XMM-LSS INTERNAL REPORT n°9-Mi

November 24, 2011

The ultimate 2XLSS deep and 10ks catalogues

The final report in advance of publication

L.Chiappetti[1]

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

Abstract. I report on the latest releases (which supersede any previous releases marked 2XLSS*) of the "deep" 2XLSSd catalogue (using "deep" i.e. full exposures) and of the candidate official 2XLSS catalogue (using 10ks truncated exposures). The report gives reference information similar to the one provided in previous internal reports, and elements which could be useful for the 2XLSS paper to be written. The catalogue tables will contain X-ray results deriving from the reprocessing with the 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 *supersedes* the reports issued in October 2009 (Chiappetti , 2009) (hereafter Report VI, describing an earlier, preliminary version of the 2XLSS catalogue) and in September 2010 (Chiappetti , 2010) (hereafter Report VII, presenting the so-called extensions to 2XLSS inclusive of the addition of the SXDS (Ueda et al., 2008) to our X-ray data from GTO to AO7). As for Report VII, it describes also the work done on the catalogues in *ranking* identifications in the optical and other wavebands.

All the X-ray data described here (as well as in Report VII) were reprocessed in an uniform way at Saclay with the *pro tempore* latest (Py3.2) XAMIN pipeline. In particular the input data used for 2XLSSd is exactly the same used for older 2XLSSe. However there are some important differences, consistent with the agreement taken at a teleconf in Jan 2011, namely:

- for both newer catalogues band merging and overlap removals were performed using a radius of 10″ instead of 6″
- for efficiency reason our own data (table jun09) and SXDS (table subaru), the *same* used for 2XLSSe in Report VII, are no longer concatenated in an *union*, but were concatenated *physically* in table jan11, the basis for 2XLSSd
- instead the 10ks data were reprocessed afresh in Saclay with XAMIN 3.2 (with the exception of SXDS 10ks chunks which were already processed), and form table may11, the basis of the new 2XLSS

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 so far 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; Adami et al., 2011).

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), an earlier 2XLSS catalogue released for internal use in Oct 2009 in Chiappetti (2009), and the earlier 2XLSSe extension released internally in Aug 2010 in Chiappetti (2010).

I present here two catalogues: the "deep" (full exposure) catalogue called 2XLSSd (released to the consortium in February 2011) and the 10ks 2XLSS catalogue (just released in early September 2011, and superseding any previous catalogue with same name). 2XLSS is more conservative and uses an uniform exposure time for all pointings. It is planned to be the official published catalogue, but it has not yet been widely used, so a final assessment will be made when people have used it. 2XLSSd instead supersedes the previous 2XLSSe version (of which uses the same input data - in particular for field S01 it uses the 40ks exposure, not the full one, as it does for all other pointings - but with a 10″ band merging).

Superseded catalogues were retired from access. The only older catalogue available is the published XLSS.

Table	Update	Content	History	(5)	(6)
jun09b/cd	Jul 09	X-ray sources from 3.2 Saclay pipeline, individual bands, full exposures	unchanged since Jul 09		
subarub/cd	Apr 10	X-ray sources from the Saclay pipeline, individual bands	unchanged since Apr 10, inclusive of 10ks chunks and S01 40ks		
jan11*	**Feb 11**	X-ray sources from 3.2 Saclay pipeline, band merged within 10″	**jun09 and subaru physically concatenated, remerged at 10″**	**10″**	a
may11*	**Apr 11**	X-ray sources from 3.2 Saclay pipeline on 10ks chunks, band merged within 10″	**reprocessed afresh** but for subaru	**10″**	a
d1t4	**May 11**	CFHTLS D1 field release T004	in use since Jan 088; **added objects**	6″	
w1t4	**May 11**	CFHTLS W1 fields release T004 both supplied by Saclay	In use since Jan 08; **added objects**	6″	
swiredr6	**May 11**	SWIRE DR6 supplied by IPAC	in use since Jan 08; **added objects**	6″	
ukidssdr5	**May 11**	UKIDSS DR5plus public release	in use since Aug 09; **added objects**	6″	
galex	**May 11**	GALEX GR4/5 public release	in use since Nov 08 for XMDS; **added objects**	6″	
simbad	**May 11**	SIMBAD sources	present since 2003 and regularly updated	20′	b
ned	**May 11**	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 2XLSSd and 2XLSS catalogues

(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)

The 2XLSSd catalogue includes 6723 entries while 2XLSS has just 5548.

Section 2 lists input database tables, namely X-ray (2.1) and optical-IR-UV data (2.2), while ancillary technical details are given in 2.3, and the astrometric correction in 2.4. The procedure used to create the 2XLSSd and 2XLSS catalogues is described step-by-step in the 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 among releases is presented in 3.4, namely for raw database tables in 3.4.1, for the old published catalogue in 3.4.2, and for the two releases presented here in 3.4.3. Section 3.6 introduces the 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 (section 4). Section 5 gives summary statistics on the catalogues, the X-ray ones (5.1) and the X-ray/optical ones (5.2), while a comparison between the 2XLSSOPTd and 2XLSSOPT catalogues is shown in 5.3. A special section (6) is dedicated to a comparison with XMDS A final section (7) lists the open points for publication.

2. Data sources

The starting point for the X-ray catalogues have been the X-ray tables (described in 2.1). For the 2XLSSOPT* virtual tables (and the astrometric correction, see 2.4 !) some other recently updated 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, as usual for previous releases and 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 X-ray catalogues proper were two families of physical tables (constituted, as usual, by the two single-band tables, and by the band-merged table, 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), and of SXDS observations.

The tables for the full exposure analysis (jun09 for our data, and subaru for SXDS) were described in Report VI. The 10ks "chunks" (and the 40ks chunk for S01) for SXDS were described in Report VII (and were already part of subaru). 10ks chunks for our fields were instead recently reanalysed by Saclay and were ingested in the `may11` table family (`may11`b and `may11`cd for the individual bands, and `may11` as band-merged).

For the deep catalogue the data used in input is exactly the same described in report VII except for the band merging radius. Band tables `jan11`b and `jan11`cd were created by the physical concatenation of respectively `jun09`b with `subaru`b or the corresponding CD-band tables. However table `jan11` was recreated afresh applying the band-merging procedure at 10″ to `jan11`b and `jan11`cd.

Instead 10ks data were ingested afresh into `may11`b and `may11`cd for our own fields. The (first) 10ks chunk of S01 to S07 in `subaru`b and `subaru`cd were then copied into the band tables, and finally the band-merging procedure at 10″ created `may11`.

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, *with the exception of the new 10″ merging radius*. 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 had been used in the past 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`), in order to prevent confusion. This was intended to allow `jun09` and `subaru` to be concatenated, which is what we did creating `jan11`, and in fact the `jan11` sources use the *same* sequence numbering since *the sources are actually the same*. *Sources in* `may11` *have instead all a new distinct sequence numbering (starting at 40000)* inclusive of the SXDS 10ks chunks (whose numbers were never official and in spite of the fact they were already present in the `subaru` table).

The remainder of the section is virtually identical to what included in Reports VI and VII, but is included for self-completeness. Changes are highlighted like this. 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 was in `subaru` a complete sequence of 10 ks chunks for each SXDS exposure, which vary in number from 7 for S01 to 1 for S03). The Saclay abbreviated name for chunk m of field S0n is S0n_cm, and the equivalent number is 2000+100*m+n (e.g. S05_c3 is 2305). *This numbering did not make its way into* `may11` *as explained below.*
- the 40 ks chunk of S01 (Saclay short name S01_40) is numbered 2901
- *Fields in* `may11`*, although event files, catalogue files and image files are relevant to a specific 10ks chunk and may have a specific name, are not differentiated from full pointings. The shorter exposure is implicit and* **not** *reflected in field number or name.*

Fields flagged as bad (typically those with the - full - pn exposure under 7ks) are marked by a boolean flag column `badfield=1`. Such column name is for the physical tables. The `2XLSS`* catalogues use instead `Xbadfield=1`. *A field is marked bad according to its status in the full exposure, even for the case of 10ks chunks*

For `subaru` no pointings were actually bad, and the `badfield=1` was used to *flag the chunks or exposures which are not used for the catalogues*. The choice made at the time of `2XLSSe` and applied also to `2XLSSd` 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. *This convention applies both to* `2XLSSd` *and* `2XLSS`.

2.2. Optical, IR and other data

This section is virtually identical to what included in Report VI and VII, but is included for self-completeness. The non-X-ray tables were however updated with the sources in the surrounding of latest X-ray ingestions (in particular for *sources in* `may11` *which are either new or displaced*).

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.

Since Report VII, 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 may11 X-ray source positions, and can be accessed in correlation with the 2XLSS* or 2XLSSd* 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).*

A correlation with the 2XLSS* and 2XLSSd* 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, is available in the database as subaruspec..

2.3. Database technicalities

This section is virtually identical to what included in Report VI and VII, 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.jan11* or may11*). There are two *individual band* tables (e.g. jan11b and jan11cd 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. jan11) with the most relevant information. Band merging is described in section 2.3.5 of the XLSS paper, *except that now it is performed with a radius of 10″.*

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* described in Report VII (section 3.1) and views like the groups of four *virtual tables* 2XLSSd, 2XLSSBd, 2XLSSCDd, 2XLSSOPTd, or the equivalent for the 2XLSS* family, 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,*nor for the displaced position in* may11*,* but merely *applied* those already generated using the full exposures. *We report here most of the text included in Report VI and VII 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 on the website in page .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 2XLSS sources (89% for 2XLSSd) have both RA and Dec offsets lower than 4″, and 55% (57%) 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 78% within 2″ (93% and

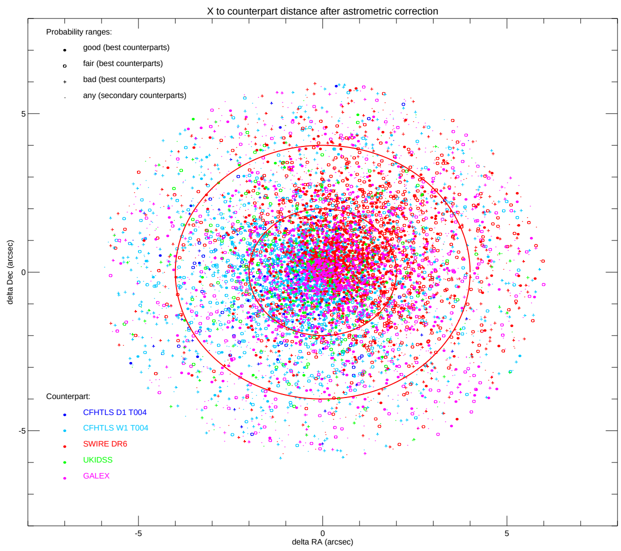

Fig. 1. Distances in RA and Dec between the X-ray corrected position and the counterpart position. Different symbols indicate the identification quality. A circle is plotted when the counterpart is the best one, and the chance probability is good or fair (filled in case of good probability). A cross 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. *This figure refers to* 2XLSS*; the equivalent figure for* 2XLSSd *is extremely similar.*

65% respectively including those with good or fair probability). For 2XLSSd the figures are 97% and 80%; 94% and 67%.

In terms of true distance 83% (84%) of the total is within 4″, which makes 90% (91%) of the good-or-fair associations (the circles in Fig. 1) and 95% (95%) 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. -0.40″,-0.07″for W1), while the one for SWIRE clusters around a point in the first quadrant (0.81″,0.58″). For 2XLSSd (-0.41″,-0.07″) and (0.80″,0.53″).

3. The procedure

The final procedure described in Report VI was applied to jan11 data (leading to the 2XLSSd* catalogues) and to may11 data for the new 2XLSS* catalogues.

3.1. Table concatenation

The concatenation into unions described in section 3.1 of Report VII is no longer applicable to the newer catalogues. Actually I originally tried to use unions concatenating single band jun09 and subaru data into jan11b and jan11cd, but later abandoned it in favour of a physical concatenation which is more efficient, database-wise (a factor 30 speed up for some queries).

3.2. Overlap removal

This section is adapted and amended from Report VI and VII. 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 10″(see discussion in section 6.1 of Report VII)*
- 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

A side effect of the increase to 10″ of the overlap removal radius (for consistency with the one used in band merging) is a slight increase in the number of merging ambiguities, which is discussed in 3.3.2.

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 2XLSSd* for fields 2001 (the really full exposure S01, about 80ks) and 2901 (the 40ks chunk which is preferred for uniformity with the other longer full exposures). *Note that* 2XLSSd *includes 59 sources (flagged* `Xbadfield=1` *in field 2001, i.e. detected only in the full exposure S01, and which one should prudentially exclude. The rest of the sources in S01 (ascribed to 2901 in* 2XLSSd*) are detected in both cases.*

2XLSS includes 124 pointings (inclusive of 7 SXDS pointings), of which 30 are flagged as bad fields. 2XLSSd includes respectively 125 and 31 because of the existence of S01 occurs in two quoted incarnations (40ks and full), flagged respectively good or bad.

The removal procedure for 2XLSSd removes 1563 entries, leaving 6723 sources in the GCT with a loss of 580 sources w.r.t. the old 2XLSSe at 6″). For the 2XLSS catalogue, based on the totally new detections in may11, the removed entries are 1199, leaving 5548 sources in the GCT.

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. Also there are differences between sources in the full exposure tables (jan11 or 2XLSSd catalogue) and in the 10ks ones (may11 or 2XLSS), which are discussed in 3.4.3.

3.3. The 2XLSS and 2XLSSd X-ray catalogues

This section is adapted and amended from Report VI and VII. For analogy with the published XLSS catalogue (see Table 11 of the XLSS paper) and previous working catalogues (see Report VI-VII) I provide *for each family* three virtual tables for the X-ray data: a merged catalogue 2XLSS or 2XLSSd and two single band ones 2XLSSB and 2XLSSCD or 2XLSSBd and 2XLSSCDd, analogous of

XLSSB and XLSSCD. The X-ray/optical tables 2XLSSOPT and 2XLSSOPTd 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 as page XLSS.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.
– only in the GCTs glorlss11 and gloropt11, i.e. those underlying 2XLSSd and 2XLSSOPTd, which can be compared directly with the old 6″ merging with which they share the input single-band detections, there is an *hidden column*, boolean upgradeflag, which assumes value 1 for the cases where the 2XLSSd source is considered now a two-band detection, while it was a soft-only or hard-only in 2XLSSe.
– *conversely only in the GCTs* glormay11 *and* gloroptmay11*, i.e. those underlying* 2XLSS *and* 2XLSSOPT*, there is an hidden column* deep*, which, when non-zero, points to the* Xseq *of the closest full-exposure (*2XLSSd*) source. For historical reasons* deep *is negative (changed of sign) in 35 cases in which the association is not well established (to be compared with 5079 regular associations and 435* 2XLSS *sources without* 2XLSSd *equivalent).*

The number of sources in the merged 10ks catalogue is 5548 for 2XLSS (4900 in 2XLSSB and 1912 in 2XLSSCD). The number of sources in the deep catalogue is 6723 for 2XLSSd (5882 in 2XLSSBd and 2647 in 2XLSSCDd).

3.3.1. Source naming

This section is adapted and amended from Report VI and VII. 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 (already since jun09) the raw input coordinates 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 2XLSSd and *a fortiori* 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, or respectively 2XLSSd Jhhmmss.s-ddmmss where coordinates are based on the corr set
– Pending registration with IAU of the 2XLSS and 2XLSSd prefixes 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 is also recommended for retired catalogues (which do implement it) in the rare case one may need to quote sources contained therein.
– 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 or 2XLSSd 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.
– there is presently no immediate way to cross-reference 2XLSS and 2XLSSd. However, when accessing 2XLSS it is possible to use the *hidden* column glormay11.deep (gloroptmay11.deep for 2XLSSOPT), which points to the value of the Xseq closest source in 2XLSSd as described above.

3.3.2. Ambiguous band merging

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 was discussed at the end of section 2.3.7 of the XLSS paper (column Xlink with the eventual addition of an a|b suffix to the catalogue name to disambiguate it).

There is only one a|b ambiguous couple in 2XLSSd and 3 in 2XLSS. All old ambiguous cases in XLSS are now present but unambiguous or (for one couple) no longer present in 2XLSSd or 2XLSS.

However a more detailed discussion of ambiguities is required now that the band merging radius is 10″. We remind that band merging combines the positions from pointlike and extended fits in both bands at the same time of the extended/pointlike classification.

First of all the distance between the selected soft and hard positions is kept in column maxdist in the physical tables (jan11, may11), and shall ordinarily be (well) below 10″. There are *only* 3 cases each in 2XLSSd and 2XLSS (one in common to both) where maxdist is (marginally) above 10″: these cases are flagged suspect=1 (same convention used also for previous catalogues) and correspond to *unambiguous* soft-hard associations where however the two associated detections are one pointlike and the other extended and therefore are "reclassified" according to the soft-band classification (so called PE reclassified pointlike and EP reclassified extended). Let us consider the most extreme case (an EP): for this source the coordinates in jan11 are stored like this (I give the case for RA, the same applies to declination): the chosen coordinates are the soft extended jan11.ra=jan11._rab=jan11b.ra_ext, the hard coordinates are taken from the reclassification jan._racd=jan11cd.ra_ext and the final distance maxdist is computed from _rab,_decb, _racd,_deccd which is 14″. However the band merging criterion was applied *before* the reclassification and therefore tested the distance jan11b.ra_ext,jan11b.dec_ext to jan11cd.ra_pnt,jan11cd.dec_pnt which is just 9.6″.

Now let us consider *real ambiguities* which occur when a soft detection is within 10″ from two different hard detections (case of identifiers sssaaa and sssbbb), or one hard w.r.t two soft (identifiers like ccchhh and dddhhh). These cases in the old catalogues (up to Report VII inclusive) were flagged suspect=2 and included but not exhausted the a|b cases described above. Of course these cases occur in couples in the physical table, but not necessarily in a catalogue (where one of the components, if spurious, may not enter). XLSS had 3 a|b couples, 1 other couple and 1 single suspect=2. The 6″-merged 2XLSSe had just one a|b couple and 1 other couple. Now, with the larger merging radius, one has more ambiguous cases (20 couples and 5 singles for 2XLSSd and 16 couples and 7 singles for 2XLSS).

However they are not all the same. It may happen that both maxdist in a couple are below 6″ (i.e. they would have been classified ambiguous even with the old criterion), or both are above 10″ (irremediably ambiguous), or one above and one below 6″. In the latter case we *divorced* the couple. The component with maxdist<6″ remains classified as a merged two-band detection (usually a PP), while the other component is reset to an only-hard or only-soft source. However, to keep track of the association, the following convention has been adopted: column suspect takes a positive value corresponding to the seq of the other component (in case of ambiguous couples, and for the two-band detection of a divorced couple), and a negative value corresponding to the seq *changed of sign* of the other component in case of the component reset as single band in a divorced couple. The resulting statistics is

- 2 couples for 2XLSSd and 5 couples for 2XLSS (all inclusive of the a|b cases) are really ambiguous with both maxdist<6″
- 2 couples for 2XLSSd and 2 for 2XLSS have both maxdist>6″
- the remainder of the couples are divorced
- in addition, since the divorced is applied at jan11 or may11 level, there are cases of singles in the catalogue which may have a positive (3 for 2XLSSd, 6 for 2XLSS) or negative (2 for 2XLSSd, 3 for 2XLSS) suspect.

In conclusion the number of ambiguities is limited and within reasonable tolerances.

3.4. Comparison between latest and with earlier releases

3.4.1. Comparison with physical tables

This section applies to the comparison of "raw" data, where by "raw" we mean here spurious and non-spurious sources, and before overlap removal and astrometric correction.

See the next sections 3.4.2 and 3.4.3 for a comparison with the published XLSS catalogue, the earlier releases and between 2XLSS and 2XLSSd.

The comparison of the "raw" data in jun09 with the combination of the earlier releases (nov06, jul07 and feb09), was reported in section 3.4.1 of Report VI and is not repeated here. Similarly the comparison of subaru 10ks chunks and 40ks chunks vs full exposure data was reported in correspondence exchanged with Saclay, whose references are given in section 3.4.1 of Report VII and are not repeated here.

3.4.1.1 Comparison of 6 to 10″ merging

There are no differences between jan11 and sdscombo (see Report VII) for what concerns source parameters like counts, likelihood or flux, because the *starting point* (the individual band tables) are the same, i.e. 10350 soft-band detections (7332 non-spurious) and 7121 hard-band ones (3232 non-spurious), with the same seq and id.

What is different is the *"numerology"* of merged sources, which were 14703 in sdscombo and are 14218 in jan11. The band merging procedure took special care to ensure that the same seq number was applied to sources resulting from the merging of the same soft and hard detections, or alike. Namely I originally got:

- 13707 cases (7824 non-spurious) *"preserved"* in the sense they are either unmerged (single band detection) or merged in the same way. They have all the same data, and extended/pointlike classification.
- 505 cases (457 non-spurious) *"upgraded"* i.e. they were previously detected in a single band, and are now the result of the merging of two detections in two bands within the new merging radius. They inherit the seq from the old source in the "best band".
- To these cases there is a correspondence of a *loss* of 492 old single-band detections.
- 6 cases were *newly numbered* because of ambiguous band merging (see previous section 3.3.2).

Condition		within 6″			6-10″	
		non-spurious	spurious		non-spuriuos	spurious
Soft band						
Common objects total	7639			210		
Extended confirmed		88			8	
Pointlike confirmed same spuriosity		5425	1416		60	55
`jan11` extended `may11` pointlike		16	10		7	9
`jan11` pointlike `may11` extended		16	11		10	5
`jan11` pointlike non-spurious			455			48
`jan11` pointlike spurious		202			8	
Only in `jan11` total	2501					
`jan11` extended not in `may11`		44				
`jan11` pointlike not in `may11`		1138	1321			
Only in `may11` total	644					
`may11` extended new, not in `jan11`		24				
`may11` pointlike new, not in `jan11`		140	480			
Hard band						
Common objects total	3801			182		
Extended confirmed		21			3	
Pointlike confirmed same spuriosity		2205	1713		24	99
`jan11` extended `may11` pointlike		2	11		1	4
`jan11` pointlike `may11` extended		3	5		3	4
`jan11` pointlike non-spurious			364			37
`jan11` pointlike spurious		135			7	
Only in `jan11` total	2674					
`jan11` extended not in `may11`		23				
`jan11` pointlike not in `may11`		731	1920			
Only in `may11` total	940					
`may11` extended new, not in `jan11`		17				
`may11` pointlike new, not in `jan11`		63	860			

Table 2. Comparison between detections in `jan11` and `may11` single band tables. The spurious/non-spurious condition in columns 2-5 refers to the `may11` table unless marked in blue when it refers to `jan11`.

After the application of the "*divorce*" procedure outlined in 3.3.2, all sources in `jan11` have the same `seq` as in `sdscombo`: 13726 are "preserved", and 492 are "upgraded". For the latter the following is the breakdown in the change of classification:

- 387 soft pointlike (P-) and 84 hard pointlike (-P) are now detected as such in both bands (PP)
- 7 soft pointlike are now PE (merged with nominal hard extended, still pointlike)
- 13 soft extended (E-) are now EP (merged with hard pointlike, considered extended)
- 1 soft E- is confirmed as such in both bands (EE)

3.4.1.2 Comparison of 10ks to full exposure

We make here a comparison between the raw single-band and band-merged physical tables (`jan11`* vs `may11`*) before overlap removal. For the comparison after such step (i.e. between the catalogues), see 3.4.3.2 below.

For the single band data we try to associate `jan11` and `may11` objects *in the same field* which are closer than 10″ (privileging those closer than 6″), and to compare the classification (extended vs pointlike) and "spuriosity" (detection likelihood below or above 15 for pointlike detections). The "numerology" is tabulated in Table 2. One clearly see that a majority of objects detected in the full exposures are confirmed, usually with the same classification and within 6″, in the 10 ks exposures, and that the differences are concentrated within the objects with poorer likelihood. However there is a significant number of detections, not necessarily spurious, which are either present only in the `jan11` full exposures (not surprising) or even only in the `may11` 10ks exposures. In particular 24% of the soft detections and 40% of the hard detections in `jan11` are not confirmed in `may11`, while 8% (soft) and 19% (hard) `may11` detections are new.

Coming to the band-merged tables (remember the merging was done for both `jan11` and `may11` at 10″), one associates sources *in the same field* using the coordinates corresponding to the classification (`ra_corr, dec_corr`). The "numerology" is tabulated in Table 3. We remind that `jan11` has 14218 sources (8286 non-spurious) and `may11` has only 11572 (6747 non-spurious). The common objects are 10007.

For extended object the *same classification* means objects are classified in the same way in both bands, so they

Condition	within 6″		6-10″	
	non-spurious	spurious	non-spuriuos	spurious
Common objects total	9683		324	
Extended confirmed same classification	96		7	
Extended confirmed compatible	9		1	
Extended confirmed different classes			1	
Extended jan11 pointlike may11	17	16	7	11
Extended may11 pointlike jan11	15	15	8	9
Pointlike confirmed same classification and spuriosity	5300	2706	56	128
idem confirmed and classification but degraded spuriosity		444		41
idem confirmed and classification but upgraded spuriosity	218		7	
Pointlike confirmed compatible class and same spuriosity	693	35	14	10
idem confirmed compatible but degraded spuriosity		93		21
idem confirmed compatible but upgraded spuriosity	6			
Pointlike confirmed incompatible class but same spuriosity	11	6		5
idem confirmed incompatible and degraded spuriosity		2		3
idem confirmed incompatible and upgraded spuriosity	1		1	
Only in jan11 total	4213			
jan11 extended not in may11	68			
jan11 pointlike not in may11	1354	2791		
Only in may11 total	1567			
may11 extended new, not in jan11	46			
may11 pointlike new, not in jan11	211	1310		

Table 3. Comparison between detections in jan11 and may11 band merged tables. The spurious/non-spurious condition in columns 2-5 refers to the may11 table unless marked in blue when it refers to jan11.

can be extended in both (EE), detected in both and extended in soft (EP), or detected as extended in a single band (E- and -E). 67% of the cases are soft-band only. A *compatible classification* means the object is extended in both jan11 and may11 in the (prevailing) band where is detected, and undetected (or pointlike in the hard band). The single case with *different classification* is hard extended in jan11 and soft extended in may11.

A limited number of objects changes from extended to pointlike (sometimes spurious) or viceversa.

For pointlike objects one has to consider both the classification and the "spuriosity". Degraded means non-spurious in jan11 but spurious in shallower may11. Upgraded means spurious in jan11 and unexpectedbly non-spurious in may11.

The *same classification* in both bands and both tables includes a majority of both band (PP) or soft-only (P-) detections for non-spurious objects (in both tables). For objects spurious in both, the same classification is equiparted between single-band soft and hard detections (only 29 out of 2706 PP). Upgraded or degraded cases prevail in the soft band.

Compatible classification includes cases which are PP (very very rarely PE) in one table and single band detection in the other. The majority of cases are jan11 PP demoted to single.

Different classification means usually soft-only vs hard-only, but also a limited number of PP vs PE or v.v. Anyhow the total number of different classification cases is rather limited with respect to the rest.

29% of the merged detections in jan11 are not confirmed in may11 (mainly single soft non-spurious, or equiparted between hard and soft for spurious), while 14%of the may11 detections are new (the majority are

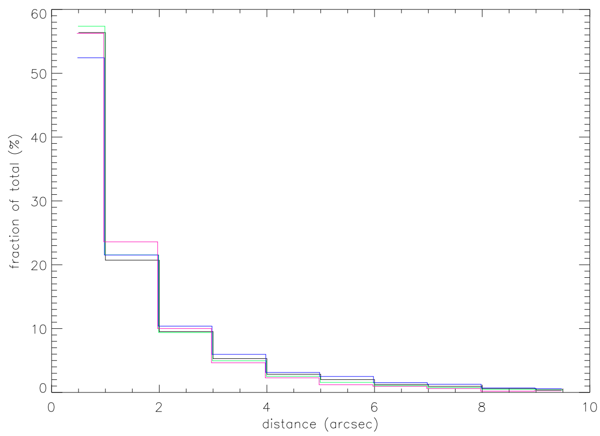

Fig. 2. Histogram of the distances between the positions of associated sources in the 10ks and full exposure catalogues. All histograms are normalized as percentages of the total number of entries in the respective dataset. Black colour for comparison between band merged tables (jan11 vs may11). Green and blue for individual band tables (green for soft band and blue for hard band). Magenta for comparison between catalogues (2XLSSd vs 2XLSS).

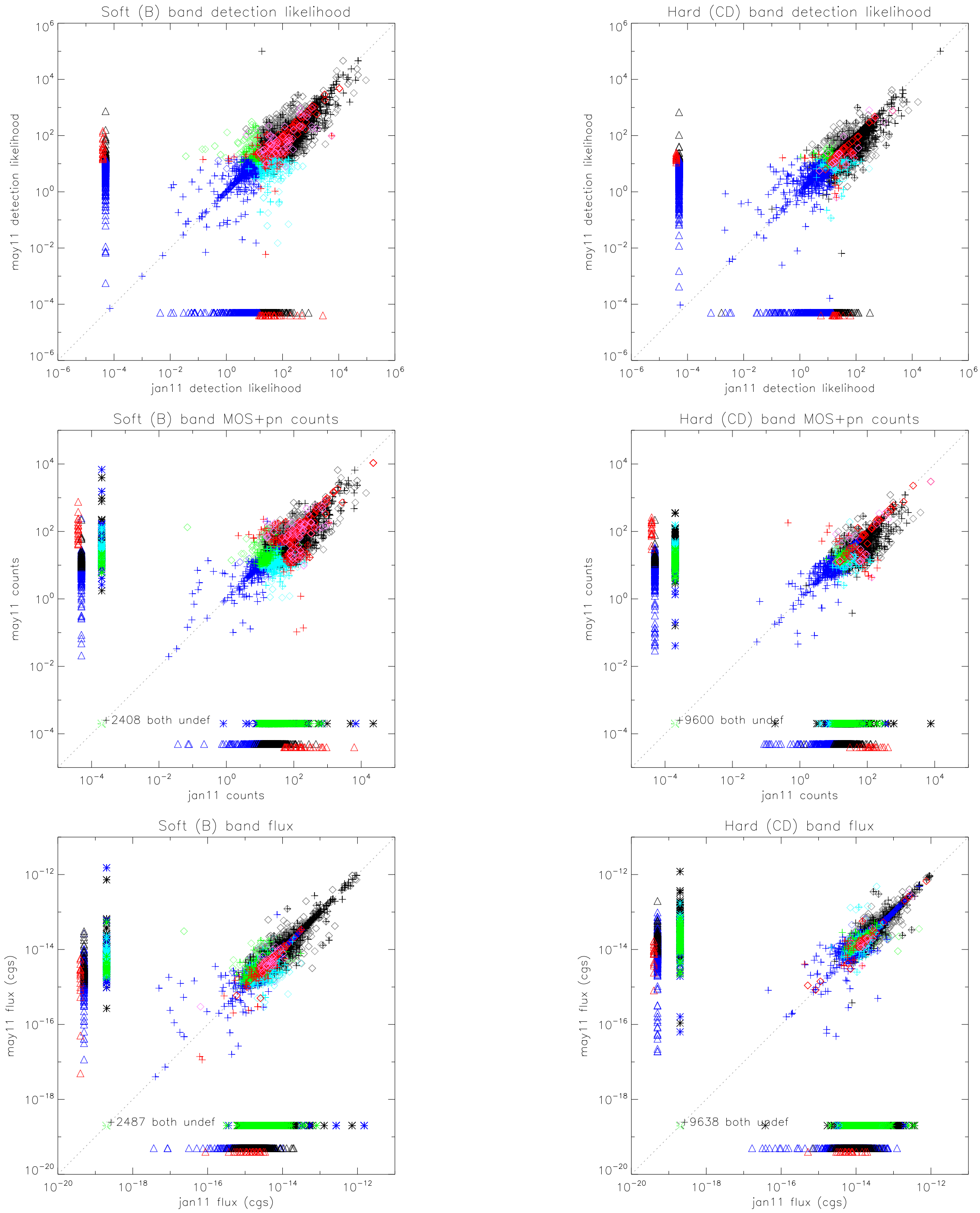

Fig. 3. Comparison of the detection likelihood (top row), of the number of MOS+pn counts (middle row) and of the flux (bottom row) in the soft (left column) and hard (right column) energy bands. between `jan11` and `may11` physical tables. Crosses indicate an *unambiguous association* (single or best) 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 `may11` and `jan11`; cyan objects demoted from non-spurious to spurious in `may11`; 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 `jan11` or `may11` extended sources classified pointlike non-spurious in the other dataset; thin red crosses are `jan11` or `may11` extended sources classified pointlike spurious, and pink diamonds are ambiguous associations to extended objects.

Condition	within 6″ non-spurious	6-10″ non-spuriuos
Common objects total	4980	103
Extended confirmed same classification	86	8
Extended confirmed compatible	8	1
Extended confirmed different classes		1
Extended 2XLSSd pointlike 2XLSS	18	7
Extended 2XLSS pointlike 2XLSSd	15	5
Pointlike confirmed same classification	4243	59
Pointlike confirmed compatible	603	20
Pointlike confirmed incompatible class	9	2
Only in 2XLSSd total	1640	
2XLSSd extended not in 2XLSS	84	
2XLSSd pointlike not in 2XLSS	1556	
Only in 2XLSS total	465	
2XLSS extended new, not in 2XLSSd	68	
2XLSS pointlike new, not in 2XLSSd	397	

Table 4. Comparison between sources in 2XLSSd and 2XLSS catalogues.

spurious, but mainly single detections prevail, in the soft band when non-spurious).

Fig. 2 shows the distribution of distances between jan11 and may11 common objects, both for individual bands and band merged data (this section) as well as for the catalogues (see 3.4.3.2). The majority of distances are extremely well compatible (for the band merged tables 86 within 2″, 95% within 4″ and 98% within 6″).

Comparison of detection likelihoods, number of counts and fluxes for sources associated within 10″ in band-merged tables are plotted in Fig. 3. This figure can be *directly* compared with Fig. 2 of Report VI i.e. includes all possible associations *irrespective of field* preferring the closest one in astrometrically-corrected coordinates. One notes a large scatter, particularly for likelihood and counts. Also both the latter quantities are often lower (below the fiducial diagonal line) in may11 than in jan11 because of the shorter exposure, while fluxes are recovered to similar values.

In lack of error bars, and considered the dubious nature of the association, it is not immediate to compare the fluxes in quantitative manner. However one can say that, when the flux in a band is defined for both jan11 and may11 associated sources, the percentage of cases having a flux within 10, 20 or 50% from each other are 57%, 78% and 95% for the soft band and 53%, 74% and 92% for the hard band, taking into account all pointlike detections (in both releases). Limiting to those which are non-spurious in jan11, the figures are 58%, 80%, 97% (soft) and 53%, 77%, 95% (hard), and considering those non-spurious in both 60%, 82%, 97% (soft) and 55%, 80%, 96% (hard).

3.4.2. Comparison with XLSS

The comparison between XLSS and an older provisional 2XLSS at 6″ was reported in section 3.4.2 of Report VI. The reliability flag tabulated in Report VI is available for 2XLSSd and 2XLSS (with *identical* content) in the GCTs as the hidden column glorlssmay11.reliable or glorlss11.reliable.

A comparison of the likelihood and flux between 2XLSSd and XLSS is almost identical (except for the few missing objects due to the new 10″ merging) to the one reported in Fig. 3 of Report VI. For 2XLSS vs XLSS the results are rather similar with a larger scatter, and with lower likelihoods as expected, as can be inferred from the comparisons between the present releases reported in 3.4.3.2.

The relevant figures are reported in an Appendix provided as a separate document.

3.4.3. 2XLSS vs 2XLSSd and earlier releases

The comparison of old 2XLSSe with even earlier releases was reported in section 3.4.3 of Report VII, to which the reader is referred, particularly for the effect of overlaps with the SXDS fields.

We compare here only the result of the latest 10″ processing on full exposures (2XLSSd) with the previous 6″ processing of the same data, and also compare both 10″ processings on full and 10ks exposures (2XLSSd vs 2XLSS).

3.4.3.1 Comparison with earlier releases

2XLSSd contains 6723 sources instead of 7083 as in 2XLSSe. Once again by construction all of them have the same Xseq and coordinates in the new 10″ and in the old 6″ catalogue, and the same values for rates and flux when defined. The *"numerology"* is the following:

- 6394 sources are *"preserved"* (2164 were and are detected in both bands, the rest were and are detected in a single band).
- 329 cases are *"upgrades"* (see 3.4.1.1) which survive the overlap removal at 10″. They are flagged by the hidden column `upgradeflag`, described in 3.3.
- 360 2XLSSe sources are *lost*, i.e. no longer present in 2XLSSd because of the combined effect of band merging and overlap removal at 10″. All but 33 are single-band detections (i.e. disappeared because of new band merging). Of the latter, 20 are detected at more than 3σ (13 at more than 4σ) in at least one band and only 3 at more than 4σ in both bands.

3.4.3.2 Comparison of 2XLSSd *with* 2XLSS

While sources in a *catalogue* are by construction non-spurious, they can be detected as such in both bands, detected as non-spurious in one band and spurious in the other, or detected in a single band. The breakdown in percentage is reported here below. The deep catalogue is marginally better for what concerns full-fledged both band detections.

Case	2XLSSd	2XLSS
Total number of sources	6723	5548
non-spurious in both bands	27%	23%
non-spurious in soft band only	8%	8%
non-spurious in hard band only	2%	1%
Detected only in soft band	52%	57%
Detected only in hard band	11%	10%

Having said that, a "numerology" similar to the one presented in 3.4.1.2 can be tabulated in table 4, considering however that all sources in a *catalogue* are nominally non-spurious, and that the association by distance does *not* necessarily imply sources are in the same field.

Fig. 2 shows the distribution of distances between common objects, together with the physical table data described in 3.4.1.2. The distances for the catalogues are in even better agreement than for the other tables: 90% within 2″, 97% within 4″, and 99% within 6″.

A comparison of likelihoods and fluxes for sources associated via the hidden column `deep` described in 3.3. is reported in Fig. 4. As expected the likelihood in the shorter exposure 2XLSS catalogue is compatible but lower than the one in 2XLSSd (the points lay below the diagonal fiducial line of equal values).

For fluxes they are generally rather well consistent (with exceptions for a few extended sources), with only a moderate scatter for fainter objects. In lack of error bars, one can compare the compatibility of fluxes for the 5113 sources associated between 2XLSSd and 2XLSS. For the 4629 with a soft-band detection in both catalogues, 60%, 81% and 96% of the sources have fluxes within 10%, 20% or 50%. The equivalent percentages for the 2127 with an hard-band detection are 54%, 78% and 95%.

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 two variants (2XLSSOPT and 2XLSSOPTd) are mimicked on the XLSSOPT table described in the XLSS paper, but provide information on the (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 and VII, 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 and VII.

- a preliminary step is to create a GCT and initialize it. The member tables of such GCT are the three X-ray tables (`jan11`, `jan11b`, `jan11cd`) or (`may11`, `may11b`, `may11cd`) used respectively for 2XLSSd or 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 the corresponding X-ray-only catalogue 2XLSSd or 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

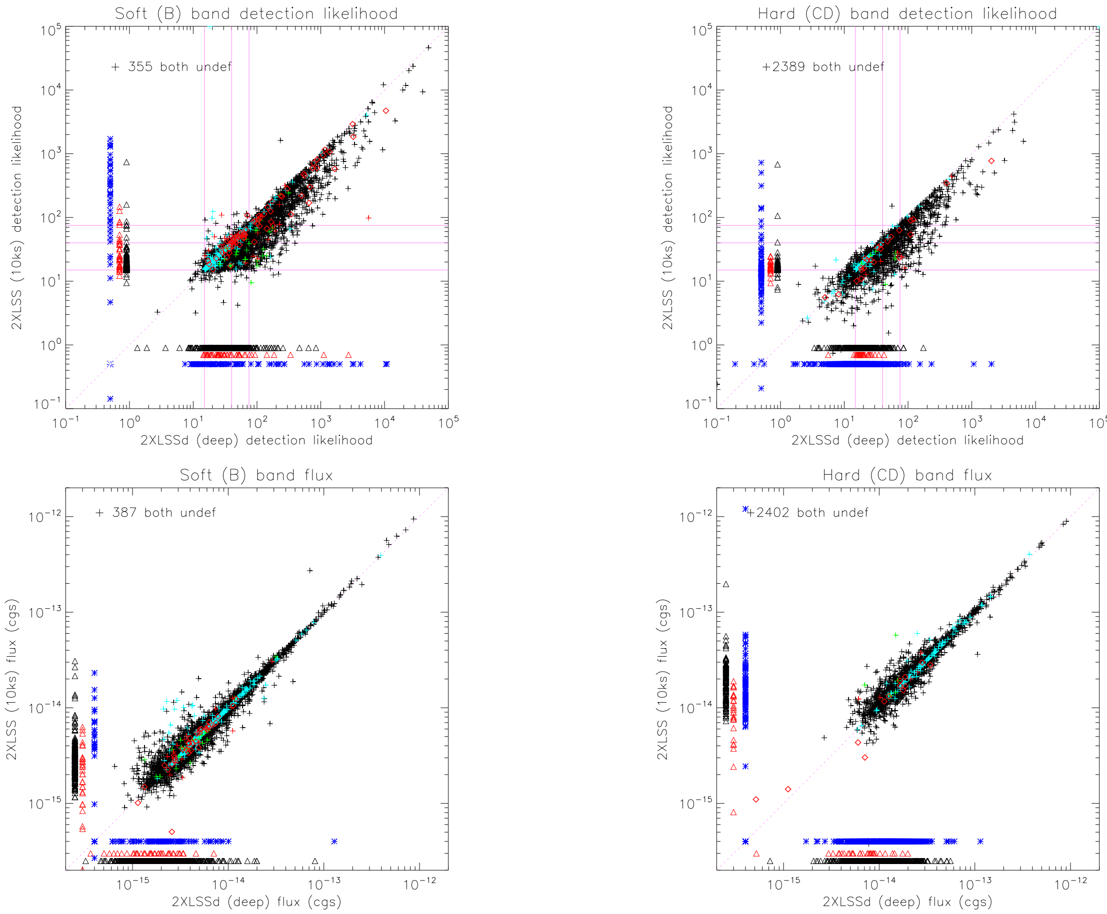

Fig. 4. Comparison of the detection likelihood (top row) and of the flux (bottom row) in the soft (left column) and hard (right column) energy bands. between 2XLSS and 2XLSSd. Crosses and diamonds indicate pointlike or extended objects associated in the two catalogues (see text). Blue asterisks indicate 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 2XLSSd extended object pointlike in 2XLSS; viceversa for red cross; red diamond for extended sources in both 2XLSS and 2XLSSd. Triangles are black or red for pointlike or extended sources which are either new in 2XLSS or present in 2XLSSd but lost in the shallower 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.

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

- *There is a peculiarity for what* 2XLSSOPTd *is concerned. The procedure to populate* 2XLSSOPT *required a preliminary update to the other waveband tables (*d1t4, w1t4, swiredr6, ukidssdr5, galex*) for objects in the surrounding of the new or displaced* may11 *X-ray positions. If everything had been done correctly before, this should not have added any new objects in the surrounding of* jan11 *positions. This was indeed the case for all catalogues, but* d1t4*, for which 109 objects in the surroundings of 33 X-ray sources had been missed somehow. They were recovered a posteriori (i.e. after the ranking procedure described in 3.6.1) and included 59 cases of new "D1 only" counterpart sets, 3 of which replaced former "blank fields", and 50 cases in which the missed D1 counterpart had to be inserted into an existing (usually at least W1) counterpart set. These objects are flagged by a special value of the hidden column* gloropt11.upgradeflag=10*. It shall be noted that most of the changes affected counterpart sets which are anyhow rejected, and that ranking was affected in a relevant way only for less than a handful of sources (list in circular mail to database users dated 28 Aug 2011)*

3.5.2. The 2XLSSOPT* tables

This section has been only minimally amended w.r.t. what included in Report VI and VII. 2XLSSOPT and 2XLSSOPTd loosely mimic XLSSOPT as described in Table 10 of the XLSS paper, but provides a number of additional columns (see page 2XLSS.html or the main database interface for details). For hidden columns in the GCT see section 3.3 above. 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 and 2XLSSOPTd contains respectively 16722 and 20849 counterpart sets, which on average means that an X-ray source has 3 possible optical or IR *not validated* associations within 6″. De facto 46% (44%) of the X-ray sources have from none to two possible counterparts, and only 18% (19%) more than 4. *Note that the above figures are affected by some manual editing (described at the end of 3.6.2).*

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 an older preliminary release. 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 and VII since the same computation was used for 2XLSS *and* 2XLSSd.

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 5. 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 in Report VI and VII.

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. 5 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. 5 top left panel, however only the

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 5. Parameters used for probability computation. *This table unchanged since reports VI-VII*

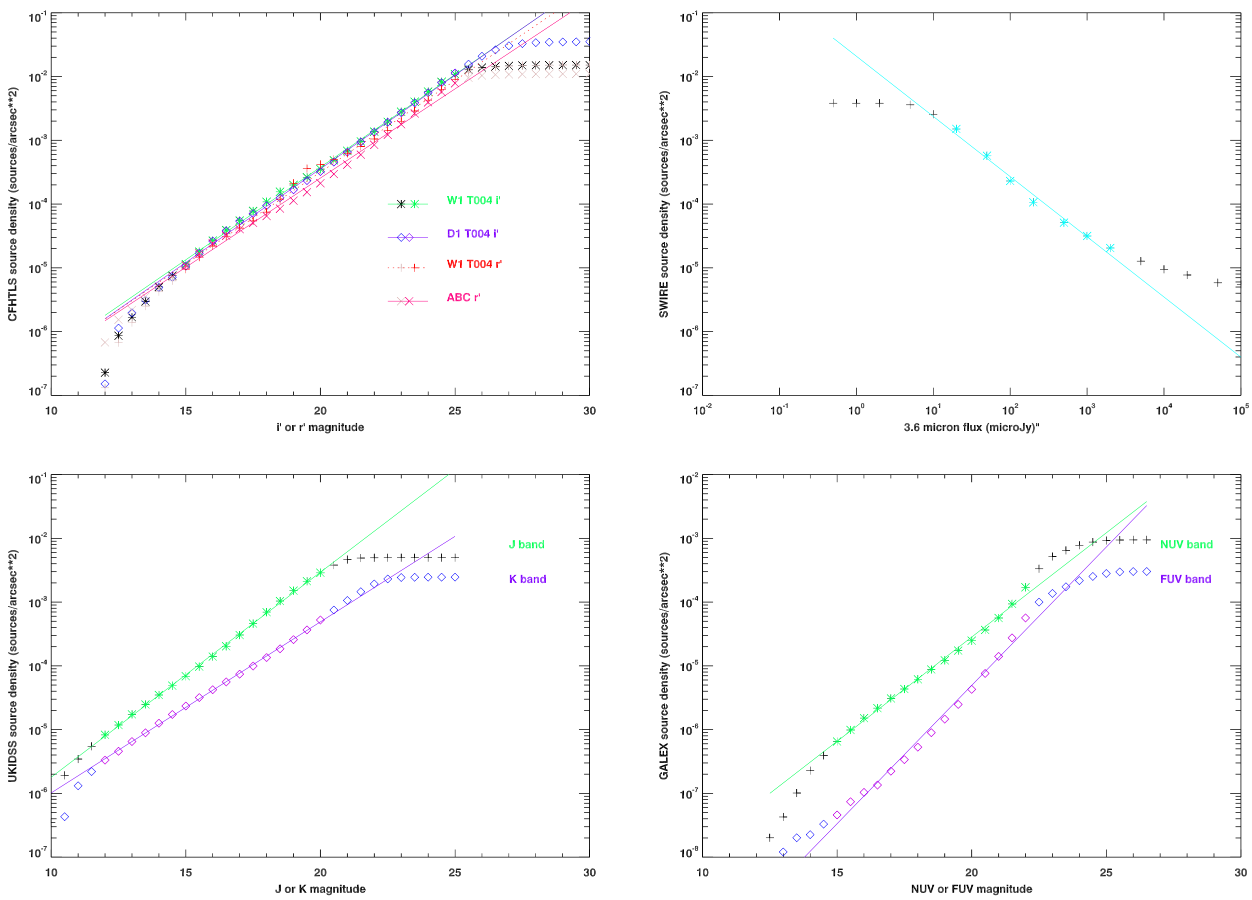

Fig. 5. Source count density for the CFHTLS D1 (asterisks) and W1 (diamonds) fields i' band, as well as for the W1 (crosses) and ABC (X) fields r' band (top left panel) ; for SWIRE DR6 at $3.6\mu m$(aperture 2) fluxes (top right panel); for UKIDSS J (crosses) and K band (diamonds) (bottom left panel); 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 5 are shown in (lighter) colour. Note the GALEX Y-axis is displaced by one decade. *This figure unchanged since reports VI-VII*

fit for the ABC fields is reported in Table 5 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

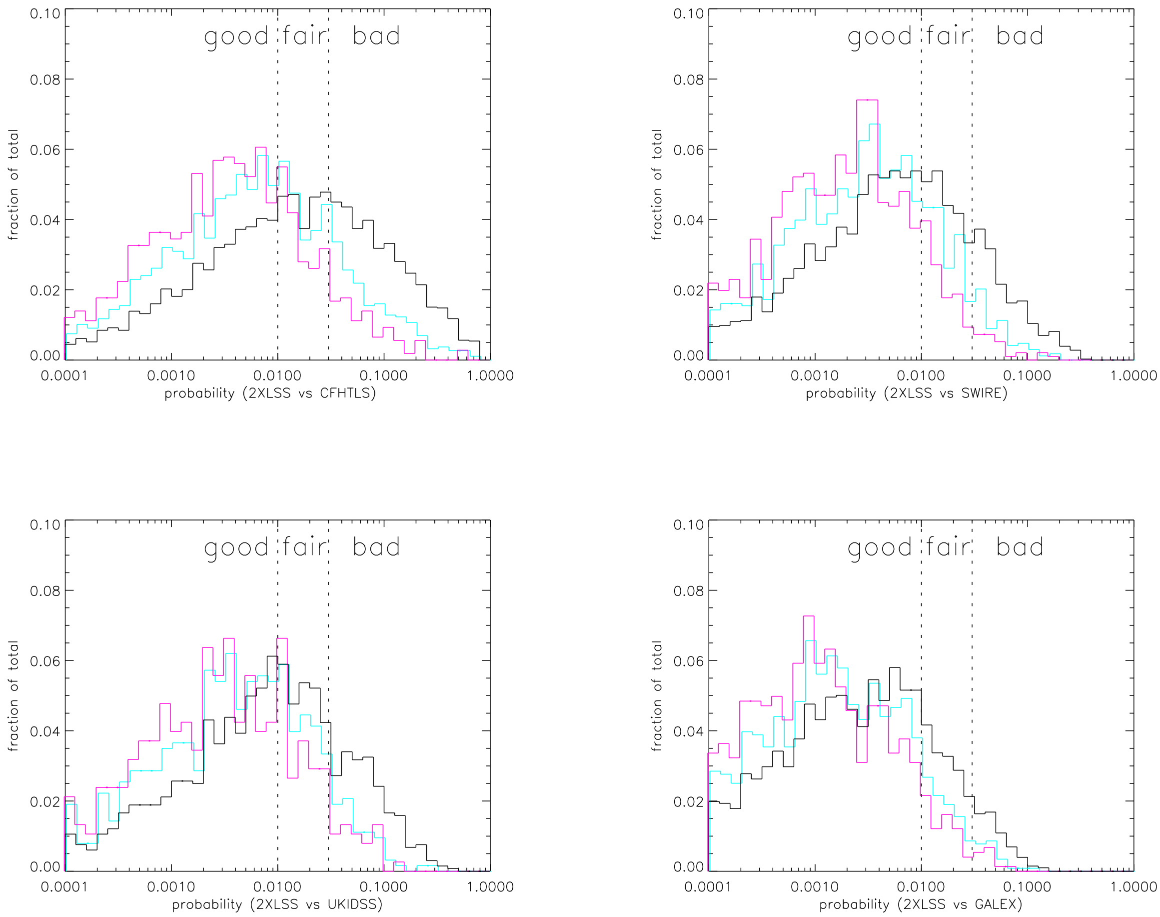

Fig. 6. Histograms of the four *uncapped* probabilities ($probXO$, $probXS$, $probXU$ and $probXG$) normalized to the total number of best counterparts with not undefined probability in the total sample (black), with a detection likelihood of at least 40 (3σ) in the best band (cyan), or of at least 75 (4σ, magenta). The dashed fiducial lines identify the loci with good, fair, or bad probability. *This figure refers to* `2XLSS`*; the equivalent figure for* `2XLSSd` *is extremely similar.*

was not possible for data retrieval for the lack of the so-called "xpf" files) using aperture 2 fluxes; see Fig. 5 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. 5 bottom left panel.

The density of GALEX sources was derived from the GR4 release using MAST CasJobs in the sky region $30^\circ \leq \alpha \leq 40^\circ$ and $-10^\circ \leq \delta \leq 0^\circ$: mode: see Fig. 5 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. 6).

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 $autorank = 4$ $rank = 1$ Note that such definition of "blank field" can be affected by artifacts, e.g. a very bright counterpart sometimes does not show up in any catalogue and this result in a fake blank field. Also the presence or absence of a "near" counterpart can be affected by displacements between the `2XLSS` and `2XLSSd` positions (see 5.3), resulting in a source being "blank" or not blank in either catalogue.
- 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 source 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 usually inspected. They are very few (29 of 3064 unambiguous, 77 of 1584 ambiguous for 2XLSSOPT and respectively 24 of 4749 and 81 of 1974 for 2XLSSOPTd). Also the high score not selected cases may be 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 both 2XLSSOPTd and 2XLSSOPT as it had been applied to all the preliminary catalogues described in section 3.6 of Report VII. As described in Report VII, as a result of the visual inspection on such preliminary tables 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 *at the time of Report VII on its relevant table* 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).

The edits applied to such first preliminary table were then propagated to the extended catalogues described in Report VII whenever applicable. About a dozen of other similar edits, resulting from correspondence with O.Melnyk in July 2010, were instead applied only to 2XLSSOPTe, as also said in Report VII.

The edits of 2XLSSOPTe and the relevant comments could therefore be safely propagated to 2XLSSOPTd. In particular the comments are all the same (actually, softlinked) since the seq number is the same.

A selection of such edits was applied (as applicable) also to 2XLSSOPT but so far it is not recorded in comments.

3.7. Data Products

This section is to be regarded as preliminary and subject to change without notice.

Currently the *X-ray data products* associated to the 2XLSSd* X-ray tables, are the same associated to the jan11 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. 2XLSSBd) are associated only to data products in the relevant energy band. *There are presently no accessible data products associated to the 2XLSS** (or may11 with the exception given below) tables, although the relevant files have been supplied by Saclay. They are considered uninteresting (10ks exposure) w.r.t. the deeper data supplied with 2XLSSd*, but *if they are requested, I could consider to put them online*.

The jan11 (and may11) tables alone are associated also with the original XAMIN FITS catalogues.

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* and 2XLSSd* 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 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 2XLSSOPTd (so far only for it) 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 were retrieved already at the time of Report VII and are presently associated only to 2XLSSd and 2XLSSOPTd. Thumbnails and comments are *physically the same* which were used for 2XLSSe (since the same X-ray sources with the same sequence number and position are present in jan11 vs jun09+subaru). In the case of 2XLSS source position and sequence numbers are different, and presently *there are no thumbnails or comments*. It would be possible to soft-link the existing thumbnails for sources common to jan11 and may11 (at the price of a possible displacement), but one will have to retrieve afresh those for sources which exist only in may11. *The issue of thumbnails and comments for 2XLSS is likely to be re-evaluated in the future.*

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 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 the figures included in the Appendix provided as a separate document).

4. Identification support tools

As anticipated in Report VI, I created at the time of Report VII 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 could originally support only one catalogue at a time, with changes done in the java and HTML code, and requiring recompilation. *I recently modified the java code to read the name of the GCT associated to a given catalogue from the login HTML form, so one can now support more catalogues, provided a menu to choose them is written in the login HTML form, and separate support HTML pages are written for each GCT depending on their format and content.* Currently both 2XLSSOPT and 2XLSSOPTd (and only them) are 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 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, according to ranks and probabilities.

I have presently several production variants derived adjusting the demo prototype: "combo26" and "combo26b" are described in Report VII, while *the newer* "combo26d" works on 2XLSSOPTd. All variants support i' or g' images, as well as gzipped FITS images for all bands.

Since *presently there are no thumbnails for* 2XLSS there is no variant with direct support to 2XLSSOPT counterparts. However I have a temporary variant "combo26dx" which uses thumbnails and counterpart sets from 2XLSSOPTd but allows to superimpose X-ray positions from may11 i.e. 2XLSS and verify the effect of the displacement of X-ray positions onto the counterparts (see 5.3).

5. Catalogue statistics

We report here results for 2XLSS *and* 2XLSSd*, but none of the earlier (abandoned) versions. One may refer to section 5 in Report VII for a comparison.*

5.1. The X-ray catalogues

The 10ks 2XLSS table contains a total of 5548 X-ray sources, of which 1807 are detected in both energy bands, 3170 only in the soft band, and 571 only in the hard band. The corresponding figures for 2XLSSd are 6723 total, 2493 in both bands, 3509 only soft and 721 only hard.

In 2XLSS of a total of 190 extended sources (52 C1 and 138 C2), there are 12 extended sources classified C1, and 14 classified C2 detected in both bands (of these only 4 C1 are detected as *extended in both bands*); there are 35 extended sources classified C1, and 86 classified C2 detected only in the soft band; there are 5 extended sources nominally classified C1, and 38 classified C2 detected only in the hard band. The corresponding figures for 2XLSSd are 211 total extended, 57 and 154 C1 and C2, 15 and 17 C1 and C2 detected in both bands (again only 4 C1 detected as *extended in both bands*), 38 and 92 soft C1 and C2, and 4 and 45 nominal hard C1 and C2.

For 2XLSS the number of pointlike sources (5358 total) is 1781 (99%) detected in both bands, 3049 (96%) in the soft band and 528 (93%) in the hard band. The corresponding figures for 2XLSSd are 6512 total, 2461 in both bands, 3379 soft and 672 hard.

For the pointlike sources in 2XLSS, 56% 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 77% above likelihood 40 (3σ level). For 89% of the sources the best band (highest detection likelihood) is the soft band. The corresponding figures for 2XLSSd are 58% above 4σ, 79% above 3σ, and 87% best in soft band

Finally for 2XLSS, for the detections only in the soft band, only 11% are above 4σ, and 30% above 3σ. In the hard band the percentages are 2% above 4σ, and 7% above 3σ. The corresponding figures for 2XLSSd are 11% above 4σ, 32% above 3σ in the soft and 2% above 4σ, 9% 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 the breakdown between the sources flagged to be in good or bad fields (see discussion in 3.2) for 2XLSSd and 2XLSS. If one wants to cover all the XLSS area without holes, but excluding dubious sources in bad fields, excepted the AO7 repeats, one should take the sources *not* indicated in parenthesis (6441 or 5309 in total).

	Xbadfield		
Field type	0 good	1 bad	
"our" B and G fields	5573	(331)	of which
in AO7 repeats		100	
SXDS fields	768	(64)	(in S01)
"our" B and G fields	4824	(340)	of which
in AO7 repeats		101	
SXDS fields	384	none	

5.2. The joint X-ray/optical catalogue

2XLSSOPT contains nominally 16722 counterpart sets, and 2XLSSOPTd 20849.

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*. 5 figures which give the sky areas covered by the various surveys used by us overplotted on the footprint of the FoV of our fields are included in an auxiliary Appendix provided as a separate document (they aren't that different from those reported in Report VII, so we make them available offline to make the present report shorter). Two complementary figures in the Appendix 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.

X-ray sources nominally flagged as *blank fields* (i.e. having no catalogued CFHTLS, SWIRE, UKIDSS or GALEX counterpart within 6″) are 223 in 2XLSSOPT and 250 in 2XLSSOPTd. *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 43302=27601, which is very close to a R=15.6 galaxy shown in SIMBAD, or 38678 (without immediate counterpart in XLSSOPTd) whose field is spoiled by nearby bright star BD-05 427. So some of the cases flagged as blank field can instead have a bright counterpart.

986 (or respectively 1133) 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:

- 2189 (2646) sources have a *single very reliable counterpart*, i.e. rank 0 and autorank in the range 10 to 13.
- 1552 (1854) 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

- 870 (1064) 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.
- 714 (910) 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 223 (250) tentative *blank fields* described above have rank 1 and autorank 4.

I have attempted a rough characterization using the criteria defined in 3.6.1. The results are summarized in Table 6 which has to be interpreted as follows:

Looking at the row "best and single", 33 (3%) of the physically single counterparts are detected in all four optical/IR/UV catalogues with *good* probability in all of them. 177 (18%) 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. 126 (13%) are similarly detected in 2 of the 4 catalogues with a good probability in all the catalogues where they are detected. 123 (12%) are detected in only one catalogue with a good probability. 28 (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 (9, 51, 53, 47 and 72). Finally 268, 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 384 of the best *non-single* counterparts are detected in all 4 catalogues with a good probability in all of them, etc. etc. up to 865 cases which, despite being the best counterpart, are detected always with a bad probability.

Considering the secondary counterparts, 9039 (representing about 90% of the "all bad") are always bad and could surely be rejected. To be precise, for 2XLSSOPTd one of the 11497 secondaries has all undefined probabilities, because it is a single 160 μm SWIRE source. There are however e.g. 23 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, 47.9% of the sources have a best counterpart with a good probability, 28.7% a fair one, and 4.0% are nominal blank fields (for 2XLSSOPT while for 2XLSSOPTd the figures are 48.5, 30.0, 3.7%.

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).

40% of the best good counterparts are detected above 4σ; 17% of the fair ones; 4% of the bad ones and 11% of the blank fields. Or conversely, of the 1404 X-ray sources above 4σ, 76% have a good counterpart, 20% a fair one, 3% a bad one and 2% are unidentified. The equivalent figures for 2XLSSOPTd are 44, 19, 5 and 11% above 4σ in each category, or 75% good, 20% fair, 3% bad and 1% blank of 1889 above 4σ.

Similarly at 3σ 61% of the best good counterparts are detected above such level; 37% of the fair ones; 13% of the bad ones and 24% of the blank fields. Or conversely, of the 2425 X-ray sources above 3σ, 67% have a good counterpart, 25% a fair one, 6% a bad one and 2% are unidentified. For 2XLSSOPTd the equivalent figures are 63, 42, 16 amd 24%; and 65, 27, 6 and 2% of 3176.

Fig.6 gives the distribution of the probabilities in their three ranges. This figure shall be compared with the equivalent figures of Reports VI and VII and also 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 all 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 already at the time of Report VII) 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 given in detail in Report VII (the figures for 2XLSS and 2XLSSd are not terribly different).

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. The differences are discussed in the next section.

5.3. Differences between 2XLSSOPT *and* 2XLSSOPTd

The differences between the catalogues with optical identifications derive from three main reasons, the former two physiological to the different exposures:

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	223											
Best and single		33	177	126	123	28	9	51	53	47	72	268
Best		264	659	493	509	248	133	299	303	240	439	751
Secondary		23	75	93	648	128	24	81	156	484	423	9039
Blank field	250											
Best and single		46	223	131	134	34	10	47	51	64	71	322
Best		384	832	609	570	300	139	420	352	275	594	865
Secondary		26	90	109	786	142	25	82	192	605	572	11497

Table 6. Basic statistics ot the 2XLSSOPT and 2XLSSOPTd catalogues

– the X-ray source may be detected in one of the input catalogues and not in the other (jan11 vs may11 or v.v.); this is discussed in 3.4.1.2 (and also 3.4.1.1)
  We remind that out of 14218 jan11 sources 26% have no may11 counterpart; out of 8286 *non-spurious* jan11 sources 13% have no may11 counterpart; conversely out of 11572 may11 sources 12% have no jan11 counterpart; and out of 6747 *non-spurious* may11 sources only 3% have no jan11 counterpart.
– the X-ray source may be detected or classified differently (spurious or non-spurious, in one or two bands, pointlike or extended); this is discussed in 3.4.3.2
  We remind that the sources in common between 2XLSSd (total 6723) and 2XLSS (total 5548) are 5113 (92% of 2XLSS).
– the X-ray source can be detected at a displaced position

The latter displacement may result in some of the possible counterparts be outside the 6″ correlation radius, and therefore in the list of counterpart sets being partially or totally different, and in different ranks.

There are 20849 possible counterpart sets for 6723 X-ray sources in 2XLSSd, and 16722 for 5548 sources in 2XLSS (see 3.5.2). Of these, 1203 entries (for 435 X-ray sources) in 2XLSS have no obvious correspondent in the deep 2XLSSd (the hidden column pointer deep=0).

On the other hand 15519 counterpart sets are associated to 5113 X-ray sources in 2XLSS with a corresponding source in 2XLSSd. These are the X-ray sources present in both catalogues.

13365 counterpart sets (86% of the common ones) for 4945 2XLSS sources (97% of the common ones) are identical (i.e. have the same counterparts in all non-X-ray catalogues). Of these 139 are confirmed "blank fields" (no *catalogued* counterpart in any waveband). Of the remaining 13326, 11474 (87%) have the same rank in 2XLSSOPTd and 2XLSSOPT (of these ranks 3687 are primary counterparts, 1430 are secondary counterparts, and 6357 are rejected counterparts, with reference to the ranks defined in 3.6.2).

1752 (of the 13326) have the same counterpart but with a different rank. For 502 of them the rank change is irrelevant (swap rank 0 and 1 or v.v., i.e. they remain anyhow the primary counterpart in both catalogues). 201 and 173 counterpart sets rejected in 2XLSSOPTd are respectively primary and secondary choice in 2XLSSOPT. 174 and 197 primary or secondary in 2XLSSOPTd are rejected in 2XLSSOPT. 256 2XLSSOPTd primaries become secondary in 2XLSSOPTd while 249 undergo the opposite change from secondary to primary.

There are additional 37 cases (only 0.2% of the common ones) where the counterpart set is the same between a couple of 2XLSS and 2XLSSd entries, but deep does not point to the expected 2XLSSd source. I.e. two sources have the same counterpart set but are not the closest. In 10 cases this is due to ambiguous band merging, but in the rest (which is anyhow a very very small number) this probably means that there are two may11 sources both displaced from but close to a given jan11 one (or v.v. two jan11 from one may11). For instance 2XLSS.Xseq=51506 is associated (closest) to 2XLSSd.Xseq=19311 at 3″ but is actually located midway between it and 2XLSSd.Xseq=19224 at 9″, and it shares with the latter a (rejected) W1/UKIDSS counterpart set which in turn is midway between 51506 and 19224 !

Let us now consider the 2154 remaining 2XLSSOPT counterpart sets (covering 1472 individual 2XLSS X-ray sources). They can be *altogether different* from all counterpart sets in 2XLSSOPTd for the associated X-ray source (the X-ray source position moved so much that entire counterpart sets are farther than 6″ from either position), or may partially match.

One might have a count of the partial matches (from 1 to 4 out of the 5 D1, W1, SWIRE, UKIDSS or GALEX catalogues). This is a breakdown.

Case	occurrences
no matches	2134
1 match	221
2 matches	65
3 matches	24
4 matches	7

The total is larger than 2154 (the number of distinct 2XLSSOPT counterpart sets) because one specific counterpart set is compared with all the 2XLSSOPTd counterpart sets for the corresponding X-ray source. A different view (totalling 2154) can be given by the alternative breakdown:

Case	occurrences
single no match	260
only multiple no matches	1600
one counterpart set with 1-3 matches	15
one no-match and one 1-3 matches	47
two mixed matches	5
more no-match and mixed matches	227

I did a very quick check of all possible matches, and I have a list available offline, with classification codes like "nlg" (No Longer GALEX), "nlsu" (No Longer SWIRE and UKIDSS" or "nas" (Now Also SWIRE), which make reference to the case a counterpart in a particular catalogue is no longer present in 2XLSSOPT (while it was in 2XLSSOPTd), or is now present while it wasn't in 2XLSSOPTd. The figures given above (and the *examples described below*) are relevant to the cases still present in the released catalogues (the exercise was repeated twice, in one case as preliminary to the edits described at the end of 3.6.2), in a second one on the released catalogue).

A different approach fro the comparison is to consider only the *best counterparts* i.e. those ranked 0-1 (by definition one per X-ray source). Let us exclude the 1642 2XLSSOPTd (24% of 6723) X-ray sources not confirmed in the 10ks catalogue, and the 435 2XLSSOPT (8% of 5548) new in the 10ks catalogue, and let us concentrate on the common sources. Of their best counterparts, 86% are essentially confirmed in both catalogues (3% are confirmed tentative blank fields, 72% have the same counterparts and the same rank, 10% have the same counterparts and compatible ranks, and only 1.4% have partially matching counterpart sets with the same or compatible best rank). A further 5% and 4% have the same counterparts but they are ranked differently (the best counterpart in one catalogue is either secondary or rejected in the other). The remaining 5% have altogether different or partially matching counterparts which are ranked differently. *So the difference between counterparts in the deep and 10ks catalogues is confined to less than 15% of the common sources.*

Coming back to the topic of the tentative *blank fields* one has to note that, besides the 139 common ones, there are 111 2XLSSOPTd blank fields not present in the 10ks catalogue and 84 2XLSSOPT blank fields which are new in the latter catalogue (43 are new X-ray sources with no deep counterpart, the other 41 are no longer blank fields).

An extreme case of problems with the so-called blank fields is represented by 2XLSS.Xseq=51425 *corresponding to* 2XLSSd.Xseq=19100. *The deep X-ray source is clearly over a bright ($i' = 15$) galaxy. However a displacement of 2.8″ in X-ray position is enough to move the* 2XLSS *source farther than 6″ from any catalogued counterpart and therefore to misclassify it as blank field.*

Similarly for 2XLSS.Xseq=50319 *corresponding to* 2XLSSd.Xseq=35513. *The deep X-ray source has a reasonable counterpart in a bright (R=18) USNO object, but a displacement of just 1″ in the X-ray position is enough to move the chosen* 2XLSS *position farther than 6″ from any counterpart and misclassify it as blank field. In this case it could be an artifact of the overlap removal (the X-ray source* may11=46985 *is not in* 2XLSSd *but is much closer to 35513 !).*

For 2XLSS.Xseq=45583 *corresponding to* 2XLSSd.Xseq=30203 *a displacement of 2.4″ moves away from a $i' = 13$ star, into an area which is however masked by the star itself and therefore has no catalogued counterparts. This can be considered semi-physiological.*

More physiological cases are when a displacement moves the X-ray position away from some unimpressive counterpart *(like for* 2XLSS.Xseq=47453 *corresponding to* 2XLSSd.Xseq=35646 *at 3.5″) in an empty area, or, as in this case, in an area masked by a nearby bright object*.

Coming now to *non* blank fields, remember that the identification procedure is incremental. So it starts (in absence of a D1 counterpart) associating a W1 object to the X-ray source. Then it may append one SWIRE object (associated to the X-ray source and within 1″from W1) to the counterpart set, and create a new counterpart set for an other SWIRE object. This association may be different as a result of a small displacement in the X-ray position. And so on and so forth for the other wavebands.

In the most favourable case this may just prefer a particular counterpart in a counterpart set otherwise identical and identically ranked. For instance *the best counterpart of* 2XLSS.Xseq=40682 *and* 2XLSSd.Xseq=24302 *at 2/6″ are "DWSUG" with the same D1, W1, SWIRE and UKIDSS counterparts but with a different GALEX one (949 vs 550)*. In other cases counterpart sets similar but differing in one waveband may be ranked differently (primary vs secondary or even rejected).

In a case like 2XLSS.Xseq=40522 *and* 2XLSSd.Xseq=24082 *at 2.5″, the* 2XLSS *best counterpart is a DWSU. A similar WSU counterpart set, without the D1 counterpart, is instead only secondary for* 2XLSSd *. This is in fact a case of a W1 ambiguity. There are two counterpart sets, one DWSU, and the other WSU, with the same SWIRE and UKIDSS counterpart, two different W1 sources (1919 and 1920) and the D1 counterpart alternately attributed to one or the other.*

A reverse case is represented by 2XLSS.Xseq=51463 *and* 2XLSSd.Xseq=19191 *at 4.4″, where the* 2XLSSd *best case (a WSUG) is just a WUG (it loses the SWIRE counterpart) and a secondary for* 2XLSS. *The best counterpart for* 2XLSS *is totally new (it was too far from the* 2XLSSd *position). Note that this is a particularly crowded area.*

2XLSS.Xseq=40288 *and* 2XLSSd.Xseq=24528 *are rather far (6.5″). The* 2XLSSd *preferred counterpart is a DWSUG. The same SWIRE and UKIDSS with a different*

GALEX counterpart are a secondary SUG (no optical) for 2XLSS. *The nominal preferred* 2XLSS *counterpart is a single GALEX object. The area is crowded (cluster or group ?).*

2XLSS.Xseq=40333 *and* 2XLSSd.Xseq=23721 *are rather close (1.5″). The displacement is enough to take the* 2XLSSd *secondary WSU, change the SWIRE counterpart, and consider the resulting WSU as preferred for* 2XLSS. *The* 2XLSSd *primary (a SU) becomes secondary for* 2XLSS.

For 2XLSS.Xseq=40273 *and* 2XLSSd.Xseq=23484 *at 3.5″, the DWSUG preferred counterpart set for* 2XLSSd *is split into two* 2XLSS *secondaries, a GALEX only, and a DWSUG with a different GALEX component, while the* 2XLSS *nominally preferred counterpart is a single UKIDSS which was rejected in* 2XLSSd.

For 2XLSS.Xseq=40682 *and* 2XLSSd.Xseq=24302 *at 2.6″, both catalogues have as preferred counterpart a DWSUG, but the GALEX component is different in the two cases.*

For 2XLSS.Xseq=50153 *and* 2XLSSd.Xseq=35346 *at 4.7″, the WS preferred counterpart of* 2XLSSd *becomes a W rejected in* 2XLSS. *The* 2XLSS *preferred counterpart is a a WG, corresponding to a rejected G in* 2XLSSd. *The W1 counterpart, visible on image close to the GALEX position, was too far to be considered in* 2XLSSd.

For 2XLSS.Xseq=48699 *and* 2XLSSd.Xseq=36022 *are rather far (6.6″), and the displacement moves the* 2XLSS *position in the opposite direction w.r.t. the WSG* 2XLSSd *counterpart which cannot be considered in* 2XLSS. *Conversely a rejected W in* 2XLSSd *becomes the preferred WS in* 2XLSS.

2XLSS.Xseq=40784 *and* 2XLSSd.Xseq=24435 *at 4.3′ are totally disjoint: The* 2XLSSd *best DWUG has no match in* 2XLSS *, whose best counterpart is a new DWU.*

6. Comparison with XMDS

This section provides a comparison between the catalogues presented in this paper (mainly 2XLSSd) and the XMDS one. A subset of it was published as the XMDS/VVDS 4σ catalogue (Chiappetti et al., 2005), and the same paper contained also the logN-LogS of the entire catalogue. The entire XMDS catalogue is unpublished, but was described in several reports (Chiappetti , 2006a,b, 2007, 2008a).

We remind here the main differences between the XMDS and the XAMIN *pipeline.* *(a)* XMDS uses the SAS to detect candidates in 5 energy bands simultaneously (and not on 2 independent bands with later merging), operating on event files merged from all 3 XMM cameras and from the entire XMM field of view (not just the central 13′), and *(b)* then applies the Baldi et al. (2002) characterization, which is oriented to point sources only (unlike the wavelet method in XAMIN which handles better extended sources). *(c)* The event pattern selection is different (non-standard and broader in XMDS). *(d)* The removal of redundant sources is handled differently, in particular the primary detection is chosen differently, the position is inherited from the primary detection, but the flux is obtained stacking data from all overlapping pointings (Chiappetti , 2006b). *(e)* Finally the astrometric correction is also different. *(f)* Also the XMDS catalogue does not include spurious objects (but only those above a probability threshold), so the difference between the raw database table and the catalogue is only due to the overlap removal procedure.

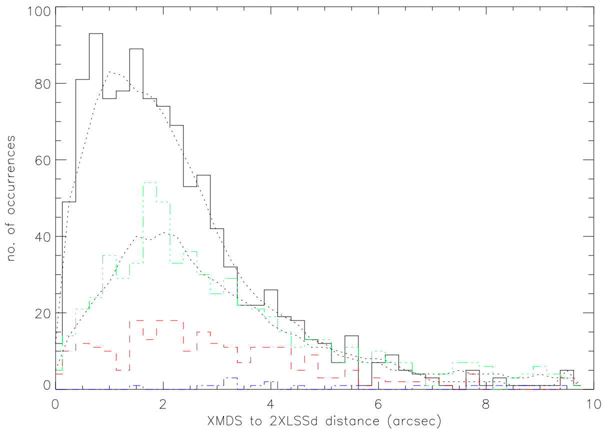

Fig. 7. Histogram of the angular distance between XMDS and 2XLSSd X-ray positions in astrometrically corrected coordinates. The solid (black) line refers to all sources. The long dash (red) line to the source having a detection likelihood below 40 in the best band in which they are detected. The dash-dot (blue) line to extended sources. For comparison we report also the distribution of the inter-band distance maxdist between 2XLSSd positions in bands B and CD, for sources detected in both bands (dotted line, green). The dotted lines are a smoothing of the corresponding histogram.

6.1. Comparison of the X-ray source lists

The XMDS catalogue includes 1168 sources, by definition all in the XMDS pointings (i.e. the G fields for XMM-LSS). The comparison may occur with detections in the jan11 or may11 tables, or in the 2XLSSd or 2XLSS catalogues, and involve also adjacent B fields if the XMM-LSS overlap removal procedure preferred those. The comparison with jan11 (or 2XLSSd) is more meaningful, because it is comparing the same event data (for the full exposures) with different pipelines and procedures. Remember that G field exposures are of the order of 20-40ks, longer than the truncated exposures used in may11 (or 2XLSS).

Of those 1168 objects, 1082 have a jan11 counterpart of which 1057 are catalogued in 2XLSSd (while 1019 have a may11 one of which 956 in 2XLSS).

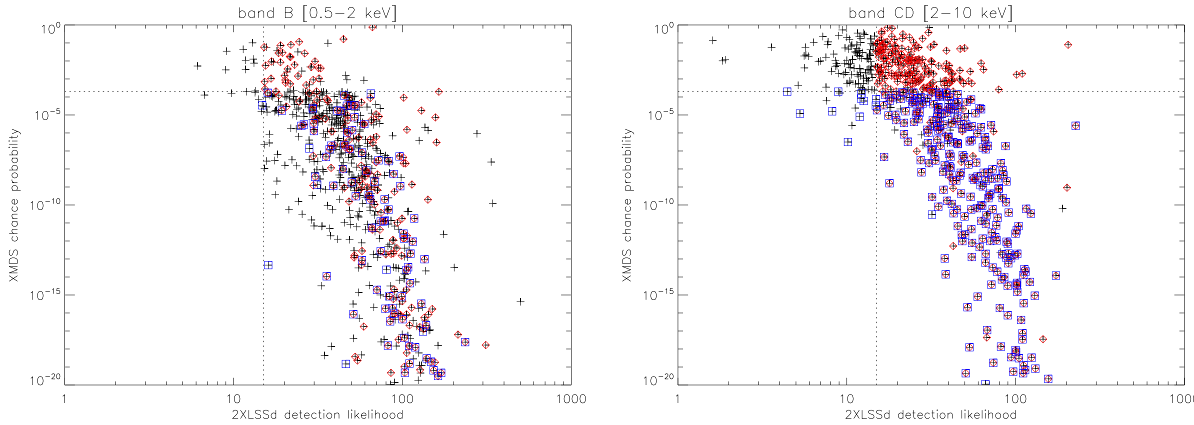

Fig. 8. Cross calibration between the `2XLSSd` (Xamin) detection likelihood and the XMDS chance detection probability. Left panel for the soft band, right panel for the hard band. The dashed lines indicate the two acceptance thresholds of $ML > 15$ and $p < 2 \times 10^{-4}$ (remember that a source can be accepted if it is above the threshold *in at least one band* but not necessarily in all). Crosses indicate all objects detected in the given band. A (red) diamond surrounds the sources detected above threshold *in both bands* in the `2XLSSd`, while a (blue) square surrounds those detected above threshold *in both bands* in the XMDS.

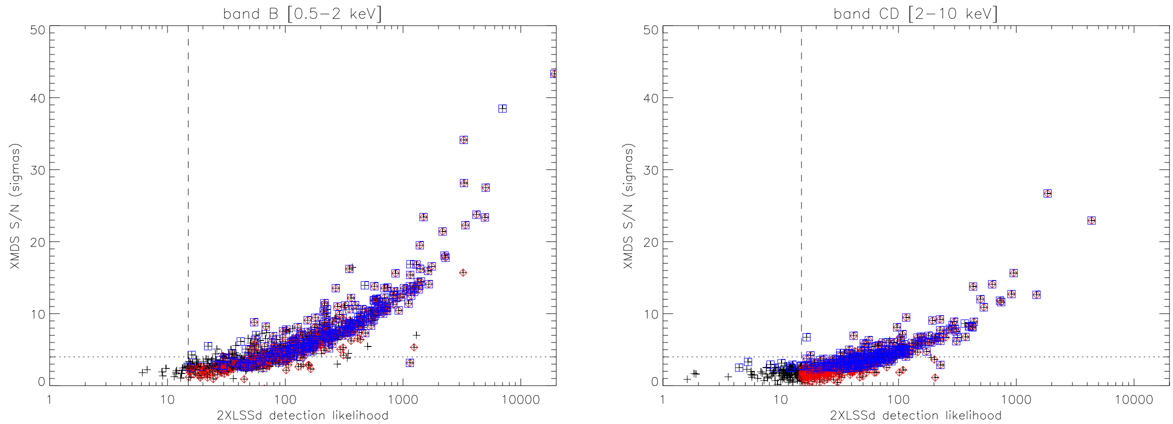

Fig. 9. Cross calibration between the `2XLSSd` (Xamin) detection likelihood and the XMDS signal to noise ratio. Left panel for the soft band, right panel for the hard band. The dashed vertical line indicates the `2XLSSd` acceptance thresholds of $ML > 15$ (as in Fig. 8), while the dotted horizontal line shows the conventional level of 4σ. Symbols as in Fig. 8.

Of the 86 or 149 XMDS sources not in any XMM-LSS physical table, 23 (22) are at off-axis angles greater than 13′ (ignored by Xamin), and 39 (45) anyhow at large off-axis angles ($> 10'$). So it is not surprising they were excluded by Xamin.

Similarly 43 (44) are potential ultrasoft sources (the band with the highest S/N ratio in XMDS is the A band (0.3-0.5 keV), which is not processed by the current release of Xamin). So they are legitimately excluded. They aren't anyhow impressive sources, only 20 have a significance above 3σ (and the best one is just at 4.1σ).

Considering the XMDS significance, 46 (94) and 72 (131) are respectively below 3σ and 4σ. If one allows the combination of different conditions, a net majority of the XMDS-only sources (76 and 126) are *either* ultrasoft, or at offaxis $> 10'$ or at $< 3\sigma$.

We associate XMDS and XMM-LSS sources within a radius of 10″. This can result in more than one association, so we verified if the counterparts are the closest, or detected in the same field, or not.

We will concentrate below on the 1057 (956) XMDS sources with a `2XLSSd` (or `2XLSS`) correspondent (which can be called *common catalogued*). For the remaining 25 (63) cases, several counterparts are possible and we considered only the closest. They are in jan (or `may11`) but not in the catalogue because they are either spurious (24 or 59), i.e. with a Xamin likelihood < 15, or because they were excluded by the overlap removal procedure.

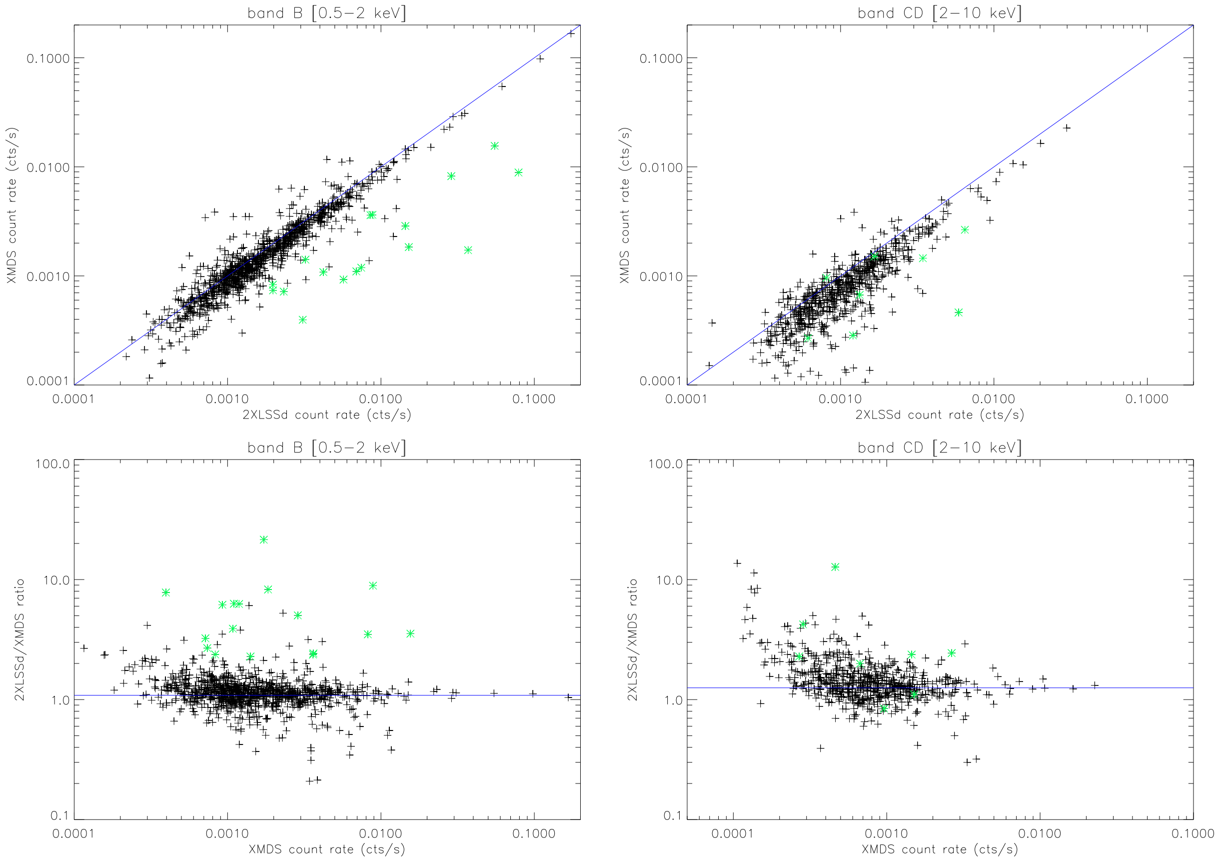

Fig. 10. Two alternate ways of comparing count rates. Top : The XMDS count rate vs the 2XLSSd camera-merged count rate for band B (left panel) and band CD (right panel). (Black) crosses mark pointike sources detected in the band, while (green) asterisks correspond to extended sources. The diagonal solid line is a fiducial line corresponding to equal XMDS and 2XLSSd rates. Bottom : The ratio of the 2XLSSd camera-merged and XMDS count rates as function of the XMDS count rate for band B (left panel) and band CD (right panel). (Black) crosses mark pointike sources detected in the band, while (green) asterisks correspond to extended sources. The horizontal solid line is a fiducial line corresponding to the actual average ratio in the band (see text).

For the common catalogued sources, 783 (684) are in the same (G) field, so should be exactly the same detection. 173 (169) are stacked XMDS entries (Chiappetti , 2006b). The remaining 103 (104) associate sources detected in different pointings : 39 (38) in another G field, 64 (66) in a B field.

Note that jan11 contains 2888 (2056 in may11) sources in the G fields. Of these 1703 (970) have no XMDS counterpart. 977 (588) of the 1703 (970) are spuriuos, so we do not care whether they aren't in the more restrictive XMDS tables. Of the 726 non-spurious, 34 are extended and 692 pointlike (382, 31 and 351). Of the latter 268 are very poor ($ML < 20$), almost all (626) have $ML < 40$, and only 17 have $ML > 75$ (172, 326 and 7).

Coming back to the 1082 (1019) common sources, we can proceed to a more detailed comparison. All text (unless otherwise stated) and figures refer to a comparison between XMDS and 2XLSSd (i.e. data from compatible full exposures). I have produced equivalent figures for 2XLSS but they are shown only in the Appendix provided as a separate document.

The distance between the (astrometrically corrected) X-ray positions in the XMDS and 2XLSSd catalogues is shown in Fig. 7. 58% of the sources are closer than 2″, 88% closer than 4″ and only 4% more distant than 6″, in general concentrated among the sources with lesser significance, and the few extended ones. The agreement between XMDS and 2XLSSd positions, peaking around 1″, is better than the typical inter-band distance between 2XLSSd detections in the two energy bands, which peaks around 2″. However it indicates a little systematic difference between XMDS and 2XLSSd positions, contrary to the fully consistent 2XLSSd and 2XLSS positions (see Fig. 2).

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

Alternatively one can use Fig. 9 to cross calibrate the detection likelihood of XAMIN with the significance in

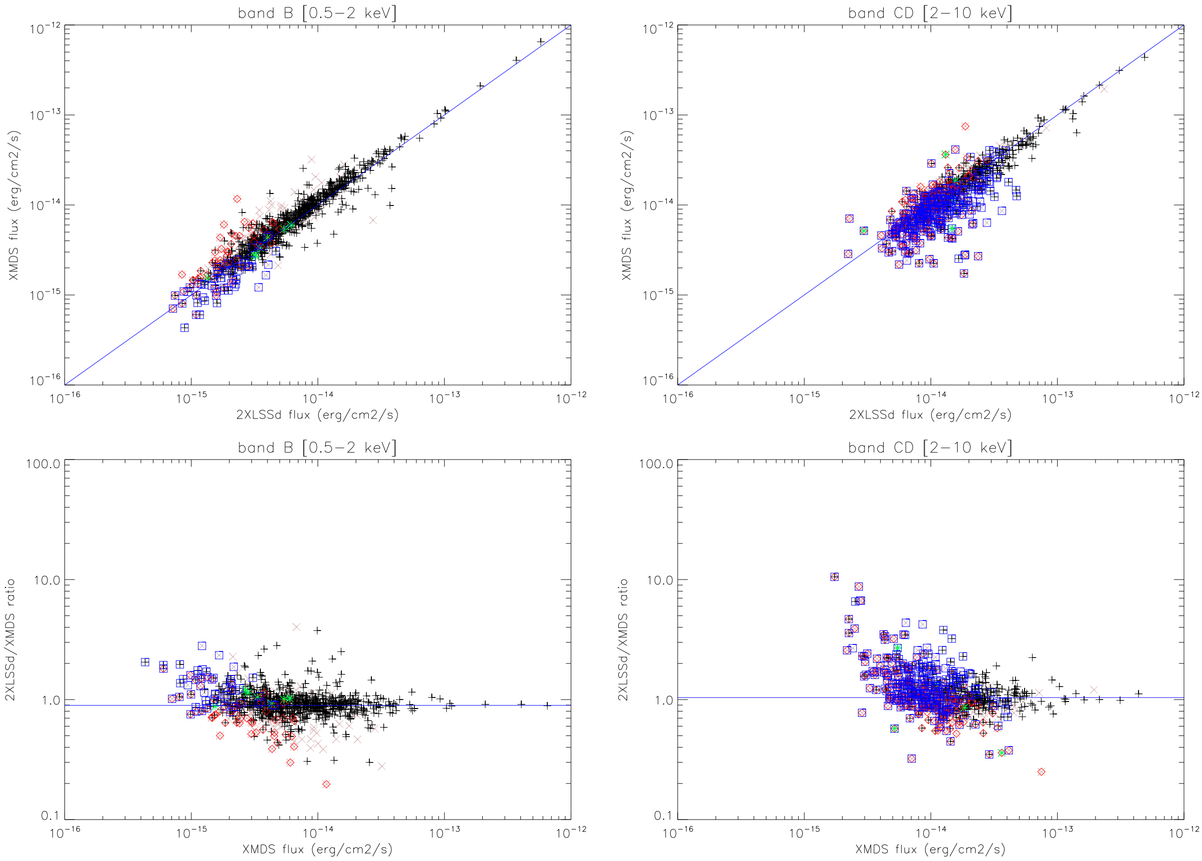

Fig. 11. Two alternate way of comparing fluxes. Top: The XMDS flux vs the 2XLSSd flux for band B (left panel) and band CD (right panel). The diagonal solid line is a fiducial line corresponding to equal XMDS and 2XLSSd fluxes. Bottom: The ratio of the 2XLSSd and XMDS fluxes as function of the XMDS flux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

terms of number of σ of the XMDS (see Chiappetti et al. (2005) and references therein). One can see that a likelihood of 75 corresponds more or less to the 4σ level, and one of 40 to the 3σ level. For 2XLSS the plot might appear more contracted, but the data are not consistent, since they won't use the same exposure times.

The common subset of 1082 (1019) sources includes 17 (17) potentially extended sources (10 (7) classified C1, of which 3 (3) detected in both bands and none detected only in the hard band; 7 (10) classified C2 all (1 only) detected in both bands).

We also note that 89% (91%) of the common sources have B as the best band (highest likelihood) in 2XLSSd (2XLSS). 96% (95%) of the sources are observed by 2XLSSd (2XLSS) in the B band, 62% (46%) are observed in the CD band, and 57% (42%) are observed in both. This can be compared with the totality of the 2XLSSd (2XLSS) catalogue, where 84% (86%) of the sources have B as the best band, 89% (90%) are observed in the B band, 48% (43%) in the CD band and 37% (33%) in both (there is no appreciable difference between the full 2XLSSd catalogue and the sources in the G fields alone). The XMDS by construction includes measurements in all 5 energy bands even if the source is above the probability threshold only in one. If we consider good detections for XMDS only those with $p < 2 \times 10^{-4}$ in the band, we have that 91% (92%) of the sources in common with 2XLSSd (2XLSS) are detected in the B band, 36% (37%) in the CD band, and 30% (32%) in both.

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

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

(**This rate is not the same as the plain summed rate used elsewhere in the XLSS paper!**)

The count rates are compared in Fig. 10. They match reasonably well, although their average ratio is not unity (which is not surprising considering they result from independent processing, and in particular the event pattern selection used in XMDS and XMM-LSS is different), but the 2XLSSd (2XLSS) rate is 1.086 (1.254) times higher than the XMDS one in the B band, and 1.253 (1.405) times higher in the CD band.

It can be seen that most outliers are either concentrated at low rates (i.e. poor significance) or correspond to extended sources, for which the XMDS obviously fails in characterizing the source.

The fluxes, computed for XMDS according to the prescriptions of Baldi et al. (2002) and for XMM-LSS as explained in the XLSS paper, are compared in Fig. 11. Extended sources classified C1 are excluded as their flux is set to undefined. The fluxes match qualitatively, although there is a systematic difference : namely the 2XLSSd (2XLSS) fluxes are 0.895 (0.943) lower than the XMDS fluxes in the B band, while they are only 1.040 (1.086) higher in the CD band. This despite the different agreement between rates quoted above.

It can be seen that a larger scatter in fluxes occurs for the sources which have poorer significance in either catalogue, while outliers are generally due to sources presumably falling near a chip gap on one detector (and as such characterized by a fluxflag of 2), or exceptionally by residual C2 extended sources (for which the XMM-LSS flux is computed from the pointlike rate).

6.2. Comparison of the optical counterparts

It is also possible, similarly to what done in 5.3, to compare the counterparts in optical (and other) bands between the XMDS catalogue and one of our XMM-LSS catalogues. In doing this one should consider

- that the identification procedure for XMDS is historically different from the one used here, in particular was done in several incremental steps, and uses capped probabilities (Chiappetti , 2006a, 2007, 2008a);
- that the catalogues used for XMDS were in larger number (for a total of 27) and included many other, including older, data sources (e.g. VVDS, radio data, SIMBAD and NED, CFHTLS T003, etc. see list in Chiappetti , 2008a)

Therefore I limited the comparison to the 2XLSSd catalogue (the one which matches better XMDS in exposures), and to *reduced counterpart sets* considering only CFHTLS T004 D1 and W1, SWIRE DR6 and GALEX. UKIDSS was not included since 2XLSSd uses release DR5, while XMDS used release DR3. On the other hand GALEX GR4 (the same used for 2XLSSOPTd) was added to XMDS in Nov 2008 after Report IV (Chiappetti , 2008a) was completed. More specifically I considered only the 1057 X-ray sources in common between XMDS and 2XLSSd, as described in 6.1.

They correspond to 4316 counterpart sets (of any rank) in 2XLSSOPTd and 4916 in XMDS. I defined a "reduced unique identifier" composed by the D1, W1 (T004), SWIRE and GALEX seqs, and found that a large fraction (3620) of the possible counterpart sets are identical (i.e. they have the same counterparts, irrespective of ranking, in both catalogues). Of these 92% have the same "UKIDSS behaviour" (namely 1563 have also an UKIDSS counterpart, though in a different release, and 1774 don't have one) in both XMDS and 2XLSSOPTd, while only 283 cases (8%) differ in this respect (one has and the other hasn't UKIDSS).

I then concentrated on the *best counterparts* (ranks 0-1; a similar ranking system, though different in details, was used also for XMDS).

For 627 cases (59% of 1057) the best counterpart is exactly the same (the counterpart sets match in D1, W1, SWIRE and GALEX).

For 16%, 15% and 4% of the cases the best counterpart is nearly the same, in the sense that 3, 2 or 1 of the non-X-ray catalogue counterparts match (the other may be different or missing, because of the displacement in the X-ray position), which makes a total of 81% with the choice of a highly compatible counterpart.

A very limited number of cases (5 and 8) are potential "blank field" respectively in 2XLSSOPTd and XMDS, with another counterpart in the other catalogue (also because of the X-ray position displacement, usually here rather large, 4″-8″).

The remaining 193 (18%) cases select an altogether different counterpart in the two catalogues. In some cases the counterpart set is present only in one catalogue and replaced fully by something else in the other. This occurs for 97 2XLSSOPTd and 77 XMDS sources. In the other cases (96 and 116) the counterpart set which is preferred in one catalogue is still present in the other with a different rank (secondary or rejected). Namely of 96 2XLSSOPTd preferred counterparts, about half (43) are secondaries in XMDS, the other half (49) are rejected in XMDS, and 4 are also preferred counterparts for *another* X-ray source. Instead of the 116 XMDS preferred counterparts 84 are secondaries in 2XLSSOPTd and 32 rejected.

In conclusion, the compatibility between the counterparts is satisfactory.

7. Work required before publication

The major items on the "to do list" in Report VII (namely the 10″band merging and overlap removal, and the 10ks exposure catalogue) have been dealt with. What can remain pending for the publication of the catalogue(s) is:

– The refinement of the (automatic) ranking of identifications, 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, via manual inspection and edit, or possibly via semi-automated procedures, in particular for 2XLSS (2XLSSd has been looked at and used more).

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 in Report VII) represent a serious concern, since they cannot be spotted a priori using the database tables, because the info is incomplete in the (optical) catalogues.

On the other hand these activities are not obstative against publication, since they essentially affect just the identification ranking. The list of X-ray sources and the composition of the counterpart sets is not affected.

The issues to be decided and eventually worked upon for publication are instead:

– a decision whether the CFHTLS, SWIRE, UKIDSS and GALEX data should be replaced by more recent releases (which will require repetition of a lot of work, particularly for what validation is concerned, and cause a substantial delay).
– a decision about *which catalogue* to publish. I.e. if we stick to the previous decision to publish the 10ks 2XLSS, despite the fact it is less tested, or we publish simultaneously also the deep 2XLSSd. We might also consider to dedicate a small section of the paper to a comparison with the XMDS catalogue, and to append such catalogue in complete form to the paper.
– the finalization of data products (see 3.7)
– the final agreement on the layout of data if any change w.r.t the XLSS paper is desired

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

Adami, C., Mazure, A., Pierre, M., et al. 2011, A&A, 526, A18
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)
Chiappetti, L., 2010, XMM-LSS Internal Report N. 7-Mi (Report VII)
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