XMM Survey Science Centre
University of Leicester
Leicester, LE1 7RH, UK

SSC-LUX-TN-0065
Issue 1.2
12 September 2007

Anja C. Schröder (acs@star.le.ac.uk), Ann-Marie Stobbart, Maite Ceballos

Target Classification for the 2XMM Serendipitous Source Catalogue

1 Introduction

The XMM-Newton satellite is used for pointed observations where one or more target objects are of scientific interest to the observers. On the other hand, the 2XMM Serendipitous Source Catalogue, which is compiled from these pointed observations, is a catalogue of all sources in the field-of-view (FOV) with the main interest in serendipitous detections. To avoid a selection bias by the target objects it is worthwhile to identify and flag these non-serendipitous detections.

The target classification scheme used for the 2XMM catalogue is a first attempt to identify and classify the targets. By the nature of the pointed observation (a group of investigators propose to observe one or more objects for a particular scientific reason) such an identification/classification process can only be complete with the help of these investigators. This, however, is not feasible with the available time and resources, and any kind of identification and classification scheme can only be incomplete and subjective. We have therefore decided to use formal information provided with the processed data as well as a manual classification scheme to give the user a choice regarding detail and reliability. The results are presented in a table of all 3491 fields from which the 2XMM catalogue was compiled (see App. A.4 in the User Guide[1]).

We describe the extraction of positions and field names from observation and proposal data in Sec. 2, and the manual identification and classification of the target objects in Sec. 3. In Sec. 4 we present a statistical analysis of the 2XMM fields and detections based on the here presented schemes.

2 Formal target information supplied with the data

There are several positions associated with an XMM observation:

- There is the pointing direction of the telescope, which is always the same position on the detector.

[1] `http://xmmssc-www.star.le.ac.uk/Catalogue/2XMM/UserGuide_xmmcat.html`

- The proposal position refers to the position given by the observer; this position is placed at a specified detector location which depends on the prime instrument (EPIC or RGS) and which avoids chip gaps, dead spots etc, unless an offset is indicated by the investigator.

- The XSA (XMM-Newton Science Archive)[2] gives the coordinates of the prime instrument viewing direction (median values [MAHFRA, MAHFDEC] in the attitude file), which are corrected for the star tracker mis-alignment.

While the pointing direction is the best choice to represent the centre of the FOV in the sky, the proposal position is in most cases a better representative of the target object as chosen by the investigators. However, there are cases where the actual target object is deliberately off-set from the proposal position, or the proposal position is not very accurate. The latter can be due to catalogue errors, positions with large uncertainties (*e.g.*, gamma ray sources), or an error by the proposer. In cases where more than one object is the target the proposal position can either be located on one of the objects or between them. In a few cases the image is not obtained at the position indicated by the proposal which cannot be explained by an offset observation. The reason for this can be a slew failure or a Target of Opportunity (ToO) observation that was not properly registered in the Observation Data File (ODF).

The XSA coordinates are in most cases near the centre of the field and/or the target but do not represent the target position as well as the proposal position.

In the table, we give proposal positions and names derived from keywords in the header of the fits-file products: `RA_OBJ` and `DEC_OBJ` (in J2000) are taken from the attitude time series file (*ATTTSR*), while `OBJECT` is taken from the calibration index file (*CALIND*). From the XSA we have taken the coordinates (transformed to degrees; J2000) as well as the proposal category and proposal program information.

3 Manual identification and classification scheme

Because of the caveats of the positions and names that are supplied by the investigators and are included in the product file and observing lists, we investigated various ways to obtain more accurate information about the target objects. One obvious way was to make use of large databases like Simbad[3] and NED[4]. In particular, Simbad is of interest since (i) it comprises all kind of objects (while NED is mainly a database of extragalactic objects), and (ii) it allows custom-formatted queries and returns by scripts.

The caveat in this case is that such a database recognizes object names only when they are given in a certain (official) structure (*e.g.*, with a registered prefix or with the correct precision of coordinates in the name). In addition, the XMM investigators are free to give their targets any name they want, which often includes offset information or abbreviated coordinates. Consequently, only about half of all targets were recognized immediately by Simbad, while others could be recognized after minor modifications.

These minor modifications were obtained in several ways:

- An offset indicator or other parts of the name that clearly were a specifier to the original name were removed.

[2] `http://xmm.esac.esa.int/external/xmm_data_acc/xsa/index.shtml`
[3] the SIMBAD Astronomical Database at `http://simbad.u-strasbg.fr/simbad/`
[4] the NASA/IPAC Extragalactic Database at `http://nedwww.ipac.caltech.edu/`

- The prefix was checked against the Simbad 'dictionary of nomenclature'[5] and, where possible, suitably altered.

- The proposal coordinates were used to search for an object with a similar name. Most of the times an object which was found at or near those coordinates was in fact the target. However, in some cases an object name could not obvioulsy be related to the proposal name, hence an unambiguous ID could not be established and no Simbad name was given.

- The proposal program and category information as well as the proposal file, available at the XSA database, were used to obtain additional information on name, coordinates or type of object; sometimes the name or the coordinates were apparently mistyped or the field was obtained with a wrong pointing (mainly due to a slew problem of the telescope).

- If no useful information could be found with Simbad, then a name and/or coordinate query at NED was performed. If an identification could be made, the NED coordinates and name were accepted in our list with the indicator '[ned]' after the name, see below.

Using the recognized Simbad (or NED) names, we extracted coordinates in degrees and the object types. In cases without a recognized Simbad (or NED) name, we made an attempt to at least give an object type according to the Simbad convention[6] based on information taken from the proposal abstract. The information about proposal category and proposal program from the XSA also proved to be very useful, *e.g.*, in distinguishing between a single source and a field observation. In a few cases we changed the object type as given automatically by Simbad to better reflect the nature of the observation as described in the proposal. For example, a galaxy that was automatically classified as a Seyfert by Simbad but was observed in the category 'Galaxies' (with, for example, the intent to study its halo), was instead given the object type 'galaxy' (note that Simbad often has a list of several object types for a single entry, from which only the most important one is returned by our simple script query).

On the other hand, the proposal category of calibration observations do not always correlate with the purpose of the observation (calibration observations are often instrument related for which there is no particular proposal category, hence the distinction through the program name).

With all this accumulated information about the probable target object(s), we could proceed to identify and classify the corresponding 2XMM detection or detections for each observation. In most cases where the target is a point source, the proposal and Simbad positions agreed very well and a 2XMM detection could be found within a radius of ~ 10 arcseconds (everything further than that was considered to be ambiguous and possible chance superposition). In case of extended sources the two positions were often different, but based on the object type information, the target object, if bright enough, could be identified. If the extent was within the limits of the fitting algorithm ($r <$ max(EXTENT) $= 80''$), the 2XMM detection is assumed to represent the target object well. If the extent is much larger then the source parameters of the 2XMM detection are less reliable. In such cases a 2XMM identification is given only to indicate the 'best'

[5] http://vizier.u-strasbg.fr/cgi-bin/Dic-Simbad?

[6] http://simbad.u-strasbg.fr/simbad/sim-display?data=otypes

position of the centre of the X-ray emission. In many cases (mostly SNR) a central detection could not be found and no 2XMM identification is given at all.

In a couple of cases, repeated observations of the same object (with different depth and/or proposal position) helped to identify the 2XMM detection in one observation even though the object was too faint or too far from the proposal and Simbad position to be safely identified on its own.

The cross-identification of the 2XMM detections with the target objects was done by hand where each case was examined individually, which means a certain subjectivity cannot be avoided. On the other hand, the two external position informations, proposal position and Simbad position, vary in which of them represents the target object better. We therefore decided to give a 'best' position, 'p' for proposal or 's' for Simbad, in an extra column where it was feasible: as a rule we preferred the Simbad position as the better one except when the proposal position was *clearly* closer to the XMM detection. In case of offset observations (often indicated in the proposal name) a dash is given in the table, and it is up to the user to decide if they prefer the position of the object itself or the offset position for the observation (which is usually near the centre of the FOV). Note that the centre of the object can be within the FOV but also a couple of degrees away in case of very large objects (*e.g.*, a dark cloud or the LMC).

At the same time, we classified the field according to the (detection) type of the target (see Table 1). The most obvious classes are point source, extended source and field observation, where every detection is potentially a target. The extended source class was further divided into three sub-classes: small extended source (*i.e.*, well within the FOV) with a radius of $< 3'$ (covering roughly 3% of the full FOV), large extended source with a radius of $> 3'$ and often extending beyond the FOV, and extended sources of undetermined radius (they are either not detected, not identified, *i.e.*, more than one object fitted the description, or beyond the edge of FOV). Furthermore, some observations were aimed at measuring emission or absorption lines in the spectra of background objects to study diffuse extended sources of low surface brightness ('X-ray shadow experiments'). In such cases none of the detections are actually a target and can therefore be considered to be serendipitous, except for the location of the field in the sky which has been chosen for the particular purpose of the study and is therefore **not** serendipitous.

A few observations had exactly two target objects (point sources or point-like) and seemed to deserve a separate class. The 2XMM detections for both objects are given in two rows in the table. Another group of observations were fully serendipitous because either the proposal coordinates were wrong or the observation was obtained at the wrong location in the sky due to a slew problem. A final class 'unknown' is used for all cases where the class was ambiguous or it could not be determined with the available information.

3.1 Guidelines for the manual classification scheme

As with many subjective classification schemes, the borderlines between the classes are not very distinct and in a few cases an object could be judged differently if it were reclassified. For that reason we give some guidelines below for problem cases. Note that a more detailed classification scheme, as used by the Simbad object types, is usually more reliable than the coarse field classification we present here, but on the other hand the object types are not complete. Furthermore, for statistical purposes a large number of classes is often more a hindrance than a help. For the best results we therefore recommend to use both the target class as well as the object type.

Table 1: Field classification

Field class	Description
p	point source or point-like source
s	small extended source ($r < 3'$)
l	large extended source ($r > 3'$)
e	extended source of unknown extent
f	'field', that is, all detections are potential targets
x	'X-ray shadow experiment' and similar, that is, only the spectra of fore- and background objects are of interest (note the location of the field should be considered as 'target')
t	two clearly identified targets (*e.g.*, a double star); in this case two source numbers are given (first and second row)
n	there is no target associated with the field (*e.g.*, the observation was not obtained at the coordinates given in the proposal and the whole field is deemed to be serendipitous)
u	unknown, *i.e.*, the target could not be classified or is of unknown nature

Galaxies: Galaxies show a large range of appearance in the X-ray. A galaxy can be very distant and point-like or close-by and resolved into point sources and/or diffuse emission; an active nucleus can dominate the total emission of a galaxy. The rule of thumb is to classify a galaxy as 'point source' when the emission from the active nucleus was the only apparent emission from the galaxy. A galaxy was classified as 'extended' when either diffuse emission was apparent or if the galaxy was large enough for discrete X-ray sources in the galaxy to be resolved (in case of doubt we compared the X-ray image with an optical image downloaded from the DSS[7]), or if the galaxy was detected as a single point source by the SAS task `emldetect` but it clearly consisted of several (unresolved) sources.

In two cases, a 'field' classification was preferred: observations of the M31 halo and offset pointings of M33. In both cases the galaxy is considerably larger than the FOV and it seemed appropriate to classify it like a halo observation of our own Galaxy. Note that the *central* observations of M31 (called M31 core) are classified as 'large extended' instead since it includes diffuse emission as well.

Galaxy clusters: Galaxy clusters show X-ray emission from the intercluster gas as well as emission from some of the galaxies within that cluster. We have not used any cluster radius information in our classifications, instead we simply assumed that we would find galaxies also outside the visible diffuse emission. That means, most galaxy clusters will be classified as 'large' except when the cluster is distant (and significantly smaller than $r = 3'$) and no point sources could be discerned within the diffuse emission.

Galaxy groups: Galaxy groups have less members than galaxy clusters. In many cases there is no measurable intercluster emission and the X-ray image will show only emission from some of the members. In some cases there is a prominent galaxy

[7] The STScI Digitized Sky Survey, see `http://stdatu.stsci.edu/cgi-bin/dss_form`

in the centre with a large X-ray halo. Despite this diversity we decided it best to classify all groups in the same way as galaxy clusters, that is, as extended emission, mixed with point or other extended sources, but where the cluster material can be above or below the detection threshold.

Extragalactic point sources: In a few cases a bright X-ray source within a galaxy was the target (*e.g.*, “superEddington” sources); these were treated like AGNs, that is, if no galaxy emission could be discerned we classified the target as ‘point source’ otherwise as ‘extended’.

Mixed targets: It is also possible to find mixed targets, *e.g.*, a particular galaxy within a galaxy cluster, or a Central Compact Object in a SNR. These were classified by the ‘larger’ target, that is, in the examples given the class would be ‘extended’. However, the Simbad object type is likely to refer to the point(-like) source. Note that there are a number of cases where such a connection was not obvious or could not be easily determined (*e.g.*, a connection between a quasar and a galaxy cluster which may be hosting the quasar or simply be superimposed in the line-of-sight) and the class refers to the quoted object only. Note also that in case of a calibration observation the object is more likely to be chosen for its own properties and not for its possible connection/interaction with the environment.

Solar system objects: There are a number of observations of planets or comets in our solar system. The Simbad object types do not list these, but for simplicity we give them a special type ‘com’ for ‘comet’ and ‘plt’ for ‘planet’. The field classification for these depended on what was visible in the image, *e.g.*, if there was visible (and detected) diffuse emission in case of a comet, or if a planet was observed long enough to produce a longish trace on the image (the pipeline processing corrects for any attitude shifts so that a fixed point in the sky is always at the same location in the image).

4 Some statistics of the 2XMM catalogue

There are 3491 fields in total in the 2XMM catalogue. For 3044 fields (87%) a Simbad name could be found, and in 53 cases (1.5%) a NED identification is given. Of the remaining 394 fields only 56 (1.6%) do not have an estimated object type.

The XSA gives information of the program as well as the category the proposal is part of. Tables 2 and 3 list the individual proposal classes together with the percentage of 2XMM fields. Note that some of the calibration observations are not properly classified into categories.

Table 4 gives the percentages for the manual field classifications. Object types and their frequencies in 2XMM are listed in Table 5.

Table 2: Proposal program from the XSA

Class	Description	Percentage
GO	Guest Observer	85%
Cal	Calibration	10%
ToO	Targets of Opportunity	3%
Cha	Co-Chandra	0.9%
ESO	Co-ESO	0.3%
Large	Large	0.8%
Trig	Triggered	0.8%
Eng	Engineering	0%

Table 3: Proposal category from the XSA

Class	Description	Percentage
I	Stars, White Dwarfs and Solar System	16%
II	White Dwarf Binaries, Neutron Star Binaries, Cataclysmic Variables, ULXs and Black Holes	15%
III	Supernovae, Supernova Remnants, Diffuse Emission, and Isolated Neutorn Stars	14%
IV	Galaxies and Galactic Surveys	9%
V	Groups of Galaxies, Clusters of Galaxies, and Superclusters	14%
VI	Active Galactic Nuclei, Quasars, BL Lac Objects, and X-ray Background	23%
VII	X-ray Background and Surveys	8%

Table 4: Field classification

Field class	Description	Percentage
p	point-like	50%
s	small extended	10%
l	large extended	22%
e	unknown extent	0.7%
f	‘field’	12%
x	‘X-ray shadow experiment’	2.5%
t	two targets	0.4%
n	no target	0.2%
u	unknown target	2%

Table 5: Simbad object types

Object type	Number	Description
Composite objects (n=1332, 38%):		
reg	301	Region defined in the sky
SCG	1	Supercluster of Galaxies
ClG	422	Cluster of Galaxies
GrG	47	Group of Galaxies
CGG	2	Compact Group of Galaxies
PaG	8	Pair of Galaxies
IG	37	Interacting Galaxies
Cl*	7	Cluster of Stars
GlC	20	Globular Cluster
OpC	30	Open (galactic) Cluster
As*	7	Association of Stars
**	19	Double or multiple star
EB*	6	Eclipsing binary
Al*	5	Eclipsing binary of Algol type
bL*	8	Eclipsing binary of beta Lyr type
WU*	3	Eclipsing binary of W UMa type
SB*	15	Spectroscopic binary
CV*	32	Cataclysmic Variable Star
DQ*	12	Cataclysmic Var. DQ Her type
AM*	54	Cataclysmic Var. AM Her type
NL*	14	Nova-like Star
No*	29	Nova
DN*	22	Dwarf Nova
XB*	33	X-ray Binary
LXB	136	Low Mass X-ray Binary
HXB	62	High Mass X-ray Binary
Galaxies (n=1082, 31%):		
G	172	Galaxy
GiC	13	Galaxy in Cluster of Galaxies
GiG	11	Galaxy in Group of Galaxies
GiP	14	Galaxy in Pair of Galaxies
DLA	3	Damped Ly-alpha Absorption Line system
rG	32	Radio Galaxy
H2G	28	HII Galaxy
LSB	6	Low Surface Brightness Galaxy
EmG	11	Emission-line galaxy
SBG	5	Starburst Galaxy
BCG	2	Blue compact Galaxy
LeQ	3	Gravitationally Lensed Image of a Quasar
AGN	14	Active Galaxy Nucleus
LIN	36	LINER-type Active Galaxy Nucleus
SyG	17	Seyfert Galaxy
Sy1	310	Seyfert 1 Galaxy
Sy2	134	Seyfert 2 Galaxy
BLL	79	BL Lac – type object
QSO	192	Quasar

Table 5: continued.

Object type	Number	Description
Nebulae (n=335, 10%):		
CGb	1	Cometary Globule
EmO	3	Emission Object
HH	6	Herbig-Haro Object
Cld	7	Cloud of unknown nature
DNe	10	Dark Nebula
RNe	3	Reflection Nebula
HI	9	H I(neutral) region
MoC	24	Molecular Cloud
HVC	3	High-velocity Cloud
HII	7	H II(ionized) region
PN	6	Planetary Nebula
sh	11	H Ishell
SR?	6	SuperNova Remnant Candidate
SNR	239	SuperNova Remnant
Stars (n=541, 15%):		
*	69	Star
*iC	2	Star in Cluster
i	7	Star in double system
Y*O	4	Young Stellar Object
Em*	21	Emission-line Star
Be*	1	Be Star
WD*	23	White Dwarf
ZZ*	1	Pulsating White Dwarf
BD*	4	Brown Dwarf ($M< 0.08\ M_{\odot}$)
pr*	6	Pre-main sequence Star (optically detected)
TT*	39	T Tau-type Star
WR*	32	Wolf-Rayet Star
PM*	11	High proper-motion Star
V*	44	Variable Star
Or*	12	Variable Star of Orion Type
Er*	1	Eruptive variable Star
Fl*	11	Flare Star
FU*	7	Variable Star of FU Ori type
RC*	1	Variable Star of R CrB type
Ro*	19	Rotationally variable Star
El*	1	Ellipsoidal variable Star
Psr	117	Pulsar
BY*	27	Variable of BY Dra type
RS*	41	Variable of RS CVn type
Mi*	4	Variable Star of Mira Cet type
dS*	1	Variable Star of delta Sct type
bC*	5	Variable Star of beta Cep type
gD*	1	Variable Star of gamma Dor type
SN*	14	SuperNova
Sy*	15	Symbiotic Star

Table 5: continued.

Object type	Number	Description
Miscellaneous (n=125, 4%):		
Rad	14	Radio-source
IR	2	Infra-Red source
X	51	X-ray source
ULX	2	Ultra-luminous X-ray source
gam	28	gamma-ray source
gB	27	gamma-ray Burst
Lev	1	(Micro)Lensing Event
Non-Simbad types (n=20, 0.6%):		
com	9	Comet
plt	6	Planet
sfr	4	Star forming region
XRN	1	X-ray reflection nebula
No type (n=56, 1.6%)		