Mon. Not. R. Astron. Soc. **350,** 785–797 (2004) doi:10.1111/j.1365-2966.2004.07703.x

XMM–Newton surveys of the Canada–France Redshift Survey fields – II. The X-ray catalogues, the properties of the host galaxies and the redshift distribution

T. J. Waskett,[1⋆] S. A. Eales,[1] W. K. Gear,[1] H. J. McCracken,[2] M. Brodwin,[3] K. Nandra,[4] E. S. Laird[4] and S. Lilly[5]

[1]*Department of Physics and Astronomy, University of Wales Cardiff, PO Box 913, Cardiff CF24 3YB*
[2]*University of Bologna, Department of Astronomy, via Ranzani 1, 40127 Bologna, Italy*
[3]*Department of Astronomy & Astrophysics, University of Toronto, 60 St. George St., Toronto, Ontario, M5S 3H8, Canada*
[4]*Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London, SW7 2AZ*
[5]*Institut für Astronomie, ETH Hönggerberg, HPF G4.1, CH-8093, Zurich, Switzerland*

Accepted 2004 February 4. Received 2004 February 4; in original form 2003 October 24

ABSTRACT

We present the X-ray source catalogues for the *XMM* surveys of the 3-h and 14-h Canada–France Redshift Survey fields (0.5–10 keV flux range $\sim 2 \times 10^{-15}$–10^{-13} erg cm^{-2} s^{-1}). We use a subset of the *XMM* sources, which have *Chandra* positions, to determine the best method of obtaining optical identifications of sources with only *XMM* positions. We find optical identifications for 79 per cent of the *XMM* sources for which there are deep optical images. The sources without optical identifications are likely to be optically fainter and have higher redshifts than the sources with identifications. We have estimated 'photometric redshifts' for the identified sources, calibrating our method using ∼200 galaxies in the fields with spectroscopic redshifts. We find that the redshift distribution has a strong peak at $z \sim 0.7$.

The host galaxies of active galactic nuclei (AGN) identified in this work cover a wide range of optical properties, with every galaxy type being represented and no obvious preference for one type over another. Redder types tend to be more luminous than blue types, particularly at lower redshifts. The host galaxies also span a wide range of optical luminosity, in contrast to the narrow range found for the starburst galaxies detected in μJy radio surveys. We find a strong correlation between optical and X-ray luminosity similar to the Magorrian relation, although selection effects cannot be ruled out.

Key words: catalogues – galaxies: active – X-rays: galaxies.

1 INTRODUCTION

Deep exposures with the most recent and powerful X-ray observatories, *XMM–Newton* and *Chandra* (e.g. Giacconi et al. 2002; Mainieri et al. 2002; Barger et al. 2003; McHardy et al. 2003; Page et al. 2003), have built on the deepest *ROSAT* X-ray surveys (e.g. Hasinger et al. 1998; McHardy et al. 1998) by going deeper and to higher X-ray energies with better positional accuracy. This has opened up the study of faint X-ray sources such as high-redshift active galactic nuclei (AGN), and has also revealed X-ray emission from otherwise normal galaxies at more modest redshifts (Hornschemeier et al. 2003). These surveys have now resolved the majority of the cosmic X-ray background (XRB) in the soft (0.5–2 keV) X-ray band with a small fraction left unaccounted for in the hard (2–10 keV) band (Moretti et al. 2003).

The nature of the XRB at these X-ray energies is well on the way to being understood but the peak in the XRB lies at a much higher energy (∼30 keV). This indicates that a population of very faint sources, with very hard spectra, make up the remaining fraction of the XRB in the hard band, and would also be expected to contribute a much greater fraction to the XRB nearer its peak (Moretti et al. 2003). Such hard sources are most likely a result of extremely high obscuration, which progressively wipes out X-ray emission from low to high energy, turning an intrinsically soft spectrum into a much harder observed one.

The radiation absorbed during this process must be re-emitted at longer wavelengths and the possibility of the far-IR/submm background being somehow connected with the XRB is discussed in many papers (e.g. Almaini, Lawrence & Boyle 1999). However, current X-ray/submm surveys suggest that the two backgrounds are

⋆E-mail: Tim.Waskett@astro.cf.ac.uk

only loosely related (e.g. Severgnini et al. 2000; Alexander et al. 2003; Waskett et al. 2003). Future instrumentation with higher energy limits is likely to be required to explain the XRB fully and the nature of the sources that dominate its peak.

At present, though, the emphasis must be turned to those sources that we can observe easily with the current instrumentation. Quasi-stellar objects (QSOs) and type I AGN dominate the softest X-ray energies, with an increasing contribution from more obscured type II AGN becoming important at higher energies (e.g. Gilli, Salvati & Hasinger 2001). Identifying the optical counterparts to these sources is crucial for a full understanding of their properties, and a great deal of effort has been expended in obtaining this information (e.g. Barger et al. 2003; McHardy et al. 2003).

For example, one of the most useful quantities that can be derived from a source list is the luminosity function. This reveals much about the nature of a population, and determining its evolution with redshift can shed light on how the population as a whole changes over time. The X-ray luminosity function (XLF) has begun to be investigated in depth by several groups (Cowie et al. 2003; Steffen et al. 2003; Ueda et al. 2003). Both Ueda et al. (2003) and Steffen et al. (2003) find that the evolution of the XLF is a function of luminosity. The population of X-ray sources with L_X (2–10 keV) $>3 \times 10^{43}$ erg s^{-1} is dominated by type I AGN, and the number density of these sources increases with redshift out to $z \sim 2$–3. At lower X-ray luminosities, however, the fraction of type II AGN increases rapidly with decreasing X-ray luminosity. The number density of these sources appears to peak at $z < 1$.

Although *Chandra* is better suited for identifying X-ray sources with optical counterparts [*XMM* has a resolution of ∼6 arcsec full width half maximum (FWHM); cf. ∼0.5 arcsec for *Chandra*], *XMM* has greater sensitivity and a larger field of view (FoV), making it better for large area surveys. In this paper, we report the results of a medium-deep *XMM* survey composed of two separate exposures (∼0.4 deg^2). We quantify the ability of such a survey to identify X-ray sources with optical counterparts by comparing the IDs for a subset of the *XMM* sources with the IDs obtained using *Chandra* positions for the same sources. We estimate redshifts for our identified sources using photometric redshift codes. These allow a quick, and reasonably reliable, way of obtaining redshifts for objects with multiband photometry. Although not as accurate as spectroscopy, these techniques are becoming widely used as a short cut for large surveys, where statistical properties are fairly insensitive to the accuracy of individual redshift measurements (Fontana et al. 2000; Csabai et al. 2003; Kashikawa et al. 2003). These methods can also be used on objects fainter than the spectroscopic limit, where many X-ray source counterparts reside (Alexander et al. 2001). We test two photometric redshift estimation codes on our X-ray source IDs and obtain a robust redshift distribution for those sources that could be identified reliably, while placing limits on the properties of those that could not.

Ultimately, we will use our identified AGN, and their redshifts, to construct the XLF for different populations, and calculate its evolution with redshift. The results of this study will be reported in Paper III, the next in this series.

We assume an H_0 of 75 km s^{-1} Mpc^{-1} and a concordance universe with $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.

2 X-RAY DATA

Two *XMM* surveys are considered in this work, X-ray surveys of the CFRS 3-h and 14-h (also known as the Groth Strip) fields (Lilly et al. 1995a). The data reduction for these surveys, together with the comparison between SCUBA and *XMM* data, are described in detail in Waskett et al. (2003) (Paper I). The 14-h *XMM* data was first presented in Miyaji & Griffiths (2001). Both surveys are of ∼50 ks duration. In this section, we summarize some of the key points of the data analysis.

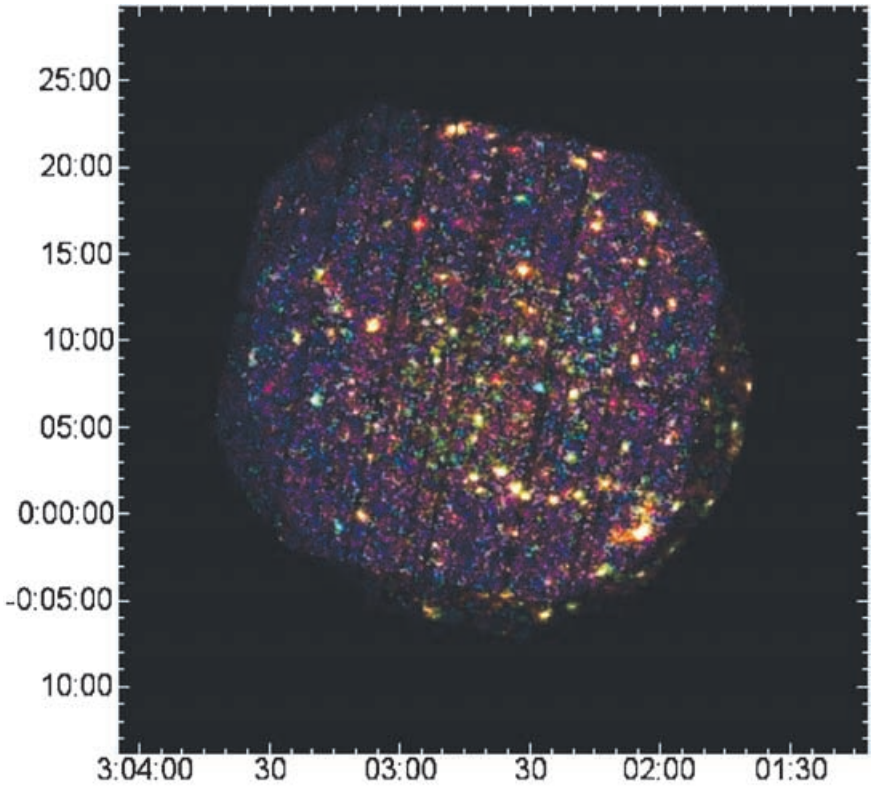
The raw X-ray data were reduced using v5.3 of the SAS software for *XMM*. Because *XMM* has a large spectral range, the data were divided into two energy bands: the soft band includes photons in the range 0.5–2 keV, whereas the hard band covers 2–10 keV. *XMM* has three X-ray cameras that operate simultaneously, so in total six images were used for the source detection: soft and hard bands for each of the two MOS cameras and also the PN camera. The source detection was performed simultaneously on all six images using the sliding box and maximum likelihood detection procedure within the SAS software, with the source extent fitting turned on. A photon index $\Gamma = 1.7$ was assumed for the counts-to-flux conversion in both bands. The thresholds for the source detection were set to 10 for the sliding box part and 15 for the maximum likelihood part, ensuring sources were only detected at greater than about 4 σ above the local background. Final source parameters were derived using data from both bands and all three instruments, for maximum accuracy and to minimize spurious detections from any single camera, while probing fainter fluxes. Using both soft and hard bands simultaneously also allows the detection routine to calculate a more accurate full 0.5–10 keV flux. The final source list contains, amongst other parameters: source positions, fluxes in the soft, hard and full combined bands and the vignetting-corrected hardness ratios for each source.

For this work, the hardness ratio is defined as $HR = \frac{N(H)-N(S)}{N(H)+N(S)}$, where $N(H)$ and $N(S)$ are the counts observed for a source in the hard and soft bands, respectively, after correction for vignetting. Higher values indicate a harder spectrum.

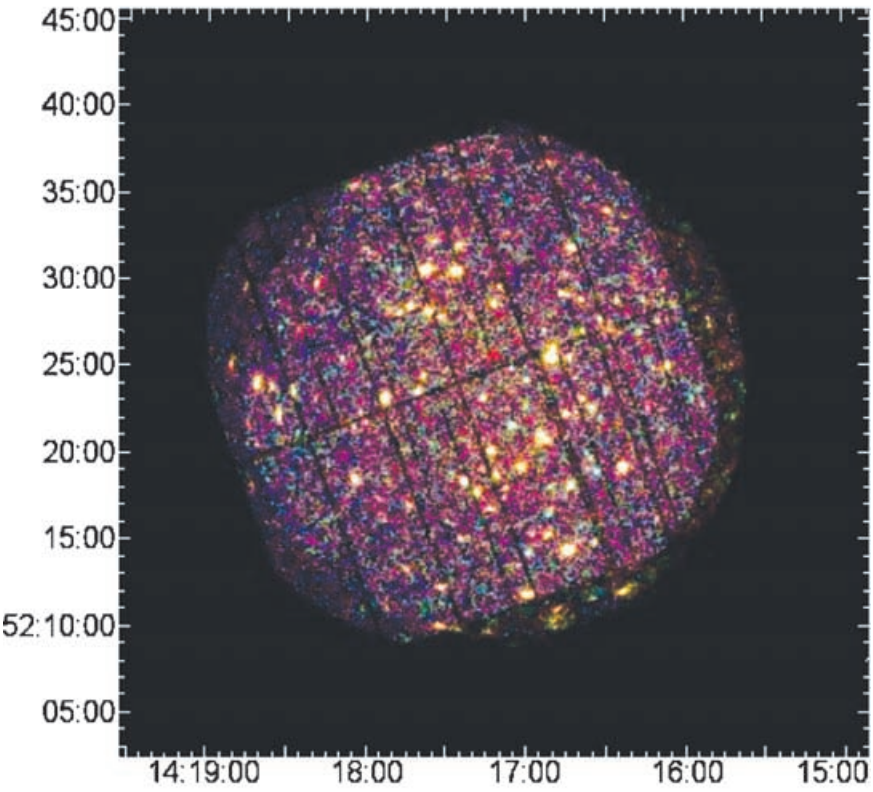
In total, there are 146 sources detected in the 3-h field and 154 in the 14-h field. Most are point sources. Tables A1 and A2 (see Appendix A) list the basic properties of a sample of the X-ray sources, in the two fields. Throughout this paper sources labelled with 3.* refer to 3-h field sources and those labelled with 14.* refer to sources in the 14-h field.

Fig. 1 shows false-colour images of the two fields considered in this work. Lowest-energy X-rays are coloured red, with progressively higher-energy X-rays being coloured green and then blue. Sources with hard spectra therefore show up blue in these images and soft sources appear red. All the extended sources detected are in the 3-h field and the majority are concentrated in the diffuse red patch visible in the lower right-hand corner of the 3-h image, surrounding a bright QSO (source 3.1 in Tables A1 and A3). This could be indicative of a galaxy cluster, and if the QSO is part of the cluster then the cluster has a redshift of 0.641. Unfortunately, because the QSO is so bright, it is hard to tell if it actually lies within a cluster, or whether the diffuse emission is simply an effect due to the broadening of the *XMM* point spread function towards the edge of the map. It is also unfortunate that this particular source lies off the edge of the deep optical map we use to identify the X-ray sources (see below), and so an optical cluster search of this region is not possible at this time. Digitized Sky Survey images of this region do not show any evidence for a galaxy cluster but do show the optical counterpart for the QSO.

Fig. 2 shows the differential source counts versus both soft- and hard-band flux, for all the sources detected in the two fields. These plots clearly demonstrate the effect of incompleteness at lower fluxes where the source counts drop off dramatically. This effect begins to become important at fluxes of 1.5 and 6×10^{-15} erg cm^{-2} s^{-1} for the soft- and hard-band sources, respectively. Above these fluxes, we



(a)



(b)

Figure 1. False-colour X-ray image of the 3-h (1a) and 14-h (1b) fields. Each field has one exposure but the images show data from all three X-ray cameras on *XMM*. Soft X-rays are red (0.5–1.5 keV), medium are green (1.5–3.5 keV) and hard are blue (3.5–10 keV).

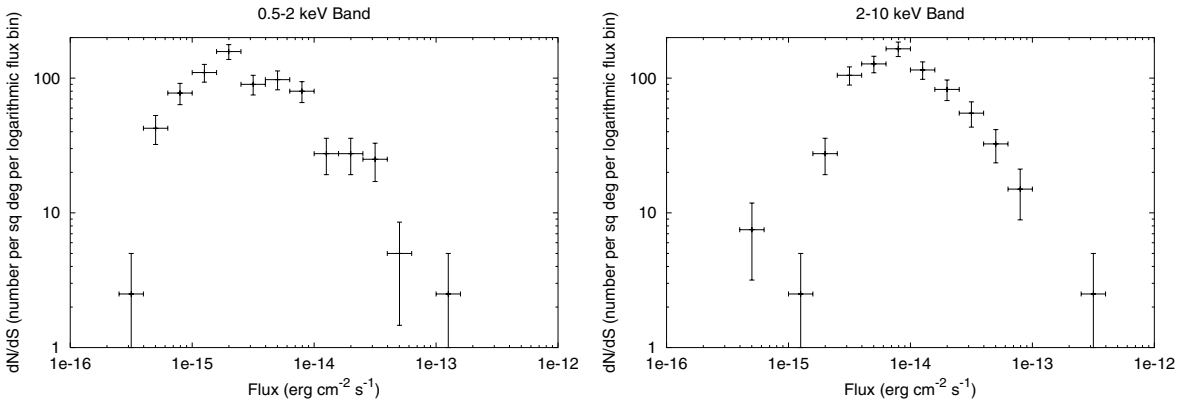


Figure 2. Differential source counts for the combined 3-h and 14-h field X-ray sources. Incompleteness causes the source counts to turn over at ∼1.5 and ∼6 × 10^{-15} erg cm^{-2} s^{-1} in the soft and hard bands, respectively. Source fluxes are calculated from the source detection procedure outlined in the text. Horizontal bars indicate the logarithmic flux ranges over which the sources are binned. Vertical error bars are the square root of the number of sources in each bin.

are effectively 100 per cent complete. This is comparable in depth to, for example, the HELLAS2XMM survey (Baldi et al. 2002), the early *XMM* Lockman Hole observations (Hasinger et al. 2001) and serendipitous *Chandra* observations (e.g. Gandhi et al. 2003), while reaching slightly deeper than the serendipitous *XMM* Survey in the AXIS field (e.g. Barcons et al. 2002).

3 OPTICAL IDENTIFICATIONS

3.1 *XMM*

After correcting the *XMM* astrometry against known bright QSOs, the process of identifying the X-ray sources with optical counterparts can be carried out. This process is important for the procedures in the following sections, especially in obtaining the redshift distribution of the AGN.

Both survey fields are coincident with the Canada–France Redshift Survey (CFRS) (Hammer et al. 1995; Lilly et al. 1995b) and the Canada–France Deep Fields survey (CFDF) (McCracken et al. 2001). The former covers a 10 × 10 arcmin^2 section in each field with spectroscopic redshifts for many of the galaxies, whereas the latter covers almost the entirety of both and reaches 3 mag deeper [$I_{AB}(3\sigma$, 3 arcsec) ∼ 25.5], albeit with no spectroscopic follow-up. We therefore use the CFDF catalogue as the basis for our identification process and extract CFRS redshifts as appropriate to monitor the accuracy of the photometric redshift determination (see Section 5). The CFDF data were taken with the Canada–France–Hawaii Telescope using the UH8K mosaic camera in *B*, *V* and *I*, with *U* data supplied by either the Cerro Tololo Inter-American Observatory (CTIO: 3-h field) or the Kitt Peak National Observatory (KPNO: 14-h field). Total exposure times were typically ∼5 h for *B*, *V* and *I*, and ∼10 h for *U*. The lengthy data reduction process is described in detail in McCracken et al. (2001). Of the 146 (154) X-ray sources in the 3-h (14-h) fields, 115 (149) lie within the CFDF regions.

To determine the optical identifications of the X-ray sources, we have used the frequentist approach of Downes et al. (1986). Because *XMM* has a positional accuracy of ∼2 arcsec (this is a worst-case situation for large off-axis angles; on-axis positional accuracy is more like 1.5 arcsec), only 1 per cent of *XMM* sources will have positions which are >6 arcsec away from the object that caused the emission. The first step in our ID procedure was thus to find all CFDF objects within 6 arcsec of the *XMM* position. We then calculated the following statistic for each object: $S = 1 - \exp[-d^2\pi n(<m)]$, where d is the offset between the *XMM* position and that of the optical object, and $n(<m)$ is the surface density of optical objects brighter than the magnitude (m) of the possible association. It may appear that this statistic gives the probability that the candidate object is a foreground or background object and is not physically related to the *XMM*. However, S is not a probability because it does not take into account galaxies that are fainter than the magnitude of the candidate galaxy, and that *might* have had a lower value of S. Therefore, this possibility needs to be taken into account when deriving the sampling distribution of S. Downes et al. (1986) describe an analytic way to do this. The end result is a true probability value, P'. Typically, a value of P' is several times higher than the equivalent S value. In all but two cases, we chose the CFDF object with the lowest value of P' as the most likely association. In these two exceptions, the galaxy with the lowest value of P' was close to 6 arcsec away from the *XMM* position, and we preferred the candidate with a slightly higher value of P' but which was much closer to the *XMM* position (these two IDs are confirmed by the *Chandra* X-ray positions, sources 14.15 and 14.50). Table 1 gives the statistics for our candidate identifications.

Table 1. Summary of ID statistics for both *XMM* fields. P' values given are for the best ID where more than one candidate lies within the 6-arcsec search radius.

	3-h	14-h
$P' < 0.05$	59	82
$0.05 < P' < 0.1$	16	24
$0.1 < P' < 0.2$	16	18
$0.2 < P' < 0.5$	13	16
$P' > 0.5$	12	8
Blank field	2	1
Outside CFDF	28	5

A consequence of this method is that because fainter objects are more numerous, they will have higher P' values than brighter objects at the same offset. Therefore, relatively optically faint objects are seldom identified with X-ray sources, unless they are very close to the X-ray position. For example, at the optical completeness limit

of $I_{AB} = 25.5$, an object at an offset of 0.8 arcsec will have $P' = 0.15$, which is the same P' as a 20.6-mag object at 6-arcsec offset.

3.2 The *Chandra* training set

We initially chose a P' value of 0.1 as being our dividing line between identifications and objects that are likely to be physically unrelated to the X-ray source. The number of spurious identifications can be estimated by simply adding up the values of P' for objects with $P' < 0.1$. This is ∼2 in the 3-h field and ∼3 in the 14-h field. In the two fields, 181 sources have $P' < 0.1$, which is 68 per cent of the *XMM* sources for which there are deep CFDF images. The error rate of false associations is $5/181 \sim 3$ per cent.

We were able to refine our identification criteria using the fact that part of the 14-h field has also been surveyed with *Chandra* (the north-east quadrant). The *Chandra* data are not the focus of this paper but they are summarized here. The data were taken in 2002 August using the ACIS-I instrument and were reduced using the standard CIAO v2.3 data reduction software. The total good exposure time after screening was 158 ks. Source detection was performed using the CIAO wavdetect algorithm (Freeman et al. 2002), run on images in the 0.5–8, 0.5–2, 2–8 and 4–8 keV bands, using a false source probability of 10^{-7}. Full details of the *Chandra* observations are given in Nandra et al. (in preparation).

Within the *Chandra* FoV, there are 63 *XMM* sources, 55 of which were also detected by *Chandra* within 10 arcsec of the *XMM* position. We performed a similar ID process to that employed above using these new positions, and succeeded in identifying 51 of the 55 *Chandra* sources. Two unidentified sources were also unidentified in the *XMM* analysis, and are essentially blank fields with no CFDF objects lying within 6 arcsec of either the *XMM* or *Chandra* position (sources 14.54 and 73). Of these 51 sources, 42 had previously been identified by *XMM*. 40 were identified as the same object by both *XMM* and *Chandra*; the remaining two had different IDs (sources 14.10 and 149). However, in one of these two cases, the *Chandra* ID was the second-best *XMM* ID (14.149) (the *XMM* IDs are listed in Tables A4 and A6). The other nine sources were securely identified by *Chandra* but not by *XMM*, so these are considered 'new' IDs (sources 14.65, 80, 85, 90, 102, 114, 115, 122, 129).

Given that the expected number of spurious *XMM* IDs for the whole 14-h field (106 identified sources) is ∼3, we would expect 1–2 spurious IDs in the subsample covered by the *Chandra* FoV. We found two IDs that were wrong in this sample and so feel confident that our estimate of ∼3 spurious *XMM* IDs in this field is accurate.

We relaxed the selection criteria for the *XMM* ID candidates to see if we could find more identifications for the *XMM* sources without significantly increasing the number of false associations. By increasing the cut-off to $P' < 0.15$, a further five *XMM* sources within the *Chandra* FoV are identified. Four of these are judged to be correct (14.85, 90, 102 and 114), given the *Chandra* ID, and one is incorrect (14.115). Extrapolating to our entire survey, we estimate that by relaxing our P' criterion we gain 22 additional identifications, of which probably ∼5 are inaccurate. For the rest of this work, IDs with $P' < 0.15$ are considered secure.

To summarize: with this new P' threshold, we identify 84 out of 115 sources in the 3-h field and 119 out of 149 sources in the 14-h field. One extra QSO lies outside the 3-h CFDF map but is coincident almost exactly with an *XMM* source and so is identified as such. An additional QSO lies on a chip boundary in the 3-h field and is assumed to be responsible for the X-ray emission detected to either side of the boundary (sources 3.7 and 19, see Table A1 and the very top of Fig. 1a; hereafter referred to as source 7), so in total 86 3-h sources are identified. In the 14-h field, the *Chandra* positions succeeded in identifying an extra four sources (14.65, 80, 122 and 129), bringing the total number of identified sources in this field to 123. Out of the *XMM* sources within the area of the CFDF, we have identified 75 per cent of the sources in the 3-h field and 83 per cent of the sources in the 14-h field. Only a small part of the difference between the two fields are the *Chandra* positions that exist for some of the 14-h *XMM* sources. We expect that, of our 209 identifications, 10 are incorrect.

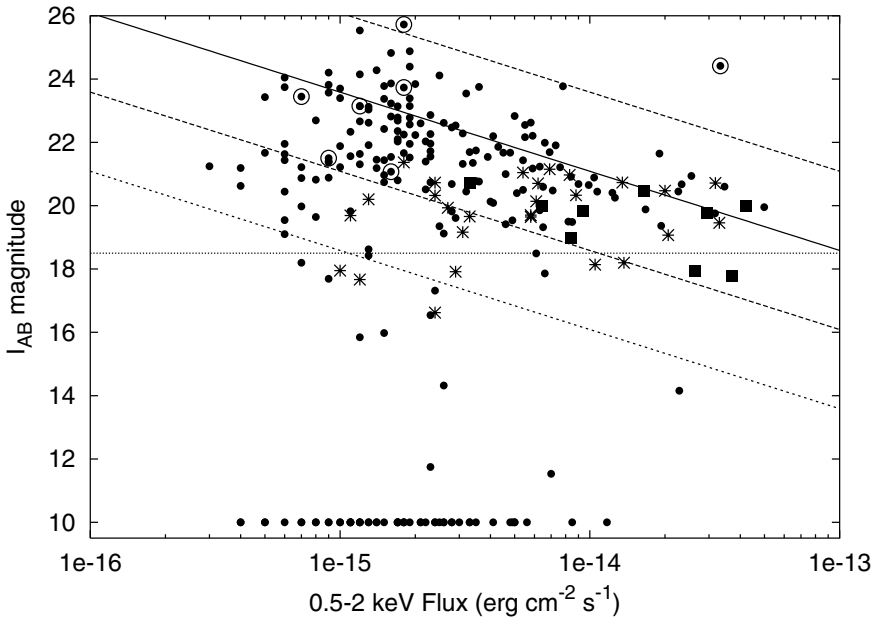
There are four new *Chandra* IDs and three IDs that were changed when *Chandra* positions were used rather than *XMM* positions (after increasing the P' limit to 0.15). These seven sources, which were not possible to identify using *XMM* positional data but which were possible to identify using *Chandra* positions (sources 14.10, 65, 80, 115, 122, 129 and 149), give us an insight into the properties of the remaining 58 unidentified *XMM* sources. The X-ray fluxes of the unidentified *XMM* sources cover a large range of fluxes (see Fig. 3), but the median I magnitude of the new IDs is 23.6; the median I magnitude for the other *XMM* IDs is 21.2 (range 11.5 to 25.5), nearly 10 times brighter. This is a consequence of the effect described at the end of Section 3.1. Section 5 describes the redshift information obtained for the IDs and it appears that these *Chandra* IDs lie, in general, at higher redshifts than most of the *XMM* IDs, which would partially explain their relative optical faintness.

4 X-RAY-TO-OPTICAL FLUX RATIOS

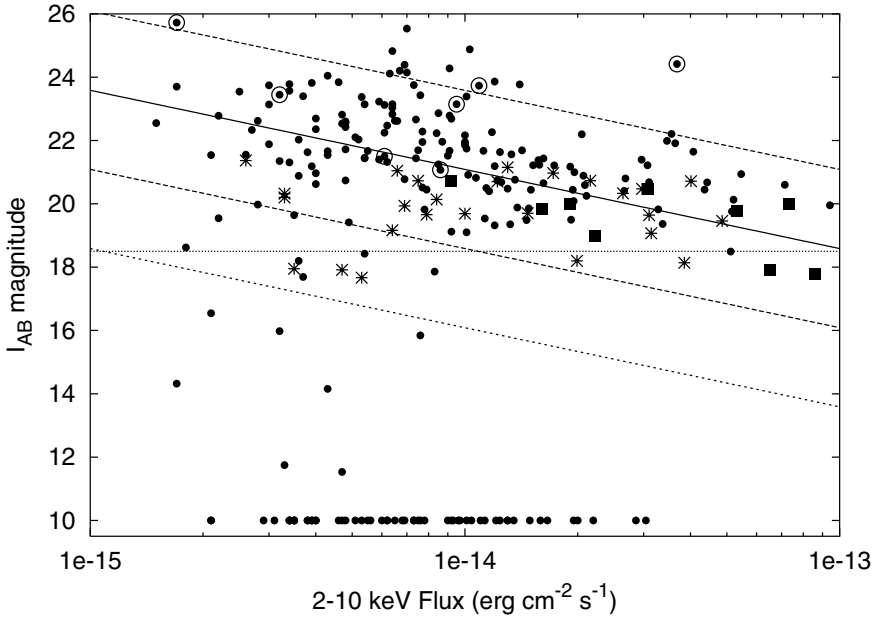
A convenient way of discriminating between different classes of X-ray source is the ratio between their X-ray and optical flux. Fig. 3 shows the total I-band magnitude (measured using a variable aperture to encompass the total flux of each object) versus the X-ray flux for all the identified sources in the 3-h and 14-h fields. The X-ray flux is calculated assuming a photon index $\Gamma = 1.7$. The I-band magnitude is related to the flux in this band, f_I, by $\log f_I = -0.4 I_{AB} - 5.57$, where f_I has the units erg cm^{-2} s^{-1}. Lines of constant X-ray-to-optical flux are plotted for comparison. AGN tend to occupy the space between the $\log(f_X/f_I) = \pm 1$ lines, whereas quiescent galaxies mostly lie below the $\log(f_X/f_I) = -2$ line, with a mixture in between.

Barger et al. (2002, 2003) have plotted similar diagrams for the *Chandra Deep Field*–North survey, an X-ray sample approximately 10 times fainter than our own. In the *Chandra* survey, the median optical apparent magnitude of X-ray sources flattens off at low X-ray fluxes, bringing the majority of sources below the AGN region on the $\log(f_X/f_I)$ plot. However, at the flux limit of our survey we are still predominantly detecting AGN with only a minor contribution from quiescent galaxies. Additionally, the redshift distribution of our identified sources (see Section 5) places the majority of the AGN in our survey at $z < 1$, which is the period of peak formation of supermassive black holes with low accretion rates (Cowie et al. 2003). These two facts mean that medium-deep surveys such as ours are well placed to study this important period of growth for intermediate luminosity AGN, without the need for very deep surveys which are able to probe much earlier times in the evolution of AGN and study the X-ray properties of more 'normal' galaxies.

In Fig. 3, the extra sources identified by *Chandra*, but not by *XMM*, in the 14-h field all reside in the higher $\log(f_X/f_I)$ regions. This suggests that they are AGN rather than starbursts or quiescent galaxies. Given that the *XMM* unidentified sources are, in general, optically fainter than the identified ones ($I_{AB} > 22$, see Section 3.2 and the end of Section 3.1), and that their X-ray fluxes are similar,



(a)



(b)

Figure 3. X-ray-to-optical flux ratio for all sources detected in both the 3-h and 14-h fields. Lines of constant flux ratio are plotted; the solid line is $\log(f_X/f_I) = 0$; the longer-dashed lines are ± 1; and the shorter-dashed line is -2. AGN tend to occupy the region between the longer dashed lines, quiescent galaxies lie mostly below the shorter dashed line, while a mixture, including starburst galaxies, occupy the region in between. The dotted line at $I_{AB} = 18.5$ shows the saturation limit of the CFDF and so magnitudes brighter than this are likely to be underestimated. Solid squares are known QSOs; asterisks are identifications with a stellar profile. Unidentified sources, within the optical coverage, are placed at $I_{AB} = 10$. The sources identified using *Chandra* positions, including the three with alternative *Chandra* IDs, are ringed with larger circles (sources 14.10, 65, 80, 115, 122, 129 and 149, see the end of Tables A4 and A6).

this implies that the unidentified X-ray sources are most likely AGN too, with high f_X/f_I ratios. One interesting point to note is that source 14.10 has a different *Chandra* ID to the one given by the *XMM* position; it is the *Chandra* ID that is plotted in these figures. However, the *Chandra* ID is significantly fainter than the *XMM* ID ($I_{AB} = 24.4$, cf. 19.0) and so this source now has an extreme $\log(f_X/f_I)$ value of $\sim$2 (cf. ~ -0.3 for the *XMM* ID). We assume in this paper that the *Chandra* ID is the correct one, but given this extreme flux ratio, it is possible that *XMM* has correctly identified this source, rather than *Chandra*.

In addition to the known QSOs in these fields, 27 of the identifications have stellar optical profiles. Fig. 3 shows that most of these

lie in the AGN part of the diagram, suggesting that they are QSOs rather than stars.

5 PHOTOMETRIC REDSHIFTS

Only a handful of the CFDF IDs have spectroscopic redshifts. Including known QSOs outside the CFRS regions, there are 13 (6) X-ray sources with spectroscopic redshifts in the 3-h (14-h) fields. The vast majority of the non-broad-line AGN do not have spectroscopic redshifts and so we turn to photometric techniques to estimate redshifts for these.

The optical spectra of broad line AGN (QSOs) are contaminated by light from the central engine, and so obtaining photometric redshifts for them is problematic. However, Gonzalez & Maccarone (2002) have shown that, for the majority of X-ray sources which are non-broad-line AGN, the optical spectrum is not significantly contaminated and so photometric techniques work just as well as they do with 'normal' galaxies. As long as the QSOs can be identified, they should not affect the rest of the sample. We therefore only use the estimated redshifts for the identifications which do *not* have a stellar profile. We use two photometric redshift estimation codes in this work, a Bayesian template fitting code called Bayesian Photometric Redshift (BPZ) (Benítez 2000) and a code developed specifically for the CFDF (Brodwin et al. 2003), calibrated against CFRS spectroscopic data. See Appendix B for the details and a comparison of the two codes.

The photometry for all the IDs is listed in Tables A5 and A6 and the results for both codes are shown in Tables A3 and A4. Fig. 4 shows the redshift distribution, as measured by each code, of all the reliable IDs that also have reliable redshift estimates, with a bin size of $\Delta z = 0.2$. Reliable *photometric* redshifts are defined here as unsaturated objects that have 95 per cent ($\sim 2\sigma$) redshift confidence limits $<0.4(1+z)$ (CFDF code) or $P_{\Delta z} > 0.9$ (BPZ code), otherwise spectroscopic redshifts are used where they exist; in total, 129 (120) estimates are reliable for the BPZ (CFDF) code. Despite the differences between the distributions measured by the two different codes, the overall shape of the distribution is clear, with a peak at around $z = 0.7$. In both distributions, nearly 60 per cent of the objects lie in the range $0.4 \leqslant z < 1$. The median redshifts are significantly different, however: 0.62 for BPZ and 0.79 for the CFDF code. For the rest of this work, the CFDF code is assumed to be more accurate (see Appendix B) and so all further quoted photometric redshifts are those given by this code.

An interesting point to note here is that the extra sources identified by *Chandra* and not by *XMM* (see the end of Table A4, sources 14.10, 65, 80, 115, 122, 129 and 149) lie, in general, at higher redshifts than the majority of the *XMM*-identified sources. If all the unidentified sources lie at higher redshifts than all of the other sources, then the median redshift of the total increases to $z \sim 1.1$.

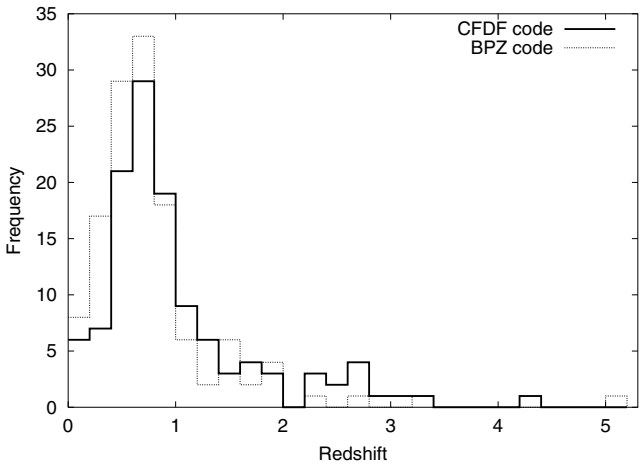


Figure 4. Redshift distribution of the identified X-ray sources as measured by the two different photo-z codes. Where a spectroscopic redshift exists, it is used in preference to the photometric one in both histograms. All unreliable redshifts are excluded (i.e. those indicated with * or s in Tables A3 and A4, and saturated objects with $I_{AB} < 18.5$). The overall shape of this distribution remains unchanged if the less reliable redshift estimates are also included; only the normalization increases.

6 ABSOLUTE MAGNITUDES – GALAXY TYPES

Fig. 5 shows the absolute I_{AB} magnitude plotted against redshift. The different symbols represent the best-fitting template determined for each galaxy by the CFDF code, using six-band photometry. Although the code uses 15 templates for greater accuracy, each symbol here represents a small range of templates for clarity. In general, the two photometric redshift (photo-z) codes agree reasonably well as to the best-fitting galaxy type.

The X-ray sources all lie in a band defined at the faint limit by the limiting magnitude of the optical data, and at the bright limit by the saturation magnitude. Objects brighter than this magnitude do not have reliable photometric redshifts and so do not appear in this plot. Aside from these selection effects, there are several other trends apparent here. Apart from QSOs, in general, at lower redshifts, the bluer galaxy types occupy the region near the faint limit, whereas progressively redder galaxies occur at brighter magnitudes (see Fig. 6). However, this trend breaks down at higher redshift where there are fewer sources, and errors in the photometry are likely to be more important. There is no clear domination of one galaxy type over any other, indicating that AGN have no preference

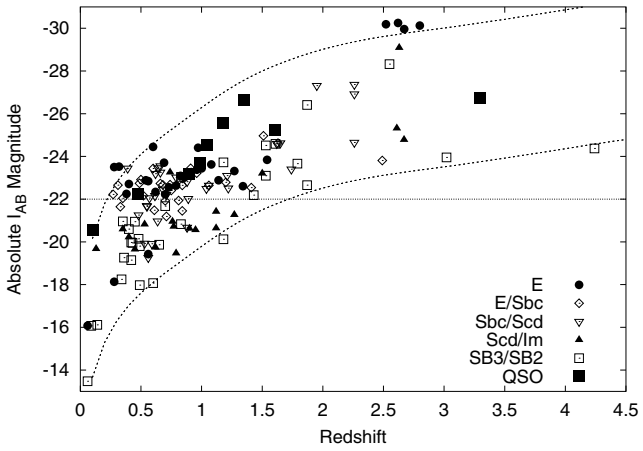


Figure 5. Absolute I_{AB} magnitude versus redshift for the identified sources (148 sources after removal of stellar and saturated objects). The horizontal line is the approximate position of M_I^*. The first four galaxy types are taken from Coleman, Wu & Weedman (1980), although some interpolation is used to create intermediate templates, and the starburst symbol represents both the SB3 and SB2 types from Kinney et al. (1996). QSOs have been plotted using the best-fitting template for the K-correction, in general the bluest starburst. Upper and lower curves are the approximate saturation limit and completeness limit of the optical data ($I_{AB} = 18$ and 24.5 respectively), calculated for an Scd galaxy (due to larger K-corrections, some ellipticals lie above the bright limit for Scd galaxies).

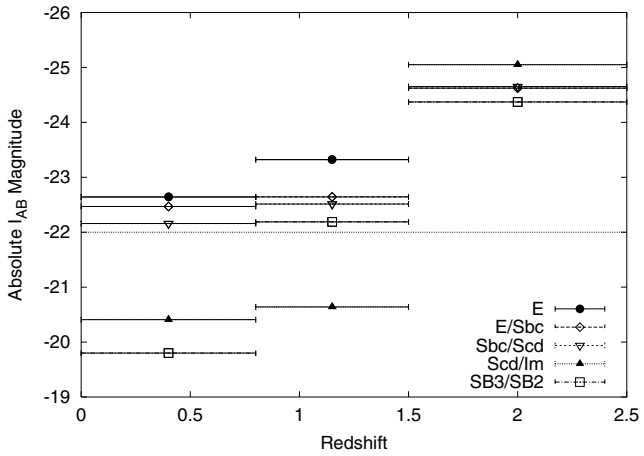


Figure 6. Median values for the galaxy types in Fig. 5. The highest z bin includes all sources with $z > 1.5$. The group of elliptical galaxies at $M \sim -30$ in Fig. 5 is off the vertical scale in this plot. The horizontal line is the approximate position of M_I^*.

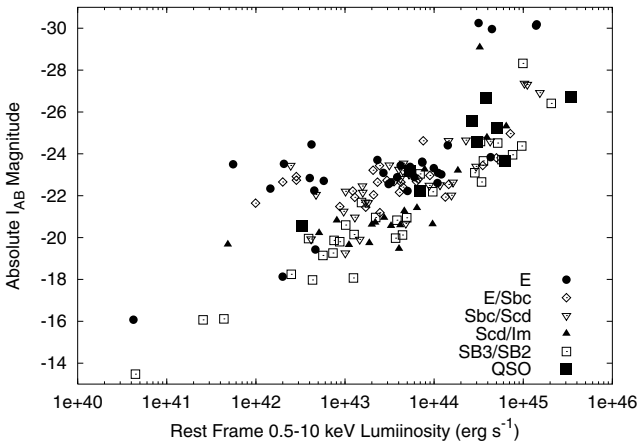


Figure 7. Absolute I_{AB} magnitude versus X-ray luminosity calculated from the 0.5–10 keV flux for the same sources as Fig. 5. The X-ray luminosity is K-corrected assuming an intrinsic power law slope with photon index $\Gamma = 1.7$.

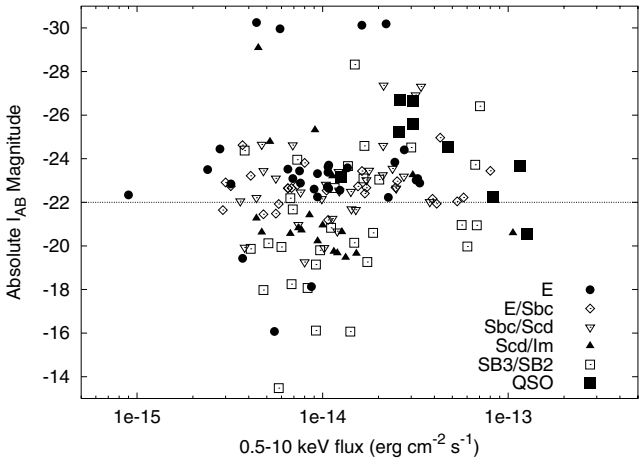


Figure 8. Absolute I_{AB} magnitude versus total X-ray flux for the same sources as Fig. 5. The horizontal line is the approximate position of M_I^*.

when it comes to the morphology of their host galaxies. Nor is there any apparent preference for optical luminosity of the host galaxy, unlike the narrow absolute magnitude range preferred by the starburst galaxies detected in μJy radio surveys (Chapman et al. 2003). X-ray sources occupy the whole optical luminosity range available to them in this plot. There are four apparently very luminous ellipticals at $z > 2.5$ (sources 3.32, 3.90, 3.92 and 14.31) which may be erroneous identifications. The CFDF code becomes less reliable above a redshift of 1.3 (Brodwin et al. 2003) and so it is possible that these sources actually lie at lower redshifts (in fact, the BPZ code places three of these sources at $z < 1$, see Tables A3 and A4, and classifies them as spirals; this is likely a consequence of this code using a magnitude-based prior) and so are consequently of less extreme luminosity. With this in mind, high-redshift sources should be viewed with some caution.

The extra sources identified by *Chandra* but not by *XMM* in the 14-h field also cover a wide range in galaxy types. The two higher-redshift sources are the bluest galaxy types, whereas the two lowest-redshift sources are the reddest types. Three of the four hug the lower luminosity limit, a consequence of their relative optical faintness.

Converting the 0.5–10 keV X-ray flux of the identified sources into X-ray luminosity gives us Fig. 7. Although the striking correlation here is possibly dominated by the same selection effects seen in Fig. 5, it is rather reminiscent of the Magorrian relation (Magorrian et al. 1998), with black-hole mass represented by X-ray luminosity and bulge mass represented by optical luminosity. Whether this correlation is real or not depends on exactly where the optically faint and saturated sources lie in this plot. We would expect optically faint sources to fall in the lower right-hand part of this plot and the saturated sources to fall in the upper left-hand part, effectively smearing out the correlation. However, if the optically faint sources are *not* at much higher redshifts than the identified sources (contrary to our arguments above), then both their X-ray and optical luminosities will be low, placing them amongst the sources plotted here. The very luminous ellipticals, mentioned above, also appear in this plot, slightly above the general trend, again suggesting that they have been misclassified (as have, potentially, a group of lower luminosity ellipticals, also lying away from the trend).

Plotting these sources in a different way illustrates what type of objects contribute to the XRB. For this discussion, we assume the XRB to have a spectrum of $I(E) = 11E^{-0.4}$ keV s^{-1} cm^{-2} sr^{-1} keV^{-1} (McCammon & Sanders 1990; Fabian & Barcons 1992), although the overall normalization is still somewhat uncertain. Fig. 8 shows absolute I_{AB} magnitude versus X-ray flux with the same symbols as in Fig. 5. The 300 sources in this survey (assuming that the majority of the 'stars' are misidentified QSOs) contribute ∼51 per cent to the XRB in the 0.5–10 keV range, whereas the 148 sources included in these figures (ie. the securely identified sources with redshift estimates) provide ∼27 per cent. Of this 27 per cent, sources brighter than M_I^* contribute 69 per cent, whereas fainter sources contribute 31 per cent. This calculation shows that the XRB is not dominated by the most optically luminous galaxies, but that a significant contribution comes from galaxies with fairly low optical luminosity.

7 CONCLUDING REMARKS

We have presented source catalogues for a survey, using the *XMM–Newton* X-ray telescope, of ∼0.4 deg^2 of sky. We show that reliable identifications can be obtained for ∼75 per cent of the *XMM* sources using *XMM* positions alone. Those sources that cannot be identified

using *XMM* positions alone are optically fainter ($I_{AB} > 22$) than most of the identified ones, and are likely to be AGN at generally higher redshifts. We have obtained the following results.

(i) The flux ratio f_X/f_{opt} of the sources in our survey show that they are predominantly AGN.

(ii) The optical properties of the AGN span a large range of absolute magnitudes, in contrast to the result found for the starburst galaxies detected in μJy radio surveys, which tend to have a very narrow range of absolute magnitudes (Chapman et al. 2003).

(iii) AGN are found in host galaxies spanning the full range of Hubble types, with no clear preference.

(iv) For the identified X-ray sources with good redshifts, there is a strong correlation between optical and X-ray luminosity, reminiscent of the Magorrian relation between black-hole mass and bulge mass. However, this may be due to selection effects.

(v) The redshift distribution of the AGN shows a clear peak at $z \sim 0.7$.

The last result supports other recent studies (Barger et al. 2003) that show the peak formation of supermassive black holes occurred at relatively recent times ($z < 1$). Medium-deep X-ray surveys such as ours, which resolve a large fraction of the XRB but are still dominated by AGN, are able to probe this epoch effectively.

We will use the results from this paper to calculate the X-ray luminosity function, and determine its evolution with redshift, in the next paper in this series.

ACKNOWLEDGMENTS

TJW acknowledges the support of a departmental postgraduate grant. SAE thanks the Leverhume Trust for a research fellowship. We thank the referee for being so thorough, which greatly clarifyied the paper and also helped us to better organise our thoughts for further work. This paper was based on observations obtained with *XMM–Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA.

REFERENCES

Alexander D. M., Brandt W. N., Hornschemeier A. E., Garmire G. P., Schneider D. P., Bauer F. E., Griffiths R. E., 2001, AJ, 124, 1839
Alexander D. M. et al., 2003, AJ, 126, 539
Almaini O., Lawrence A., Boyle B. J., 1999, MNRAS, 305, L59
Baldi A., Molendi S., Comastri A., Fiore F., Matt G., Vignali C., 2002, ApJ, 564, 190
Barcons X. et al., 2002, A&A, 382, 522
Barger A. J., Cowie L. L., Brandt W. N., Capak P., Garmire G. P., Hornschemeier A. E., Steffen A. T., Wehner E. H., 2002, AJ, 124, 1839
Barger A. J. et al., 2003, AJ, 126, 632
Benítez N., 2000, ApJ, 536, 571
Brodwin M. et al., 2003, ApJ, submitted (astro-ph/0310038)
Chapman S. C. et al., 2003, ApJ, 585, 57
Coleman G. D., Wu C.-C., Weedman D. W., 1980, ApJS, 43, 393
Cowie L. L., Barger A. J., Bautz M. W., Brandt W. N., Garmire G. P., 2003, ApJ, 584, L57
Csabai I. et al., 2003, AJ, 125, 580
Downes A. J. B., Peacock J. A., Savage A., Carrie D. R., 1986, MNRAS, 218, 31
Fabian A. C., Barcons X., 1992, ARA&A, 30, 429
Fontana A., D'Odorico S., Poli F., Giallongo E., Arnouts S., Cristiani S., Moorwood A., Saracco P., 2000, AJ, 120, 2206
Freeman P. E., Kashyap V., Rosner R., Lamb D. Q., 2002, ApJS, 138, 185
Gandhi P., Crawford C. S., Fabian A. C., Johnstone R. M., 2003, MNRAS, 348, 529
Giacconi R. et al., 2002, ApJS, 139, 369
Gilli R., Salvati M., Hasinger G., 2001, A&A, 366, 407
Gonzalez A. H., Maccarone T. J., 2002, ApJ, 581, 155
Hammer F., Crampton D., Le Fèvre O., Lilly S. J., 1995, ApJ, 455, 88
Hasinger G., Burg R., Giacconi R., Schmidt M., Trümper J., Zamorani G., 1998, A&A, 329, 482
Hasinger G. et al., 2001, A&A, 365, L45
Hornschemeier A. E., Bauer F. E., Alexander D. M., Brandt W. N., Sargent W. L., Vignali C., Garmire G. P., Schneider D. P., 2003, Astron. Nachr., 324, 12
Kashikawa N. et al., 2003, AJ, 125, 53
Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, ApJ, 467, 38
Lilly S. J., Le Fèvre O., Crampton D., Hammer F., Tresse L., 1995a, ApJ, 455, 50
Lilly S. J., Hammer F., Le Fèvre O., Crampton D., 1995b, ApJ, 455, 75
McCammon D., Sanders W. T., 1990, ARA&A, 28, 657
McCracken H. J., Le Fèvre O., Brodwin M., Foucaud S., Lilly S. J., Crampton D., Mellier Y., 2001, A&A, 376, 756
McHardy I. M. et al., 1998, MNRAS, 295, 641
McHardy I. M. et al., 2003, MNRAS, 342, 802
Magorrian J. et al., 1998, AJ, 155, 2285
Mainieri V., Bergeron J., Hasinger G., Lehmann I., Rosati P., Schmidt M., Szokoly G., Della Ceca R., 2002, A&A, 393, 425
Miyaji T., Griffiths R. E., 2001, in Neumann D. M., Van J. T. T., eds, XXXVIth Rencontres de Moriond, Recent results of XMM–Newton and Chandra, XXIst Moriond Astrophys. Meeting., p. 65. Available online at http://www-dapnia.cea.fr/Conferences/Morion_astro_2001/pdf/miyaji.pdf
Moretti A., Campana S., Lazzati D., Tagliaferri G., 2003, ApJ, 588, 696
Page M. J. et al., 2003, Astron. Nachr., 324, 101
Severgnini P. et al., 2000, A&A, 360, 457
Steffen A. T., Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., 2003, ApJ, 596, L23
Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886
Waskett T. J. et al., 2003, MNRAS, 341, 1217 (Paper I)

APPENDIX A: CATALOGUE

The tables following are a sample of the full catalogue, which can be found in the electronic version of this paper (http://www.blackwellpublishing.com/products/journals/suppmat/MNR/MNR7703/mnr7703sm.htm). It is split into three sections for each of the two fields in this survey. The first set of tables for each field (A1 and A2) contain the positions and fluxes of the X-ray sources as measured by the source detection software. The second set of tables (A3 and A4) have the identification information for all of the good ID candidates ($P' < 0.15$) including the CFDF catalogue number, the ID position, the distance between the ID and its corresponding X-ray source, P' value and redshift information. The final set of tables (A5 and A6) show the photometry for each good ID. The X-ray sources are ordered by total number of counts detected in the full X-ray band, with the greatest first. Due to vignetting, this order is approximately but not exactly the same as the flux order. Source 23 in the 14-h field is detected by SCUBA at 850 μm and is discussed in Paper I in more detail.

All six tables can also be found in their entirety, in electronic text format, at the following address: http://www.astro.cardiff.ac.uk/pub/Timothy.Waskett/

APPENDIX B

The work in this paper depends heavily on the reliability of the photometric redshift estimation codes we use. There are two codes whose results are presented in this work, one by Benítez (2000), which uses a Bayesian approach and template fitting technique,

Table A1. X-ray properties of the 3-h field *XMM* sources. Sources 7 and 19 are in fact the same source split into two owing to its lying on a PN chip gap (*).

XMM	R.A.[fk5]	Dec.[fk5]	0.5–2 keV[a]	2–10 keV[a]	0.5–10 keV[a]	PN HR[b]	M1 HR[b]	M2 HR[b]	Notes[c]
1	45.52820	−0.02260	119.0 ± 2.6	262.9 ± 9.7	382.9 ± 10.1	−0.4 ± 0.0	−0.4 ± 0.0	−0.4 ± 0.0	- q
2	45.78054	0.17228	33.0 ± 1.2	48.6 ± 3.8	81.7 ± 4	−0.6 ± 0.0	−0.5 ± 0.1	−0.5 ± 0.1	s
3	45.64490	0.01902	23.3 ± 1.0	44.3 ± 3.2	67.7 ± 3.4	−0.5 ± 0.0	−0.5 ± 0.1	−0.4 ± 0.1	
4	45.51813	0.27387	33.4 ± 1.5	76.3 ± 6.1	109.7 ± 6.3	−0.5 ± 0.1	−0.3 ± 0.1	−0.3 ± 0.1	-
5	45.63815	0.22543	16.5 ± 0.8	30.7 ± 2.7	47.2 ± 2.9	−0.6 ± 0.1	−0.2 ± 0.1	−0.4 ± 0.1	q
6	45.73817	0.26816	22.8 ± 1.2	4.3 ± 2.2	28.2 ± 2.6	−1.0 ± 0.0	−0.8 ± 0.1	−1.0 ± 0.4	
7	45.70564	0.35812	37.1 ± 2.3	85.8 ± 10.1	126.3 ± 10.3	−0.4 ± 0.1	1.0 ± 109.0	? ± ?	q_*
8	45.58501	0.32717	20.6 ± 1.2	31.4 ± 4.2	53.6 ± 4.4	−0.5 ± 0.1	−0.6 ± 0.1	−0.6 ± 0.1	s
9	45.65876	0.03438	7.0 ± 0.5	23.0 ± 2.3	30.5 ± 2.4	−0.1 ± 0.1	−0.3 ± 0.1	−0.3 ± 0.1	q
10	45.59179	0.10849	6.4 ± 0.5	19.1 ± 2.0	26.1 ± 2	−0.2 ± 0.1	−0.5 ± 0.1	−0.1 ± 0.1	q

[a]Flux in units of 10^{-15} erg s^{-1} cm^{-2}, based on a photon index of 1.7.
[b]Hardness ratio given by source detection procedure, one for each X-ray camera. Marked with '?' if not detected or a bad measurement.
[c] – = lies outside CFDF map; e = extended source (X-ray property); q = known QSO; s = object with a stellar profile, from $P' < 0.15$ list (q and s are optical properties).

Table A2. X-ray properties of the 14-h field *XMM* sources.

XMM	R.A.[fk5]	Dec.[fk5]	0.5–2 keV[a]	2–10 keV[a]	0.5–10 keV[a]	PN HR[b]	M1 HR[b]	M2 HR[b]	Notes[c]
1	214.2072	52.42472	34.6 ± 0.9	71.4 ± 3.1	106.0 ± 3.2	−0.5 ± 0.0	−0.4 ± 0.0	−0.4 ± 0.0	c d
2	214.4009	52.50781	42.2 ± 1.2	73.4 ± 4.0	115.9 ± 4.1	−0.5 ± 0.0	−0.5 ± 0.0	−0.5 ± 0.0	c d q
3	214.1816	52.24290	49.9 ± 1.5	94.2 ± 5.3	144.3 ± 5.5	−0.5 ± 0.0	−0.4 ± 0.1	−0.4 ± 0.1	
4	214.3536	52.50655	29.3 ± 1.0	53.2 ± 3.5	82.6 ± 3.7	−0.5 ± 0.0	−0.4 ± 0.1	−0.5 ± 0.1	c d q
5	214.0966	52.32077	25.5 ± 1.0	54.6 ± 3.6	80.1 ± 3.8	−0.4 ± 0.0	−0.3 ± 0.1	−0.4 ± 0.1	
6	214.4645	52.38579	19.3 ± 0.8	33.7 ± 2.7	53.1 ± 2.8	−0.5 ± 0.0	−0.4 ± 0.1	−0.5 ± 0.1	c d
7	214.2442	52.20099	31.7 ± 1.3	51.6 ± 4.8	83.8 ± 5.0	−0.6 ± 0.0	−0.4 ± 0.1	−0.5 ± 0.1	
8	214.2543	52.32128	16.7 ± 0.7	13.8 ± 1.7	30.6 ± 1.8	−0.8 ± 0.0	−0.6 ± 0.1	−0.7 ± 0.1	c d
9	214.2152	52.34575	12.6 ± 0.6	21.1 ± 2.0	33.9 ± 2.1	−0.6 ± 0.1	−0.5 ± 0.1	−0.4 ± 0.1	c d
10	214.6612	52.39937	33.3 ± 1.5	36.8 ± 5.0	70.4 ± 5.2	−0.7 ± 0.1	−0.6 ± 0.1	−0.6 ± 0.1	c d

[a]Flux in units of 10^{-15} erg s^{-1}cm^{-2}, based on a photon index of 1.7.
[b]Hardness ratio given by source detection procedure, one for each X-ray camera. Marked with '?' if not detected or a bad measurement.
[c] – = lies outside CFDF map; c = lies within Chandra map; d = detected by Chandra (c and d are X-ray properties); q = known QSO; s = object with a stellar profile, from $P' < 0.15$ list (q and s are optical properties).

Table A3. ID properties of the 3-h field *XMM* source IDs ($P' < 0.15$). Coordinates are for the CFDF objects, *not* for the *XMM* sources. The offset is between the *XMM* source and the CFDF object. The first two photo-z columns ($UBVI$ & $UBVIK$) are BPZ estimates, * indicates that $P_{\Delta z} < 0.9$ and so may be less reliable. $UBVRIZ$ is the photo-z estimate given by the CFDF code, and here * also indicates a less reliable estimate because of multiple likelihood peaks or broad errors. Note that not all IDs are included in the recent $UBVRIZ$ catalogue, so these sources do not have redshift estimates in this column. Photo-z estimates are always unreliable for QSOs (and potentially misidentified stars) and saturated objects ($I_{\rm Tot} < 18.5$), regardless of any other reliability measure. $z_{\rm sp}$ is the spectroscopic redshift. Notes have the same meaning as in Table A1.

XMM	CFDF	R.A.[fk5]	Dec.[fk5]	Offset (arcsec)	P'	$z_{\rm sp}$	$UBVI$	$UBVIK$	$UBVRIZ$	Notes
1		45.52829	−0.02246			0.641				q
2	48603	45.78065	0.17224	0.4	8.63E-04		0.01		2.01	s
3	80878	45.64505	0.01880	1.0	8.64E-03		0.40		0.45	
5	36830	45.63823	0.22535	0.4	1.65E-03	1.048	0.30	0.20	2.22	q
6	27229	45.73788	0.26803	1.1	9.46E-05		0.36		3.09 *	
7	9684	45.70320	0.35877			0.107	0.08		0.38	q
8	15331	45.58520	0.32705	0.8	1.95E-03		0.04		0.51	s
9	78735	45.65896	0.03432	0.7	7.20E-04	1.350	0.19 *			q
10	63707	45.59203	0.10864	1.0	5.86E-03	3.300	0.27	0.02	3.27	q

called BPZ, and another that is developed by one of us specifically for the CFDF (Brodwin et al. 2003) utilizing the CFRS (Hammer et al. 1995; Lilly et al. 1995b) to calibrate the template fitting. These two codes are slightly different and each have their strengths. This appendix is concerned with the reliability testing of these two codes. For a more detailed discussion of the CFDF code, and its reliability when compared against the CFRS spectroscopic sample, refer to Brodwin et al. (2003).

For reasons of timing, the input to BPZ is from an older version of the CFDF catalogues than that employed for the specific CFDF code. Therefore this should be taken into account when comparing the two codes.

Table A4. As Table A3, but for the 14-h field. All IDs are for *XMM* sources, except the last seven sources, of which four are the extra *Chandra* IDs and three are the alternative *Chandra* IDs for sources 10, 115 and 149 (assumed to be the correct IDs in this work). The two BPZ columns have been corrected for the systematic error found in this field (see Appendix B). Notes have the same meaning as in Table 2.

XMM	CFDF	R.A.[fk5]	Dec.[fk5]	Offset (arcsec)	P'	z_{sp}	*UBVI*	*UBVIK*	*UBVRIZ*	Notes
1	34649	214.2061	52.42517	0.7	5.05E-03		0.33		0.35 *	c d
2	56149	214.3996	52.50816	0.1	8.83E-05	0.985	0.12	0.01	0.35	c d q
3	32209	214.1803	52.24317	0.4	9.65E-04		1.11			
4	50800	214.3523	52.50681	0.4	8.15E-04	0.479	0.01 *	0.01	1.16	c d q
5	22314	214.0946	52.32109	1.4	1.85E-02		0.79		0.91	
6	62713	214.4630	52.38622	0.4	7.77E-04		0.35		0.35	c d
7	38711	214.2426	52.20126	0.6	1.81E-03		0.25 *			
8	39972	214.2527	52.32180	0.8	3.59E-03		0.60		0.74	c d
9	35492	214.2133	52.34607	1.1	8.10E-03		0.05 *		1.95	c d
10	83085	214.6597	52.39960	0.5	6.75E-04		0.01		1.87	c d

Table A5. Optical properties of the 3-h field *XMM* source IDs, as extracted from the original CFDF *UBVI* catalogues, including CFRS *K* photometry where available. All magnitudes are AB and measured in a 3-arcsec diameter aperture, except for the $I_{\rm Tot}$ magnitude which is the total magnitude measured using a variable aperture. This total magnitude is used as the prior in the BPZ photometric redshift code, while the 3-arcsec aperture magnitudes and errors are used as the input catalogue.

XMM	CFDF	$I_{\rm Tot}$	U	ΔU	B	ΔB	V	ΔV	I	ΔI	K	ΔK
2	48603	19.460	19.983	0.010	20.049	0.007	19.534	0.003	19.631	0.002		
3	80878	20.675	21.677	0.017	21.995	0.009	21.185	0.008	20.895	0.005		
5	36830	20.470	21.978	0.021	22.340	0.022	21.442	0.007	20.755	0.004	19.05	0.05
6	27229	14.156	19.903	0.008	18.154	0.002	16.867	0.001	15.310	0.000		
7	9684	17.771	20.683	0.012	19.184	0.004	19.143	0.002	18.295	0.001		
8	15331	19.071	20.343	0.010	19.986	0.007	19.575	0.003	19.217	0.002		
9	78735	17.924	18.909	0.005	18.654	0.003	18.481	0.002	18.082	0.001		
10	63707	19.996	24.437	0.071	21.355	0.006	20.537	0.005	20.243	0.003	19.93	0.11

Table A6. As in Table A5 but for the 14-h field. All are *XMM* IDs except the seven sources at the end, of which four are the extra *Chandra* IDs and three are the alternative *Chandra* IDs for sources 10, 115 and 149 (assumed to be the correct IDs in this work).

XMM	CFDF	$I_{\rm Tot}$	U	ΔU	B	ΔB	V	ΔV	I	ΔI	K	ΔK
1	34649	20.604	22.083	0.027	21.833	0.008	20.906	0.005	20.887	0.004		
2	56149	19.996	21.011	0.011	20.814	0.004	20.206	0.003	20.146	0.002	19.25	0.04
3	32209	19.955	21.499	0.016	21.650	0.012	21.687	0.009	20.375	0.003		
4	50800	19.762	20.918	0.011	20.549	0.003	20.338	0.003	19.932	0.002	18.62	0.02
5	22314	20.940	23.925	0.052	23.580	0.023	22.796	0.014	21.181	0.005		
6	62713	19.365	22.842	0.031	21.814	0.007	20.807	0.004	19.632	0.002		
7	38711	19.756	20.889	0.012	20.609	0.007	20.529	0.005	19.986	0.003		
8	39972	19.882	21.883	0.019	21.622	0.009	21.131	0.006	20.266	0.003		
9	35492	20.252	21.836	0.019	21.512	0.008	21.141	0.006	20.690	0.004		
10	83085	19.029	19.341	0.006	19.402	0.003	19.050	0.002	19.163	0.002		

B1 BPZ photometric redshift estimation code

This code has been tested by Gonzalez & Maccarone (2002) and Benítez (2000) and has proven to be highly successful, but we use a different filter set and photometry from these studies, and so it was prudent to re-test the code for our specific needs.

Fig. B1 shows a comparison of the BPZ photometric redshift estimates with spectroscopic measurements from the CFRS catalogue, for those sources covered by both surveys. Photometric estimates with low reliability (quantified by an in-code statistic, $P_{\Delta z} < 0.95$) are removed from these plots leaving the most reliable estimates. We achieve reasonably good results both with and without the inclusion of *K*-band photometry from the CFRS (only about half the objects here have *K* data), although the 14-h field suffers from a slight systematic underestimation, which can be seen in the figures. All BPZ redshifts have been corrected for this effect in the main body of this paper. In general, the scatter of the photometric redshifts is of the order of $\sigma \sim 0.1$ (see Fig. B1 for details).

Equivalent plots for the CFDF code can be found in Brodwin et al. (2003) (their fig. 2). In comparison to BPZ, the CFDF code redshifts fair rather better when compared with the CFRS spectroscopic sample, with fewer outliers and a smaller scatter ($\sigma \sim 0.04$ to $I_{AB} = 22.5$, $\sigma \sim 0.06$ to $I_{AB} = 24$). There are also no systematic effects, as seen in the BPZ 14-h sample.

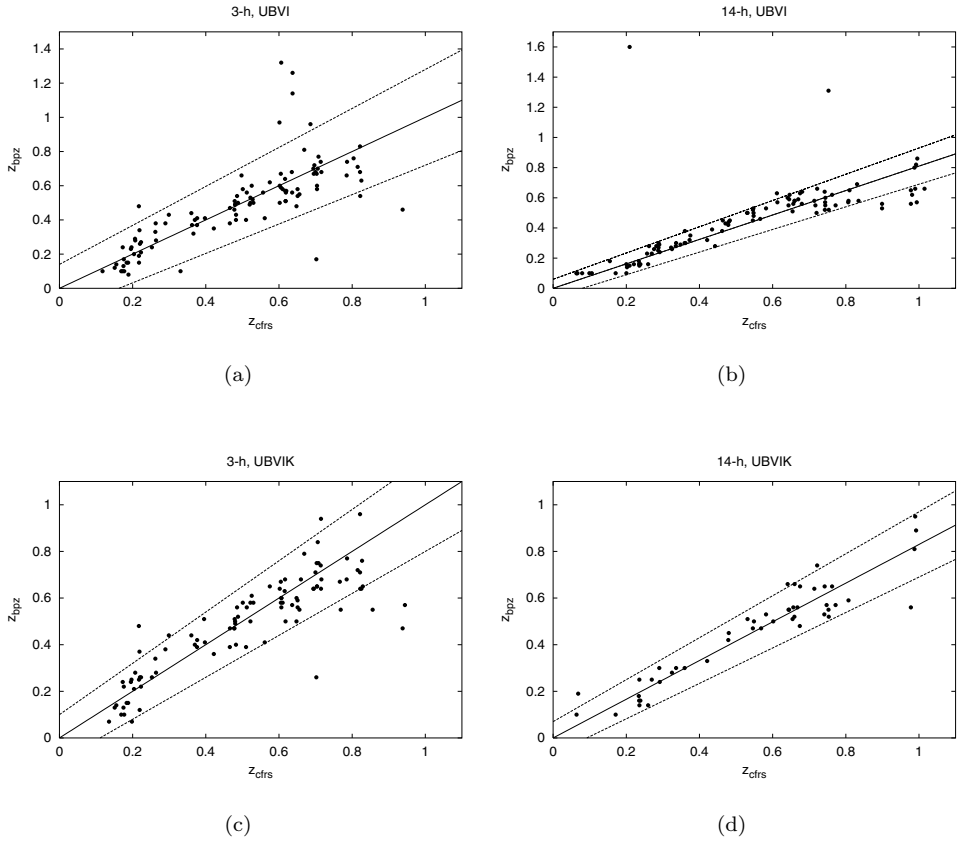


Figure B1. Photometric versus spectroscopic redshift for the 3-h [Fig. B1(a), B1(c)] and 14-h [Fig. B1(b), B2(d)] fields. The best-fitting gradients for the 14-h field are 0.81 for $UBVI$ and 0.83 for $UBVIK$. The error lines shown are of the form $\sigma(1+z)$, where $\sigma = 0.14, 0.06, 0.1, 0.07$ for the sequence of plots (for the 14-h $UBVI$ plot, this error ignores the two outliers, and is $\sigma = 0.19$ if they are included).

B2 CFDF photometric redshift estimation code

The CFDF is currently extending beyond the original $UBVI$ survey to include additional R and Z photometry. These extra filters remove some potential redshift degeneracies in certain galaxy templates and so should provide more reliable photometric redshift estimates. The CFDF photometric redshift program is now underway with the full $UBVRIZ$ photometry using a code developed by Mark Brodwin (Brodwin et al. 2003, which includes a thorough analysis of its reliability). As an additional check on the original estimates we made with BPZ, we obtained photometric redshifts for our identified X-ray sources from the CFDF photometric redshift program utilizing these new catalogues.

Fig. B2 shows the comparison between the redshift estimates made by the two codes for only those objects which had reliable estimates as judged by both of the in-code measures. Star-like objects, known QSOs and saturated objects ($I_{AB} < 18.5$) (see catalogue tables) are excluded because photometric redshifts are unreliable for these objects. In general, the agreement is good, with 79 per cent of objects agreeing to within a factor of 1.7. The agreement is also better for $z < 1$, where the majority of objects lie (72/94 CFDF; 77/94 BPZ) and where the peak in the number density of intermediate luminosity AGN is (Cowie et al. 2003). Assuming that the CFDF redshifts are correct, $7/72$ $z < 1$ objects are given poor redshifts by BPZ, whereas assuming that the BPZ redshifts are correct, $13/77$ $z < 1$ objects are given poor redshifts by the CFDF code.

For the actual X-ray sources considered in this work, only a handful have spectroscopically measured redshifts, with half of these being previously known QSOs. Fig. B3 shows the results of the photometric redshift codes for all these objects. There is a clear problem in obtaining photometric redshifts for QSOs, both codes struggling to pin them down with any accuracy. However, for more optically normal AGN, both BPZ and the CFDF code cope quite well for the greater part. The 95 per cent confidence limits are slightly better for

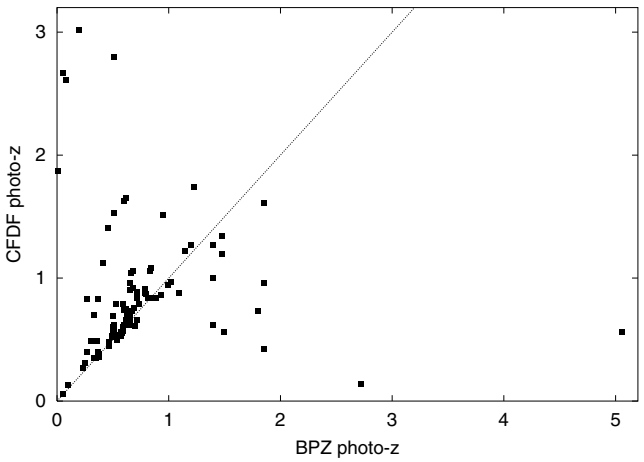


Figure B2. CFDF versus BPZ photometric redshift estimates for objects that have reliable estimates from both codes (see text for details). All possible stars and known QSOs have been excluded, as have saturated objects with $I_{AB} < 18.5$.

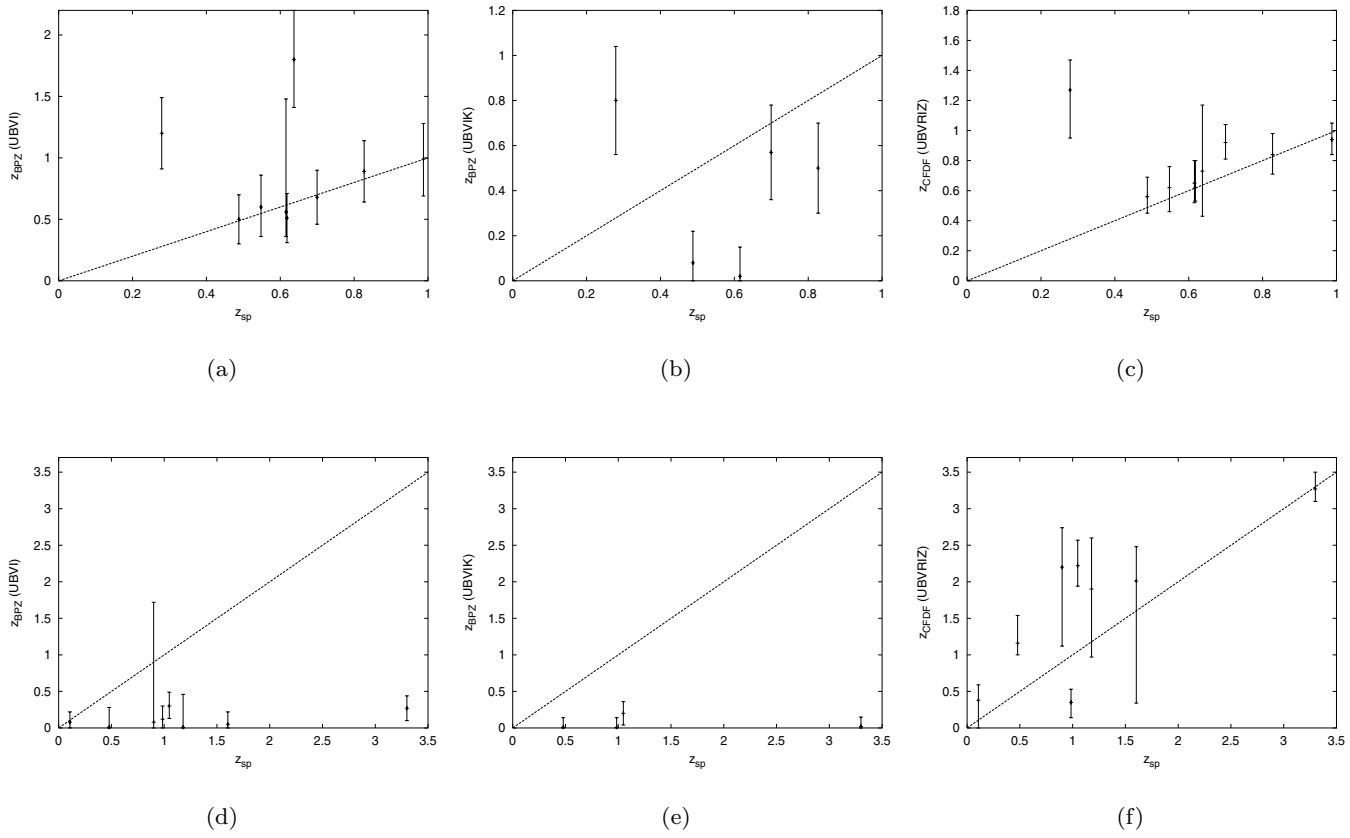


Figure B3. A series of plots to demonstrate the effectiveness of the two photo-z codes when compared to the handful of X-ray sources that have spectroscopic redshifts in our surveys. The top row [B3(a), (b) and (c)] shows the results for the non-QSO X-ray sources. The bottom row [B3(d), (e) and (f)] shows the results for the known QSOs. The left-hand plots [B3(a) and (d)] are BPZ results with $UBVI$ photometry, the middle [B3(b) and (e)] are BPZ results for the few $UBVIK$ objects and the right-hand plots [B3(c) and (f)] show the results for the CFDF code with $UBVRIZ$ photometry. The error bars in all cases are the 95 per cent confidence limits around the best-fitting redshift. The CFDF code is marginally more effective in our tests, and seems to handle QSOs more effectively, although still with larger error bars than non-QSOs. In general, for the CFRS (see main text and Fig. B1), the addition of K data appears to improve the BPZ results; however, this effect is questionable for the five X-ray sources shown in Fig. B3(b).

the CFDF code, however, and it also wins out over BPZ, with fewer unreliable redshifts in our X-ray sample. Interestingly enough, the inclusion of K-band data to the BPZ code does not improve the estimates, as we see in our tests of BPZ on the CFRS sources (above). Although the sample here, is small it actually appears to have an adverse effect on the redshift estimations (Fig. B3(b)) rather than improving them as expected.

The CFDF code, being specifically designed for the objects used as IDs in this survey, seems the logical choice for obtaining redshifts for the X-ray sources. This is especially true given that it also takes full advantage of the more recent extension in the number of filters for the CFDF. This code does allow for the use of a Bayesian prior, like BPZ, although none was used in obtaining these particular results (instead, the data itself is used to derive a prior for statistical analysis of the full sample). This may be seen as a slight disadvantage, because priors have been shown to be effective in reducing the number of catastrophic errors for individual galaxies (Benítez 2000). However, the extra photometry used in the CFDF code should compensate for this to some degree, and our tests and those in Brodwin et al. (2003) show the CFDF code to be superior to BPZ in this situation.

This paper has been typeset from a TEX/LATEX file prepared by the author.