

The XMM deep survey in the CDF-S

III. Point source catalogue and number counts in the hard X-rays^{★,★★}

P. Ranalli^{1,2,3}, A. Comastri², C. Vignali^{3,2}, F. J. Carrera⁴, N. Cappelluti², R. Gilli², S. Puccetti⁵, W. N. Brandt^{6,7}, H. Brunner⁸, M. Brusa^{3,2,8}, I. Georgantopoulos^{1,2}, K. Iwasawa⁹, and V. Mainieri¹⁰

¹ Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, Palaia Penteli, 15236 Athens, Greece

e-mail: piero.ranalli@oabo.inaf.it

² INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy

³ Università di Bologna, Dipartimento di Fisica e Astronomia, via Bertini Pichat 6/2, 40127 Bologna, Italy

⁴ Instituto de Física de Cantabria (CSIC – UC), 39005 Santander, Spain

⁵ ASI Science Data Center, via Galileo Galilei, 00044 Frascati, Italy

⁶ Department of Astronomy and Astrophysics, Pennsylvania State University, University Park, PA 16802, USA

⁷ Institute for Gravitation and the Cosmos, Pennsylvania State University, University Park, PA 16802, USA

⁸ Max –Planck – Institut für extraterrestrische Physik, 85748 Garching, Germany

⁹ ICREA and Institut de Ciències del Cosmos (ICC), Universitat de Barcelona, (IEEC – UB), Martí y Franquès 1, 08028 Barcelona, Spain

¹⁰ ESO, Karl-Schwarschild-Strasse 2, 85748 Garching bei Munchen, Germany

Received 31 January 2013 / Accepted 17 April 2013

ABSTRACT

Nuclear obscuration plays a key role in the initial phases of AGN growth, yet not many highly obscured active galactic nuclei (AGN) are currently known beyond the local Universe, and their search is an active topic of research. The *XMM-Newton* survey in the *Chandra* Deep Field South (XMM-CDFS) aims at detecting and studying the spectral properties of a significant number of obscured and Compton-thick ($N_{\text{H}} \gtrsim 10^{24} \text{ cm}^{-2}$) AGN. The large effective area of *XMM-Newton* in the 2–10 and 5–10 keV bands, coupled with a 3.45 Ms nominal exposure time (2.82 and 2.45 Ms after light curve cleaning for MOS and PN, respectively), allows us to build clean samples in both bands, and makes the XMM-CDFS the deepest *XMM-Newton* survey currently published in the 5–10 keV band. The large multi-wavelength and spectroscopic coverage of the CDFS area allows for an immediate and abundant scientific return. In this paper, we present the data reduction of the XMM-CDFS observations, the method for source detection in the 2–10 and 5–10 keV bands, and the resulting catalogues. A number of 339 and 137 sources are listed in the above bands with flux limits of 6.6×10^{-16} and $9.5 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. The flux limits at 50% of the maximum sky coverage are 1.8×10^{-15} and $4.0 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. The catalogues have been cross-correlated with the *Chandra* ones: 315 and 130 identifications have been found with a likelihood-ratio method, respectively. A number of 15 new sources, previously undetected by *Chandra*, is found; 5 of them lie in the 4 Ms area. Redshifts, either spectroscopic or photometric, are available for ~95% of the sources. The number counts in both bands are presented and compared to other works. The survey coverage has been calculated with the help of two extensive sets of simulations, one set per band. The simulations have been produced with a newly-developed simulator, written with the aim of the most careful reproduction of the background spatial properties. For this reason, we present a detailed decomposition of the *XMM-Newton* background into its components: cosmic, particle, and residual soft protons. The three components have different spatial distributions. The importance of these three components depends on the band and on the camera; the particle background is the most important one (80–90% of the background counts), followed by the soft protons (4–20%).

Key words. catalogues – surveys – galaxies: active – methods: data analysis – X-rays: general

1. Introduction

The understanding that the X-ray background (XRB) is due to unresolved emission by discrete extragalactic sources (Giacconi & Zamorani 1987; Schwartz et al. 1976; Maccacaro et al. 1991), and that obscured active galactic nuclei (AGN) are a key ingredient to produce the observed shape of the XRB spectrum

(Setti & Woltjer 1989; Comastri et al. 1995) has put a strong focus on the role of absorption as a driver of the observable properties. This idea has shaped the original model of AGN unification (Antonucci 1993; Urry & Padovani 1995) and keeps its importance unaltered in the current modelling of AGN and of their evolving populations (Treister & Urry 2005; Gilli et al. 2007; Treister et al. 2009a).

Nuclear obscuration might be associated with the initial phases of AGN emission (Page et al. 2004; Hopkins et al. 2006; Menci et al. 2008). The idea is that a large gas reservoir is available at high redshift to feed (and obscure) an accreting super-massive black hole. The same gas reservoir would also sustain high star formation rates in the host galaxy, leading to the broad

* Based on observations obtained with *XMM-Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA.

** The catalogue is only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/555/A42>

similarity between the cosmic histories of accretion and star formation.

Deep X-ray surveys are a primary tool for the census of AGN and the study of the properties of their populations. Since the launch of the *Einstein* Observatory, the first imaging X-ray telescope, many surveys have been performed with different combinations of area and flux limit and in many locations on the sky.

The *Chandra* Deep Field South (CDFs) survey started with 1 Ms of observations (Giacconi et al. 2002) and was afterwards extended to 2 Ms (Luo et al. 2008) and finally to 4 Ms (Xue et al. 2011, hereafter X11). It currently reaches the deepest X-ray fluxes ever probed, with a flux limit of $9.1 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ in the soft (0.5–2 keV) band. However, given the energy dependence of its effective area, *Chandra* is much less sensitive in the hard and very hard bands (2–10 keV and 5–10 keV, respectively) which are the most important for obscured objects. The flux limit for 2–10 keV is currently $5.5 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ (X11). The only analysis for the 5–10 keV band was done for the initial 1 Ms survey (Rosati et al. 2002) and yielded a flux limit of $1.2 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$; however, the 4 Ms survey was analysed in the 4–8 keV band, moving the flux limit to $6 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ (Lehmer et al. 2012). A larger area centred around the CDFs was surveyed with *Chandra* and with much shorter exposure times as the Extended CDFs (ECDFs; Lehmer et al. 2005).

XMM-Newton is better suited to collect photons in the hard and very hard bands, because at energies ≥ 5 keV the *XMM-Newton* effective area drops less sharply than *Chandra*'s. However, it has a larger point spread function (PSF), which makes it suffer more from source confusion, and it has a higher instrumental background. While *XMM-Newton* surveys cannot probe very faint fluxes as *Chandra*, they provide good quality spectroscopy in the whole 0.5–8 keV band for a large number of sources. The large effective area of *XMM-Newton* has been exploited in several surveys, including the Lockman Hole (with 1 Ms of exposure, Brunner et al. 2008), the ELAIS-S1 field (Puccetti et al. 2006), COSMOS (Cappelluti et al. 2007, 2009), and 2XMM (Mateos et al. 2008).

The *XMM-Newton* survey in the CDFs (hereafter XMM-CDFs) has been started with the main aim of finding obscured AGN, especially the Compton-thick ones, and studying their properties. With a nominal exposure of 3.45 Ms centred on a single point on the sky it is currently the deepest *XMM-Newton* observation ever performed. Moreover, it has the advantage of having almost pan-chromatic coverage and a large number of available optical spectra and redshifts.

The first results of the XMM-CDFs survey have been published in Comastri et al. (2011). Different approaches to the selection and study of obscured sources are being employed (Iwasawa et al. 2012; Georgantopoulos et al. 2013; spectral stacking Falocco et al. 2013; Castelló-Mor et al. 2013) and spectral analysis of a flux limited sample (Comastri et al., in prep.) are also ongoing. Extended sources will be discussed in Finoguenov et al. (in prep.).

In the present paper, we present the hard and very hard catalogues, and the Log N –Log S in the two bands. A number of 339 and 137 sources are detected, respectively, with a significance threshold roughly equivalent to 4.8σ . Supplementary catalogues are provided for sources with lower significance (4σ).

In Sect. 2 we present the observation details and the data reduction procedure. In Sect. 3 we describe the source detection procedure and present the catalogues in the hard and very hard bands. The survey coverage and number counts are presented in Sect. 4. Simulations of mock XMM-CDFs fields are

instrumental in determining the coverage, and are presented in Sect. 5. The simulations also used in Sect. 6 for a first estimate of the number of spurious sources. We present in Sect. 7 an analysis of the source confusion in the hard band. In Sect. 8 we identify the *Chandra* counterparts to the *XMM-Newton* sources. We summarize our conclusions in Sect. 9. Finally, in Appendix A we present the simulator that we have developed.

2. Observations and data reduction

The CDFs was initially observed by *XMM-Newton* in the years 2001–2002 (P.I.: J. Bergeron) with an exposure of 541 ks. It was observed again, for the proposal which led to this series of papers, in the years 2008–2010 (P.I.: A. Comastri), to reach a nominal exposure of 3.45 Ms.

Table 1 shows the observation identification number (hereafter obsid), date, pointing and exposure for all observations. The number of obsid s is 33; considering the MOS1, MOS2 and PN cameras, the XMM-CDFs survey comprises a total of 99 event files. We used the *XMM-Newton* SAS software for our analysis (version 10 for the initial processing and the catalogue; version 11 for the simulations; there have not been significant changes between the two versions which could affect this work), with the help of many custom-developed scripts to automate most of the data processing.

2.1. Background flares and quiescent level

Each one of the 99 event files was screened individually for background flares. Full-field light curves were generated in the 10–13.5 keV interval. A 3σ -clipping procedure was applied to the light curves to identify and reject the high-background periods. This procedure worked well most times, though it failed for some event files which were severely flared. In such cases, we adopted the nominal count rate thresholds given by the *XMM-Newton* documentation¹ to identify and reject the high-background periods; these observations are marked in Table 1.

A relevant and unexpected feature is an increase by a factor of ~ 2 in the quiescent background level in the 2008–2010 observations with respect to the years 2001–2002 (Fig. 1). The increase is seen in all three cameras and is due to the instrumental (“particle”) component of the background (see Sect. 5). The reason for this increase is not clear, though one possibility is that it is related to the Solar cycle.

The images in the 2–10 and 5–10 keV bands, showing the total counts from the MOS and PN cameras, are presented in Fig. 2. The total exposure, after the high-background period filtering, is thus 2.82 and 2.45 Ms for the two MOS and the PN, respectively. The exposure maps for the 2–10 and 5–10 keV bands are shown in Fig. 3.

2.2. Copper lines complex

For the PN camera, the strongest background feature in the 2–10 keV interval is a complex of lines (Cu $K\alpha$, Ni $K\alpha$, Zn $K\alpha$, the Cu being the strongest one) around 8 keV, due to fluorescence from the mirror holding structure, and which alone can make $\sim 30\%$ of the total counts in the 2–10 keV band. This

¹ http://xmm.esac.esa.int/sas/current/documentation/threads/EPIC_filterbackground.shtml

Table 1. XMM-CDFS observation log.

OBS_ID	Date	RA	Dec	PA	MOS1 EXP	MOS2 EXP	PN EXP	BORES.GR.
0108060401	2001-07-27	53.0812	-27.7929	58.74	^a 24.6	^a 24.6	^a 18.8	A
0108060501	2001-07-27	53.0919	-27.7981	58.75	43.6	44.6	35.8	A
0108060601	2002-01-13	53.1457	-27.8252	238.71	^a 52.6	^a 52.6	43.6	B
0108060701	2002-01-14	53.1406	-27.8218	238.72	78.5	78.7	69.4	B
0108061801	2002-01-16	53.1450	-27.8195	238.72	55.1	54.9	54.4	B
0108061901	2002-01-17	53.1456	-27.8133	238.71	42.9	42.8	39.1	B
0108062101	2002-01-20	53.1511	-27.8168	238.72	44.4	44.3	42.6	B
0108062301	2002-01-23	53.1462	-27.8147	238.71	83.7	82.5	72.3	B
0555780101	2008-07-05	53.1476	-27.7368	61.73	106.6	109.3	88.8	C
0555780201	2008-07-07	53.1488	-27.7457	61.72	116.7	118.7	^a 91.5	C
0555780301	2008-07-09	53.1387	-27.7448	61.73	^a 109.0	108.6	96.7	C
0555780401	2008-07-11	53.1385	-27.7371	61.74	89.9	91.5	77.7	C
0555780501	2009-01-06	53.1325	-27.8371	241.73	97.9	99.2	90.7	D
0555780601	2009-01-10	53.1343	-27.8439	241.71	104.9	108.8	79.3	D
0555780701	2009-01-12	53.1322	-27.8545	241.72	101.5	102.5	95.7	D
0555780801	2009-01-16	53.1233	-27.8359	241.71	88.4	89.1	77.0	D
0555780901	2009-01-18	53.1230	-27.8445	241.74	96.0	96.3	80.6	D
0555781001	2009-01-22	53.1234	-27.8528	241.73	106.4	108.2	101.3	D
0555782301	2009-01-24	53.1235	-27.8531	241.74	106.5	107.1	96.4	D
0604960101	2009-07-27	53.1486	-27.7447	62.73	101.5	105.1	86.8	C
0604960201	2009-07-17	53.1404	-27.7448	62.73	104.7	104.4	^a 76.9	C
0604960301	2009-07-05	53.1496	-27.7550	62.72	^a 101.4	^a 102.3	^a 89.0	C
0604960401	2009-07-29	53.1393	-27.7544	62.73	119.2	120.2	112.5	C
0604960501	2010-01-18	53.1307	-27.8384	245.73	46.3	46.5	45.0	D
0604960601	2010-01-26	53.1312	-27.8463	245.73	^a 106.9	^a 106.8	95.6	D
0604960701	2010-01-12	53.1224	-27.8388	245.73	^a 91.7	^a 92.6	^a 51.1	D
0604960801	2010-02-05	53.1315	-27.8569	245.73	90.3	90.3	82.3	D
0604960901	2010-02-11	53.1221	-27.8573	245.72	^a 84.7	^a 84.8	75.4	D
0604961001	2010-02-13	53.1211	-27.8476	245.72	91.4	93.7	86.1	D
0604961101	2010-01-04	53.1316	-27.8301	245.72	104.2	105.9	100.6	D
0604961201	2010-01-08	53.1197	-27.8294	245.71	114.0	115.0	^a 96.9	D
0604961301	2010-01-19	53.1307	-27.8379	245.71	^a 11.2	^a 11.9	^a 5.8	D
0604961801	2010-02-17	53.1213	-27.8474	245.73	94.9	95.5	88.8	D

Notes. The columns show: the obsid number; the date of observation start; the pointing coordinates (RA and Dec in deg and in J2000 coordinates; position angle in deg); the exposure times in ks for the MOS1, MOS2 and PN cameras after background filtering; and the boresight group (see Fig. 3). ^(a) Highly flared observations for which the 3σ -clipping failed (see text).

complex has a well-defined spatial pattern with a central region where it is virtually absent².

Throughout this paper, in PN data we have excluded the energy intervals pertaining to this complex. We have chosen different trade-offs for the 2–10 and 5–10 keV bands. For the 2–10 keV band, we have excluded the Cu complex (energy intervals 7.2–7.6 and 7.8–8.2 keV) across all the field, to favour a uniform coverage. For the 5–10 keV band, given the smaller energy band and the lower signal-to-noise ratio (S/N), we have been more aggressive in excluding the Cu complex (energy intervals 7.35–7.60, 7.84–8.28 and 8.54–9.00 keV), but we have done so only in the outer detector regions where the line is present. Our aim for 5–10 keV has been to retain the photons around 8 keV in the centre of the field of view (FOV) at the expense of a less uniform coverage of the field.

2.3. Pointings and astrometry

The XMM-CDFS observations have been taken with some differences in the pointing direction, in order to have a field coverage as uniform as possible with the PN detector. Because

of the field visibility, the observations have been performed in the months of January, February and July. The orbit of *XMM-Newton* modulates the position angle, so that the satellite rotates along its optical axis by about 180 deg every six months. Therefore, when looked at in detector coordinates, observations performed in summer appear “upside-down” with respect to the winter ones. While the MOS detectors are centred on the respective telescope optical axes, the PN is shifted by 2′. The pointing coordinates have thus to be shifted by twice that amount in order to position the PN on the same sky region in the two seasons. This, however, has the effect of making the MOS detectors offset with respect to the nominal CDFS coordinates.

Minor shifts of $\leq 1'$ were added to all observations, to smooth the chip gaps and to avoid the possibility that any source could fall in a gap for a major part of the survey.

On top of these patterns, there are smaller random errors ($\sim 1''$, with 4 obsids in which they were $\sim 4''$) in the pointing direction. These have been estimated by running a source detection on the inner 5′ of each obsid, and cross-correlating the results with the *Chandra* catalogue (X11) using a modified version of *Chandra*’s routine `align_evt`. The event and attitude files have been corrected for these shifts. All coordinates have been aligned to those from X11 (which was itself aligned to VLA radio frame).

² *XMM-Newton* Users Handbook, Issue 2.10, Sect. 3.3.7.2; http://xmm.esac.esa.int/external/xmm_user_support/documentation/uhb/epicintbkgd.html

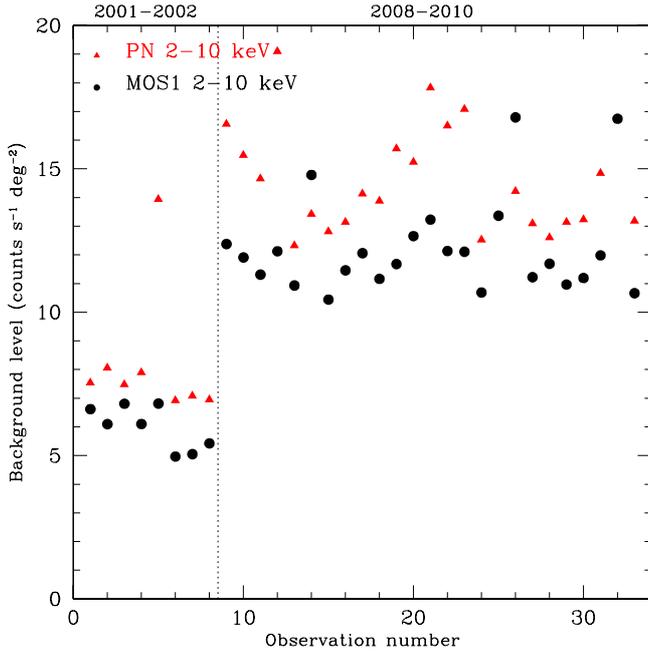


Fig. 1. Total surface brightness in the 2–10 keV band for the 33 obsids of the XMM-CDFS, after excluding the positions around the 30 brightest sources. The increase in background intensity between 2002 and 2008 is evident. The obsids are numbered from 1 to 33 according to observation date; the vertical dotted line marks the separation between the 2001–2002 and 2008–2010 observations. Filled triangles: PN; filled circles: MOS1. The MOS2 camera has values very similar to the MOS1.

The *XMM-Newton* PSF has a $\sim 5''$ FWHM on-axis, which degrades at large off-axis angles³. We measured the observed 2–10 keV FWHM of the 50 brightest sources for the total 3 Ms image and we found that the median FWHM, across the whole field, is $8.5''$.

2.4. Colour image

A colour image of the XMM-CDFS has been derived from the cleaned, astrometry-aligned data. Three images in the 0.4–1, 1–2 and 2–8 keV energy intervals have been produced. To better enhance the sources, the background has been suppressed up to different thresholds in the three bands. Finally, the non-linear scaling and colour-mixing procedure by Lupton et al. (2004), whose main advantage is that bright sources are not “whited-out”, has been applied producing the colour image shown in Fig. 4.

3. Catalogue

The 2–10 and 5–10 keV catalogues have been produced independently. A two-stage process has been used: the first step is a detection with a wavelet code (PWxDetect), followed by maximum likelihood fitting (EMLDetect) of the sources. This detection method can be regarded as a variant of the standard *XMM-Newton* detection process, which uses a sliding box detection algorithm for the first stage instead of the wavelet (see e.g. Brunner et al. 2008 and Cappelluti et al. 2009 for applications to the Lockman Hole and COSMOS surveys, respectively).

³ The *XMM-Newton* Users Handbook. http://xmm.esac.esa.int/external/xmm_user_support/documentation/uhb/XMM_UHB.html

The advantage of using a wavelet-based tool is that it produces more accurate source positions (Puccetti et al. 2009), and that it is especially suited for crowded fields, as happens in a few small areas of the CDFS. Another difference with respect to the Lockman Hole and COSMOS applications is that we chose a higher threshold (i.e., more conservative) for the first stage than for the second (see below); doing so means that the selection is mostly done at the first stage, and that the second is used mainly for the counts estimate.

First, we ran the Palermo Wavelet XMM Detect (PWxDetect) software to identify source candidates. PWxDetect is a version of the PWDetect tool (Damiani et al. 1997), specially tuned for *XMM-Newton* data. It can combine data from the three EPIC detectors and from different observations, even taken with different pointings, provided that the pointing boresights are close enough that a single model for the off-axis angle dependence of the PSF can be used (in practice, this amounts to a few arcmins). We used this feature to perform the detection on all obsids and cameras. We chose a significance threshold for source detection at the 4σ level, corresponding⁴ to a probability of 6.3×10^{-5} for a random fluctuation to be detected as a source (“false-positive”). The first stage source lists include 411 and 196 sources for the 2–10 and 5–10 keV bands, respectively. The source identification numbers (ID210 for 2–10 keV and ID510 for 5–10 keV) were assigned at this stage.

For the second stage we used EMLDetect, a standard tool from the *XMM-Newton* SAS software. Originally developed for ROSAT (Cruddace et al. 1988; Hasinger et al. 1993), the current version of EMLDetect⁵ includes many improvements in terms of PSF models, ability to simultaneously fit data from different cameras and pointings, deblend sources, and fit extended sources⁶.

EMLDetect gives more accurate estimates of net source counts, with lower systematic and random uncertainties (Puccetti et al. 2009). Since EMLDetect was not designed to run on a large number of input files, we grouped the observations in four sets according to the pointing direction (see Table 1), in order to have homogeneous PSFs across a single set. We summed together the images from the MOS1, MOS2 and PN cameras, obtaining four images which were used as the input to EMLDetect. The background maps for EMLDetect were calculated with the method of Cappelluti et al. (2007).

We ran EMLDetect using the PWxDetect coordinates as input and without re-fitting the positions; thus the EMLDetect source list is a subset of the PWxDetect one.

EMLDetect uses a maximum-likelihood estimator to assign a significance value to a detection. The EMLDetect source list was cut at a likelihood value of $L = 4.6$. This likelihood obeys the law $L = -\ln(p)$, where p is the probability of a false-positive detection. The threshold $L = 4.6$ corresponds to a probability of 1.01×10^{-2} . The number of sources detected at the second stage is 337 and 135 for the 2–10 and 5–10 keV bands, respectively.

After visual inspection, we identified two cases of source blending, which correspond to ID210s 228 and 276. These blends were detected in the 2–10 keV band at 29.1σ and 22.3σ

⁴ The correspondence between probability and “number of sigmas” only applies for Gaussian probability distributions. Our use should just be regarded as a handy mnemonic.

⁵ The XMM SAS reference manual for EMLDetect; <http://xmm.esac.esa.int/sas/11.0.0/doc/emldetect/node3.html>

⁶ While the CDFS field contains some extended sources, none of them was present in the PWxDetect output.

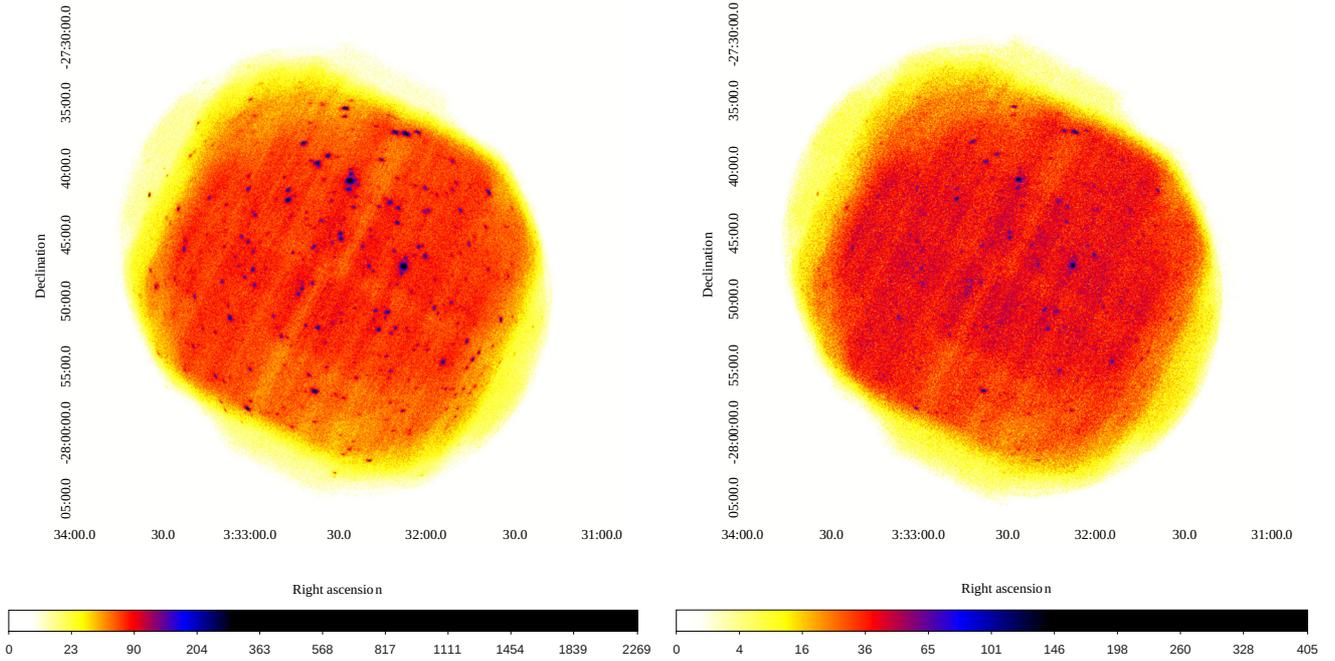


Fig. 2. Images of the XMM-CDFS in the 2–10 (*left*) and 5–10 keV (*right*) bands. The colour wedges show the total counts as the sum of data from the MOS1, MOS2 and PN cameras.

by PWXDetect, and with likelihoods $L = 565.5$ and 340.1 by EMLDetect, respectively. In both cases, two sources whose cores can be visually separated were identified by PWXDetect as a single one. The separation between the two components is $11.5''$ and $18.1''$ for ID210s 228 and 276, respectively (see also Sect. 7). Both blends are detected also in the 5–10 keV band, as ID510 1109 and 1134, respectively. In the catalogue, we do not list the blends; instead we provide counts, rates and fluxes estimated with aperture photometry for the four individual sources, which in the 2–10 keV catalogue are numbered 501, 502 (components of 228) and 503, 504 (components of 276). These sources have been de-blended also in the 5–10 keV band, and their components are numbered with ID510s 1501–1504. The main catalogues are defined as the sources detected at the second stage, minus the two blends, plus the four blend components, and thus contain 339 and 137 sources in the 2–10 and 5–10 keV bands, respectively.

A comparison of the source counts derived by PWXDetect and EMLDetect showed a 1-to-1 correlation, as expected, but with a few outliers. Three of these sources (ID210: 4, 105; ID510: 1001) lie at the south-west rim of the field and are probably just fluctuations in an area with a strong exposure gradient. A few others (ID210: 154, 184, 280, 290, 328, 406; ID510: 1066, 1089, 1100, 1124, 1146) lie in a crowded area, or very close to brighter sources, and they were also flagged as “extended” by EMLDetect. Their EMLDetect-derived counts are about one order of magnitude larger than the PWXDetect estimate, probably because EMLDetect fits them as tails of the PSF of a larger and brighter source. Aperture photometry showed agreement with the counts derived by PWXDetect. Therefore, in the catalogue these 6 sources are listed with the PWXDetect counts; count rates and fluxes were calculated accordingly.

The flux to count rate conversion factors have been calculated for a power-law spectrum with slope $\Gamma = 1.7$ (Mainieri et al. 2007); their values are 1.86×10^{11} and 0.91×10^{11} counts $\text{cm}^{-2} \text{erg}^{-1}$ for the 2–10 and 5–10 keV bands, respectively. Had we used a different slope (for example, 1.4 or 2.0), the conversion factors would differ by 8–9%. The Galactic

column density in the CDFS direction is $8 \times 10^{19} \text{ cm}^{-2}$, so that the bands 2–10 and 5–10 keV are not significantly affected and no correction for Galactic absorption is needed. A histogram of the fluxes of the XMM-CDFS 2–10 keV catalogue is shown in Fig. 5. The number of detected sources drops at fluxes $\leq 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$, with the faintest source at $6.6 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ giving the survey flux limit. For comparison, we also plot the flux distribution from X11 and Brunner et al. (2008). While the *Chandra* 4 Ms survey is deeper as expected, the XMM-CDFS contains a considerable larger number of sources at medium-faint fluxes ($2 \times 10^{-15} - 2 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$), most of which are anyway detected in the ECDFS. The Lockman hole survey has a distribution similar to the XMM-CDFS, though a larger number of sources (mainly at faint fluxes) is detected in the latter than in the former.

3.1. Catalogue significance

It is not easy to compute analytically the joint probability of a false-positive detection in the two stages, though some hints can be obtained using conditioned probability:

$$\mathcal{P}(\text{detection}) = \mathcal{P}(P) \mathcal{P}(E|P) \quad (1)$$

where $\mathcal{P}(P)$ is the probability of a PWXDetect false-positive detection (6.3×10^{-5}), and $\mathcal{P}(E|P)$ is the probability of an EMLDetect false-positive after the PWXDetect selection. The value of $\mathcal{P}(\text{detection})$ can be bounded considering the two extreme cases of $\mathcal{P}(E|P) = 1$ (i.e., accepting all PWXDetect sources) and $\mathcal{P}(E|P) = \mathcal{P}(P) \mathcal{P}(E)$ (i.e., considering the two detection stages as independent experiments; $\mathcal{P}(E) = 1.01 \times 10^{-2}$):

$$\mathcal{P}(P) \mathcal{P}(E) < \mathcal{P}(\text{detection}) < \mathcal{P}(P) \quad (2)$$

which, in terms of σ , give a threshold significance for the catalogues in the interval $4\sigma - 5\sigma$.

The number of spurious detections can be used to obtain an approximation for $\mathcal{P}(E|P)$. This number will be first estimated in

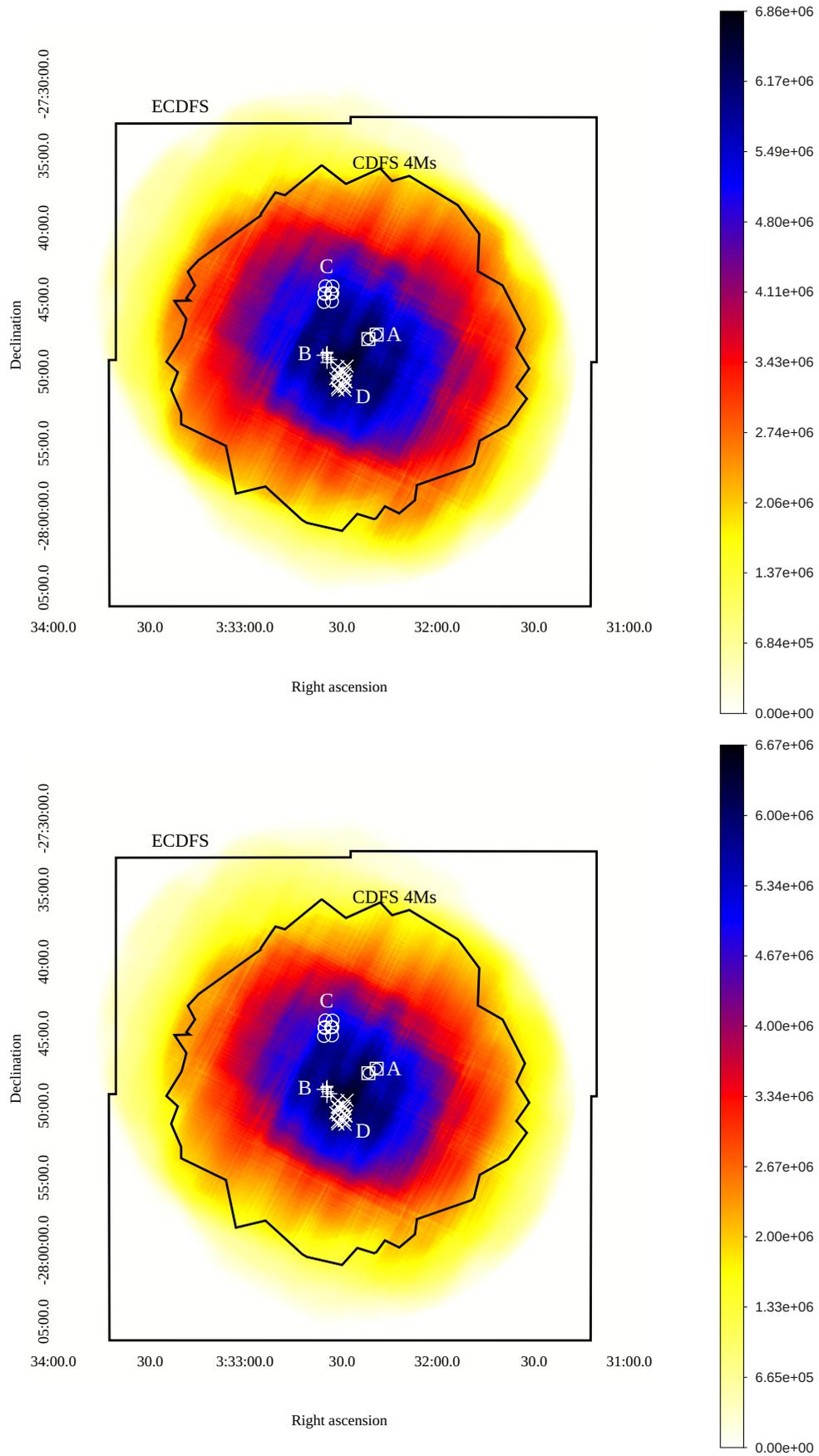


Fig. 3. Exposure maps of the XMM-CDFS in the 2–10 (*left*) and 5–10 keV (*right*) bands. The *Chandra* 4 Ms and ECDFS areas, and the pointing positions and groups are superimposed with different symbols and letter marks according to the observing season (see Table 1). The colour wedges show the exposure time in s as the sum of the MOS1, MOS2 and PN exposures.

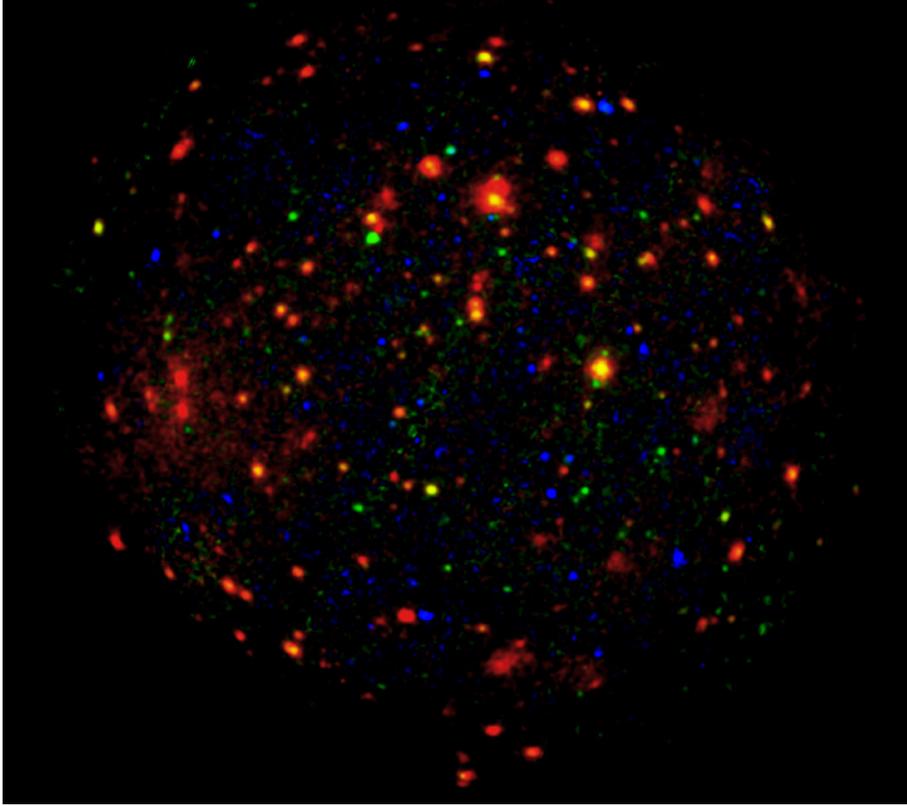


Fig. 4. Colour image of the XMM-CDFS. Red: 0.4–1 keV; green: 1–2 keV; blue: 2–8 keV. The colour scaling is non-linear, and is done first by using a quadratic polynomial to suppress the background, and then by following the Lupton et al. (2004) procedure.

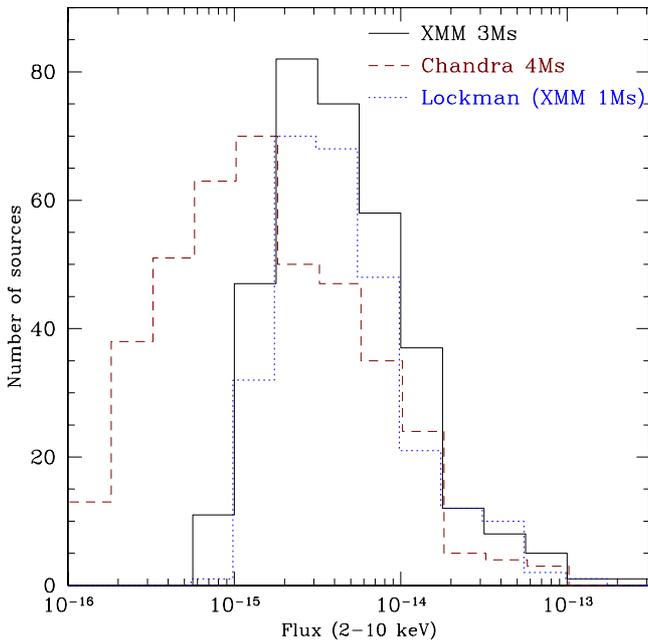


Fig. 5. 2–10 keV fluxes of the sources detected in the XMM-CDFS, compared with other surveys. Solid black histogram: main XMM-CDFS catalogue; dashed red histogram: *Chandra* catalogue (X11); dotted blue histogram: Lockman hole (Brunner et al. 2008).

Sect. 6 using simulations, and then refined in Sect. 8.3 by comparing with the *Chandra* catalogues and by inspecting the S/N of the sources. Therefore, considering 12 (4) spurious sources in the 2–10 (5–10) keV band:

$$\mathcal{P}(\text{detection}; 2-10) \sim 6.3 \times 10^{-5} \times 12/411 \sim 1.8 \times 10^{-6} \quad (3)$$

$$\mathcal{P}(\text{detection}; 5-10) \sim 6.3 \times 10^{-5} \times 4/196 \sim 1.3 \times 10^{-6} \quad (4)$$

which may be quoted as the approximate significance threshold for the catalogues presented here, and roughly correspond to 4.77σ and 4.84σ for the 2–10 and 5–10 keV bands, respectively.

3.2. Main catalogues

The XMM-CDFS 2–10 keV and 5–10 keV catalogues include 339 and 137 sources, respectively, detected in both stages. The catalogues are available in electronic form from the Centre de Données Astronomiques de Strasbourg (CDS), and from the XMM-CDFS website⁷, where we will publish updates to the redshifts whenever they become available. The column description is for both as follows:

1. IAU_IDENTIFIER – source identifier following International Astronomical Union conventions;
2. ID210 (2–10 keV catalogue) or ID510 (5–10 keV catalogue) – PWDetect source number;
3. RA – right ascension (degrees);
4. Dec – declination (degrees);
5. RADEC_ERR – error on position (arcsec; 1σ);
6. COUNTS – EMLDetect sum of the net source counts from MOS1, MOS2 and PN;
7. COUNTS_ERR – error on the COUNTS (1σ);
8. BKG – (EMLDetect) background counts/arcsec²;
9. EXPOSURE – exposure time, averaged over the three cameras;
10. RATE – (EMLDetect) net count rate, summed on the three cameras;
11. RATE_ERR – error on the RATE (1σ);
12. FLUX – (EMLDetect) flux (erg s⁻¹ cm⁻²);
13. FLUX_ERR – error on the FLUX (1σ);

⁷ <http://www.bo.astro.it/xmmcdfs/deeprime/>

14. SIGNIFICANCE – (PWxDetect) detection significance (no. of sigma);
15. DET_ML – (EMLDetect) detection likelihood;
16. DET_SCALE – (PWxDetect) wavelet detection scale;
17. CID – *Chandra* source number (see Sect. 8). If the letter E appended to the number, then the ID is from Lehmer et al. (2005); otherwise, it is from X11;
18. REDSHIFT;
19. REDSHIFT_REF – reference for choice of redshift;
20. ID510 (5–10 keV catalogue) or ID210 (2–10 keV catalogue) – source number in the other band;
21. NOTES – mainly about source blends and counterparts.

3.3. Supplementary catalogues

The catalogues of the 74 2–10 keV, and 61 5–10 keV sources detected only in the first stage are also available in electronic form. These supplementary sources are on average detected at low significance; many of them are on the borders of the FOV; and a few are in crowded fields where EMLDetect has trouble separating the different PSF components. Nonetheless, 4 of these sources are bright enough that a spectrum could be extracted (Comastri et al., in prep.).

The columns whose values in the main catalogue were derived with EMLDetect have been replaced, where possible, with analogous quantities derived with PWxDetect. The main differences between the output columns of PWxDetect and EMLDetect are *i*) that PWxDetect rescales the MOS counts to the PN response before summing them, while EMLDetect does not, and *ii*) that PWxDetect sums the exposure times while EMLDetect averages them. For consistency, we renormalized the source counts and exposure times to the same scale of EMLDetect.

The column description is as follows:

1. IAU_IDENTIFIER – source identifier following International Astronomical Union conventions;
2. ID210 (2–10 keV catalogue) or ID510 (5–10 keV catalogue) – PWxDetect source number;
3. RA – right ascension (degrees);
4. Dec – declination (degrees);
5. RADEC_ERR – error on position (arcsec; 1σ);
6. COUNTS – net source counts from MOS1, MOS2 and PN;
7. COUNTS_ERR – error on the COUNTS (1σ);
8. BKG – background counts/arcsec²;
9. SUMMED_EXPOSURE – exposure time, summed on the three cameras;
10. RATE – net count rate;
11. RATE_ERR – error on the RATE (1σ);
12. FLUX – flux (erg s⁻¹ cm⁻²);
13. FLUX_ERR – error on the FLUX (1σ);
14. SIGNIFICANCE – detection significance (no. of sigma);
15. DET_SCALE – wavelet detection scale;
16. CID – *Chandra* source number (see Sect. 8). If the letter E appended to the number, then the ID is from Lehmer et al. (2005); otherwise, it is from X11;
17. REDSHIFT;
18. REDSHIFT_REF – reference for choice of redshift;
19. ID510 (5–10 keV catalogue) or ID210 (2–10 keV catalogue) – source number in the other band;
20. NOTES – mainly about source blends and counterparts.

4. Coverage and number counts

The survey coverage has been estimated with extensive simulations, which are described in Sect. 5. In Fig. 6 we show the coverage in deg², defined as the average fraction of sources per simulation with a successful match between the output and input catalogues, times the surveyed area, as function of the flux. The confidence intervals for the coverage, estimated according to the Fieller (1954) formula for the ratio of two average quantities, are of the order of the line width and thus they are not plotted. The coverage is normalized to 0.2466 deg², i.e. the area with exposure $\geq \min(T)$, where T is the exposure column in the 2–10 keV catalogue.

In Fig. 6 we also show, for comparison, the coverage for the *Chandra* 4 Ms survey in the 2–10 keV band (X11) and that for the *XMM-Newton* Lockman Hole 1 Ms survey for the 5–10 keV band (Brunner et al. 2008).

The number counts can be derived from the 2–10 and 5–10 keV catalogues and the coverage curves derive above. The cumulative number counts are the sum of the inverse areas in which the sources could be detected:

$$N(>S) = \sum_{S_i > S} \frac{1}{A_i} \quad (5)$$

where S_i and A_i are the flux and area for a given source i . With this definition, cumulative counts are not binned.

For the differential counts, we bin the sources according to their fluxes, and define

$$n(S) = \frac{1}{\Delta S} \sum \frac{1}{A_i} \quad (6)$$

where the sum is performed on all sources with flux $S_i \in [S - \Delta S/2, S + \Delta S/2]$. The Poissonian error on the counts in any bin can be computed with the Gehrels (1986) approximation.

The number counts are shown in Fig. 7 (upper panels: cumulative number counts; lower panels: differential). The cumulative and differential number counts observed by *Chandra* (Lehmer et al. 2012) and the cumulative model for AGN (Gilli et al. 2007) are also shown. The *Chandra* number counts have been converted from the 2–8 and 4–8 keV bands to the 2–10 and 5–10 keV bands, respectively, assuming the same spectrum used for the XMM-CDFS catalogue.

In both bands there is, as expected, a near-perfect agreement between the *Chandra* and *XMM-Newton* estimates. Some low-significance deviations at fluxes around 6×10^{-15} – 10^{-14} (2–10 keV) and 2 – 6×10^{-15} erg s⁻¹ cm⁻² (5–10 keV) might be due to the larger FOV of the XMM-CDFS with respect to *Chandra*.

A cross-correlation analysis between the *Chandra* ACIS-S3 detector chip and the three *XMM-Newton* cameras in the 2–8 keV band (Tsujimoto et al. 2011) showed that *Chandra* finds fluxes which are larger than the *XMM-Newton* ones by 10–20%. The 4 Ms survey however was performed with the ACIS-I detector, which was not analysed by Tsujimoto et al. (2011). If the ACIS-S3 results also hold for the ACIS-I detector, than the larger fluxes found by *Chandra* would be consistent with the minor deviations found in the Log N –Log S . As to the 5–10 keV band, Fig. 7 (upper right panel) would suggest that any deviation would go in the opposite direction (*Chandra* fluxes being lower than the *XMM-Newton* ones); however the very hard band was not considered by Tsujimoto et al. (2011).

In both bands the number counts are also in agreement with the model.

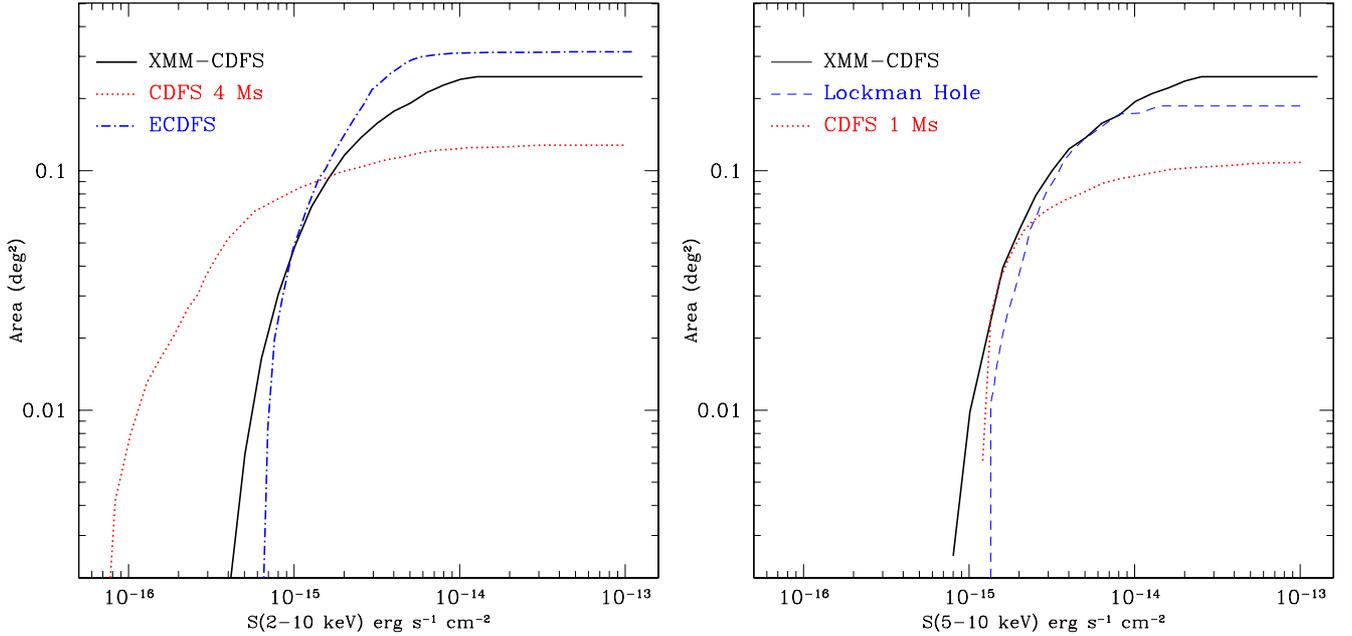


Fig. 6. Coverage of the XMM-CDFS survey (solid black lines) in the 2–10 keV (*left panel*) and 5–10 keV band (*right panel*), defined as the fraction of simulated sources which were successfully detected as a function of their output flux, and normalized to the area probed by the catalogues. The dotted red lines show, for comparison, the same quantity for *Chandra* (*left panel*: 4 Ms survey, X11; *right panel*: 1 Ms survey, Rosati et al. 2002). The dashed blue line in the *right panel* shows the coverage for the *XMM-Newton* Lockman Hole (Brunner et al. 2008). We have not shown in the *right panel* the CDFS 4 Ms (Lehmer et al. 2012) since formally it is in a softer band (4–8 keV).

The Lockman Hole was also observed with *XMM-Newton*; Brunner et al. (2008) presented a 5–10 keV Log N –Log S which is in agreement, within poissonian errors, with the XMM-CDFS. At brighter fluxes, we compare with the counts from the Hellas2XMM (Baldi et al. 2002) and 2XMM (Mateos et al. 2008) surveys; all of them are in agreement, within errors, with the XMM-CDFS. Poissonian errors are plotted only for the CDFS and Hellas2XMM, to avoid cluttering the figure.

One interesting thing to note is that in the interval 7×10^{-15} – 1.2×10^{-14} , the CDFS and the XMM-CDFS 5–10 keV number counts stay below all other determinations. This may be a low-significance feature of the CDFS field (though formally all number counts are consistent with each other within Poissonian errors).

Finally, we show the differential number counts derived by Georgakakis et al. (2008) from a combination of deep (CDFS and *Chandra* Deep Field North) and shallow surveys. Georgakakis et al. (2008) report as their best-fit model a broken power-law of the form:

$$\frac{dN}{dS} \propto \begin{cases} S^{\Gamma_1}, & S \leq S_{\text{break}} \\ S^{\Gamma_2}, & S \geq S_{\text{break}} \end{cases} \quad (7)$$

with slopes $\Gamma_1 = -1.56 \pm 0.04$ and $\Gamma_2 = -2.52^{+0.07}_{-0.09}$, and with a break flux $\text{Log}(S/\text{erg s}^{-1} \text{cm}^{-2}) = -13.91^{+0.08}_{-0.05}$ for the 2–10 keV band ($-1.70^{+0.08}_{-0.06}$, $-2.57^{+0.07}_{-0.09}$, $-14.09^{+0.08}_{-0.05}$ respectively for the 5–10 keV band). The XMM-CDFS data points are consistent within 1 or 2σ with the broken power-law model. A binned χ^2 fit yields for the XMM-CDFS (quoted errors are 1σ): $\Gamma_1 = -0.85 \pm 1.2$, $\Gamma_2 = -2.29 \pm 0.36$ and break flux $\text{Log}(S/\text{erg s}^{-1} \text{cm}^{-2}) = -14.46 \pm 0.30$ for the 2–10 keV band, and -0.77 ± 2.6 , -2.49 ± 0.48 and -14.47 ± 0.41 for the 5–10 keV band. The best-fit parameter are in agreement, within errors, with the Georgakakis et al. (2008) ones, with the possible exception for the 2–10 keV break flux which differs by 2σ . This

may be the result of the smaller number of data points at fluxes fainter than the break in the XMM-CDFS than in Georgakakis et al. (2008), and of the larger error on the faintest data point.

5. Simulations

We have performed extensive simulations to address the XMM-CDFS properties in terms of survey coverage and source confusion. For each simulation we built a mock survey, and ran the two-stage detection process.

We have used a simulator developed with the aim to reproduce the background and the PSFs of any given observation as closely as possible. Though currently tuned for *XMM-Newton*, the simulator can be adapted to other missions. While more details are given in the Appendix, the main features are:

- the simulator attempts to reproduce the source and background distribution for each of the 99 exposures (33 obsids times 3 cameras) of the XMM-CDFS independently;
- the background level is decomposed into its constituents (cosmic, particle and residual soft protons), whose spatial distributions are reproduced separately;
- the cosmic X-ray background is modelled after Gilli et al. (2007); the particle and proton backgrounds are resampled from maps distributed as part of the *XMM-Newton* calibration;
- the distribution of the XMM-CDFS background counts is reproduced by the simulator with an error of $\sim 10\%$;
- the two-stage source detection is run on each simulation;
- a sizable number of simulations is run so that average quantities (number of sources, coverage, number of confused sources) can be derived.

Given the long exposure time (~ 3 Ms) of the XMM-CDFS survey, the average number of counts per pixel is much larger than

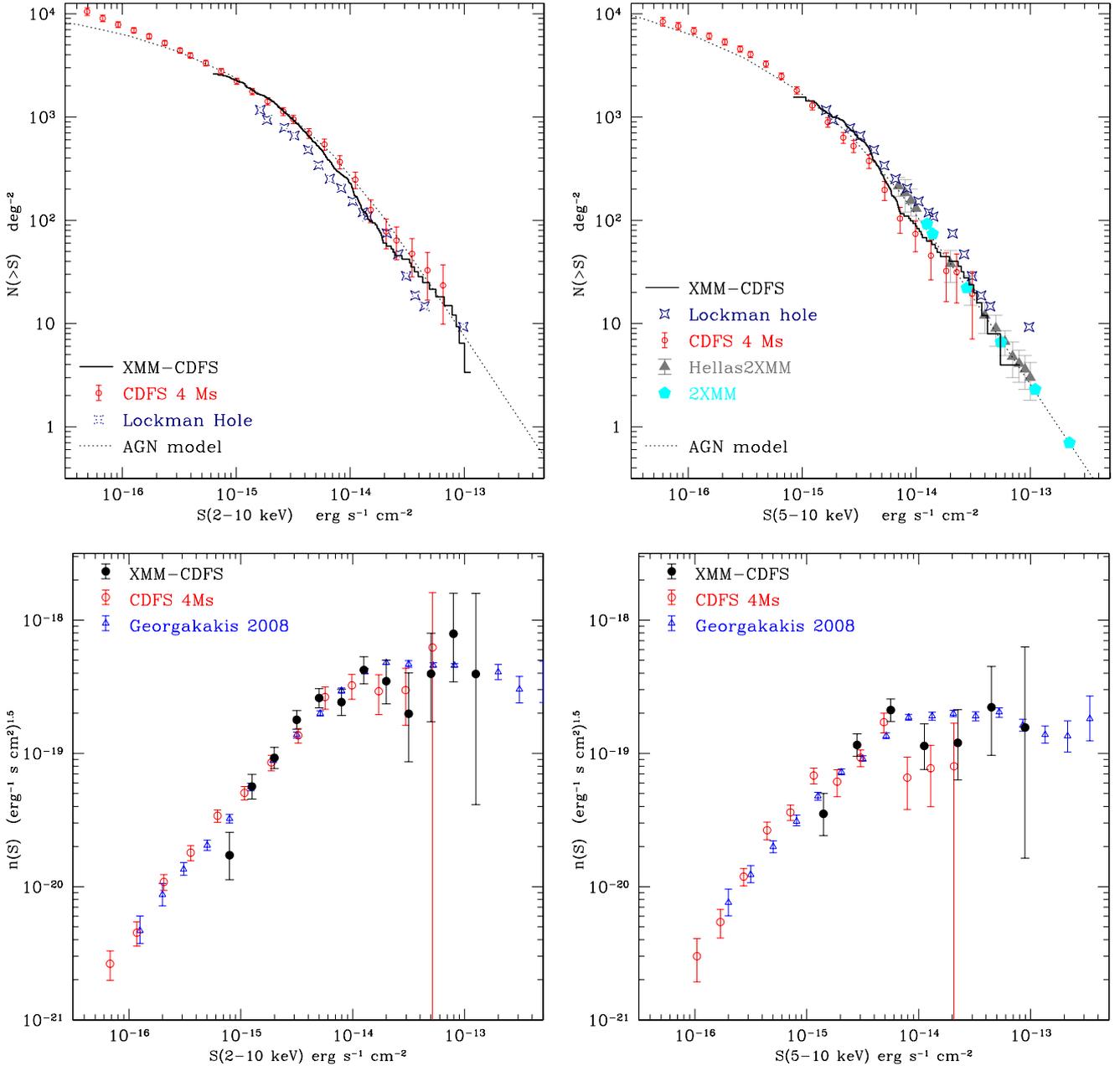


Fig. 7. *Upper panels:* cumulative number counts in the 2–10 keV (*left*) and 5–10 keV (*right*) bands. *Lower panels:* differential number counts (*left*: 2–10; *right*: 5–10 keV band) in Euclidean scale (i.e., $S^{2.5}dN/dS$). For comparison, in *all panels* we plot the *Chandra* number counts (Lehmer et al. 2012, red circles). In the *upper panels*, we also show the cumulative number counts from the *XMM-Newton* survey in the Lockman Hole (Brunner et al. 2008, blue stars), and the model cumulative counts for AGN (Gilli et al. 2007, black dotted line). In the *upper right panel*, we also plot the number counts from the Hellas2XMM (Baldi et al. 2002) and 2XMM (Mateos et al. 2008). In the *lower panels*, we also show the differential number counts from Georgakakis et al. (2008).

that for other *XMM-Newton* survey projects. Because of the difference in the pointing positions among the 33 obsids, the background counts have a complex spatial pattern, mainly given by the chip gaps and borders occurring in slightly shifted positions. In any *XMM-Newton* observation, the following background components are present in addition to the detectable sources:

- cosmic X-ray background (CXB), produced by the many sources which are below the survey flux limit;
- particle background (PB; sometimes also called “non X-ray background”), which basically includes the electronic noise

and a few fluorescent lines produced by the telescope assembly;

- residual soft protons (SP): a quiescent background due to soft proton clouds;
- solar wind charge-exchange, negligible at energies ≥ 2 keV and therefore not considered here.

The modelling of each simulated component will be discussed in detail in the next sections. Our treatment of the different background components builds directly on the work by De Luca & Molendi (2004), Snowden et al. (2004, 2008); see also Snowden & Kuntz (2012).

5.1. X-ray sources and cosmic background

For each simulation, an input catalogue of cosmic sources is generated from a $\text{Log } N\text{--}\text{Log } S$ obtained as the sum of: i) the 2–10 keV $\text{Log } N\text{--}\text{Log } S$ of AGN, calculated with the [Gilli et al. \(2007\)](#) model over its full range of parameters, and ii) the 2–10 keV galaxies number counts computed from the [Ranalli et al. \(2005\)](#) model. This $\text{Log } N\text{--}\text{Log } S$ extends from 2.8×10^{-20} to $1.0 \times 10^{-11} \text{ erg s}^{-1} \text{ cm}^{-2}$, thus allowing simulation of both the detectable sources, and the cosmic X-ray background. It is sampled on a square area slightly larger than the CDFS, assuming uniform distributions in RA and Dec, thus giving a number of $\sim 50\,000$ sources per simulation.

Most of these sources are too faint to contribute, on average, even a single photon to an observation; however when they are taken together, the sources below the survey flux contribute to the total (cosmic+particle+proton) observed background.

The flux to count rate conversion factors have been calculated for a power-law spectrum with slope $\Gamma = 1.7$, as done for the real catalogue (considering $\Gamma = 1.4$ or 2.0 would change the fluxes by $\sim 8\%$); in doing this, we have assumed a single average spectral slope for all sources, which is acceptable since we are interested only in broad-band detections.

5.2. Particle background

The PB is mainly due to the instrument electronics and to reflection lines from the telescope and its mounting. The spatial distribution of the PB has been taken from the filter wheel closed (FWC) observations provided by the ESAS CALDB⁸. The FWC were made with the filter wheel in the closed position, and without the calibration source on⁹.

The actual PB level in any observation may be determined by considering the corner areas, i.e., the parts of the MOS and PN detectors which lie outside of the FOV and are neither exposed to the CXB nor to the protons. The data for the corner areas are normally filtered out by the *XMM-Newton* pipeline, but can be obtained by reprocessing the “raw data” (i.e., the observation data files, ODF). The surface brightness of the corner areas needs to be appropriately rescaled to obtain the PB brightness inside of the FOV.

Let us define the total average surface brightnesses which are observed inside and outside of the FOV as Σ_{in} and Σ_{out} , respectively. The value of both Σ s is different for any obsid and camera, and can also be measured for the FWC. To minimize uncertainties due to small number statistics for bright sources, the positions corresponding to the 30 brightest sources have been masked out during the determination of the above surface brightnesses. This corresponds to removing sources brighter than the $2 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$ threshold; at this flux level, 43% of the background is resolved (using the HEAO-1 value, [Gruber et al. 1999](#)).

On the one hand, the FWC only contain PB, therefore:

$$\begin{cases} \Sigma_{\text{in,FWC}} = \Sigma_{\text{PB,FWC}} \\ \Sigma_{\text{out,FWC}} = k \Sigma_{\text{in,FWC}} = k \Sigma_{\text{PB,FWC}} \end{cases} \quad (8)$$

where k accounts for spatial variations of the PB surface brightness.

⁸ The Extended Source Analysis Software (ESAS), developed by S. Snowden and K. Kuntz, is a part of the standard *XMM-Newton* SAS software and includes several calibration files.

⁹ Conversely, the more common CALCLOSED observations do have the calibration source visible, in addition to the PB.

Table 2. Average percentage contribution of the different background components to the overall level.

Camera	Band (keV)	CXB	Particle	Proton
MOS1	2–10	1.1 ± 0.4	81 ± 10	19 ± 10
MOS2	2–10	1.3 ± 0.4	80 ± 9	18 ± 9
PN	2–10	2.2 ± 0.8	90 ± 6	7.4 ± 6.0
MOS1	5–10	0.5 ± 0.2	88 ± 8	12 ± 8
MOS2	5–10	0.6 ± 0.2	87 ± 8	12 ± 8
PN	5–10	1.5 ± 0.5	95 ± 4	$3.9_{-3.9}^{+4.4}$

Notes. All values are per cent. When calculating the brightness levels, the areas around the 30 brightest sources have been excluded. Thus, the cosmic background refers to sources with fluxes $\leq 2 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$.

On the other hand, for a generic observation, the FOV will contain PB, proton and cosmic background components, while the corners will only have PB:

$$\begin{cases} \Sigma_{\text{in}} = \Sigma_{\text{PB}} + \Sigma_{\text{SP}} + \Sigma_{\text{CXB}} \\ \Sigma_{\text{out}} = k \Sigma_{\text{PB}} \end{cases} \quad (9)$$

where Σ_{PB} , Σ_{SP} , and Σ_{CXB} are surface brightnesses.

The important assumption here is that the k is the same for the FWC and for all the observations, which is equivalent to assuming that the spatial distribution of the PB is the same in all observations. Taking k from Eq. (8), from Eq. (9) we get

$$\Sigma_{\text{PB}} = \Sigma_{\text{out}} \frac{\Sigma_{\text{in,FWC}}}{\Sigma_{\text{out,FWC}}} \quad (10)$$

The average fractions of PB with respect to the total of CXB+PB+protons are reported in Table 2. The number of PB photons to be simulated is $\Sigma_{\text{PB}} \times A \times E$ where A is the FOV area and E the exposure time.

5.3. Soft protons

Clouds of soft protons are often encountered by *XMM-Newton*, especially when the satellite is close to perigee, and cause strong background flares. When such a flare is detected, the time intervals in which it occurs are usually excluded from further analysis. However, there is also a quiescent component of soft proton background, with much smaller brightness, that may occur for the entire length of an observation, and for which the standard recipe of identifying and excluding high-background periods is not applicable – or which may remain even after the flares are excluded. The quiescent soft protons are subject to the telescope vignetting, though to a lesser extent than X-ray photons (see Fig. 17 in [Kuntz & Snowden 2008](#)). Both [De Luca & Molendi \(2004\)](#) and [Snowden et al. \(2004\)](#) have considered this problem and we refer to these papers for further details.

Here we estimate the residual soft proton level from Eq. (9):

$$\Sigma_{\text{SP}} = \Sigma_{\text{in}} - \frac{\Sigma_{\text{out}}}{k} - \Sigma_{\text{CXB}} \quad (11)$$

which is also shown in Fig. A.1 and Table 2.

We assume the value $\Sigma_{\text{CXB}} = 7.12 \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ deg}^{-2}$ for the CXB brightness, which results from integrating the AGN and galaxy number counts from zero to the $2 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$ threshold (Sect. 5.2). The number counts are normalized to the HEAO-1 value for the X-ray background ([Gruber et al. 1999](#)); had we normalized to any of the higher values derived

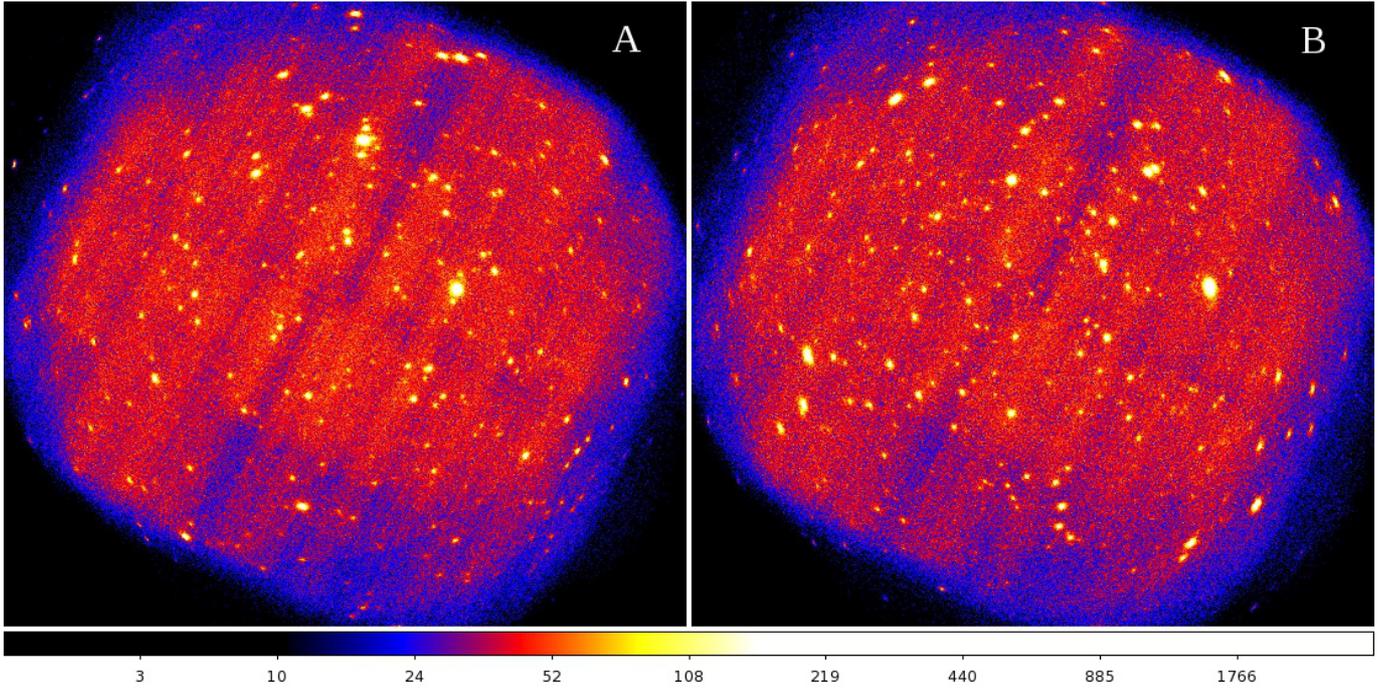


Fig. 8. Comparison of the real XMM-CDFS A) with one simulation B). The colour scale is the same for both panels and maps the 2–10 keV intensity.

from *XMM-Newton* data (e.g., Lumb et al. 2002 or De Luca & Molendi 2004), this number would have been proportionally larger. However, the increase would have been compensated by a decrease in the soft proton component, with the net effect of a slight change in the amount of background vignetting. This effect is probably within the uncertainties with which the background brightness is measured.

We now can determine the number of photons contributed by the soft protons as $\Sigma_{\text{SP}} \times A \times E$, where A and E are again the FOV area and the exposure time.

A simulation produced with this recipe is shown in Fig. 8, where an image of the real XMM-CDFS is shown with the same colour scale for comparison.

5.4. Simulated fields and catalogues

A comparison between the distribution of counts reproduced by the simulator and that of the XMM-CDFS is shown in Fig. 9, where the histograms of the counts per arcsec² are plotted in the 2–10 and 5–10 keV bands. The counts include both the background (mainly the peak of the distribution) and the sources (the right tail). In both bands, the agreement seems good, with a horizontal shift between simulations and the XMM-CDFS of $\sim 10\%$ around the peaks.

We ran about 150 simulations in the 2–10 and 5–10 keV bands, respectively, to assess the survey properties. For each simulation, the two-stage detection process was performed in the same way as for the (real) XMM-CDFS main catalogue. The output catalogues were cross-correlated with the (input) list of simulated sources. For each detected source, its input counterpart is defined as the one which minimizes the score S (Cappelluti et al. 2009):

$$S = \left(\frac{\Delta \text{RA}}{\sigma_{\text{RA}}} \right)^2 + \left(\frac{\Delta \text{Dec}}{\sigma_{\text{Dec}}} \right)^2 + \left(\frac{\Delta \text{RATE}}{\sigma_{\text{RATE}}} \right)^2 \quad (12)$$

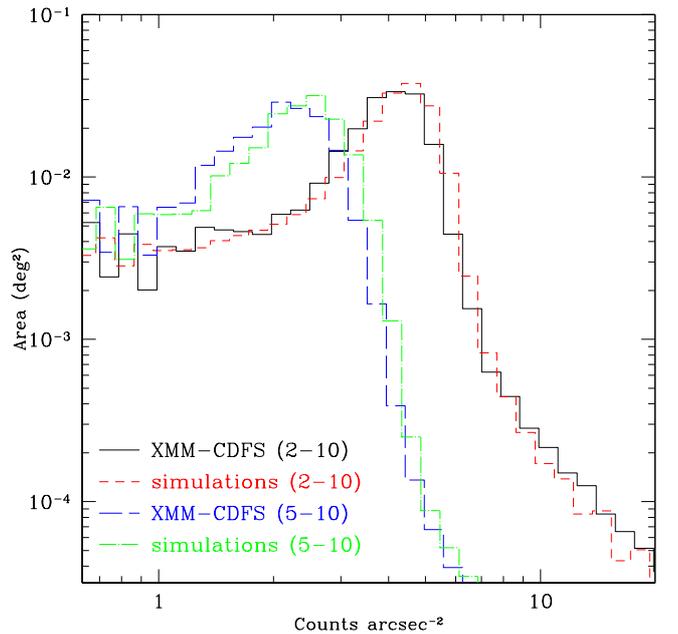


Fig. 9. Distribution of the counts per arcsec² in the simulations (average over 5 simulations), compared with the XMM-CDFS, for the 2–10 and 5–10 keV bands. The vertical axis shows the area covered by each bin.

where $\Delta(\text{RA}, \text{Dec}, \text{RATE})$ are the differences between the output and input right ascension, declination, and count rate, respectively, and the σ are the corresponding errors.

We show in Fig. 10 the histogram of the spatial separation between the output and input positions. Most sources show a displacement of $\sim 1.25''$, which was already highlighted by Cappelluti et al. (2009) and Pineau et al. (2011), and therefore it seems to be a characteristic of *XMM-Newton*. The displacement does not occur in any particular direction and occurs even

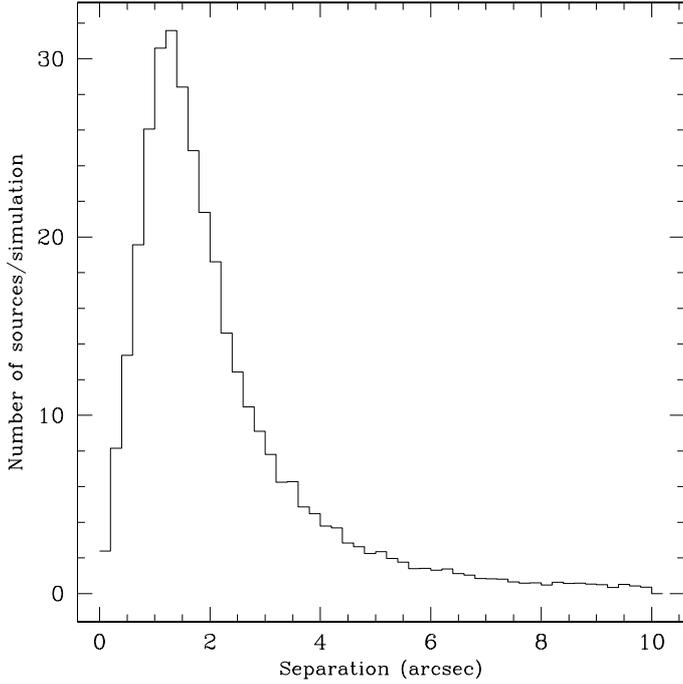


Fig. 10. Spatial separation between the input and output positions of simulated sources.

if the source detection is run on just one single camera, or on just one obsid. The number of matches drops sharply for displacements $\geq 2''$.

The input and output fluxes of the simulated sources are shown in Fig. 11; only sources with displacements $\leq 6''$ are considered here. A few outlier per simulations are present, which is compatible with the amount found in the XMM-CDFS (Sect. 3). These outliers are mainly due to EMLDetect identifying some sources, especially in crowded areas, as extended sources. In the compute of the coverage, we identify the outliers has having an EMLDetect flux larger than 3 times the PWXDetect flux; the latter is then used for the calculation. At bright fluxes ($\geq 10^{-14}$ erg s $^{-1}$ cm $^{-2}$), the output values are systematically larger by about 20% than the input ones. This uncertainty does not affect the flux range where most of the sources (82%) are detected.

6. Number of spurious sources

We estimate the number of spurious sources in the 2–10 keV band by running the detection process on the simulated observations. The total number of sources detected in the simulations is reported in Table 3. In the 2–10 keV band, the simulations produce catalogues which contain on average 358 ± 16 src sim $^{-1}$, i.e. a larger number of sources than the XMM-CDFS, though the difference is less than twice the simulations' standard deviation; in around 10% of the simulations a number of sources lower than or equal to the XMM-CDFS is detected. The number of detected sources is in general dependent on the background level; residual variations between the simulations and XMM-CDFS background (see Sect. 5.4) may help to explain the difference observed in the 2–10 keV band.

In the *Chandra*-COSMOS survey, which used the same detection software as this paper, Puccetti et al. (2009) found that a large number of spurious sources, only present in simulations, were characterized by source sizes smaller than the PSF sizes – a clear indication of fluctuations detected as sources – and removed them. PWXDetect uses the extent column to express

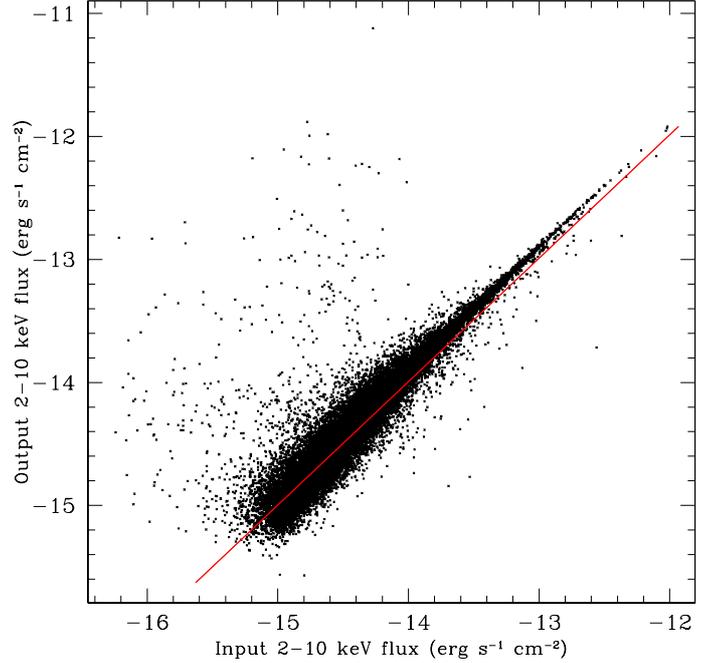


Fig. 11. Input vs. output fluxes of the simulated sources, from 100 2–10 keV simulations. Only sources with a match within $6''$ are considered. Some outliers are present, due to EMLDetect incorrectly associating the source with extended emission especially in crowded regions, and for which we rather use the PWXDetect fluxes; the number of outliers is a few per simulation. The red line shows the identity relation. At bright fluxes ($\geq 10^{-14}$ erg s $^{-1}$ cm $^{-2}$), the output values are about 20% larger than the input ones.

Table 3. Number of spurious sources in the 2–10 keV band simulations.

Exp.	CDFS	Total sim.	W.count.	Spurious	Fract.
any	335	358 ± 16	328 ± 17	30 ± 5	8.4%
≥ 300	302	306 ± 15	296 ± 15	10 ± 3	3.3%

Notes. The column “CDFS” gives the number of sources found in the XMM-CDFS which respect the same selection of the simulations (i.e. extent <0.5 and no de-blending; see text). The columns “Total sim.”, “W.count.”, “Spurious” give the average and standard deviation of the number of sources per simulation, the number of sources with a counterpart in the input catalogues, and the number of spurious, respectively. The column “Fract.” gives the fraction of spurious sources over the total number. Exposure thresholds are expressed in ks.

the source size in “PSF units” (i.e., adimensional); in particular, all sources with extent <0.5 are likely to be spurious. Thus, in all simulations we filtered the catalogue to remove source candidates satisfying this criterion.

Spurious sources may arise mainly from background fluctuations, or from an incorrect reproduction of the spatially-dependent background level. For instance, the simulations made use of spatial maps of the soft proton background, which were derived by time filtering of highly flared observations, and may still contain small local enhancements originally due to astronomical sources. Local enhancements are also present in the particle background maps, and further investigation would be needed to determine whether these are representative of any given observation. In general, any background enhancement in the proton and particle maps is likely to appear in many simulations, since these maps were used as probability distributions

without smoothing on significantly large scales (e.g., larger than a few PSF cores). The number of this kind of spurious sources is non-trivial to estimate; for this reason, the numbers of spurious sources quoted in Table 3 should be regarded as upper limits if applied to the XMM-CDFS; and in Sect. 9 the comparison of the *Chandra* and *XMM-Newton* catalogues will be used to further refine the estimate of the spurious fraction.

A check on the position of the candidate spurious sources shows that about half of them (i.e. 20 ± 4 in the 2–10 keV band) lie in the outer parts of the field, where the exposure is ≤ 300 ks. Under the latter threshold, the number of sources detected in the simulations with a counterpart in the input catalogues, and the number of sources in the XMM-CDFS, match more closely.

The main reason for the larger fraction of spurious sources at low exposure (≤ 300 ks) is that fluctuations are more important at the borders of the FOV than in the inner area – or, in other words, that fluctuations are “averaged out” when the exposure is large. The spatial distribution of spurious sources has well-defined peaks, falling in the low-exposure areas, which can be traced back to localized high background features.

For the rest of this Sect., we define spurious sources as those detected in the simulations, but without any counterpart in the input catalogues within $6''$ from their position; the numbers are reported in Table 3. The threshold of $6''$ was chosen after the check for *Chandra* counterparts (Sect. 8.2), in which a tail of reliable associations is present up to $6''$ especially for sources at large off-axis angles. For the 2–10 keV, the average number of spurious sources is 30 ± 5 src sim $^{-1}$. The number of XMM-CDFS sources to which the comparison can be done is 335 (337 sources detected at the second stage, without deblending sources no. 501–504, and with 2 sources not considered because they have extent < 0.5).

Confused sources may also give rise to spurious sources, in the case that a pair of input sources is detected as a single one, and that the detected position is distant more than $6''$ from both the components of the input pair (see Sect. 7, “double-detection pairs”). The number of this kind of spurious sources is ~ 3 per simulation, and does not depend on the position in the field.

A check on the number of spurious sources, based on the probability of association with *Chandra* counterparts and on the S/N in the *XMM-Newton* images, is discussed in Sect. 8.2.

The fraction of spurious sources is higher in the 5–10 keV band than in the 2–10 keV, probably because localized background features contribute with a larger proportion to the total counts. Improving this would require a deeper study of the *XMM-Newton* background and an update of the simulator, which we defer to a future paper.

7. Source confusion

Because of the size of the XMM-CDFS average angular resolution ($8.5''$ FWHM, which is the median across the FOV; see Sect. 2.3) and the faint flux limit of the XMM-CDFS, it is important to understand how the detection process treats close sources. Depending on the distance of its components, a pair of sources may be mistakenly detected as a single source (“source confusion”); the off-axis angle dependence of the PSF and local small-scale variations of the background also play a role. This may also impact the estimated number of spurious sources, if the detected position is distant enough from that of the pair components. Here, we use the simulations to estimate the amount of source confusion in the XMM-CDFS survey, and compare at end with *Chandra* data. In this Section, we only consider

the 2–10 keV band; the 5–10 keV range is expected to be less affected by source confusion.

We start our analysis by considering input sources with fluxes greater than 10% that of the faintest source detected in the real XMM-CDFS, and scaled for lower sensitivity areas through an exposure map (i.e., $0.1 \times 6.6 \times 10^{-16} / \text{Exp}(\text{RA}, \text{Dec})$ erg s $^{-1}$ cm $^{-2}$, where $\text{Exp}(\text{RA}, \text{Dec})$ is normalized so that $\max(\text{Exp}(\text{RA}, \text{Dec})) = 1$). This avoids cluttering the results with sources too faint to have a significant impact on the detection process. For simplicity, we also do not consider triplets of sources, nor groups with a larger number of members. With these constraints, we identify an average of 61 (108) pairs per simulation with separation $\leq 10''$ ($\leq 15''$) in the input catalogues.

The number of pairs per simulation, as function of the flux ratio between the two components of the pair, is shown in Fig. 12 (left panel) as the solid histogram (“input pairs”). Most pairs have a flux ratio ≤ 10 between the two components. Pairs with a larger difference between the fluxes of the components are considerably more rare.

The right panel of Fig. 12 shows the number of pairs as function of the separation between the pair components. To be more representative of what is observable in practice in the XMM-CDFS, we restrict this panel to pairs with a flux ratio < 3 , and we also drop any pair with a component fainter than the XMM-CDFS flux limit (6.6×10^{-16} erg s $^{-1}$ cm $^{-2}$). In the following we will refer to this sub-selection as the “detectable pairs”. The number of pairs with very small separation ($\leq 2''$) is quite small (~ 0.4 pairs/sim). It grows rapidly at larger separations, stabilizing at $\geq 8''$ with a value of 1.7 pairs/sim/arcsec (i.e., *every arcsec of separation adds 1.7 pairs*).

We consider the following possibilities for the detection of a pair of input sources:

1. detected as two distinct sources;
2. detected as one single source.

7.1. Pairs detected as two distinct sources

We have searched for source pairs where both components are individually detected in the output catalogues of simulated sources (“double-detections”).

The number of double-detection pairs, when seen as a function of the pair flux ratio (Fig. 12, left panel, dashed histogram), roughly follows the number of input pairs, though it stays lower by ~ 1.5 orders of magnitude and has a sharper drop (a factor 10 vs. a factor 2 for the double-detections and input pairs, respectively, at flux ratios < 10). Thus, double detections are more probable when the flux ratio is small.

While the number of detectable input pairs rises with the separation from 0 to $8''$ and remains constant afterwards, the number of double-detections is consistent with 0 up to $11''$ and only starts to rise afterwards (Fig. 12, right panel). This means that input pairs are not recognised as such by the detection process if the separation is $\leq 11''$.

7.2. Pairs detected as a single source

Input pairs detected as a single output source are what is usually referred to as confused sources. We have searched for them by requiring: i) one output source within $6''$ from the pair centroid; ii) no other output sources within $6''$ from the positions of the two input sources. The pair centroid is defined as the weighted mean of the pair coordinates, using the count rates

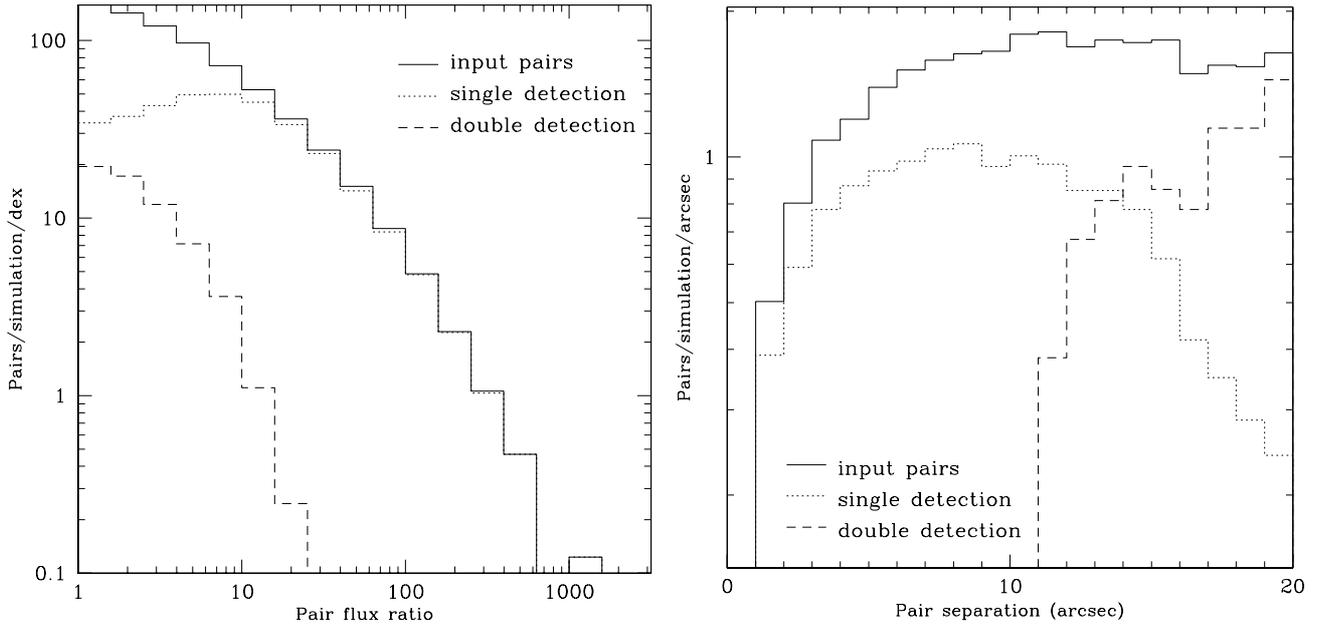


Fig. 12. *Left panel:* number of pairs of sources, as function of the flux ratio between pairs of sources. Sources with fluxes 10 times fainter than the local (i.e. position-dependent) flux limits are included. Solid histogram: pairs of input sources. Dotted histogram: pairs detected as a single output source. Dashed histogram: pairs where both components have a detection. *Right panel:* number of pairs of sources, as function of the spatial separation between the two pair components. Line styles as in the *left panel*. Only pairs with flux ratios lower than 3, and whose components are brighter than $6.6 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ are considered. The larger number of input pairs with respect to output ones in the $3''$ – $10''$ range is due to pairs where the distance between the centroid and the position of the detection is larger than $6''$; this happens for ~ 3 sources per simulations.

as the weights. The number of single-detection pairs is also shown in Fig. 12 (both panels, dotted histograms). By considering only pairs with separation ≤ 15 arcsec, their average number is 13 pairs/simulation. When comparing with the XMM-CDFS (see Sect. 8.2), this number should be taken as a lower limit, since the simulations do not include clustering; in Sect. 8.2, the number of XMM-Newton sources with two or more Chandra counterparts is estimated to be 21.

Almost all pairs with a large flux ratio (≥ 10) are detected as a single source (Fig. 12, left panel), probably because the fainter source cannot be distinguished from a fluctuation in the wings of the PSF of the brighter source. Conversely, at small flux ratios the fraction of single-detections drops, and $\leq 1/4$ of the input pairs with flux ratio ≤ 2.5 are single-detected. The reason is probably that it is easier to have a double-detection at small flux ratios, where the cores of the two PSFs can be better recognised. Also, part of the explanation might involve the inclusion of very faint sources in the input list.

Considering the detectable pairs, the number of single-detections as function of the pair separation (Fig. 12, right panel) peaks at ~ 9 arcsec, and it always stays lower than the number of input pairs, probably because our selection of “detectable pairs” is conservative and allows for pairs where one of the components is just below the local flux limit. At separations $\geq 9''$ the number of single-detections starts to decline, and matches the number of double-detections at $14''$. The total number of single-detections with separation $\leq 20''$ is 14, which may taken as an estimate of the expected number of confused sources in the XMM-CDFS.

Summarizing, sources closer than $11''$ cannot be individually detected. Sources separated by larger angles *may* be individually detected, the odds for it increasing with the separation, though the PSFs may still be blended. Confusion is no longer an issue at angles larger than $\sim 16''$.

8. Chandra and optical identification and redshifts

In the catalogue tables we include the most probable association between our catalogue and the Chandra 4 Ms and ECDFS ones, determined with the procedure explained in Sect. 8.1 (blends and other possible associations are reported in the NOTES column).

Optical identifications and redshifts are primarily taken from X11 (i.e., spectroscopic redshifts from Treister et al. 2009b; Popesso et al. 2009; Balestra et al. 2010; Silverman et al. 2010; Le Fèvre et al. 2004; Szokoly et al. 2004; Mignoli et al. 2005; Ravikumar et al. 2007; Vanzella et al. 2008; and photometric redshifts from Zheng et al. 2004; Luo et al. 2010; Cardamone et al. 2010; Rafferty et al. 2011). To search for more redshifts and updates, we checked spectroscopic redshifts in recent works (Cooper et al. 2011; Coppin et al. 2012; Kurk et al. 2013). Finally, in four cases, redshifts have been estimated by the clear presence of an iron emission line (Iwasawa et al. 2012; Georgantopoulos et al. 2013).

8.1. Theory

We have cross-correlated the XMM-CDFS 2–10 keV catalogue with the Chandra 4 Ms (X11) and ECDFS (Lehmer et al. 2005) catalogues. We have used the likelihood-ratio estimator with a simple prior as described in Pineau et al. (2011)¹⁰. Briefly, this method works in two steps. First, for each XMM-Chandra source association we define the likelihood-ratio between the probability of the association being real vs. it happening by chance:

$$LR(r) = \frac{e^{-r^2/2}}{2\lambda} \quad (13)$$

¹⁰ This method has been implemented in a plugin to Aladin, and is available from <http://saada.u-strasbg.fr/docs/afx/plugin/>

(Eq. (9) in Pineau et al. 2011) where both r and λ are adimensional. r is the angular distance between both sources divided by the combined (in quadrature) positional error (i.e. $r = d / \sqrt{\sigma_X^2 + \sigma_C^2}$ where d is the angular distance and the σ are the positional errors), and λ is the angular density of *Chandra* sources with flux higher or equal to that of the *Chandra* member of the association (using the number counts in Luo et al. 2008¹¹) multiplied by the combined square positional error (i.e., $\lambda = (\sigma_X^2 + \sigma_C^2) \times N(>S)$). Second, to obtain the probability of real association we use Eq. (11) in Pineau et al. (2011).

$$P(H_{\text{cp}} | r) = \left(1 + \left(\frac{P(H_{\text{cp}})}{1 - P(H_{\text{cp}})} LR(r) \right)^{-1} \right)^{-1} \quad (14)$$

where $P(H_{\text{cp}})$ is the (unknown) prior probability of a *Chandra* source to be the counterpart of an *XMM-Newton* source. Given the different characteristics and depths of the X11 and ECDFS catalogues, we assume $P(H_{\text{cp}})$ with two constant values: one if the source comes from X11, another if it comes from the ECDFS.

The (unknown) number of real *Chandra* counterparts is the sum of the probabilities of association, over all associations:

$$N_{\text{real}} = \sum P(H_{\text{cp}} | r) \quad (15)$$

which can be approximated by using the total number of candidate *Chandra* counterparts N :

$$P(H_{\text{cp}}) \sim \frac{N_{\text{real}}}{N} \quad (16)$$

We finally arrive at

$$N_{\text{real}} = \sum \left(1 + \left(\frac{N_{\text{real}}}{N - N_{\text{real}}} LR(r) \right)^{-1} \right)^{-1} \quad (17)$$

combining Eqs. (14), (16) and (15) above. This equation can be resolved iteratively for N_{real} (starting with $N_{\text{real}} = N/2$ for instance) to get $P(H_{\text{cp}})$ and hence $P(H_{\text{cp}} | r)$ for each association.

8.2. Cross-correlation of XMM-Newton and Chandra sources

We have used the method outlined above to cross-correlate independently our *XMM-Newton* 2–10 keV and 5–10 keV catalogue with the *Chandra* 4 Ms (using Tables 3 and 6 in X11) and the ECDFS (using Tables 2 and 6 in Lehmer et al. 2005) catalogues.

Then the cross-correlations were merged; when a source was present in both X11 and the ECDFS, we only considered the X11 position. The cross-correlation was performed considering all *XMM-Newton* associations with $r \leq 5$. For each *XMM-Newton* source with more than one possible *Chandra* counterpart, we dropped the association whose probability was $\leq 3\%$ of that of the most probable one. We flagged initially as “Good” the associations with probabilities of association $P(H_{\text{cp}} | r) \geq 90$ per cent.

The restrictions in the cross-correlation catalogue above were chosen to provide a manageable size, while likely keeping all possible associations of interest. This restricted catalogue

¹¹ The 2–8 keV number counts were used if the *Chandra* source was detected in that band, otherwise the 0.5–2 keV number counts were used. If the source was only an upper limit in both bands, the upper limit in the 2–8 keV band was assumed to be the flux of the source and the number counts in that band were used.

was extensively inspected by eye (also taking the *XMM-Newton* and *Chandra* fluxes into account) to get our definitive cross-correlation catalogue. The visual inspection was carried out by superimposing contours of S/N on top of the *XMM-Newton* and *Chandra* images, to help locate the peaks of emission. The S/N contours were also the basis for the annotations on the source blending present in the NOTES column of the catalogue.

As a result of this inspection, 10 *Chandra* counterparts to 2–10 keV *XMM-Newton* sources had their “Good” flag revoked (2 counterparts to 5–10 keV sources), because either the corresponding *XMM-Newton* source was spurious, the *Chandra* member of the association had only an upper limit to its 2–8 keV flux or the counterpart was from the ECDFS and there was a better X11 counterpart. Conversely, 13 (5) counterparts were considered “Good” despite having probabilities of association below 90 per cent: this was changed because the position of the *XMM-Newton* source from our source detection did not coincide with the peak emission on the corresponding *XMM-Newton* image, or blending with nearby *XMM-Newton* or *Chandra* sources. In all cases we have kept the positional information from PWXDdetect.

Finally, we obtained 336 “Good” *XMM-Newton* associations for *XMM-Newton* sources in the main 2–10 keV catalogue (137 in the main 5–10 keV catalogue), corresponding to 339 (137) unique *XMM-Newton* sources. Out of the 315 (130) *XMM-Newton* sources with one or more *Chandra* counterparts, 295 (124) had a single counterpart, 19 (5) had two and 1 (1) had three. 24 (7) *XMM-Newton* sources did not have any *Chandra* counterpart.

8.3. New sources

To identify candidate real XMM-CDFS sources not previously detected by X11 or Lehmer et al. (2005) with *Chandra* (as opposed to candidate spurious *XMM-Newton* detections) we adopt the following baseline criterion: a source is considered real if there is at least a S/N contour where $S/N > 3$ in the band where it was detected.

The S/N was computed, in the Bayesian framework, by treating a source’s count rate c as a random variable constrained to be $c \geq 0$. Since the background images have more than 10 counts over the area where the exposure map $t_{kl} > 10^6$ s, we use a Gaussian approximation to the Poisson statistics, so that for a pixel (i, j) the probability that a source is in that pixel with count rate c is

$$P_{ij}(c) \propto \prod_{kl} \frac{\exp\left(\frac{-(ct_{kl}w_{kl} - T_{kl} + B_{kl})^2}{2(T_{kl} + B_{kl})}\right)}{\sqrt{2\pi(T_{kl} + B_{kl})}} \quad (18)$$

where the (k, l) indices run over a circular aperture of radius $16''$ around pixel (i, j) , w_{kl} is the value of a PSF centered in (i, j) at pixel (k, l) (approximated here by a Gaussian of $FWHM = 8.5''$), and T_{kl} and B_{kl} are the total and background image values. This can be rewritten as

$$P_{ij}(c) = \frac{\exp\left(\frac{-(c - c_{0,ij})^2}{2\sigma_{ij}^2}\right)}{\sigma_{ij} \sqrt{\frac{\pi}{2}} \operatorname{erfc}\left(\frac{c_{0,ij}}{\sigma_{ij} \sqrt{2}}\right)} \quad c \geq 0 \quad (19)$$

where $\operatorname{erfc}(x) = 2\pi^{-1} \int_x^{+\infty} e^{-t^2} dt$ is the complementary error function. $P_{ij}(c)$ is normalized taking into account that $c \geq 0$ and $c_{0,ij}$ and σ_{ij} are functions of t_{kl} , w_{kl} , T_{kl} and B_{kl} around pixel (i, j) . Since $P_{ij}(c)$ is symmetric around $c_{0,ij}$ for $c \geq 0$, by analogy

with the “pure” Gaussian case, the probability of having detected a source at pixel (i, j) is then

$$S/N_{ij} = \int_0^{2c_{0,ij}} P_{ij}(c) dc = \frac{2 - 2 \operatorname{erfc}(c_{0,ij}/\sigma_{ij} \sqrt{2})}{2 - \operatorname{erfc}(c_{0,ij}/\sigma_{ij} \sqrt{2})}. \quad (20)$$

We have drawn the contours in terms of “Gaussian sigmas” taking into account that, for example, $S/N = 3$ corresponds to $S/N_{ij} = 0.9973$.

Of the 24 (7) sources in the 2–10 (5–10) keV main catalogue without a *Chandra* counterpart, a $S/N > 3$ contour could be drawn on the 2–10 (5–10) keV images for 11 (2) of them, namely ID210 5, 85, 176, 189, 207, 224, 280, 348, 381, 402, 407, and ID510 1098 and 1150. Of this list, ID210 207 and ID510 1098 correspond to the same source. Two sources from the supplementary 2–10 keV catalogue (ID210 186, 410) also satisfy this criterion and are thus considered real. The $S/N > 3$ contour of source ID510 1150 includes a blend of ID210 304 (from the supplementary sample) and 306 (main), but it peaks around ID210 304; ID210 304 has no *Chandra* counterpart, while 306 does (source 443 in X11, whose position coincides with a peak of 0.5–2 keV emission in *XMM-Newton*). Thus we consider ID210 304 as the counterpart to ID510 1150, and include 304 among the candidate real sources.

We also considered a source as real, if it has at least a $S/N > 2$ contour in the same band of its detection, plus a $S/N > 3$ contour in the other band. This happens for one source: ID510 1149 (not detected in the 2–10 keV band).

Thus the total number of candidate “new”, *XMM-Newton*-only sources, is 15, divided as follows: 11 main 2–10 keV, 3 supplementary 2–10 keV, 3 main 5–10 keV (two sources were detected in both bands). Conversely, the number of candidate spurious sources is 13 in the main 2–10 keV sample and 4 in the 5–10 keV. Thus, the spurious fraction in the main 2–10 keV catalogue is $(24-11)/339 \sim 3.8\%$.

8.4. *Chandra* sensitivity limits and optical/IR counterparts for the new sources

Among the 15 candidate new sources, 2 are outside the 4 Ms and ECDFS areas (ID210: 407, 410) and were not observed by *Chandra*. Of the remaining 13, 8 are only covered by the ECDFS (ID210/ID510: 5, 186, 189, 207/1098, 224, 348, 381, 402), and 5 have also data from the 4 Ms survey (ID210/ID510: 85, 176, 280, 304/1150, 1149; though source 280 falls in an area with a strong exposure gradient). The 4 Ms flux limits are fainter than the ECDFS ones by about two order of magnitudes, therefore we will focus on the 5 sources falling in the 4 Ms area. The source positions are shown in Fig. 14 on an *XMM-Newton* exposure map with the *Chandra* survey boundaries superimposed.

Visual inspection of *Chandra* 4 Ms data showed low-significance local enhancements in the positions of the new sources. The *XMM-Newton* fluxes are compared to the *Chandra* sensitivity in the soft and hard bands in Table 4. The *Chandra* limits in 2–8 keV band for sources 85, 280 and 119 inside the 4 Ms area are a factor 2–3 lower than the flux observed by *XMM-Newton* in the 2–10 keV band, while for source 176 the *Chandra* limit is 15 times lower than the *XMM-Newton* observed flux. Only for source 304/1150 the *Chandra* limit is above the *XMM-Newton* flux. Therefore, sources 85, 176, 280 and 119 should, in principle, have been detected by *Chandra*.

All the sources inside the *Chandra*-covered areas are hard sources, because no signal is present in *XMM-Newton* data in the

Table 4. *XMM-Newton* fluxes and *Chandra* sensitivities for the five candidate new sources falling in the 4 Ms area.

ID210/ID510	XMM flux	XMM HR	<i>Chandra</i> limits	
			0.5–2 keV	2–8 keV
85	1.2	0.22	0.080	0.40
176	8.3	0.88	0.089	0.53
280	3.0	0.31	0.22	1.1
304/1150	0.21	0.30	0.073	0.42
1149	2.4	1	0.21	1.2

Notes. *XMM-Newton* fluxes are for the 2–10 keV band (except for source ID510 1149, which is in the 5–10 keV band). *Chandra* sensitivity limits in the 2–8 and 0.5–2 keV bands from the 4 Ms survey. All fluxes in units of 10^{-15} erg s $^{-1}$ cm $^{-2}$. Hardness ratio are from PN data.

0.5–2 keV band. Also, the hardness ratios¹² for the 5 sources inside the 4 Ms area are ≥ 0.3 (Table 4) and consistent with an obscured spectrum with $N_{\text{H}} \gtrsim 10^{22}$ cm $^{-2}$. The *Chandra* flux limits were calculated for an unabsorbed spectrum with $\Gamma = 1.4$; for obscured sources the limits would be higher depending on the column density. Thus it is possible that the differences between the *Chandra* and *XMM-Newton* effective area curves (*Chandra* having a higher drop at high energies than *XMM-Newton*) may favour the detection by *XMM-Newton* of very hard sources close to the nominal *Chandra* flux limit in the hard band. Consistently source 176, which has the largest ratio between observed *XMM-Newton* flux and *Chandra* limit, is also the hardest source with a $HR \sim 0.88$ (this HR may be produced by a power-law with $\Gamma = 1.7$ and $N_{\text{H}} \sim 5 \times 10^{22}$ cm $^{-2}$; source 176 is not detected in the 5–10 keV band probably because of its faintness, though an excess is still visible in the 5–10 keV image). If the hardness of these sources is due to obscuration, then the non-detection by *Chandra* in the 0.5–2 keV band would also be justified. None of these sources was included in the [Iwasawa et al. \(2012\)](#) paper because their PWXDetect significance is below the threshold set in that paper (except for source 407 which has however a low number of counts). Further study of these sources may help shed light on the obscured AGN population responsible for the missing fraction of the cosmic X-ray background, and will be presented in a further paper of this series ([Ranalli et al., in prep.](#)).

For the 15 new sources, we have searched for identifications in the *K* and 3.6 μm bands from the MUSYC ([Taylor et al. 2009](#)) and SIMPLE ([Damen et al. 2011](#)) surveys. The closest possible matches are presented in Table 5; in a few case we list more than one possibility due to the crowdedness of the field. Finally, in Fig. 13 we present *K* and 3.6 μm image cutouts in the with *XMM-Newton* contours superimposed.

8.5. Number of chance associations

We have used the simulations detailed in Sect. 5 to assess how many of these *XMM-Newton-Chandra* associations are due to chance: i.e., we check the reliability of the match, not that of the individual sources. We took the output XMM source lists from five of those simulations (for each band) and cross-correlated them with the CDFS 4 Ms and the ECDFS catalogues following the procedure outlined above (cross-correlation with CDFS 4 Ms and ECDFS independently, removal of the

¹² Calculated following [Park et al. \(2006\)](#) as the median of the posterior probability distribution of the ratio $(H - S)/(H + S)$ where H and S are the counts in the 2–10 and 0.5–2 keV band, respectively.

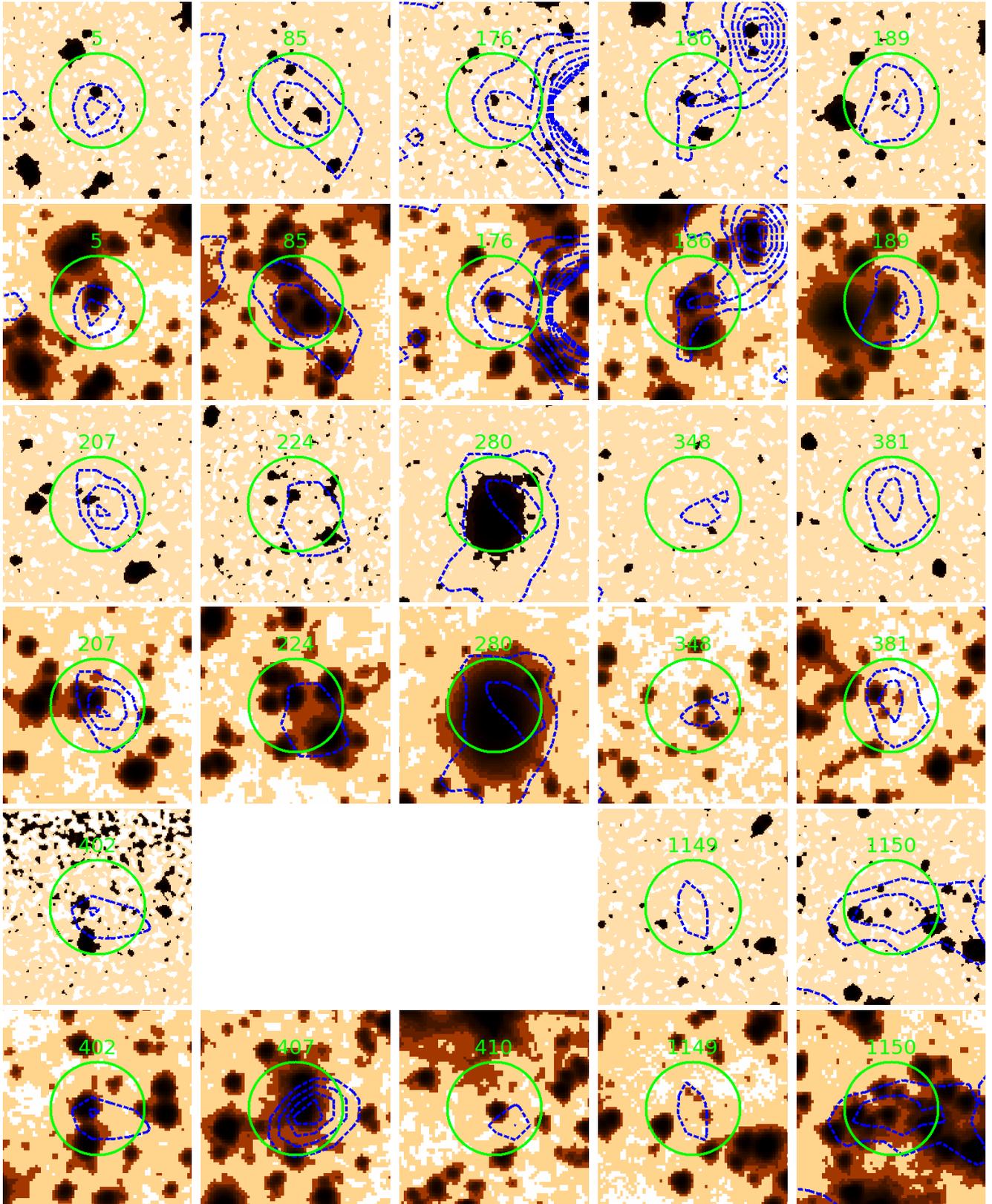


Fig. 13. *K*-band and $3.6\ \mu\text{m}$ cutouts of the 15 XMM-CDFS sources not previously detected by *Chandra* (images from the MUSYC and SIMPLE surveys). The *XMM-Newton* contours from the 2–10 keV band (5–10 keV for sources 1098 and 1149) are superimposed in blue. A green circle with radius $10''$ is drawn around the *XMM-Newton* position. The first, third and fifth rows show *K*-band images; the second, fourth and sixth rows show $3.6\ \mu\text{m}$ images. From left to right, and from top to bottom: ID210 5, 85, 176, 186, 189, 207, 224, 280, 348, 381, 402, 407, 410, and ID510 1149, 1150. *K*-band data are not available for ID210 407 and 410. Each cutout covers an area of $40'' \times 42''$. The *K*-band images have been smoothed with a Gaussian kernel of size 3 pixel ($0.75''$). The images are shown in histogram-equalized scaling, ranging from light (fainter fluxes) to dark brown (brighter fluxes).

Table 5. *K*-band and 3.6 μm counterparts to new XMM-CDFS sources.

ID210/ ID510	Location	RA _{NIR} J2000	Dec _{NIR} J2000	Dist arcsec	ID MUSYC	<i>K</i> AB	ID SIMPLE	3.6 μm AB	Redshift	Notes
407	out	53.25516	-27.54088	1.46	–	–	54179	18.58	1.13	<i>a,b</i>
410	out	53.22289	-27.51963	0.83	–	–	56022	21.33	–	<i>c</i>
5	ECDFS	53.09970	-28.02297	1.92	1613	20.56	11010	20.31	0.68	
186	ECDFS	53.89292	-27.80909	1.09	8374	19.71	30699	19.67	–	
189	ECDFS	52.90288	-27.80730	1.60	8413	21.35	30653	20.02	1.43	
207/1098	ECDFS	52.88460	-27.78947	1.81	8998	20.94	32579	20.91	0.62	
224	ECDFS	53.32407	-27.77321	1.86	9603	20.51	34104	19.80	1.38	<i>d</i>
348	ECDFS	53.36649	-27.64710	3.31	–	–	46367	23.17	–	<i>e</i>
		53.36634	-27.64582	2.71	–	–	46310	21.71	–	<i>e</i>
		53.36575	-27.64786	6.89	13769	21.81	46252	22.63	–	
381	ECDFS	53.03626	-27.59340	4.2	–	–	50549	21.56	–	<i>e</i>
		53.03574	-27.59456	2.40	–	–	50687	23.64	–	<i>e</i>
		53.03471	-27.59362	2.32	–	–	50766	23.69	–	<i>e</i>
402	ECDFS	53.18346	-27.56021	3.14	15482	20.51	52934	19.79	–	<i>f</i>
		53.18321	-27.56246	7.74	16344	18.95	52602	19.53	–	
85	4 Ms	53.07197	-27.90447	1.93	4907	21.37	21475	20.53	1.1125	<i>c</i>
176	4 Ms	53.25321	-27.81607	0.12	8087	21.49	30251	21.17	1.75	
280	4 Ms	53.26146	-27.72112	2.64	10952	15.63	37523	16.42	–	<i>d</i>
1149	4 Ms	52.99465	-27.70056	3.2	–	–	41386	23.93	–	<i>e</i>
		52.99755	-27.70113	7.4	12085	21.23	41190	20.53	–	
304/1150	4 Ms	53.13756	-27.70008	2.49	12132	21.54	41087	20.57	1.2357	<i>c</i>

Notes. The columns show: ID210 (ID510 if the number is >1000); if the source falls inside or outside the 4 Ms and ECDFS surveys; J2000 positions from the *K* band, unless in case of a missing *K* counterpart, in which case we list the IRAC 3.6 μm position; distance in arcsec between the *XMM-Newton* and NIR position; ID number and *K* magnitude from the MUSYC survey; ID number and 3.6 μm magnitude from the SIMPLE survey; redshift; notes. Magnitudes are in the AB scale. ^(a) No *K*-band imaging. ^(b) Bright *XMM-Newton* source. ^(c) Possible extended NIR source, only closest *XMM-Newton*/NIR match listed. ^(d) *K*-band empty field. ^(e) Another source is present at 7.7". ^(f) Bright galaxy.

duplicated sources, merging, subselection out to $r \leq 5$, removal of further counterparts with low relative probability and selection of associations with high probability). We then merged the five resulting catalogues and calculated the fraction of simulated XMM sources that had one or more counterparts with probability above 90 per cent: out of the 1793 (685) sources in those five simulations, 71 (30) had one “Good” *Chandra* counterpart and 6 (3) had two, there were none with three or more. These correspond to a total of chance associations with one or more *Chandra* sources of 4.3 ± 0.5 per cent¹³ ($4.8_{-0.8}^{+0.9}$). Therefore, for the 2–10 keV main catalogue of 339 sources, we would expect between about 13 and 16 chance associations of sources, while for the 5–10 keV main catalogue of 136 sources there would be between about 5 and 8 spurious pairs.

9. Conclusions

We have presented the observations, data reduction, catalogues and number counts of the XMM-CDFS survey. Currently the deepest *XMM-Newton* survey, the XMM-CDFS observations pose several challenges in their reduction. The large number of independent exposures (33 obsids times three cameras) centred on the same field was unprecedented for *XMM-Newton*.

- The very large time span of the observations allowed us to find an increase by a factor ~ 2 in the instrumental background in the years 2008–2010 with respect to the years

¹³ We have estimated the uncertainties in those fractions using a Bayesian approach and the binomial distribution (Wall & Jenkins 2003) with the shortest 68 per cent condence interval (S. Andreon, priv. comm.).

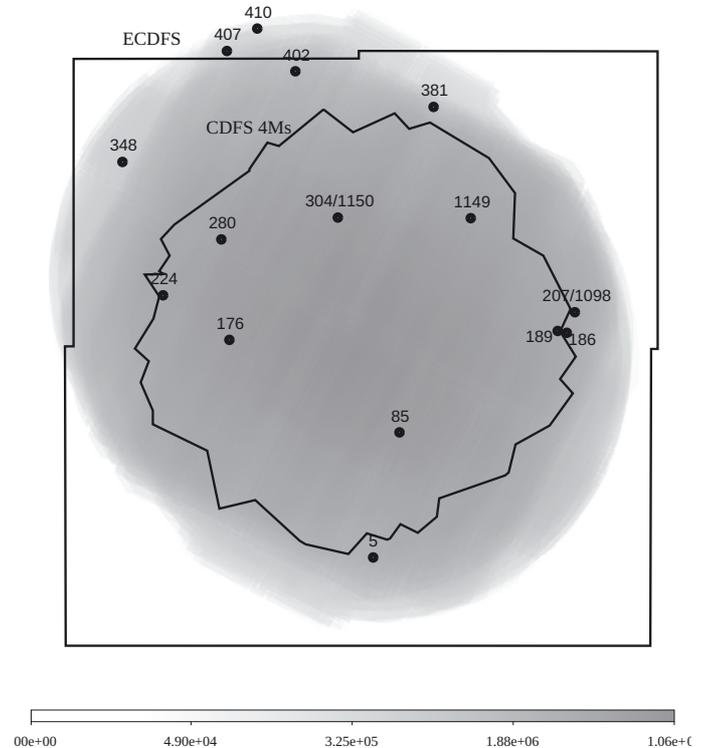


Fig. 14. Positions and ID210/ID510 numbers of the 15 candidate real sources not previously detected by *Chandra*, superimposed on the *XMM-Newton* 2–10 keV exposure map. The exposure map is drawn in logarithmic scale to show also the field border where the exposure is very low. The areas previously surveyed by *Chandra* are also shown.

2001–2002, and estimate the variations of the various background components.

- A careful study of the instrumental background of *XMM-Newton* during the XMM-CDFS observations was done to produce simulated observations, whose properties were designed to be as close as possible to the real XMM-CDFS. By using the count ratio between in-FOV and out-of-FOV data, and the FWC observations, the background was decomposed in cosmic, particle, and residual soft proton. The simulated observations reproduce the spatial details of the background (chip gaps, a missing MOS1 CCD, vignetting), and include detected sources distributed according to the [Gilli et al. \(2007\)](#) and [Ranalli et al. \(2005\)](#) model $\text{Log } N - \text{Log } S$. Mock catalogues were produced by running the source detection procedure on the simulated observations, and were used to calculate the coverage and the number of chance *XMM-Newton-Chandra* associations.
- We derived the catalogues in the 2–10 and 5–10 keV bands with a two-step procedure. First, the `PWXDetect` wavelet software was used to identify candidate sources with a significance equivalent to 4σ , and to find their coordinates. Next, we used the `SAS EMLDetect` tool to further check the significance of the sources, and obtain counts, count rates and fluxes. The final catalogues contain 339 and 137 sources in the 2–10 and 5–10 keV bands respectively, and have a significance equivalent to 4.8σ . The faintest sources have fluxes of 6.6×10^{-16} and $9.5 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. The flux limits at 50% of the sky coverage are 1.8×10^{-15} and $4.0 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. The catalogues were cross-correlated with the *Chandra* 4 Ms (X11) and ECDFS ([Lehmer et al. 2005](#)), using a likelihood-ratio method. Simulations provided an upper limit to the number of spurious sources, whose number is better estimated with comparison with *Chandra* data and inspection of the *XMM-Newton* S/N. The spurious fraction is thus $13/339 \sim 3.8\%$ in the 2–10 keV band, and $4/137 \sim 2.9\%$ in the 5–10 keV.
- Despite the high background level, *XMM-Newton* was able to detect a few extremely hard sources. Further study of these objects will help understanding the most obscured AGN population responsible for the missing fraction cosmic X-ray background ([Ranalli et al.](#), in prep.).
- The number counts were derived in both the 2–10 keV and 5–10 keV bands, and extend from 6.6×10^{-16} and 9.5×10^{-16} to 1.1×10^{-13} and $6.7 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. They are in agreement with previous derivations with the *Chandra* and *XMM-Newton* telescopes, and in different fields (the CDFS, the Lockman Hole, and the Hellas2XMM and 2XMM surveys). The XMM-CDFS currently reaches the faintest flux limit in the 5–10 keV band ($9.5 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$), ~ 3 times fainter than the Lockman Hole (the *Chandra* 4 Ms survey reaches fainter fluxes, though the band is formally softer, 4–8 keV, [Lehmer et al. 2012](#)).
- Finally, the simulations were used to study the source confusion in the XMM-CDFS relative to the 2–10 keV band. An average of 14 confused sources per simulation (i.e. detected as a single source, but corresponding to two input sources separated by $<15''$) was found; when comparing to the XMM-CDFS, this number should be considered as a lower limit because the simulations do not include source clustering. The cross-correlation with *Chandra* yielded 20 *XMM-Newton* sources associated with two or three *Chandra* counterparts.

Acknowledgements. We thank an anonymous referee whose comments have contributed to improve the presentation of this paper. We thank F. Damiani for support and help with the `PWXDetect` software, and K. Kuntz, S. Molendi, and S. Sciortino for useful discussions. PR thanks R. Di Luca for very valuable assistance in recovering data from a problematic disc; and A. Longinotti and all the *XMM-Newton* helpdesk team for their kind, focused and prompt support of the XMM-SAS software. We acknowledge financial contribution from the agreement ASI-INAF I/009/10/0 and from the PRIN-INAF-2011. PR acknowledges a grant from the Greek General Secretariat of Research and Technology in the framework of the programme Support of Postdoctoral Researchers. The XMM-CDFS simulator was developed by PR as per contract with the Astronomy Department of the University of Bologna. F.J.C. acknowledges support by the Spanish ministry of Economy and Competitiveness through the grant AYA2010-21490-C02-01. W.N.B. acknowledges the NASA ADP grant NNX11AJ59G. This research has made use of the Perl Data Language (PDL) which provides a high-level numerical functionality for the Perl programming language ([Glazebrook & Economou 1997](#); <http://pdl.perl.org>).

Appendix A: The XMM-CDFS simulator

We have developed a general purpose simulator of X-ray CCD observations, tuned for the *XMM-Newton* EPIC camera, but easily extensible to other missions. Although its primary use has been the validation of the XMM-CDFS catalogue, the simulator has in fact been written with future missions in mind (from eROSITA to concepts such as the Wide Field X-ray Telescope, and the project formerly known as XEUS/IXO/Athena). The simulator reproduces both point-like sources and the (cosmic, instrumental, and soft proton) background. As to sources and cosmic background, the only assumptions are that a library of position- and energy-dependent PSFs is available, that the effective area of the telescope and detector is known, and that an exposure map is provided to account for vignetting and eventual chip gaps. As to the instrumental background, the levels for the particle and soft proton contributions are needed, which in our case have been estimated by analysing the existing XMM-CDFS observations.

The main features are:

- it produces event files, to be analysed with common X-ray data analysis software;
- it supports an arbitrary number of sources;
- arbitrary spectra can be assigned to the sources;
- it calculates the most appropriate PSF for each source, according to position and energy, by interpolating from the PSF library.

The simulator is written in the Perl language, making use of the Perl Data Language¹⁴ libraries (PDL; [Glazebrook & Economou 1997](#)) for numerical computation and FITS input/output. The `Ftools` package is used to process the simulated event file headers and make it readable by the SAS; the SAS tool `merge` is used to join the source and background event files.

The computing time for a single-camera, single-obsid simulation is around 4–5 min on a 2.80 GHz Xeon processor running Linux; the time needed to simulate the whole XMM-CDFS on a single CPU is therefore around 7.5 h. The simulations of different cameras/obsid can however be run in parallel.

The simulator is released under the terms of the General Public License (GPL) as published by the Free Software Foundation, and is available on the XMM-CDFS website¹⁵.

¹⁴ Available at <http://pdl.perl.org>

¹⁵ <http://www.bo.astro.it/xmxcdfs/deeprime/index.html>

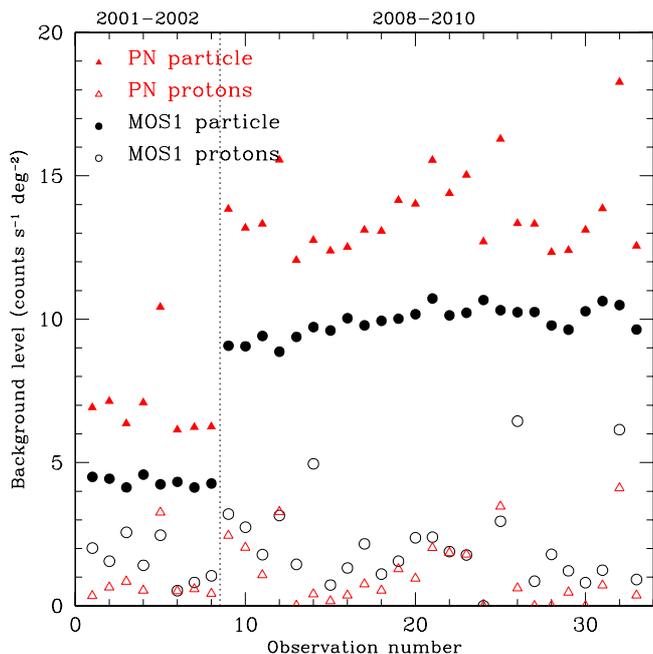


Fig. A.1. Intensity of the particle and proton components of the quiescent background in the 2–10 keV band for the 33 obsids of the XMM-CDFS. The obsids are numbered from 1 to 33 according to observation date; the vertical dotted line marks the separation between the 2001–2002 and 2008–2010 observations. Filled triangles: PN particle background; empty triangles: PN residual soft protons (after light curve screening; see Sect. 5.3); filled circles: MOS1 particle; empty circles: MOS1 protons. The MOS2 camera has values very similar to the MOS1.

A.1. Point-like sources and cosmic X-ray background

The goal is to produce an event list with the same properties (camera, pointing, exposure) of a real observation. This stage of the simulator acts basically as a PSF sampler. Given a list of coordinates and count rates, for each source it calculates the correct PSF according to the off-axis angle and the spectrum, and samples it for a number of photons depending on the count rate and exposure time. The PSF library is provided, for *XMM-Newton*, by three calibration files (CCF), one for each of the MOS1, MOS2 and PN, which contain normalized images of the PSFs at 11 energies (from 0.1 keV to 15 keV) times 6 off-axis angles (from 0' to 15'). Photon times and energies are also assigned to the events, to obtain an output file which looks like a real *XMM-Newton* event file.

The cosmic X-ray background is simulated by generating many faint sources according to the number counts models of AGN (Gilli et al. 2007) and galaxies (Ranalli et al. 2005).

The simulator accepts the following input: a list of RA, Dec, and count rates of the sources to simulate; a model spectrum formatted as an ASCII table; an event file, which is used to get the information about the boresight and camera and (optional) an exposure map.

The output is an event list in FITS format, which is readable by both PWXDetect and the SAS.

The algorithm works as follows. For each source, the expected counts are calculated, and rounded if they are fractional. The rounding needs to conserve the sum of the counts from all sources, in order to correctly reproduce the level of the cosmic X-ray background. The model spectrum is sampled, obtaining the event energies, which are binned according to the same sampling scheme of the PSF. Finally, the PSF is sampled, according

to the event energies; if needed, PSFs at different off-axis angles are interpolated. The positions of the sampled events are placed at the RA and Dec coordinates specified for the source, with the correct position angle.

A.2. Particle and soft proton background

The input for the simulation of the background are: the levels of the particle and soft-proton components (as an example, we show in Fig. A.1 the levels for the XMM-CDFS), which can be estimated from the in/out of FOV count ratio (see Sects. 5.2 and 5.3); an energy interval; and an event file used to derive the pointing coordinates, the position angle and the exposure time.

The FWC observations, which consist of event files, are filtered for the energy interval resampled with the bootstrap method (Efron & Tibshirani 1986) to obtain the simulated particle background event file.

The soft proton are simulated using a series of images, part of the ESAS CALDB files. These were originally obtained from severely flared *XMM-Newton* observations, by filtering out the low-background periods and retaining only the flares. The appropriate image is sampled, producing a simulated soft proton event file.

The two event files of the background components are finally merged with that for the source and cosmic X-ray background.

References

- Antonucci, R. 1993, *ARA&A*, 31, 473
 Baldi, A., Molendi, S., Comastri, A., et al. 2002, *ApJ*, 564, 190
 Balestra, I., Mainieri, V., Popesso, P., et al. 2010, *A&A*, 512, A12
 Brunner, H., Cappelluti, N., Hasinger, G., et al. 2008, *A&A*, 479, 283
 Cappelluti, N., Hasinger, G., Brusa, M., et al. 2007, *ApJS*, 172, 341
 Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, *A&A*, 497, 635
 Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al. 2010, *ApJS*, 189, 270
 Castelló-Mor, N., Carrera, F. J., Alonso-Herrero, A., et al. 2013, *A&A*, submitted
 Comastri, A., Setti, G., Zamorani, G., & Hasinger, G. 1995, *A&A*, 296, 1
 Comastri, A., Ranalli, P., Iwasawa, K., et al. 2011, *A&A*, 526, L9
 Cooper, M. C., Yan, R., Dickinson, M., et al. 2012, *MNRAS*, 425, 2116
 Coppin, K. E. K., Danielson, A. L. R., Geach, J. E., et al. 2012, *MNRAS*, 427, 520
 Cruddace, R. G., Hasinger, G. R., & Schmitt, J. H. 1988, in *ESO Conf. Works. Proc. 28*, eds. F. Murtagh, A. Heck, & P. Benvenuti, 177
 Damen, M., Labbé, I., van Dokkum, P. G., et al. 2011, *ApJ*, 727, 1
 Damiani, F., Maggio, A., Micela, G., & Sciortino, S. 1997, *ApJ*, 483, 350
 De Luca, A., & Molendi, S. 2004, *A&A*, 419, 837
 Efron, B., & Tibshirani, R. 1986, *Stat. Sci.*, 1, 54
 Falocco, S., Carrera, F. J., Corral, A., et al. 2013, *A&A*, in press, DOI: [10.1051/0004-6361/201321083](https://doi.org/10.1051/0004-6361/201321083)
 Fieller, E. C. 1954, *J. Roy. Stat. Soc. Ser. B (Methodological)*, 16, 175
 Gehrels, N. 1986, *ApJ*, 303, 336
 Georgakakis, A., Nandra, K., Laird, E. S., Aird, J., & Trichas, M. 2008, *MNRAS*, 388, 1205
 Georgantopoulos, I., Comastri, A., Vignali, C., et al. 2013, *A&A*, 555, A43
 Giacconi, R., & Zamorani, G. 1987, *ApJ*, 313, 20
 Giacconi, R., Zirm, A., Wang, J., et al. 2002, *ApJS*, 139, 369
 Gilli, R., Comastri, A., & Hasinger, G. 2007, *A&A*, 463, 79
 Glazebrook, K., & Economou, F. 1997, *The Perl Journal*, 5, 5
 Gruber, D. E., Matteson, J. L., Peterson, L. E., & Jung, G. V. 1999, *ApJ*, 520, 124
 Hasinger, G., Burg, R., Giacconi, R., et al. 1993, *A&A*, 275, 1
 Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, *ApJS*, 163, 1
 Iwasawa, K., Gilli, R., Vignali, C., et al. 2012, *A&A*, 546, A84
 Kuntz, K. D., & Snowden, S. L. 2008, *A&A*, 478, 575
 Kurk, J., Cimatti, A., Daddi, E., et al. 2013, *A&A*, 549, 63
 Le Fèvre, O., Vettolani, G., Paltani, S., et al. 2004, *A&A*, 428, 1043
 Lehmer, B. D., Brandt, W. N., Alexander, D. M., et al. 2005, *ApJS*, 161, 21
 Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, *ApJ*, 752, 46
 Lumb, D. H., Warwick, R. S., Page, M., & De Luca, A. 2002, *A&A*, 389, 93
 Luo, B., Bauer, F. E., Brandt, W. N., et al. 2008, *ApJS*, 179, 19

- Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2010, *ApJS*, 187, 560
- Lupton, R., Blanton, M. R., Fekete, G., et al. 2004, *PASP*, 116, 133
- Maccacaro, T., della Ceca, R., Gioia, I. M., et al. 1991, *ApJ*, 374, 117
- Mainieri, V., Hasinger, G., Cappelluti, N., et al. 2007, *ApJS*, 172, 368
- Mateos, S., Warwick, R. S., Carrera, F. J., et al. 2008, *A&A*, 492, 51
- Menci, N., Fiore, F., Puccetti, S., & Cavaliere, A. 2008, *ApJ*, 686, 219
- Mignoli, M., Cimatti, A., Zamorani, G., et al. 2005, *A&A*, 437, 883
- Murdoch, H. S., Crawford, D. F., & Jauncey, D. L. 1973, *ApJ*, 183, 1
- Page, M. J., Stevens, J. A., Ivison, R. J., & Carrera, F. J. 2004, *ApJ*, 611, L85
- Park, T., Kashyap, V. L., Siemiginowska, A., et al. 2006, *ApJ*, 652, 610
- Pineau, F.-X., Motch, C., Carrera, F., et al. 2011, *A&A*, 527, A126
- Popesso, P., Dickinson, M., Nonino, M., et al. 2009, *A&A*, 494, 443
- Puccetti, S., Fiore, F., D'Elia, V., et al. 2006, *A&A*, 457, 501
- Puccetti, S., Vignali, C., Cappelluti, N., et al. 2009, *ApJS*, 185, 586
- Rafferty, D. A., Brandt, W. N., Alexander, D. M., et al. 2011, *ApJ*, 742, 3
- Ranalli, P., Comastri, A., & Setti, G. 2005, *A&A*, 440, 23
- Ravikumar, C. D., Puech, M., Flores, H., et al. 2007, *A&A*, 465, 1099
- Rosati, P., Tozzi, P., Giacconi, R., et al. 2002, *ApJ*, 566, 667
- Schwartz, D. A., Murray, S. S., & Gursky, H. 1976, *ApJ*, 204, 315
- Setti, G., & Woltjer, L. 1989, *A&A*, 224, L21
- Silverman, J. D., Mainieri, V., Salvato, M., et al. 2010, *ApJS*, 191, 124
- Snowden, S. L., Collier, M. R., & Kuntz, K. D. 2004, *ApJ*, 610, 1182
- Snowden, S. L., Mushotzky, R. F., Kuntz, K. D., & Davis, D. S. 2008, *A&A*, 478, 615
- Snowden, S. L., & Kuntz, K. D. 2012, Cookbook for analysis procedures for XMM-Newton EPIC MOS observations of extended sources and the diffuse background, version 5, Tech. Rep.
- Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, *ApJS*, 155, 271
- Taylor, E. N., Franx, M., van Dokkum, P. G., et al. 2009, *ApJS*, 183, 295
- Treister, E., & Urry, C. M. 2005, *ApJ*, 630, 115
- Treister, E., Urry, C. M., & Virani, S. 2009a, *ApJ*, 696, 110
- Treister, E., Virani, S., Gawiser, E., et al. 2009b, *ApJ*, 693, 1713
- Tsujimoto, M., Guainazzi, M., Plucinsky, P. P., et al. 2011, *A&A*, 525, A25
- Urry, C. M., & Padovani, P. 1995, *PASP*, 107, 803
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, *A&A*, 478, 83
- Wall, J., & Jenkins, C. 2003, *Practical Statistics for Astronomers*, Cambridge Observing Handbooks for Research Astronomers (Cambridge University Press)
- Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, *ApJS*, 195, 10
- Zheng, W., Mikles, V. J., Mainieri, V., et al. 2004, *ApJS*, 155, 73