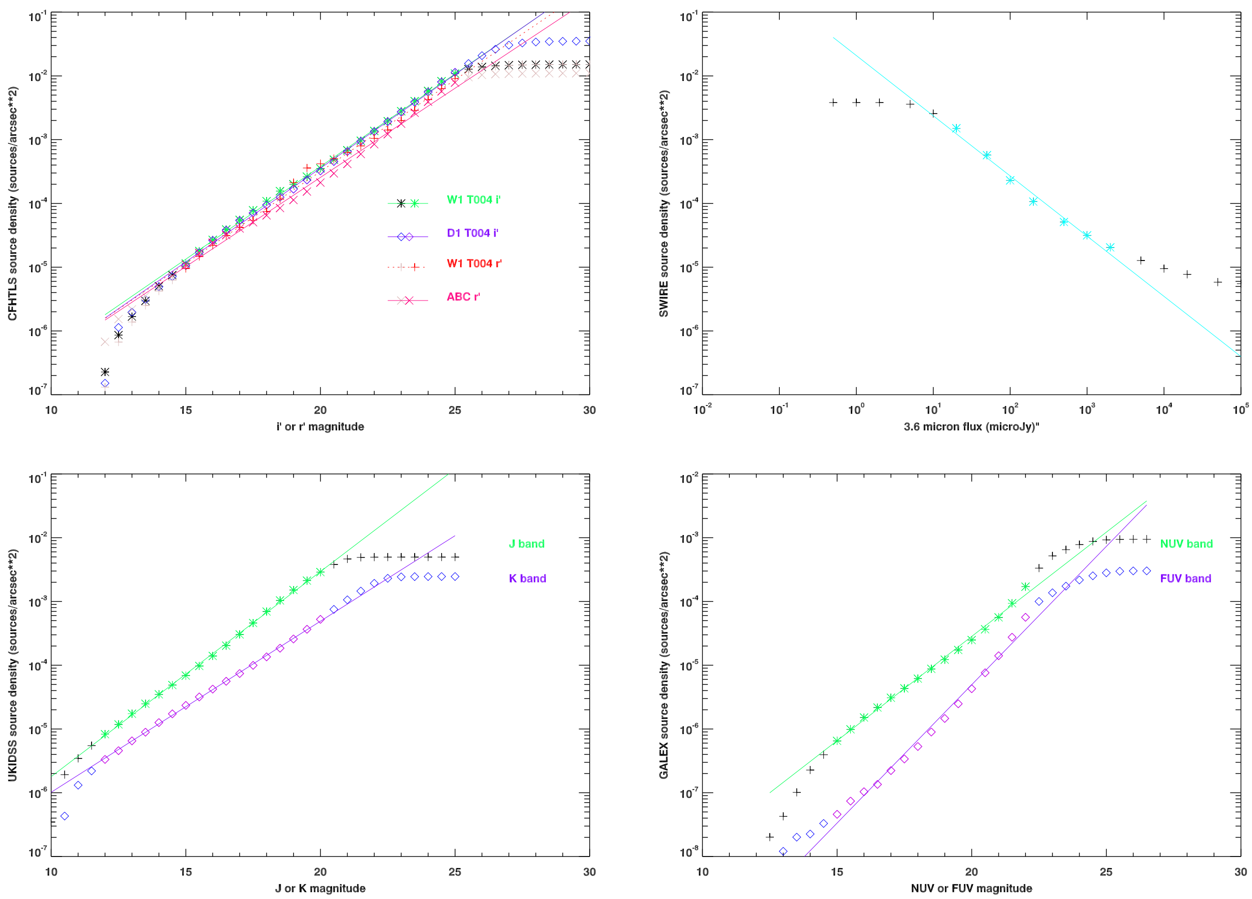

Fig. 3. Source count density for the CFHTLS D1 (asterisks) and W1 (diamonds) fields i' band, as well as for the W1 (crosses) and ABC (X) fields r' band (top left panel) ; for SWIRE DR6 at $3.6\mu m$(aperture 2) fluxes (top right panel); for UKIDSS J (crosses) and K band (diamonds) (bottom left panel); 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 2 are shown in (lighter) colour. Note the GALEX Y-axis is displaced by one decade.

The coefficients are the same used in Report VI.

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.

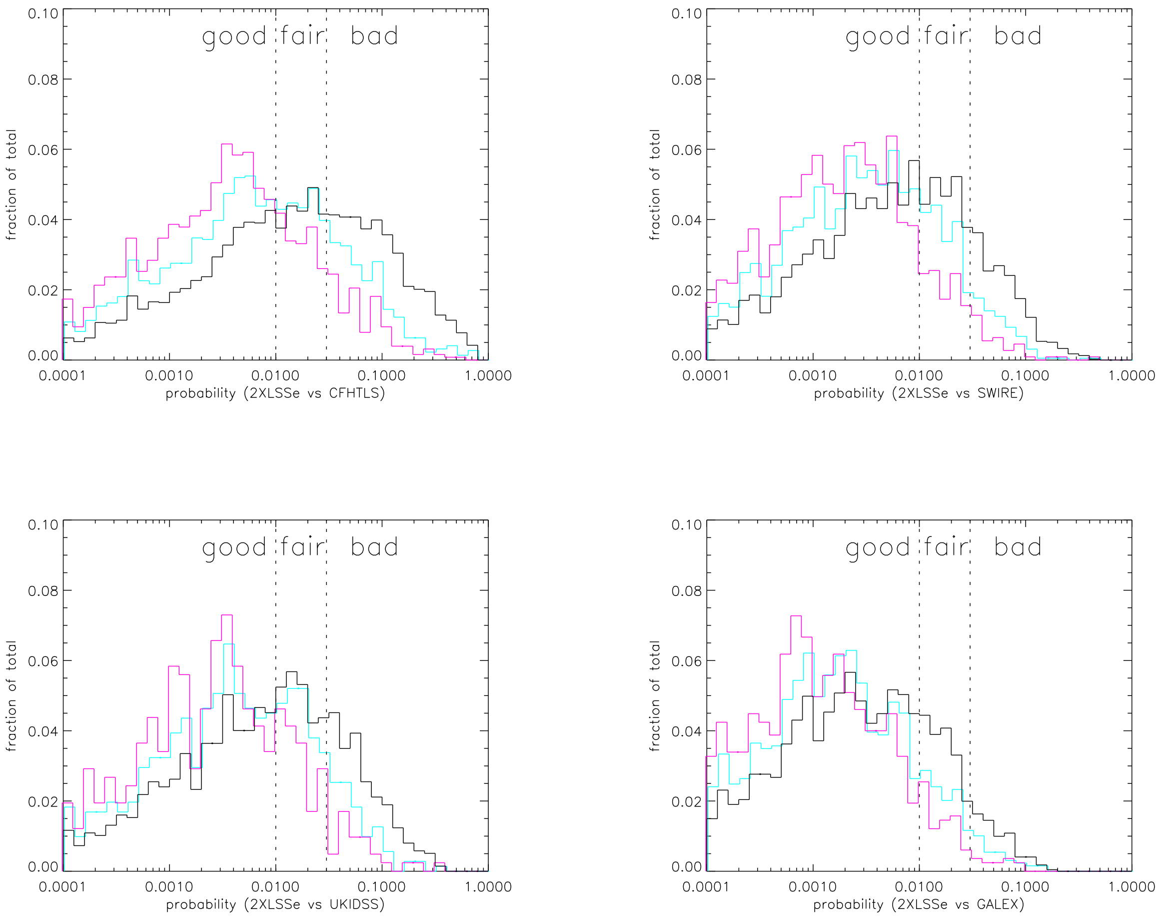

Fig. 4. Histograms of the four *uncapped* probabilities (*probXO*, *probXS*, *probXU* and *probXG*) normalized to the total number of best counterparts with not undefined probability in the total sample (black), with a detection likelihood of at least 40 (3σ) in the best band (cyan), or of at least 75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

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. 3 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. 3 top left panel, however only the fit for the ABC fields is reported in Table 2 and has been used for probability computation.

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. 3 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. 3 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. 3 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.

3.6.1. 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. 4).

3.6.2. Identification ranking

At this point one has a list of potential candidates with the associated probabilities. The ranking procedure is a multi-step euristic procedure similar to what I already used for XMDS. Most of it is automatic (i.e.objective ?) and repeatable, and is described in detail by a sequence of scripting commands. Here I give a summarized textual description.

The procedure uses a *rank* and an intermediate flag or "*autorank*", which are updated several times during the procedure. The idea is that *the rank assumes value -1 for rejected counterpart sets, and ranks 0,1 or 2 in somewhat decreasing order of preference*. The autorank also has in principle values 0, 1, 2 or 3 in decreasing order of preference, but there are other "technical" values possible. The final rank (described below) is available in catalogue column `Xrank`.

Here are the steps of the procedure:

– placeholder records are assigned $autorank = 7$ $rank = -1$ in order to be ignored in all subsequent steps
– sources which have a single counterpart set with just an X-ray entry and no entry in any other catalogue are preliminarily considered *blank fields* and assigned $autorank = 4$ $rank = 1$
– sources with an *unique* counterpart are provisionally assigned rank 0 and autorank 0 or 1 if they have at one good probability, and none or one bad, or rank 1 and autorank 2 or 3 if there is at least one fair probability, or none
– for sources with more counterparts, if one is *brightest and closest* it is assigned initially autorank 1, 2 or 3 according to the best probability being good, fair or bad. In this first step all 4 probabilities are considered. Then one refines the choice considering only optical or SWIRE intensity for brightness, while distance is considered with the closest non-X-ray counterpart in any table.
– autorank is provisionally incremented by 10 for the brightest optical, 20 for the brightest SWIRE and 100 for the closest. This results in a composite flag where for instance 132 means closest, brightest in optical and SWIRE and fair; or 21 may mean brightest in SWIRE only, not closest, good; or 119 may mean closest, brightest in optical only, and not the best in probabilities; etc.
– an interim rank 90-93 is assigned stepwise (i.e. for those not yet ranked so far) like this
  – 90 if brightest and closest is good
  – 91 if brightest and closest is fair
  – 92 if brightest and closest is not best and at least fair
  – 93 for remaining brightest and closest
– at this stage what remains with all bad (defined) probabilities is *irrevocably rejected* ($rank = -1$ $autorank - 3$)
– then one continues disposing
  – 90 remaining best and good
  – 91 remaining best and fair, or not best and at least one good probability
  – 92 remaining not bad
  – 93 remaining bad
– the interim rank is decremented by 90 and transferred to the autorank, the rank is reset to undefined (except for the unique)
– At this point a new stage begins, which consider *single* the X-ray sources which have just one *non rejected* counterpart set.
– singles with autorank 0 or 1 are assigned rank 0 (good or fair)
– singles with autorank 2 or 3 are assigned rank 1 (lower quality)
– for the multiple with best rank, autorank 0 and all other counterpart sets worse (autorank 2 or 3) the rank is assigned to 0
– for the multiple with best rank, autorank 2 and all other counterpart sets worse, the rank is assigned to 1
– the other counterpart sets of those sources are assigned rank 2
– the remaining best by rank are assigned rank 1
– the other counterpart sets of those sources are assigned rank 2

- the remaining best by probability are assigned rank 1
- the other counterpart sets of those sources are assigned rank 2
- the rank 2 with all undefined (the single 160 μm) source) or all bad probabilities are rejected ($rank = -1$)
- if at this stage an X-ray souece has more than one counterpart set ranked 0 or 1, those worse by probability are reset to rank 2
- the rank 1 counterpart sets with all probabilities good are reset to rank 0
- At this point one performs an *ambiguity analysis* where one defines
  - *unambiguous* the case where an X-ray source has one rank 0 or 1 counterpart set, and all other (if any) are rejected
  - *ambiguous* the case where an X-ray source has one rank 0 or 1 counterpart set, and at least one rank 2 (plus zero or more rejected)
- the analysis consists in computing a *score* based on the so-called 3 *Brera rules*
  - adding score 1 if either the optical or SWIRE probability is good, and score 0.5 if fair
  - adding score 1 for the presence of a SWIRE counterpart
  - adding score 1 if the best probability of the rank 0-1 counterpart set is at least 10 times better than all other counterpart sets
- A flag is set for the "selected" rank 0-1 counterpart to
  - "plus" if its score is greater or equal than any other counterpart set (this should be the normal behaviour)
  - "minus" if it is less
  - "solitary" if there is just one counterpart set (what was called above "physically single" and is obviously a subset of the "unambiguous")
- The "minus" cases are always inspected. They are very few (29 of 5014 unambiguous, 89 of 2069 ambiguous). Also the high score not selected cases are inspected visually using the tool in 4.2.
- For the ambiguous cases one considers the score difference between the selected counterpart sets and the other
  - The rank is assigned to 0 if the score difference is greater or equal to 1. This means the rank 0 counterpart set is definitely better than the other. *The ambiguity is just nominal*.
  - The rank is assigned to 1 if the score difference is less than 1 which means a *real ambiguity* among the various counterpart sets.
- In order to distinguish the ambiguous from the unambiguous, the autoranks of the latter are incremented by 10 (so they assume values 10-13 instead of 0-3)
- Visual inspection was done also for "intrinsical ambiguities", which are the cases in which two counterpart sets have an identical best probability, i.e. when two counterpart sets share the same object in one of the non-X-ray tables. E.g. if two optical objects are associated to the same GALEX counterpart.

This procedure was originally applied to 2XLSSOPT (better, to the underlying GCT), and later to 2XLSSOPTe (and 2XLSSOPTf) . As a result of the visual inspection already for 2XLSSOPT one discovered *anomalies and artifacts*, e.g. residual tiling artifacts (overlaps between different CFHTLS pointings not removed), or saturated optical sources split in two entries, or missed associations between an object observed in D1 and the same observed in W1, etc. These anomalies were cured by *manual editing* of the GCT using the tool described in 4.1, e.g. rejecting one redundant counterpart set (flagging $rank = -1$) or merging two entries (transferring one counterpart into the other entry and eventually recomputing the probabilities, and physically deleting one of the entries). When the edit was not trivial, a note was logged in the *comments* (see next section 3.7).

Edits applied already at 2XLSSOPT level were propagated to the extended catalogues whenever applicable.

About a dozen of other similar edits, resulting from correspondence with O.Melnyk in July 2010, were instead applied only to 2XLSSOPTe.

3.7. Data Products

This section is now applicable to both 2XLSS *and* 2XLSSe *unless otherwise stated.*

Currently the *X-ray data products* associated to the 2XLSS* and 2XLSSe* X-ray 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. Of course the individual band catalogues (e.g. 2XLSSBe) are associated only to data products in the relevant energy band.

The jun09 (and subaru) tables alone are associated also with the original XAMIN FITS catalogues.

For SXDS fields and sources (and subaru tables) the data products are located in the same jun09 directories.

The above "per field" data products are accessible as usual from the Query Results screen ticking on the link RETRIEVE ALL OBJECTS RELATED FILES, provided in the Tables tab one has ticked the tick box SHOW MEMBER TABLES (AND DATA PRODUCTS!) located at the very bottom of the screen.

All 2XLSS* tables provide additionally as invididual *object-related* data products (i.e. from the VIEW DATA link of the Query Results screen) the *SIMBAD and NED pointers* associated to the X-ray sources. To access them, one should in the Tables tab select the virtual table *and* one of the SIMBAD or NED tables, activate the "natural" correlation table ticking on it at top right, tick the tick box at very bottom . . . and duly follow the tree of data products

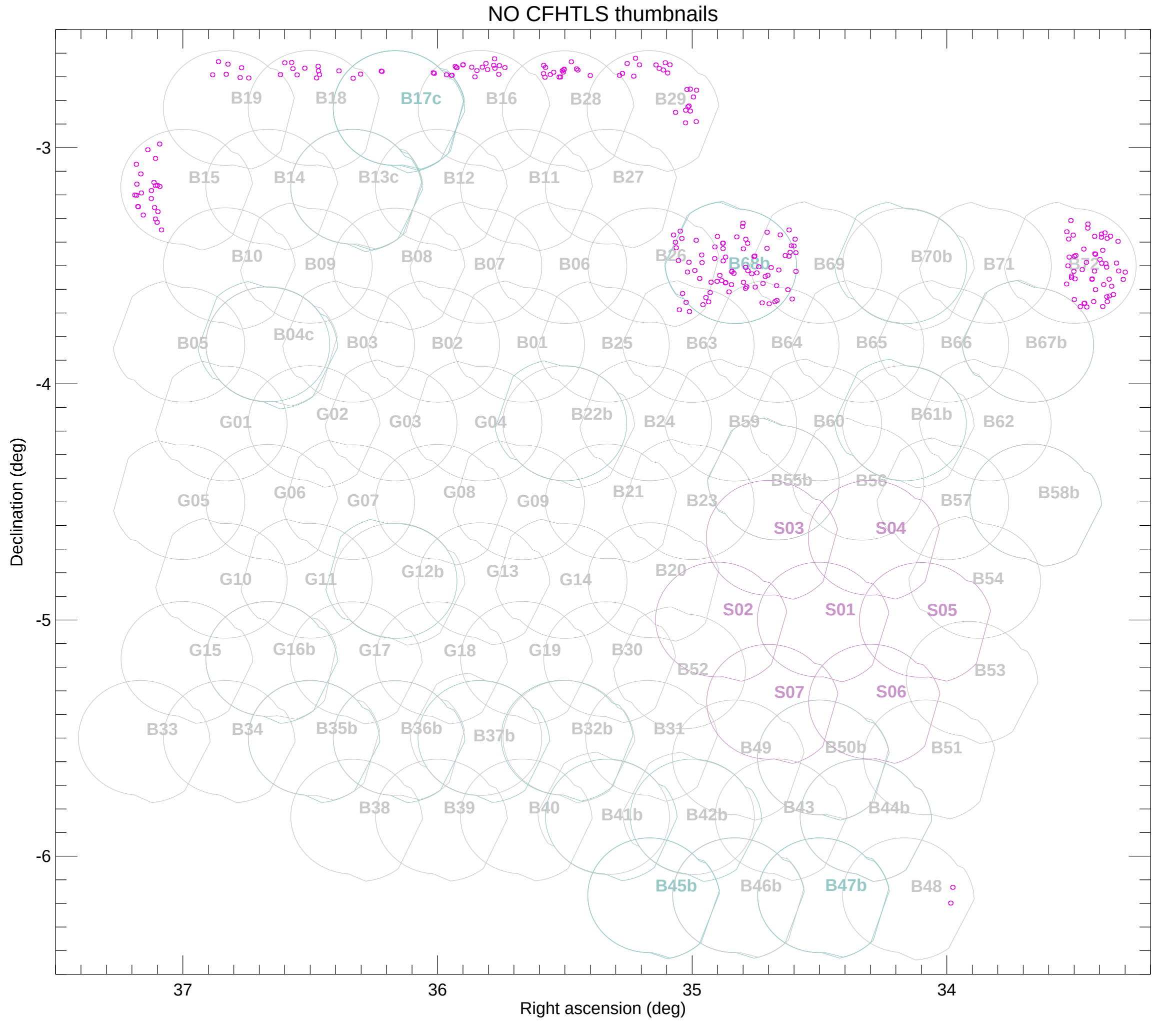

Fig. 5. Positions of the X-ray sources for which no optical thubmnails are available. For symbols see 5 in text.

reachable from the last column of the VIEW DATA page. The effect of the "online generation" of the data product is the opening on a new web page at the SIMBAD or NED site.

A further kind of *object-related* data products for the optical identification tables `2XLSSOPT*` (only) are *textual comments* manually inserted with the tool described in 4.1.

Finally we generated *thumbnail images* (which can be inspected for identification and ranking validation, using a tool like the one proposed in 4.2) from the CFHTLS and SWIRE (UKIDSS is potentially available but not supported), as anticipated in Reports IV-VI, 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).

With the exception of UKIDSS thumbnails for the reason quoted, all other optical and IR thumbnails have been retrieved and are associated to `2XLSS`, `2XLSSe`, `2XLSSOPT` and `2XLSSOPTe`. Originally I thought to retrieve the thumbnails only around X-ray sources having respectively at least one optical or a SWIRE counterpart, but then I decided to attempt the retrieval for all X-ray positions (with the idea that it could be useful to inspect the optical or IR field even when no catalogued counterpart

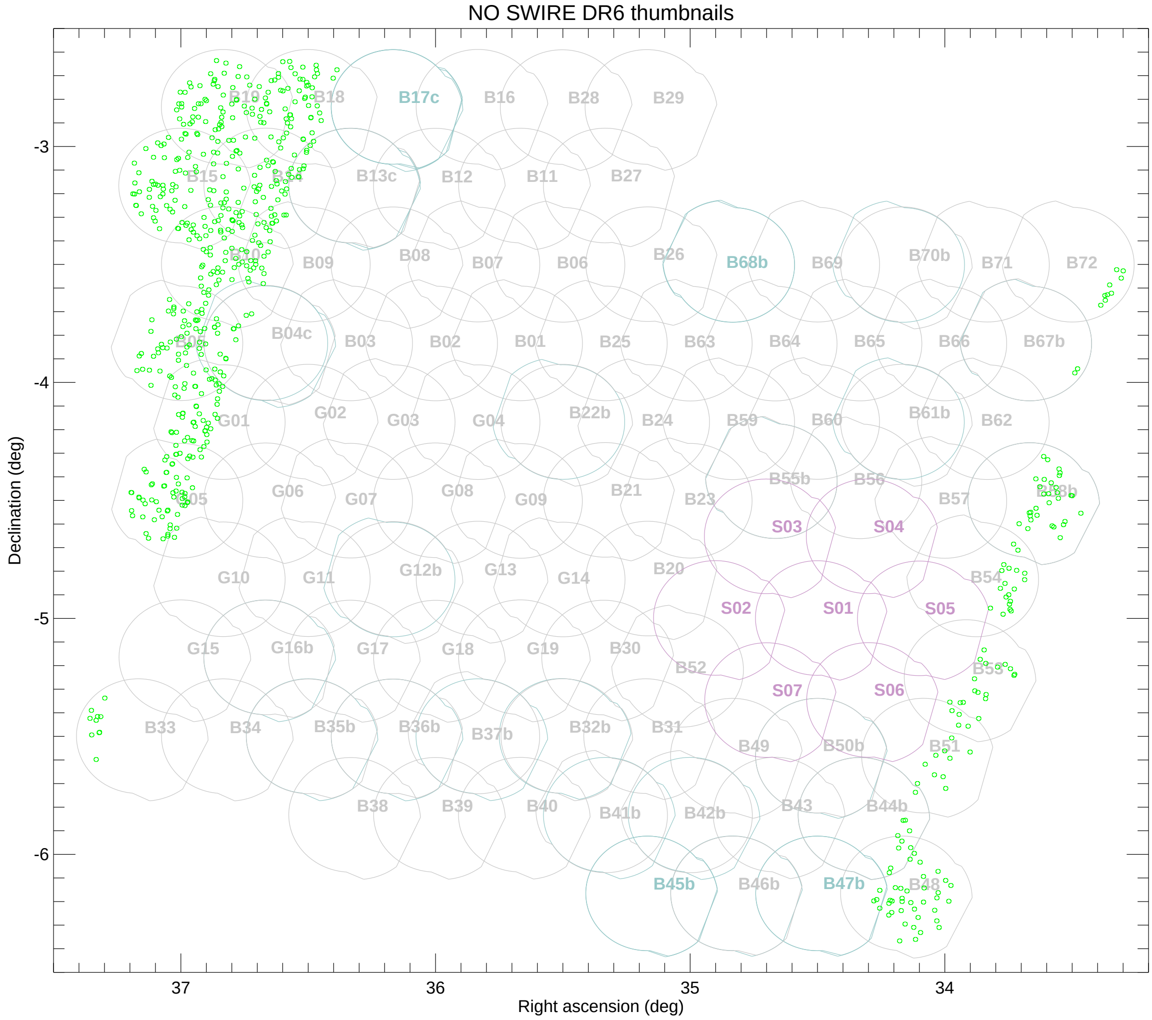

Fig. 6. Positions of the X-ray sources for which no SWIRE thubmnails are available. For symbols see 5 in text.

exists, particularly since some bright, saturated objects do not appear in the optical catalogue).

In some cases the attempted retrieval may fail for some (SWIRE) bands for which there are not data in the band. In (few) other cases it fails completely, tipically because the sky area was not observed by CFHT or Spitzer (see positions in Fig. 5 and 6).

4. Identification support tools

As anticipated in Report VI, I created web interface tools to support the procedure for the validation of optical identifications in a way similar to what I did for the XMDS (see Report IV). 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.

4.1. The validation interface

The *validation interface* is a tool which will allow to confirm or alter the automatic ranking, and at the same time to edit the GCT underlying one of the 2XLSSOPT* catalogue. Such interface could be used to manually (re)assign ranks to preferred counterpart sets and to reject unwanted ones. It allows also to clear mistakes in association of counterparts due to peculiarities in the data, and finally to insert comments about choices made.

The tool can support only one catalogue at a time, with changes done in the java and HTML code, and requiring recompilation. It currently points to 2XLSSOPTe, which means this is at present the only table supported for editing.

4.2. The graphical interface

A second tool existed already in prototypal form (see URL in Report V). Such a tool is an applet which allows 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

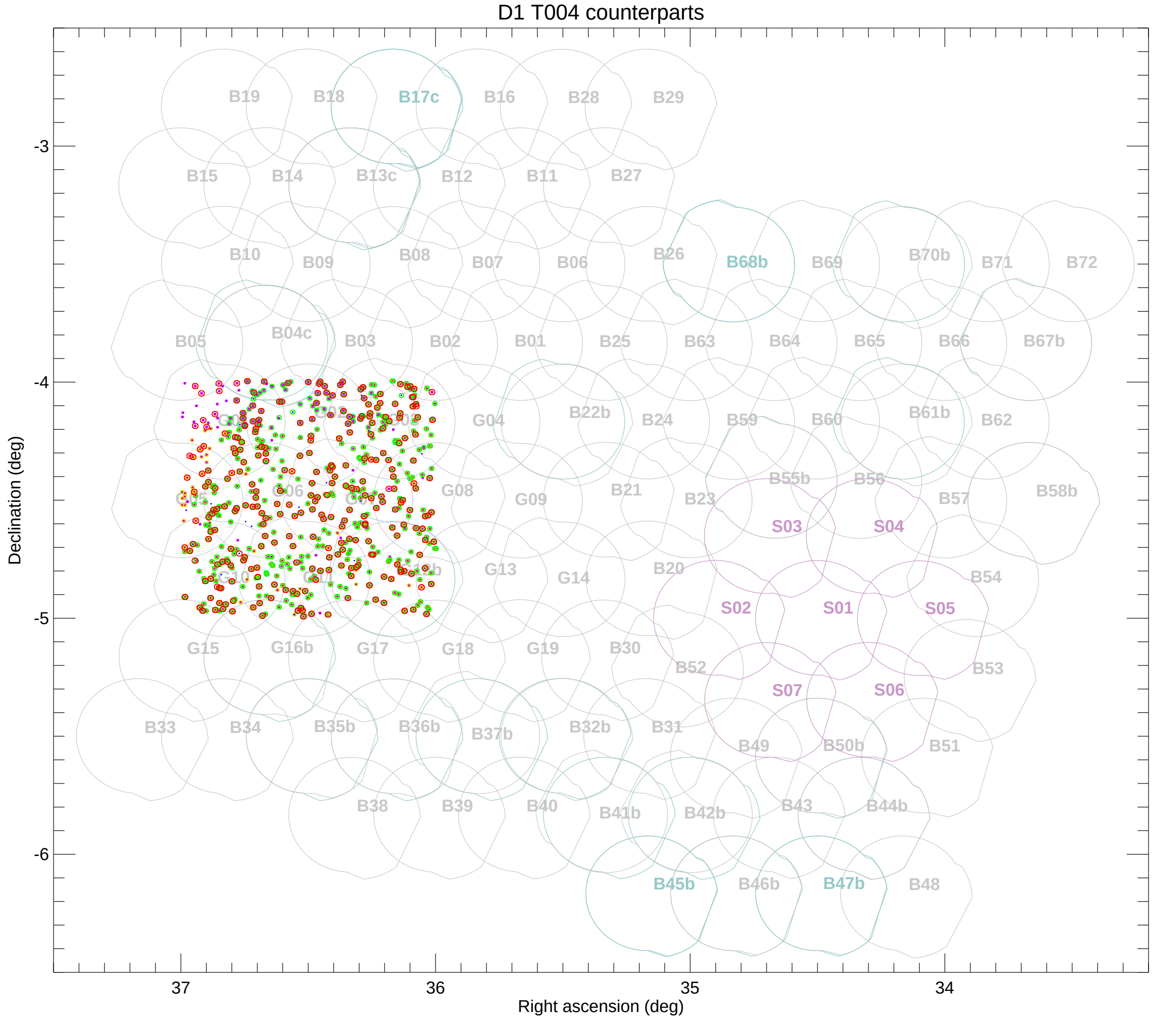

Fig. 7. Positions of the X-ray sources with a CFHTLS D1 counterpart. For symbols see 5 in text. The CFHTLS D1 covers the central part of the XMDS (G) fields. In this and in the next 4 figures the EPIC FoV footprint appears in light pink-gray for good fields, and in azure-gray for bad fields. SXDS fields are drawn in darker tone. All figures of this family were generated from `2XLSSOPTe`.

external non-member tables i.e. presently SIMBAD, NED and USNO) and interact with them.

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

I have presently two production variants derived adjusting the demo prototype: the so called "combo26" reads regions from `2XLSSOPT`, while "combo26b" allows to choose between `2XLSSOPTe` and `2XLSSOPTf`. Both support i' or g' images, as well as gzipped FITS images for all bands.

5. Catalogue statistics

We report here also some `2XLSS` *results from Report VI so that one can have at a glance a comparison with the extensions.*

5.1. The X-ray catalogues

The `2XLSS` proper 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. The corresponding figures for `2XLSSe` (or respectively `2XLSSf`) are 7083 (7082) total, 2197 (2215) in both bands, 3945 (3932) only soft and 941 (935) only hard.

In `2XLSS` proper of a total of 197 extended sources (57 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 band; there are 5 extended sources nominally classified C1, and 44 classified C2 detected only in the hard band. The corresponding figures for `2XLSSe` (or respectively `2XLSSf`) are 225 (225) total extended, 60 and 165 (59 and 166) C1 and C2, the same

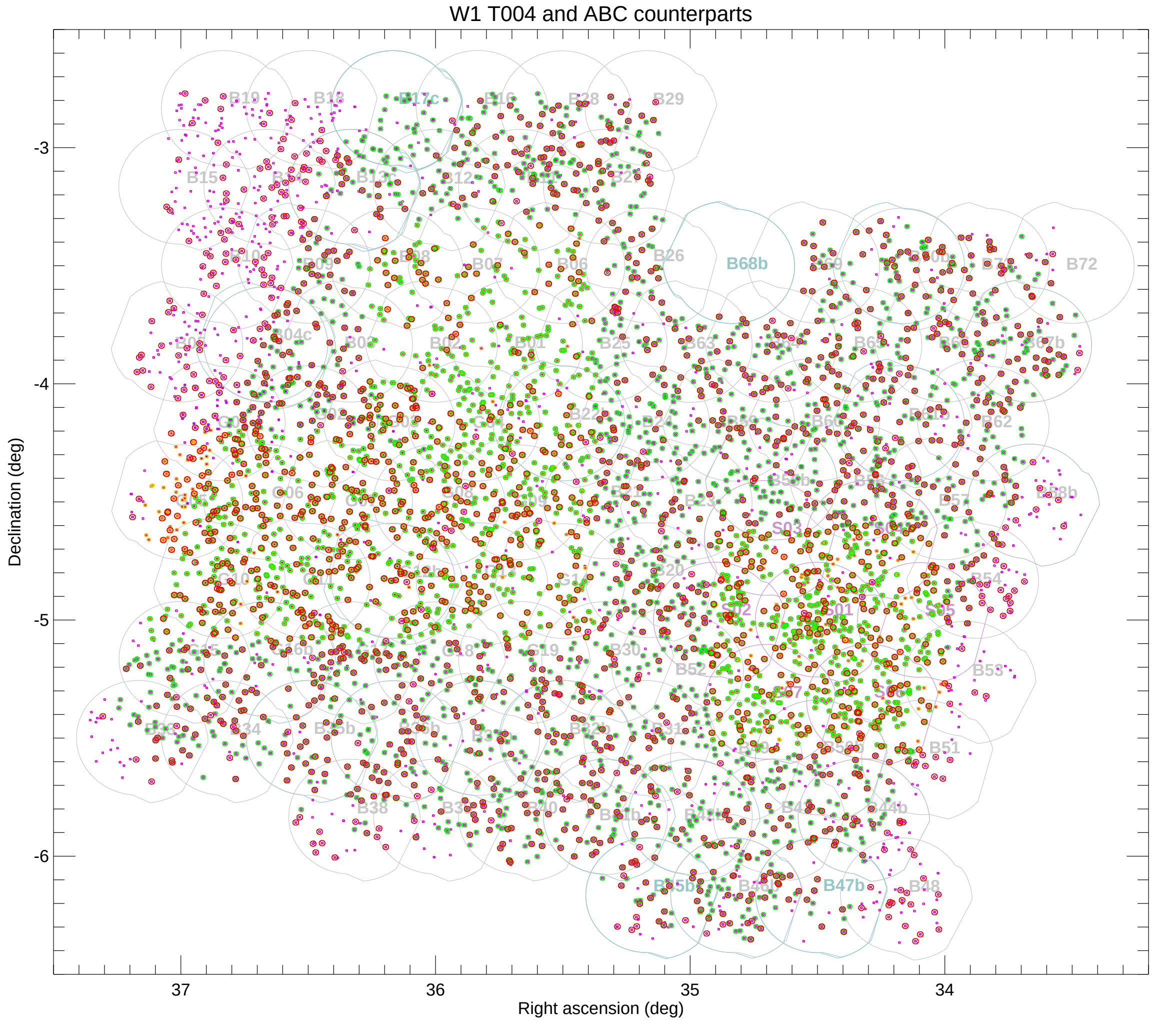

Fig. 8. Positions of the X-ray sources with a CFHTLS W1 counterpart. For symbols see 5 in text. The CFHTLS W1 extended with the northern ABC fields covers almost all of our fields.

10 and 7 detected in both bands, 45 (44) and 104 (04) soft C1 and C2, and 5 (5) and 51 (52) nominal hard C1 and C2.

For `2XLSS` proper, the number of pointlike sources (6089 total) is 1862 (99%) detected in both bands, 3449 (93%) in the soft band and 778 (94%) in the hard band. The corresponding figures for `2XLSSe` (or respectively `2XLSSf`) are 6858 (6857) total, 2177 (2195) in both bands, 3796 (3784) soft and 885 (878) hard.

For the pointlike sources in `2XLSS` proper, 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. The corresponding figures for `2XLSSe` (or respectively `2XLSSf`) are 59% (59%) above 4σ, 77% (77%) above 3σ, and 85% (86%) best in soft band.

Finally for `2XLSS` proper, 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. The corresponding figures for `2XLSSe` (or respectively `2XLSSf`) are 13% (13%) above 4σ, 33% (34%) above 3σ in the soft and 3% (3%) above 4σ, 11% (11%) above 3σ in the hard band.

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

Finally it is worth reporting this breakdown between the sources flagged to be in good or bad fields (see discussion in 3.2):

	`Xbadfield`		
Field type	0 good	1 bad	
"our" B and G fields	5822	(409)	of which
in AO7 repeats		109	
SXDS fields	788	(64)	(in S01)

If one wants to cover all the `XLSS` area without holes, but excluding dubious sources in bad fields, excepted the

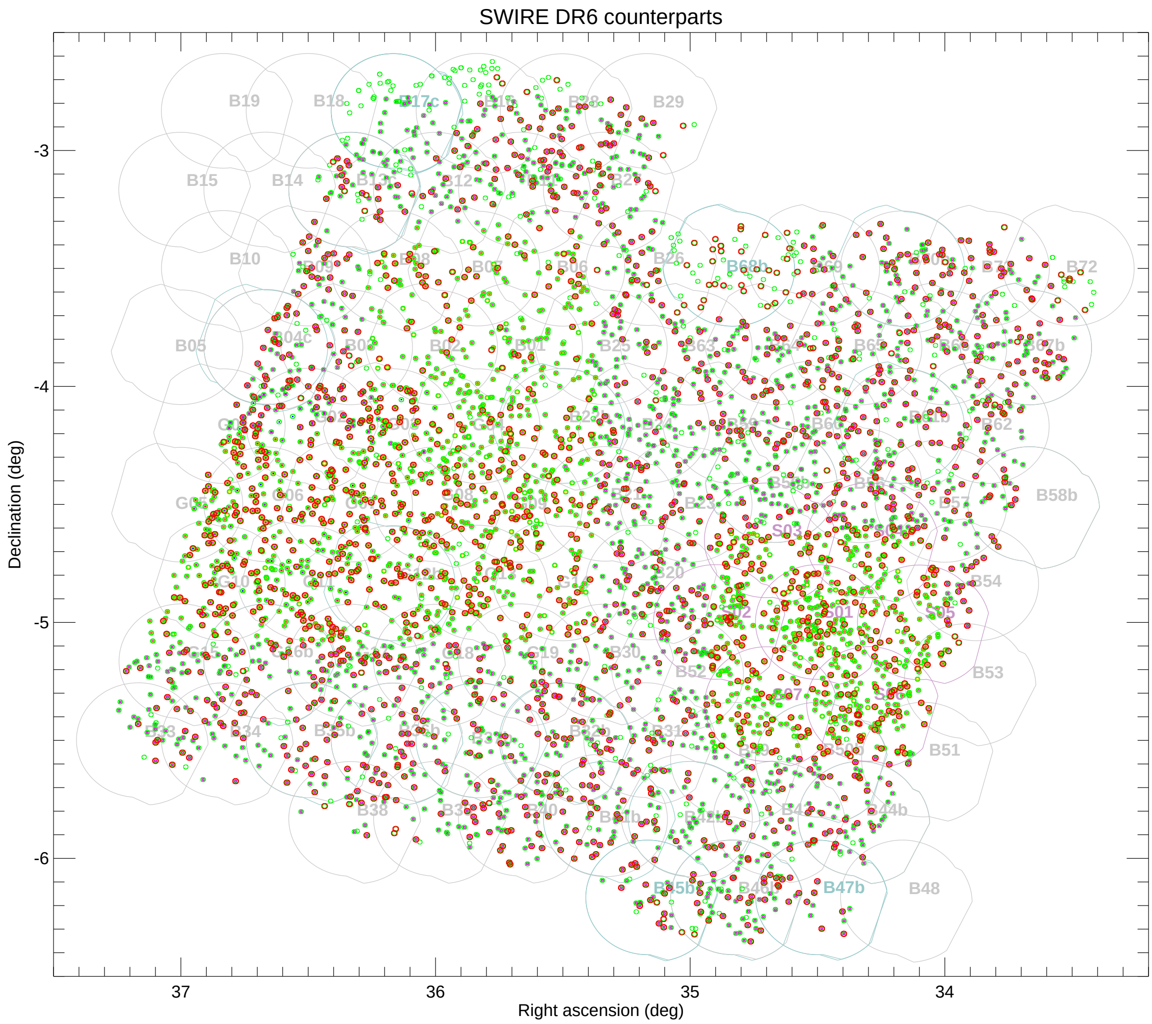

Fig. 9. Positions of the X-ray sources with a SWIRE counterpart. For symbols see 5 in text. SWIRE covers almost all the fields, except the E and W edges.

AO7 repeats, one should take the sources *not* indicated in parenthesis (6719 in total).

5.2. The joint X-ray/optical catalogue

2XLSSOPT contains nominally 19168 counterpart sets, 2XLSSOPTe 21909 and 2XLSSOPTf 21902 .

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. 7 to Fig. 11) 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
– 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).

Fig. 5 and 6 use the same notation, but indicate the X-ray sources from which we have *no CFHT or SWIRE thumbnail*. This occurs outside of the SWIRE pointings, or where we had no access to CFHTLS data or no CFHT observation was made (even around our own ABC fields) because of the presence of bright star Mira Ceti.

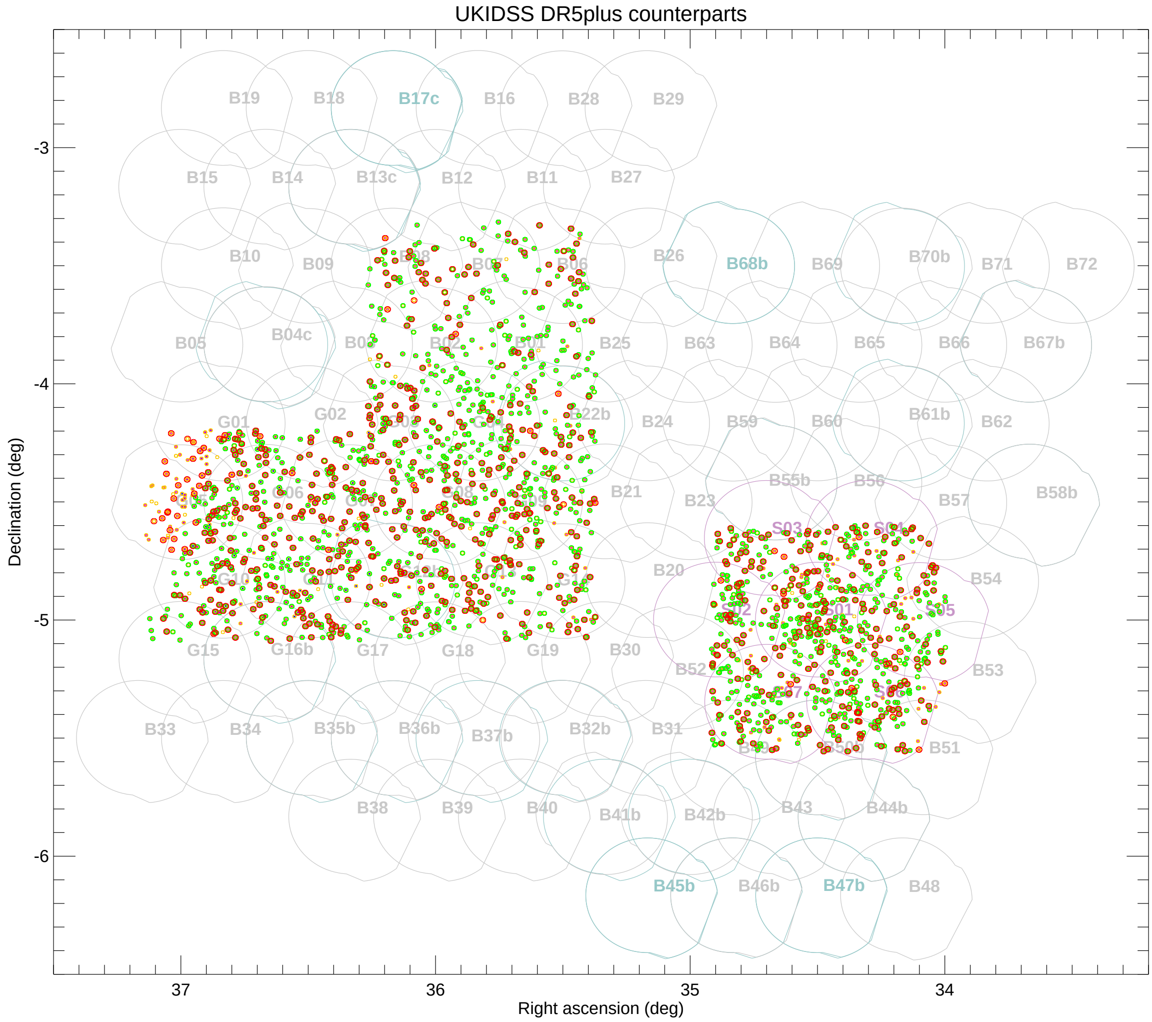

Fig. 10. Positions of the X-ray sources with a UKIDSS counterpart. For symbols see 5 in text. UKIDSS DR5plus still covers so far two disjoint areas (DXS and UDS), one of which covers the SXDS fields.

X-ray sources nominally flagged as *blank fields* (i.e. having no catalogued CFHTLS, SWIRE, UKIDSS or GALEX counterpart within 6″) are 268 in `2XLSSOPT`, 275 in `2XLSSOPTe` and 274 in `2XLSSOPTf`. *Note that the absence of catalogued sources does not mean they are necessarily real blank fields*. Often bright sources are omitted by the catalogues, but are visible if one inspects the thumbnail image. Compare for instance the cases of sources 27601, which is very close to a R=15.6 galaxy shown in SIMBAD, or 38678 whose field is spoiled by nearby bright star BD-05 427. So some of those sources can have a bright counterpart.

1182 X-ray sources have a *physically single counterpart*, while the rest has potentially more counterparts. A different count can be obtained using the rank (`Xrank`) and "autorank" described in 3.6.2, which give:

– 2729 sources have a *single very reliable counterpart*, i.e. rank 0 and autorank in the range 10 to 13.
– 2015 have a *single , but not so reliable, counterpart*, i.e. rank 1 and autorank in the range 10 to 13.
The distinction between the two groups is somewhat blurred. Anyhow they both include not only the physically single, but also cases with other *rejected* counterpart sets
– 1090 X-ray sources are *pseudo-ambiguous*, with one *definitely* preferred counterpart (rank 0 and autorank in the range 0-3), plus one or more nominal secondary counterparts with rank 2.
– 974 X-ray sources are *definitely ambiguous*, with one *nominally* preferred counterpart (rank 1 and autorank in the range 0-3), plus one or more secondary counterparts with rank 2, at least one of which is not terribly worse than the nominally preferred one.
For the latter two groups the rank is really meaningful and distinctive.
– the 275 tentative *blank fields* described above have rank 1 and autorank 4.

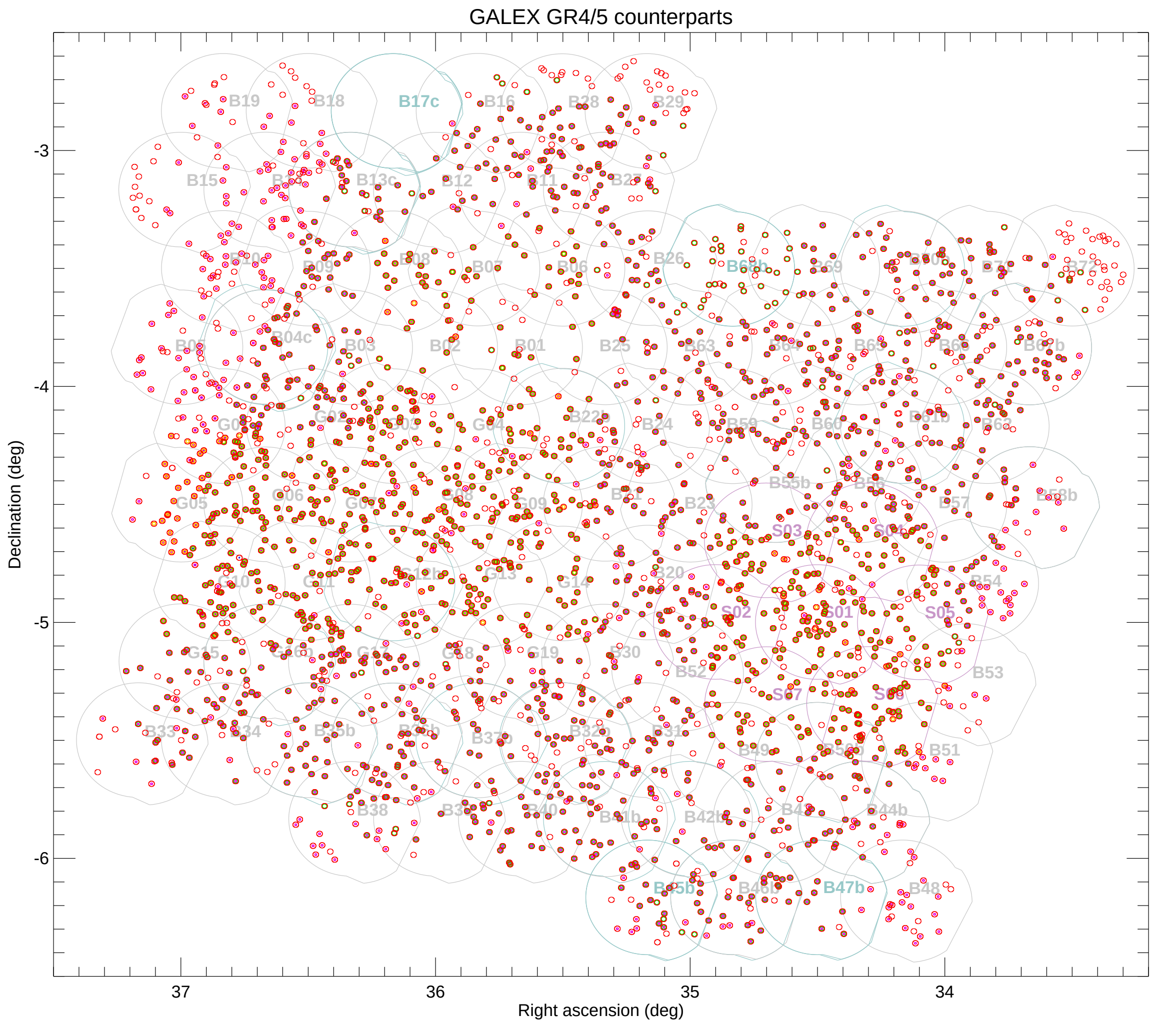

Fig. 11. Positions of the X-ray sources with a GALEX counterpart. For symbols see 5 in text. GALEX GR4/5 covers almost all of our fields.

Probability class		good $p < 0.01$					fair $0.01 < p < 0.03$					bad $p > 0.03$
in how many catalogues ?	n/a	4	3	2	1	some	4	3	2	1	some	all
Counterpart set												
Blank field	275											
Best and single		46	224	133	139	36	10	52	56	69	73	344
Best		388	848	626	623	313	145	437	369	294	624	959
Secondary		4	89	115	636	601	28	89	201	636	601	12078

Table 3. Basic statistics ot the 2XLSSOPTe extended catalogue

I have attempted a rough characterization using the criteria defined in 3.6.1. The results are summarized in Table 3 which has to be interpreted as follows:

Looking at the row "best and single", 46 (4%) of the physically single counterparts are detected in all four optical/IR/UV catalogues with *good* probability in all of them. 224 (19%) of such single counterparts are detected in 3 out of 4 catalogues (and not detected in the other) with a good probability in all three. 133 (11%) are similarly detected in 2 of the 4 catalogues with a good probability in all the catalogues where they are detected. 139 (12%) are detected in only one catalogue with a good probability. 36 (3%) are detected in 2 up to 4 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 (10, 52, 56, 69 and 73). Finally 344, despite being the only possible counterpart, are detected in a number of catalogues from 1 to 3 or exceptionally 4, but always with a bad probability.

Similarly 388 of the best *non-single* counterparts are detected in all 4 catalogues with a good probability in all of them, etc. etc. up to 959 cases which, despite being the best counterpart, are detected always with a bad probability.

Considering the secondary counterparts, 12078 (representing about 90% of the "all bad") 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. 4 cases where the secondary counterpart has a good probability in all 4 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, 47.6% of the sources have a best counterpart with a good probability, 30.0% a fair one, and 3.8% are nominal 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).

43% of the best good counterparts are detected above 4σ; 18% of the fair ones; 5% of the bad ones and 11% of the blank fields. Or conversely, of the 1922 X-ray sources above 4σ, 75% have a good counterpart, 20% 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; 42% of the fair ones; 16% of the bad ones and 23% of the blank fields. Or conversely, of the 3265 X-ray sources above 3σ, 64% have a good counterpart, 27% a fair one, 6% a bad one and 2% are unidentified.

Fig.4 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. In fact the current ranking procedure (3.6.2) (as indicated also by O.Melnyk) seems to favour sometimes as rank 0/1 a GALEX-only counterpart with a nominally very good probability, and assign rank 2 to (or reject) an optical/SWIRE counterpart, so a tuning would be desirable, as shown by a preliminary analysis. There are 493 cases in which the best counterpart is a GALEX-only one. In 174 it is the only non-rejected one (but only 74 are physically single) , while in 319 there is at least another rank 2 one (only in 8 cases the GALEX counterpart has rank 0, the remaining 311 are all definitely ambiguous). In 208 cases the best rank 2 companion has also good probability, in 248 it includes an optical counterpart (though only 93 of them have a good optical probability). In 54 cases the best rank 2 companion is another GALEX-only, and in 124 (inclusive of the 54) the best probability in the counterpart set is the GALEX one (in 115 the SWIRE one, in 60 the optical one, and in 20 the UKIDSS one). Of the 54 cases, only for 21 both best counterparts are GALEX-only (15 because they are outside the CFHT and SWIRE areas, 2 near uncatalogued bright objects), for the rest in most cases there is a "tertiary" optical/SWIRE/UKIDSS counterpart with good probabilities also.

One shall also note that the *ranking depends* on the probabilities, and these depend on the distance (see 3.6) and therefore *ultimately on the X-ray position*. If the latter changes, the rank choice will change. This is demonstrated comparing e.g. 2XLSSOPT and 2XLSSOPTe for the 51 sources replaced between B and SXDS fields, or for the 127 differences in S01 between 2XLSSOPTe and 2XLSSOPTf . While most cases are compatible, in a few the slightly different X-ray position in the different pointings causes a swap of rank (rank 0/1 in one catalogue is rank 2 in the other and v.v.), or a swap between a rank 0/1 and a rejected counterpart, or rarely an altogether different choice (usually due to largish X-ray position displacement in nearly blank field, but sometimes instead in a crowded area).

6. Future work

The following activities should be considered before the publication of a final catalogue, as well as to refine and support ongoing work. Any ideas on these and other topics will be appreciated.

- The possibility of re-doing the band merging and overlap analysis with a different radius to cope with possible missed mergers and overlaps is discussed in the next section (6.1).
- The replacement of the current SDS fields with 10 ks exposures could be considered
  Both these items will affect the overall list of sources in the X-ray catalogues. The next two items instead affect (only) the X-ray/optical catalogue descending from the primary catalogue,
- The (automatic) ranking of identifications might be refined modifying the criteria used (e.g. to compensate for the excessive bias in favour of GALEX-only counterparts as described in 5.2 above)
- The refinement of individual identifications might be required, via manual inspection and edit, or possibly via semi-automated procedures.

In particular on the latter item one may consider things like a systematic inspection of putative blank fields, or of counterpart sets with good probabilities but no CFTHLS counterparts. Ideally, to avoid manual inspection of a relatively large number of cases, one might want some form of *systematic analysis of optical image (thumbnail) files* to locate bright uncatalogued sources (or just use SIMBAD/NED presence as guideline ?).

Cases like these (and the examples listed below) represent a serious concern, since they cannot be spotted a priori using the database tables, because the info is incomplete in the (optical) catalogues. Consider e.g. source 28816 (nominally separate optical/SWIRE and GALEX counterparts to be merged over bright star BD-04 381); or source 30008 (nominally a GALEX only object over the bright galaxy NGC 894); or source 35483 (which is in the outskirts of bright galaxy MCG 01 06 080). There are surely many other similar cases.

6.1. The 6″ vs 10″ issue

O.Melnyk and A.Elyiv noted a number of X-ray sources which are apparently rather close and share the same optical counterpart, and this instigated me to review the choice of 6″ as correlation radius for band merging and overlap removal (decision done at the time of the `XLSS` paper).

The suspect cases may look as a couple of soft-only and hard-only sources detected in same field, at a distance marginally above 6″. I call them "*potential missed mergers*" since they might have escaped band-merging because of the distance. Or they may be sources detected in different pointings, with a distance marginally above 6″, but somehow compatible. These could be "*potential missed overlaps*".

At the time of the `XLSS` paper we concluded that "*it is highly unlikely (much less than 1%) that distinct sources (detected in same field and same band) are closer than 10″*" and converged on using the same, conservative, radius of 6″ for both band merging and overlap removal.

The expectation from XMDS was that sources detected in one pointing were always more distant than 18″, while the same source detected in two pointings were closer than 6″. But XMDS was based on the Milan pipeline, where detections, unlike XAMIN, were done in all bands at one time, and therefore there was no separate band-merging step.

I now repeated an analysis, considering first of all *physical, band-merged, database tables*, and comparing non-spurious band-merged detections *in the same field*, which are closer than 30″. I consider as *potential missed mergers* two detections which are: (a) in the same field; (b) detected in one energy band only; (c) one is detected in the soft band and the other one in the hard band. In Fig.12 I report the frequency histograms (normalized in percentage to the total number of entries) for all the objects within 30″ (in black), for the potential missed mergers (in blue), and for the remainder (in magenta).

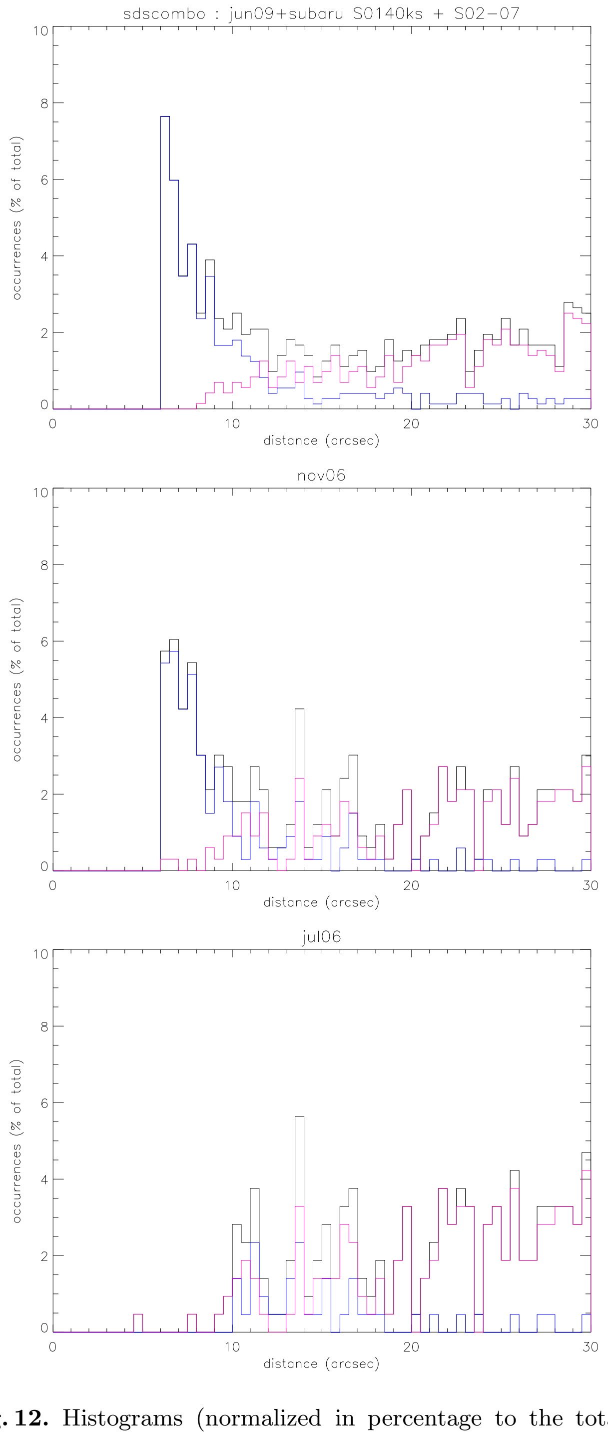

Fig. 12. Histograms (normalized in percentage to the total number of entries) of the inter-source distance for all the couples of objects in same field within 30″ (in black), for the potential missed mergers (in blue), and for the remaining objects (in magenta). The top panel refers to the band-merged table used as *input* to the present `2XLSSe` catalogue, the middle panel to the one used as input for the published `XLSS` one, and the bottom panel to an earlier version (`jul06`) using a 10″ band-merging radius.

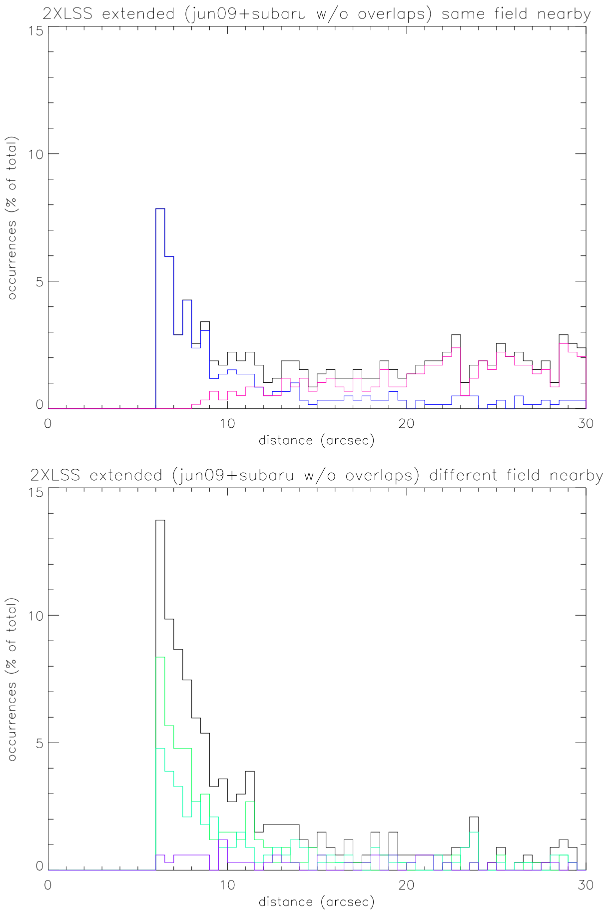

Fig. 13. Histograms (normalized in percentage to the total number of entries) of the inter-source distance from `2XLSSe`. The top panel is for the objects in the same field, and uses the same colour code of Fig.12; the bottom panel is for objects in different fields: all in black, green for couples with fully compatible classification in both bands (likely missed overlaps), cyan for 1-band compatible cases (possible missed overlap) and violet for incompatible cases (independent sources).

For `sdscombo` (the concatenation of `jun09` and `subaru` used as input for `2XLSSe`) one has the following breakdown. Of 719 detections closer than 30″, 48% are potential missed mergers. However of the 232 detections closer than 10″, 95% are flagged potential missed mergers, while only 25% of those farther than 10″. All cases closer than 8.1″are potential missed mergers. This is apparent from the top panel in Fig.12 where the blue histogram coincides with the black (total) one below 8″. The situation is not unlike `nov06` (central panel of Fig.12), which is the input table, based on the previous XAMIN release, used for `XLSS`, while it is by construction different for `jul06` (bottom panel), which is an abandoned test on the same data used in `nov06` but which used a band merging radius of 10″.

Next I considered a *catalogue after overlap removal*. Results for `2XLSSe` are shown in Fig.13. The top panel considers the catalogued sources *in the same field*, and is fully alike the cases shown in the previous figure (but for 586 sources surviving overlap removal). The bottom panel instead considers the sources *in different fields* (total histogram in black) and classifies them as: (a) *fully compatible* if they have the same classification in both fields (e.g. both `PP` pointlike detections in both bands, or both `P-` pointlike soft detection only, etc.); (b) *one-band compatible* if both were either detected or undetected in one band, and different in the other band; (c) *fully incompatible* otherwise. Of 335 cases of sources in different fields closer than 30″, 48% are fully compatible (so likely missed overlaps ?), 39% are one-band compatible (so possible missed overlaps ?) and 13% are incompatible (independent ?). For distances within 10″, one has 55% compatible, 37% one-band compatible and 8% incompatible.

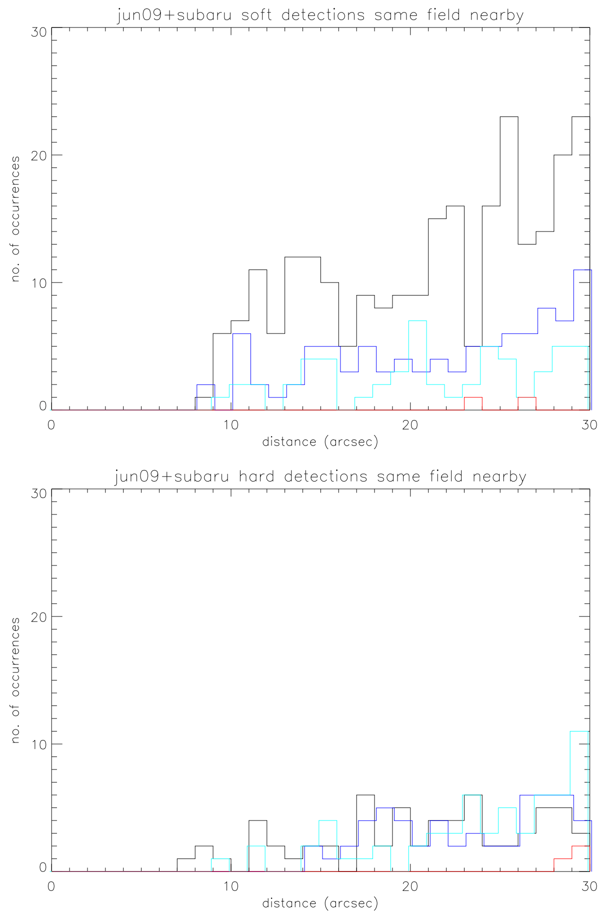

Fig. 14. Histograms (absolute counts) of the inter-object distance for individual band detections (top panel for soft band, bottom panel for hard band). In black couples of objects which are both non-spurious pointlike detections; in blue the couples where only one object is spurious; in cyan the couples where both objects are spurious; in red the very few couples where one of the elements is extended in one pointing. The various histograms are slightly displaced horizontally for clarity.

All the above looks unaffected by things like the detection likelihood or the exposure time of the different pointings or the off-axis angle.

Finally I considered the detections in the single bands (i.e. the step *before* band merging, tables `sdscombob` and `sdscombocd`). For instance in the soft band there are 318 couples of non-spurious detections in the same field closer than 30″ (of which 314 are both pointlike). Of these none are closer than 8″, only 1 is closer than 9″, and 9 are closer than 10″. If one includes also spurious detections, one gets 501 couples (just 12 within 10″). In the hard band one has 90 couples of non-spurious detections in the same field (all both pointlike), and, including also spurious detections, 236 couples. Only 7 of these (5 non-spurious) are closer than 10″, none closer than 7″.

Fig.14 gives (absolute frequency) histograms separately for the following cases: (a) in black the (314 or 90) couples of objects which are both non-spurious pointlike detections (meaning they will surely pass to the band merging stage as non-spurious sources or (soft/hard) components thereof); (b) in blue the couples (121 soft, 72 hard) where one object is spurious and the other isn't (meaning the latter will be recovered *only* if merged with a non-spurious in the *other* band); (c) in cyan the couples (61 soft, 71 hard) where both objects are spurious (so *either* could be recovered only if merged with a non-spurious in the other band); (d) in red the very few couples where one of the elements is extended in one pointing.

What is striking is the fact that in the current band-merged tables (Fig.12 and 13) there is an *excess* of couples of putative sources in the 6-10″ range with respect to the lack of such couples in the individual band detections (Fig.14). This throws doubt on their reality and strengthens the supposition they are missed mergers or missed overlaps.

It is therefore possible that a final catalogue should *repeat afresh the band merging and overlap removal* procedures (and all the rest, which is not only time consuming, but will in principle generate a new source numbering) with a radius larger than 6″, or *attempt some selective recovery* of the cases between 6″ and the new larger radius (preserving when possible source numbering, but at the price of a less clean though faster procedure).

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)
Chiappetti, L., 2009, XMM-LSS Internal Report N. 6-Mi (Report VI)
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. 2010, MNRAS, 401, 294
Tajer, M., et al. 2007, A&A, 467, 73
Ueda, Y., et al. 2008, ApJS, 179, 124