The luminosity-dependent high-redshift turnover in the steep spectrum radio luminosity function: clear evidence for downsizing in the radio-AGN population

E. E. Rigby,[1,2]★ P. N. Best,[2] M. H. Brookes,[2] J. A. Peacock,[2] J. S. Dunlop,[2] H. J. A. Röttgering,[3] J. V. Wall[4] and L. Ker[2]

[1]*School of Physics & Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD*
[2]*SUPA†, Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ*
[3]*Leiden Observatory, PO Box 9513, 2300 RA, Leiden, the Netherlands*
[4]*Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Rd, Vancouver, BC V6T 1Z1, Canada*

ABSTRACT

This paper presents a new grid-based method for investigating the evolution of the steep-spectrum radio luminosity function, with the aim of quantifying the high-redshift cut-off suggested by previous work. To achieve this, the Combined EIS–NVSS Survey of Radio Sources (CENSORS) has been developed; this is a 1.4-GHz radio survey, containing 135 sources complete to a flux density of 7.2 mJy, selected from the NRAO VLA Sky Survey (NVSS) over 6 deg^2 of the ESO Imaging Survey (EIS) Patch D. The sample is currently 73 per cent spectroscopically complete, with the remaining redshifts estimated via the K–z or I–z magnitude–redshift relation. CENSORS is combined with additional radio data from the Parkes All-Sky, Parkes Selected Regions, Hercules and Very Large Array (VLA) COSMOS samples to provide comprehensive coverage of the radio power versus redshift plane. The redshift distributions of these samples, together with radio source count determinations, and measurements of the local luminosity function, provide the input to the fitting process.

The modelling reveals clear declines, at $>3\sigma$ significance, in comoving density at $z > 0.7$ for lower luminosity sources ($\log P = 25$–26); these turnovers are still present at $\log P > 27$, but move to $z \gtrsim 3$, suggesting a luminosity-dependent evolution of the redshift turnover, similar to the 'cosmic downsizing' seen for other active galactic nucleus populations. These results are shown to be robust to the estimated redshift errors and to increases in the spectral index for the highest redshift sources.

Analytic fits to the best-fitting steep spectrum grid are provided so that the results presented here can be easily accessed by the reader, as well as allowing plausible extrapolations outside of the regions covered by the input data sets.

Key words: galaxies: active – galaxies: evolution – galaxies: high-redshift – galaxies: luminosity function, mass function.

1 INTRODUCTION

It has become increasingly apparent in recent years that radio-loud active galactic nuclei (AGN) play a key role in galaxy evolution; the interplay of their expanding radio jets and the surrounding intergalactic and intracluster medium acts to provide part, or possibly all, of the heat required to prevent both large-scale cluster cooling flows and the continued growth of massive ellipticals (e.g. Best et al. 2006,

2007; Bower et al. 2006; Croton et al. 2006; Fabian et al. 2006). Determining the evolution of the radio luminosity function (RLF) is therefore important for understanding the time-scales on which they impose these effects. Also, since radio-loud AGN are powered by the most massive black holes, their RLF can be used to investigate the behaviour of the upper end of the black hole mass function and hence the build-up of these objects in the early Universe.

The work of Sandage (1972), Osmer (1982), Peacock (1985), Schmidt, Schneider & Gunn (1988) and, in particular, Dunlop & Peacock (1990, hereafter DP90) has shown that the comoving number density of both flat- and steep-spectrum powerful radio galaxies, selected at 2.7 GHz, is greater by 2 to 3 orders of magnitude at a

★E-mail: emmaerigby@gmail.com
†Scottish Universities Physics Alliance

redshift of 2 compared with the present-day Universe. This density increase is expected to peak at some point simply because sufficient time is needed for their host galaxies to grow into the massive ellipticals, with correspondingly large central black holes, that are typically observed for radio-loud AGN (e.g. Best, Longair & Röttgering 1998). This high-redshift 'cut-off' was seen by Peacock (1985) in the flat-spectrum population and was also detected in the steep-spectrum population by DP90 beyond $z \sim 2.5$; but their results were limited by the accuracy of the photometric redshifts from their faintest radio-selected sample, so they were not able to quantify the decline. It should also be noted that the DP90 work assumed an Einstein–de Sitter cosmology, which means that the high-redshift sources in their sample were ascribed lower luminosity than for a $\Omega_\Lambda > 0$ cosmology, thus potentially making a cut-off easier to find.

Following DP90, Shaver et al. (1996) reported evidence of a sharp cut-off in space density in their sample of flat-spectrum radio sources. However, this result was disputed by Jarvis & Rawlings (2000) who showed that it could be caused by an increasing curvature of the spectral indices with redshift, and that a shallower decline was more consistent with the data. A more rigorous analysis of radio-loud quasars by Wall et al. (2005), using a larger sample, confirmed a decrease in the number density of flat-spectrum sources at $z \gtrsim 3$.

A shallow space density decline between $z \simeq 2.5$ and 4.5 was also found in low-frequency-selected steep-spectrum sources by Jarvis et al. (2001), although a constant density value was also consistent with their data; their sample lacked the depth needed for firm results. Waddington et al. (2001) used a deeper survey and saw evidence that the turnover for lower luminosity sources appeared to occur at lower redshift than that of brighter flux-limited samples. They were also able to discount some of the DP90 models, but their study lacked the volume needed for better measurements of the space density changes of powerful radio sources. Indications of a similar luminosity dependence of the cut-off redshift were also seen for radio-loud Fanaroff–Riley type I (FR I) (Fanaroff & Riley 1974) sources by Rigby, Snellen & Best (2008).

The evolution of radio-selected AGN can be linked to the behaviour of AGN selected in other bands. For example, the space density of optically selected quasi-stellar objects (QSOs) shows a strong decrease at $z > 2.1$ (Boyle et al. 2000; Fan et al. 2001; Wolf et al. 2003; Fan et al. 2004), consistent with that found for both radio-loud (Wall et al. 2005) as well as X-ray selected quasars (Hasinger, Miyaji & Schmidt 2005; Silverman et al. 2005). There is also evidence for a luminosity-dependent redshift cut-off in the optical and X-ray selected QSO samples (Ueda et al. 2003; Hasinger et al. 2005; Richards et al. 2005; Wall 2008). Since radio-loud QSOs are thought to correspond to flat-spectrum sources (e.g. Barthel 1989; Antonucci 1993), investigating the evolution of the flat and steep populations can result in a new understanding of the links between radio-loud (quasars and radio galaxies) and radio-quiet (QSOs) sources. A full review of radio-loud AGN evolution can be found in De Zotti et al. (2010).

To properly investigate the evolution of the steep-spectrum RLF, it is clear that a combination of several radio surveys of differing depth is needed to ensure a broad coverage of the radio luminosity–redshift (P–z) plane. In particular, faint radio samples are needed if the $z \gtrsim 2$ behaviour is to be determined. This has motivated the development of the 150-object, 1.4-GHz selected, Combined EIS–NVSS Survey of Radio Sources (CENSORS; Best et al. 2003a, hereafter Paper I), which has been designed to maximize the information for high-redshift, steep-spectrum, radio sources close to the break in the RLF.

In this paper the CENSORS data set, combined with additional samples, is used to investigate the nature of the high-redshift evolution of radio sources, via a new grid-based modelling technique in which no prior assumptions are made about the behaviour of the luminosity function. This is an improvement on previous investigations which have either used functional forms, or only considered pure luminosity or density evolution, or a combination of both [although Dye & Eales (2010) have recently developed a similar method to study the evolution of submm galaxies].

The layout of the paper is as follows. Section 2 describes both CENSORS and the additional data sets needed. Section 3 presents the modelling technique. Section 4 describes the results from the best-fitting model and investigates their robustness. Finally, Section 5 summarizes the findings. Throughout this paper, values for the cosmological parameters of $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$ are used and the spectral index, α, is defined as $S_\nu \propto \nu^{-\alpha}$.

2 INPUT DATA

As discussed above, several data sets are needed to constrain the radio source cosmic evolution. In addition to the CENSORS sample, therefore, four other radio samples, along with determinations of the local RLF (LRLF), and measurements of the radio source counts are used; these are described in this section. Fig. 1 illustrates the coverage of the P–z plane obtained using these radio samples.

2.1 CENSORS

The full CENSORS sample contains 150 sources with $S_{1.4\,\mathrm{GHz}} > 3.8$ mJy in a $3 \times 2 \, \mathrm{deg}^2$ field of the ESO Imaging Survey (EIS) Patch D, centred on $09^\mathrm{h}51^\mathrm{m}36\fs0$, $-21°00'00''$ (J2000). Paper I presents the radio data, along with the optical host galaxy identifications, with additional K-band imaging presented by Brookes et al. (2006, hereafter Paper II). Spectroscopic data for a subset of the sample can
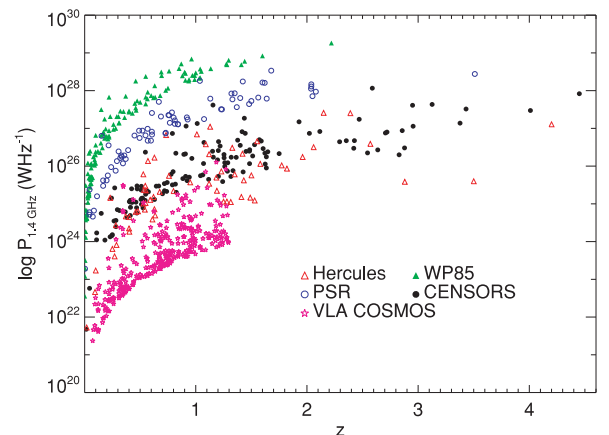

Figure 1. The Wall & Peacock (1985, WP85), Parkes Selected Regions (PSR; Downes et al. 1986; Dunlop et al. 1989), CENSORS, Hercules (Waddington et al. 2001) and VLA-COSMOS (for $z \leq 1.3$ only; Smolčić et al. 2008) samples plotted on a radio luminosity versus redshift plane to illustrate how they efficiently cover a large part of the plane without much overlap. Radio luminosities were calculated using previously published spectral indices for the WP85, PSR and Hercules samples; $\alpha = 0.8$ was assumed for the sources in VLA-COSMOS. The spectral indices for the CENSORS sample are taken from Ker et al. (in preparation). The PSR, WP85 and COSMOS samples are restricted to steep-spectrum sources only. See text for full details of sample selection.

be found in Brookes et al. (2008, hereafter Paper III), which also gives estimated redshifts for the remainder of the sample, calculated using the K–z and, for one source, the I–z magnitude–redshift relations.

Since the publication of Papers I–III some small reassessments have been made to the sample. Subsequent radio data have shown the following.

(i) CENSORS 66 and CENSORS 82 are actually the lobes of a FR II radio source whose host galaxy is located at $09^h50^m48^s.97$, $-21°32'55''.8$ (Gendre, private communication), with a K-band magnitude of 17.8 ± 0.1 in a 4.5-arcsec diameter aperture ($K = 18.2 \pm 0.1$ when corrected to the standard 63.9-kpc diameter aperture; see Paper II for details) and $I = 21.7 \pm 0.1$. This results in an estimated K–z redshift of 1.40 (calculated using $\log z = 0.0025K^2 + 0.113K - 2.74$; cf. Paper III).

(ii) Similarly, CENSORS 84 and CENSORS 85 are also the lobes of an extended double radio source whose host galaxy is located at $09^h55^m36^s.87$, $-21°27'12''.5$, with a K-band aperture-corrected magnitude of 13.1 ± 0.2 and a corresponding estimated K–z redshift of 0.15.

(iii) The host galaxy of CENSORS 64 is located at $09^h49^m01^s.60$, $-20°50'00''.7$ (Ker et al., in preparation) with K- and I-band magnitudes of 15.04 ± 0.08 (when aperture corrected) and 17.44 ± 0.02 respectively, with a corresponding estimated redshift of 0.33. This compares well to its new spectroscopic redshift of 0.403 (Ker et al., in preparation).

Spectral indices for the sources, calculated using the new radio data, will be presented in Ker et al. (in preparation), but are also included here in the analysis of Fig. 3.

Additional K-band imaging was obtained for 14 sources using IRIS2 (in service mode) on the Anglo-Australian Telescope. As a result of this, the host galaxy of CENSORS 69 has now been detected, with a K-band magnitude of 20.1 ± 0.3 (19.6 ± 0.3 when aperture corrected). New spectroscopic data subsequently showed it to have a redshift of 4.01 (Ker et al., in preparation). Finally, the K–z limits presented in Paper II were discovered to be in error and have been recalculated. The correct values and corresponding new K–z redshift limits, along with the new IRIS2 K-band data, can be found in Appendix B (with the online version of the paper – see Supporting Information), within the up-to-date CENSORS data set.

Of the original 150 CENSORS sources, 135 are deemed to be complete to a flux density limit of 7.2 mJy, and it is this subset which is used in this paper. No other selection is performed, meaning that it contains both star-forming galaxies along with steep- and flat-spectrum sources. The current host galaxy identification fraction of this is 96 per cent with a spectroscopic completeness of 73 per cent. Table 1 summarizes the salient information for this subsample.

2.2 The Wall and Peacock 2.7-GHz radio sample

The sample of Wall & Peacock (1985, hereafter WP85) includes the brightest radio sources at 2.7 GHz over an area of 9.81 sr, and is complete to 2 Jy. The original paper presents redshifts for 171 of the 233 source sample, and a further 20 were added by Wall & Peacock (1999). Since then, further spectroscopic observations have been published by a variety of groups, raising the redshift completeness to 98 per cent. As part of the back-up programme for CENSORS, two of the WP85 sources, 0407−65 and 1308−22, have been observed with FORS2 on the Very Large Telescope (see Appendix A with the online version of the article – Supporting Information). The redshift measured for 0407−65 is in good agreement with that observed by Labiano et al. (2007), but the observation of 1308−22 updates the redshift estimate quoted in McCarthy et al. (1996). For the remaining two sources without useful spectra, the estimated redshifts of WP85 are used. There is an excess of sources in this sample at $z < 0.1$, some of which may be contaminating starburst galaxies, so a minimum redshift of 0.1 is imposed here; doing this does not degrade the analysis, as this region of parameter space will be well constrained by the LRLF.

Complete spectral indices are available for this sample making it straightforward to convert the 2.7-GHz flux densities to 1.4 GHz; since it is only the behaviour of the steep-spectrum luminosity function which the model will assess (as discussed in Section 3), only steep spectrum WP85 sources are considered in this.

In order to use this sample in the modelling of the RLF, it needs to be converted into a sample with an effective 1.4-GHz flux density limit. To do this, a flux limit of $S_{1.4\,GHz} = 4$ Jy was adopted, corresponding to a spectral index $\alpha = 1.06$ for sources at the 2.7-GHz limit. It is possible that sources with a steeper spectral index than this may be below the WP85 sample flux limit at 2.7 GHz and still have a 1.4-GHz flux density above 4 Jy, thus leading to incompleteness in the sample. A search was therefore carried out for such sources, so they could be added to the sample. In the Northern hemisphere, this is simple as there is a fainter ($S_{2.7\,GHz} \geq 1.5$ Jy) sample from Peacock & Wall (1981) covering the WP85 area; this contains one source with a flux lower than 2 Jy, but a corresponding 1.4-GHz value above 4 Jy (and $z \geq 0.1$). In the Southern hemisphere, there are two 408-MHz surveys, the Best et al. (2003b) equatorial sample and Burgess & Hunstead (2006), which between them cover the rest of the WP85 survey region. Since these are lower frequency surveys, any such steep-spectrum sources will be bright and can be identified. Searching these samples reveals two $z \geq 0.1$ steep-spectrum sources brighter than 4 Jy at 1.4 GHz, but not in the WP85 sample. Oddly, calculating the 2.7-GHz flux densities of these sources reveals that they should have been detected by WP85. It is not clear why they were missed, but it is possible that they have very curved spectra. The data for these three 'missing' sources are

Table 1. Key information relating to the five radio samples used in the modelling. N(total) gives the total number of sources in each sample as used here; $N(z_{spec})$, $N(z_{photo})$, $N(z_{est})$ and $N(z_{limit})$ give the number of sources with spectroscopic, photometric, estimated (via either the K–z or I–z relations) or lower limit K–z redshifts, respectively.

Name	1.4-GHz limit (Jy)	Sky area (sr)	N(total)	$N(z_{spec})$	$N(z_{photo})$	$N(z_{est})$	$N(z_{limit})$
WP85	4.0	9.81	83	79	–	4	–
PSR	0.30	0.075	74	41	–	30	3
CENSORS	0.0072	0.0018	135	99	–	31	5
Hercules	0.002	0.000 38	64	42	19	1	2
VLA-COSMOS	0.0001	0.000 36	314	–	314	–	

Table 2. The additional sources that were added into the WP85 sample to correct for incompleteness. The z_{ref} gives the reference for the redshift – G05 for Grimes, Rawlings & Willott (2005) or BH06 for Burgess & Hunstead (2006).

Name	$S_{2.7\,GHz}$ (Jy)	$S_{1.4\,GHz}$ (Jy)	α	z	z_{ref}
3C325	1.84	4.29	1.29	1.135	G05

Name	$S_{408\,MHz}$ (Jy)	$S_{1.4\,GHz}$ (Jy)	α	z	z_{ref}
1526$-$423	17.86	5.08	1.02	0.5	BH06
1827$-$360	25.83	6.49	1.12	0.12	BH06

given in Table 2, whilst Table 1 summarizes the WP85 sample as used in this paper (including the three sources discussed above); the full data set can be found in Appendix C (with the online version of the article – see Supporting Information).

2.3 The Parkes Selected Regions 2.7-GHz radio sample

The Parkes Selected Regions (PSR; Wall, Cole & Milne 1968; Downes et al. 1986; Dunlop et al. 1989) cover 0.075 sr in six 6.5 deg^2 fields of view down to a flux density limit of 0.1 Jy at 2.7 GHz. The updated sample of Dunlop et al. (1989) presents redshifts for 82 of the 178 sources in the sample, and subsequent observations published by a variety of groups result in a further 24 redshifts. Of the remaining 72 sources, 10 have estimated values from Dunlop & Peacock (1993), whilst the remainder were estimated using the K–z relation as outlined for the CENSORS sources in Paper III. The K-band photometry for the sample (Dunlop et al. 1989) uses 12.4-arcsec diameter apertures for most objects. These are large enough such that aperture corrections are sufficiently small to be ignored. Where a K-band magnitude is not available, one is estimated via the relation: $K = -1.1(B - R) + 18.3$ (Dunlop et al. 1989). Two sources have been omitted from this sample due to unclear identification and non-detection.

The current PSR data set suffers from some incompleteness below 0.15 Jy, becoming quite large by 0.1 Jy (DP90), so there is also a desire to minimize the contribution of these faintest sources. The 1.4-GHz flux density limit is therefore set at 0.3 Jy to achieve this. All 2.7-GHz sources with a 1.4-GHz flux density greater than this are included in the final sample; the lowest of these has a 2.7-GHz value of 0.14 Jy which should be sufficiently above the incompleteness limit to avoid problems. Table 1 summarizes the PSR sample as it used in this paper, whilst the full data set can be found in Appendix C.

2.4 The Leiden–Berkeley Deep Survey Hercules sample

The Leiden-Berkeley Deep Survey covers 5.52 deg^2 over nine high-latitude fields, and was originally based upon multicolour plates from the 4-m Mayall telescope at Kitt Peak (Kron 1980; Koo & Kron 1982) alongside Westerbork Synthesis Radio Telescope 1.4-GHz radio observations (Windhorst, van Heerde & Katgert 1984b). One of these fields, in the constellation of Hercules, has subsequently been followed up by Waddington et al. (2000, 2001; see also Windhorst, Kron & Koo 1984a), who defined a complete sample of 64 radio sources (both star-forming galaxies and steep- and flat-spectrum objects) with $S_{1.4\,GHz} > 2$ mJy within 1.2 deg^2. The spectroscopic redshift completeness of the sample is 66 per cent

[41 redshifts measured by Waddington et al. (2001) with one additional value from Rigby, Snellen & Best (2007)]. Of the remaining 22 sources, 20 have photometric redshifts also from Waddington et al. (2001), but the final two only have estimated K–z lower limits due to host galaxy non-detections. Table 1 summarizes the salient information for the Hercules sample, whilst the full data set can be found in Appendix C.

2.5 The AGN subsample of the VLA-COSMOS survey

Smolčić et al. (2008) defined a sample of 601 AGN with $z \leq 1.3$ in the 2 deg^2 COSMOS field using the multiwavelength imaging available for the region. The redshift limit was imposed because beyond this the AGN/star-forming galaxy separation becomes unreliable, leading to possible contamination of the sample. The radio sensitivity varies with position across the field, but over a well-defined area of 1.17 deg^2 it is possible to define a clean sample complete to $S_{1.4\,GHz} \geq 100\,\mu$Jy; this contains 314 steep-spectrum sources. Robust photometric redshifts were calculated for the survey and are available for all of these sources. Table 1 summarizes the salient information for the Very Large Array (VLA)-COSMOS sample; however, unlike the other samples, this full data set is not available in Appendix C as it is still proprietary, and was obtained via private communication with the authors.

2.6 Source counts data

Any model of RLF evolution must match the measured radio source counts as a function of flux density and so this comparison will also be used in the modelling process. The data used for this were taken from Bondi et al. (2008), Seymour, McHardy & Gunn (2004), Windhorst et al. (1984b), White et al. (1997) and Kellermann & Wall (1987) to ensure that a sufficient range of flux densities (0.05 mJy to 94 Jy) was covered. The White et al. counts are limited to $S_{1.4\,GHz} > 2$ mJy only, as they note that they are incomplete below this. The full set of 1.4-GHz source counts used in the modelling can be found in Appendix C and the counts themselves are shown in Fig. 6, within the discussion of the modelling results (Section 4.1).

2.7 Local radio luminosity functions

The LRLF for AGN was measured by Sadler et al. (2002) using the 2 degree Field Galaxy Redshift Survey (2dFGRS; Colless 1999, Colless et al. 2001), by Best et al. (2005)) and Best et al. (in preparation), using the Sloan Digital Sky Survey (SDSS; York et al. 2000), and by Mauch & Sadler (2007) using the 6 degree Field Galaxy Survey Data Release 2 (6dFGS DR2; Jones et al. 2004). Sadler et al. (2002) compiled their LRLF from a sample of all 2dFGRS, NRAO VLA Sky Survey (NVSS) selected radio galaxies with $z \leq 0.3$. Note that while the 2dFGRS LRLF includes points at $\log P_{1.4\,GHz} = 25.9$ and 26.3 (as converted to the cosmology used here) these are not included due to the small number (1) of sources in each band. Best et al. (in preparation) use the seventh data release of the SDSS, combined with the NVSS and FIRST 1.4-GHz radio surveys [using a similar process to that carried out by Best et al. (2005) for the second data release, but now with an independent normalization], to produce a sample of 9168 radio sources, with a median redshift of ~0.1. Similarly, the Mauch & Sadler (2007) LRLF was calculated using the 7824 NVSS radio sources contained in the second incremental data release of the 6dFGS; the median redshift for their sample is 0.043. In addition, 95 per cent confidence upper limits are added at the highest radio luminosities

(log $P_{1.4\,\text{GHz}} > 27.0$) as no sources were detected at these luminosities in the three data sets. The LRLF data used in the modelling can be found in Appendix C.

3 MODELLING TECHNIQUE

Unlike previous attempts to model the evolution of the RLF (e.g. Dunlop & Peacock 1990; Jarvis et al. 2001; Willott et al. 2001, etc.), no assumptions are made about the shape of the luminosity functions here. Instead, they are determined by allowing the densities, ρ, at various points on a P–z grid of radio luminosities and redshifts to each be free parameters and then simply finding the best-fitting values to this many-dimensional problem. The P–z grid points were chosen so as to allow sensitive calculations to be made without having so many parameters involved that finding a best fit becomes a prohibitively long task (but with more powerful computers and larger data sets the technique could be expanded to sample much finer detail). The range of radio luminosities covered was $\log P_{1.4\,\text{GHz}} = 19.25$–$29.25$, equally spread in steps of 0.5 in $\log P_{1.4\,\text{GHz}}$. Densities were evaluated at the redshifts 0.1, 0.25, 0.5, 1.0, 2.0, 3.0, 4.0 and 6.0. The ρ at any (P, z) can then be interpolated from the nearest four grid points, except at $z < 0.1$ where the densities are assumed to be constant (see Section 3.1 below).

The P–z plane is constrained by the samples described in Section 2. This is demonstrated in Fig. 2, which highlights the different regions which are covered by different data types. The densities that are actually included in the fitting process are those that are constrained by the faintest redshift distribution or the source counts; grid points that are unconstrained are excluded. As Fig. 2 shows, high-redshift, low-luminosity sources do not satisfy this criterion and are therefore not included in the modelling. The total number of points fitted, and hence the dimensions of the minimization process, is 101.

The `amoeba` algorithm for downhill simplex minimization (Nelder & Mead 1965) is used to obtain the best-fitting space densities. It takes as input a set of parameters and a scaling factor and uses these to construct a geometrical object of $N + 1$ points in N dimensions called a simplex. It uses a user-defined function to cal-

culate the likelihood of each vertex of the simplex, and then reflects, contracts or expands these vertices until the function is minimized. In order to achieve this best fit in reasonable time, the algorithm was run in a multistage loop with varying scaling and tolerance for the first five steps. The maximum number of iterations allowed within `amoeba` for each stage is set to 3000, and the process ends when the likelihood ratio of successive stages is less than 1.000 0001.

The subsections below describe the calculation of the input modelling parameters, along with an explanation of the likelihood calculation used.

3.1 Construction of the input density grids

Three P–z grids are used as inputs to the modelling program containing the source densities for steep-spectrum, flat-spectrum and star-forming populations separately (the total radio source space density obviously being the sum of these three at any grid point). Considering these three grids separately is essential since the five different radio source samples used in the modelling include and exclude different populations, as well as on the good physical grounds that the different populations might well evolve differently. As the aim of this work is to investigate the evolution of the radio galaxy RLF, the star-forming grid is fixed. Its inclusion is necessary as the low flux densities of the CENSORS and Hercules samples mean that radio emission from star formation becomes significant. The flat-spectrum radio source grid is also held constant since the low numbers of this type of source in the input samples do not allow minimization; more accurate constraints come from previous surveys explicitly targeting these sources.

The star-forming grid is created by evolving the local star-forming galaxy luminosity function of Sadler et al. (2002), such that

$$\Phi(P, z) = \begin{cases} \Phi\left(\dfrac{P}{(1+z)^{2.5}}, 0\right) & z \leqq z_{\text{max}}, \\ \Phi\left(\dfrac{P}{(1+z_{\text{max}})^{2.5}}, 0\right) & z > z_{\text{max}}, \end{cases} \quad (1)$$

where Φ is the comoving density of radio sources due to star formation, P and z are the luminosity and redshift respectively, and z_{max} is the redshift where the space density plateaus. The power-law index of 2.5 is taken from Smolčić et al. (2009) who studied star-forming galaxies in VLA-COSMOS out to $z \sim 3$; their value agreed with that previously found by Seymour et al. (2004). The value $z_{\text{max}} = 2$ is adopted, following Blain et al. (1999), but its precise value is irrelevant to this work as the contribution of star-forming galaxies at $z > 1$ is negligible at the flux densities studied.

Similarly, the starting steep-spectrum grid (which the minimization then varies) is formed by evolving the Sadler et al. (2002) local AGN RLF by $(1 + z)^3$ in density. Conversely, the flat-spectrum grid is created by taking the median value from the results of the seven evolutionary models presented in DP90, after conversion to the cosmology used here. This is in broad agreement with the recent results of Wall et al. (2005), and minor variations are not critical given that this population is small compared to steep-spectrum sources.

An S–z (flux density–redshift) grid is, in many cases, more readily compared with real data than the P–z grid, as it can be converted easily into source numbers. The modelling code therefore uses the three P–z grids to populate three corresponding S–z grids containing 120 flux density bins covering the range 0.1 mJy to 50 Jy at 1.4 GHz, equally spaced in $\log S$, and 300 redshift bins covering the range 0–6, equally spaced in z. This z-range is wider than that previously used to define the P–z grid in order to cover the full range of the radio samples. To account for this, two extra bins, $z = 0.001$ and
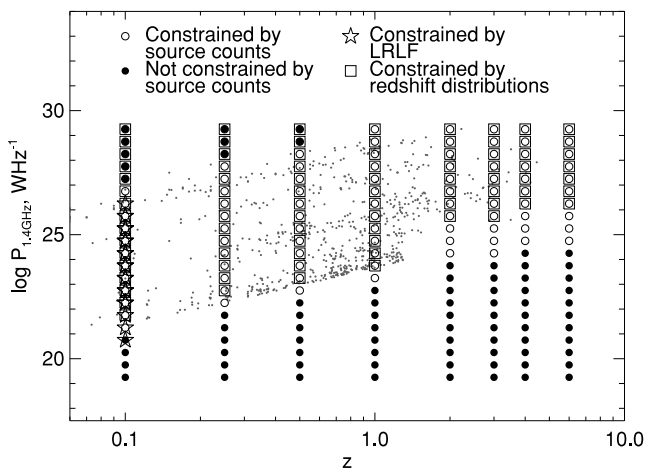

Figure 2. The various regions of the P–z plane that are constrained by the redshift distributions, the source counts and the LRLF. The redshift distribution constraint is determined by the COSMOS sample at $z \leq 1.3$ and the Hercules sample above this, as they are the deepest surveys; the positions of the individual sources in the five samples are shown as grey dots. The source counts cover the S–z grid used in the modelling process, which takes a lower limit of 0.1 mJy and an upper limit of 50 Jy.

0.05, are inserted into the *P*–*z* grids. For the steep-spectrum grid, the densities for these additional grid points are assumed to be constant with redshift, and are therefore set to the $z = 0.1$ values; for the flat-spectrum and star-forming grids, they are calculated from the input models. For a given *P*–*z* grid, the density at a given point on the *S*–*z* grid, $\rho(S, z)$, is found by linear interpolation of the density values from the four surrounding points in the expanded *P*–*z* grid. The total number of sources per steradian in each bin is then given by

$$N(S, z) = \rho(S, z) \frac{\mathrm{d}V}{\mathrm{d}z} \mathrm{d}(\log S) \, \mathrm{d}z. \qquad (2)$$

The star-forming and flat-spectrum *S*–*z* grids are only calculated once as they are not changed in the minimization process. The steep-spectrum *S*–*z* grid is recalculated in each cycle of the `amoeba` process as $\rho(S, z)$ changes when $\rho(P, z)$ changes. Although the source counts extend to lower flux densities, it is not useful to extend the modelled region because source counts alone do not provide a sufficient constraint on the RLFs.

3.1.1 Spectral index selection

The calculation of the appropriate luminosity for a given flux density and redshift bin in the *S*–*z* grid requires a value for the spectral index, α. For the flat-spectrum and star-forming grids, α is assumed to be 0.0 and 0.8, respectively; the small contribution of these populations means that a more precise value is not necessary. However, the steep-spectrum spectral index may vary with redshift and luminosity, which could significantly alter the results. This *P*–*z*–α degeneracy means that either parameter could be used to give the α variation. Unlike DP90, who adopted a *P*–α relation, it is the redshift dependence which is used here for the steep-spectrum grid; Ker et al. (in preparation) will investigate these spectral index variations in more detail. The default form for this variation is taken from Ubachukwu et al. (1996) who found that the mean spectral index increases with redshift as $\alpha = 0.83 + 0.4\log(1 + z)$; as Fig. 3 shows, this relation gives a reasonable approximation to the available steep-spectrum sample data, with a possible overestimation
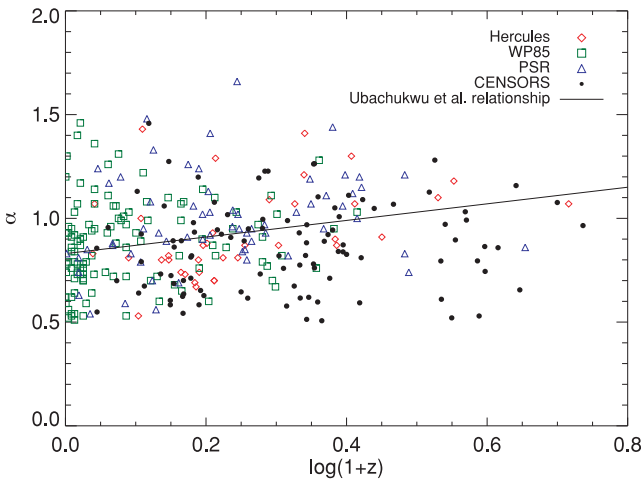
of α at high redshift. The CENSORS spectral indices in particular would also be consistent with a flat relation, but this may be because they are low frequency, making them appear flatter than the high-frequency WP85 and PSR values. Section 4.3.1 considers the effects other α assumptions, including a constant value, have on the modelling, though it should be noted that reasonable variations in α do not lead to qualitative differences in the results.

In addition to this, a Gaussian scatter in α of 0.2 is also incorporated at each redshift in the steep-spectrum grid to account for any variations in the value within a redshift bin; this is a reasonable assumption given the spread seen in Fig. 3. In practice, this is implemented by creating 21 versions of the *S*–*z* grid, extending to $\pm 2.5\sigma$ (in steps of 0.25σ), which are each assigned a weight depending on how far they are away from the mean. Redshift bins where $\alpha < 0.5$ are ignored (as sources within them would not be in the steep-spectrum sample) and their weight evenly distributed over the remainder. The *P*–*z* grid densities, $\rho(P, z)$, corresponding to each (S, z) point are then interpolated on to the bins in these 21 new *S*–*z* grids as before; the final grid carried forward into the minimization is their weighted sum.

3.2 Comparing the model predictions to data

The input data from the redshift distributions of the five different radio source samples described in Section 2, along with the LRLF and the observed source counts, now need to be compared to the model grids described in the previous section. To do this, the model LRLFs are simply read from the appropriate $(z = 0.1)$ column in the *P*–*z* grid, and the model redshift distributions, $N(z, S_{\mathrm{limit}})$, for some sample with survey area A and flux limit S_{limit}, and source counts, $N(S)$, are easily calculated from the *S*–*z* grid using

$$N(z, S_{\mathrm{limit}}) = A \times \sum_{S > S_{\mathrm{limit}}} N(S, z) \qquad (3)$$

and

$$N(S) = \sum_z N(S, z), \qquad (4)$$

respectively. Different data sets are compared to different combinations of the model grids, depending on their content. For instance, the sum of the steep- and flat-spectrum *P*–*z* grids is fitted to the LRLF data as they were created using only AGN, but the source counts are compared against the sum of all three *S*–*z* grids as they include star-forming galaxies. Table 3 summarizes which of the grids are used for comparison to each data set.

Figure 3. The changes in spectral index with redshift for the steep-spectrum sources in the four radio samples with measured α values. Also shown is the relation from Ubachukwu et al. (1996) which follows the data points reasonably well. The spectral indices were calculated using 1.4–0.6 GHz, 1.4–5 GHz, 2.7–5 GHz and 1.4 GHz to 325 MHz flux densities for the Hercules, PSR, WP85 and CENSORS samples, respectively.

Table 3. The grids used for each data set comparison in the modelling process. 'Steep', 'stars' and 'flat' indicate the grids for steep-spectrum, star-forming and flat-spectrum grids, respectively. Note that it is only the steep-spectrum grid which is allowed to vary in the modelling process.

Data set	Comparison grid	
	P–*z*	*S*–*z*
VLA-COSMOS		steep
Hercules		steep+stars+flat
CENSORS		steep+stars+flat
PSR		steep
WP85		steep
LRLFs	steep+flat	
Source counts		steep+stars+flat

A χ^2-test, $\chi^2 = \sum_i \frac{(O_i - E_i)^2}{\sigma^2}$, is used to assess the model predictions for each data set in turn. For the source counts and the LRLF, the data points and their uncertainties provide the values of O and σ. It is impractical, given the large number of data points involved (and the wish to add more as they become available), to choose model grid points that exactly match the locations (flux density or luminosity) of the data, and therefore the values of E are calculated by interpolating between neighbouring model grid points. The χ^2 comparison also includes a term for the upper limits on the high radio luminosity points in the LRLF; this is done by setting the 'data' points for these luminosities equal to zero and setting the corresponding errors equal to the upper limit of the LRLF.

When comparing the model predictions with the redshift distributions, it is important to note that there are insufficient sources at the high-redshift end of the redshift distributions (a part of the parameter space that is of particular interest) to formally allow a χ^2-test to be used. A possible solution to this would be to carry out a source-by-source likelihood analysis instead, calculated using the product of the model prediction for the redshift probability distribution at the redshift of each of the sample sources; this returns the maximum-likelihood value if the data exactly match the model distribution. A problem with this method, however, is the difficulty of accounting for the uncertain redshifts present in several of the samples.

An alternative solution is therefore adopted here, by matching the shape of the redshift distributions using polynomial fits, and comparing these with the model predictions. The smoothing that this provides at high redshifts mitigates the issue of small source numbers in the high-redshift bins. It also helps to lessen anomalous features in the redshift distributions, such as the drop in source numbers at $z \sim 1.5$ in the CENSORS, Hercules and PSR samples (Fig. 4). The latter arises because of the onset of the 'redshift desert', where spectroscopic redshifts are difficult to obtain: the K–z estimates should fill this gap, but in practice the scatter in the K–z relation, combined with the lack of spectroscopic measurements, means there is a overall bias for redshifts to lie outside of this region.

The polynomials are fitted to histograms with bin sizes of 0.1 in z, beginning at $z = 0.0$, with the exception of WP85, which, as discussed previously in Section 2.2, has a minimum redshift of 0.1. Additionally, the VLA-COSMOS sample was only defined to $z = 1.3$ (due to lack of AGN/star-formation separation beyond that redshift), so no fitting is allowed beyond this. Once the value of a polynomial fit reaches zero it is set to zero for all subsequent redshifts, to prevent it returning negative source numbers.[1] The order of polynomial selected for each redshift distribution is the one that gives the minimum reduced chi-squared – 6 for CENSORS, Hercules and COSMOS, 3 for PSR and 4 for WP85 – and these input fits are shown in Fig. 4.

The polynomial fitting takes into account the uncertainty from both estimated and photometric redshifts present in these data sets, by representing each source as a Gaussian distribution centred on the given redshift, with the width equal to either the published photometric redshift error, or, for the K–z estimates, 0.14 in $\log z$ [the 1σ spread in the 7C K–z relation (Willott et al. 2003)]. The K–z limits can similarly be modelled, but are assigned a width (in $\log z$) of 0.4 above and 0.1 below the given value to represent their increased uncertainty (see Paper III for more discussion). This distribution is

then included in the redshift histograms, thus spreading the source across different redshift bins, depending on the errors. The effect of these redshift uncertainties on the modelling results is investigated, in Section 4.3.2, by changing these assumptions.

The polynomial fits are then compared to the model predictions using a χ^2-test, with the polynomial values providing the E parameters in these equations, evaluated at redshifts and flux densities which match points in the model grid, and the O values given by the model prediction at those grid points. The corresponding σ values are taken as the uncertainty in the polynomial value, generated from the covariance matrix of the polynomial-fit coefficients.

Using polynomial fits to represent the redshift distributions in this way has one drawback: although the fitting method should return the correct minimum point, and thus identify the correct best-fitting model grid, the absolute value of χ^2 at that point will not necessarily be a true measure of the goodness-of-fit of the model. This is because the polynomial-derived errors are correlated. Any scaling error in χ^2 will then lead to a misestimate of the uncertainties on the model predictions (see Section 3.3). To account for this, the χ^2 values derived using the polynomial-fit method for the best-fitting model prediction were compared with 'true' χ^2 values. The latter were derived by comparing the best-fitting model with the observed source numbers in different redshift bins, but with all estimated redshifts treated as exact, and the bin sizes increased to ensure that each redshift bin contained a minimum of five sources; this avoids problems with small number statistics, but means that resolution is lost at the high-redshift end. This analysis showed that the polynomial-fit method provided values of χ^2 which were a factor $f = 0.3$ away from the true value. The χ^2 values from the redshift distributions are therefore scaled by this factor in order to ensure that accurate uncertainties are determined.

The scaled χ^2 values are transformed to likelihoods, and combined to produce an overall measure of the goodness of the fit:

$$L = \exp\left(\frac{-\chi^2_{\text{LRLF}}}{2}\right) \exp\left(\frac{-\chi^2_{\text{counts}}}{2}\right) \exp\left(\frac{-f\chi^2_{z-\text{dists}}}{2}\right), \quad (5)$$

which is then returned to the minimization routine `amoeba`.

3.3 Constraint in the *P*–*z* grid

It is not sufficient to find the best fit to the data without obtaining some measure of the uncertainty associated with each point on the P–z grid. In this section the intrinsic modelling limits offered for the constraint of the P–z grid are presented, whilst the effect of uncertainty in the input data is discussed later in Sections 4.3.1 and 4.3.2.

Assuming an ideal data set, uncertainty still arises due to the degeneracy across parameters in the model. The conditional error for parameter p is given by-

$$\sigma^2_{\text{cond,p}} = \left(-\frac{\partial^2 \ln L}{\partial p^2}\bigg|_{\text{peak}} \right)^{-1}. \quad (6)$$

In practice, it is determined by finding the value of the parameter for which $\ln L(p) - \ln L(\hat{p}) = 0.5$, from which $\sigma = \rho(p) - \rho(\hat{p})$. This is then a measure of the uncertainty in a given parameter whilst holding all other parameters constant.

However, the actual uncertainty may be greater when the variation of other parameters is taken into account. This is the marginalized error and comes from the diagonal of the inverse Hessian matrix,

---

[1] The polynomial fitting is not done in log space as this appeared to provide a poorer match to the total number of sources in each sample at high redshift.
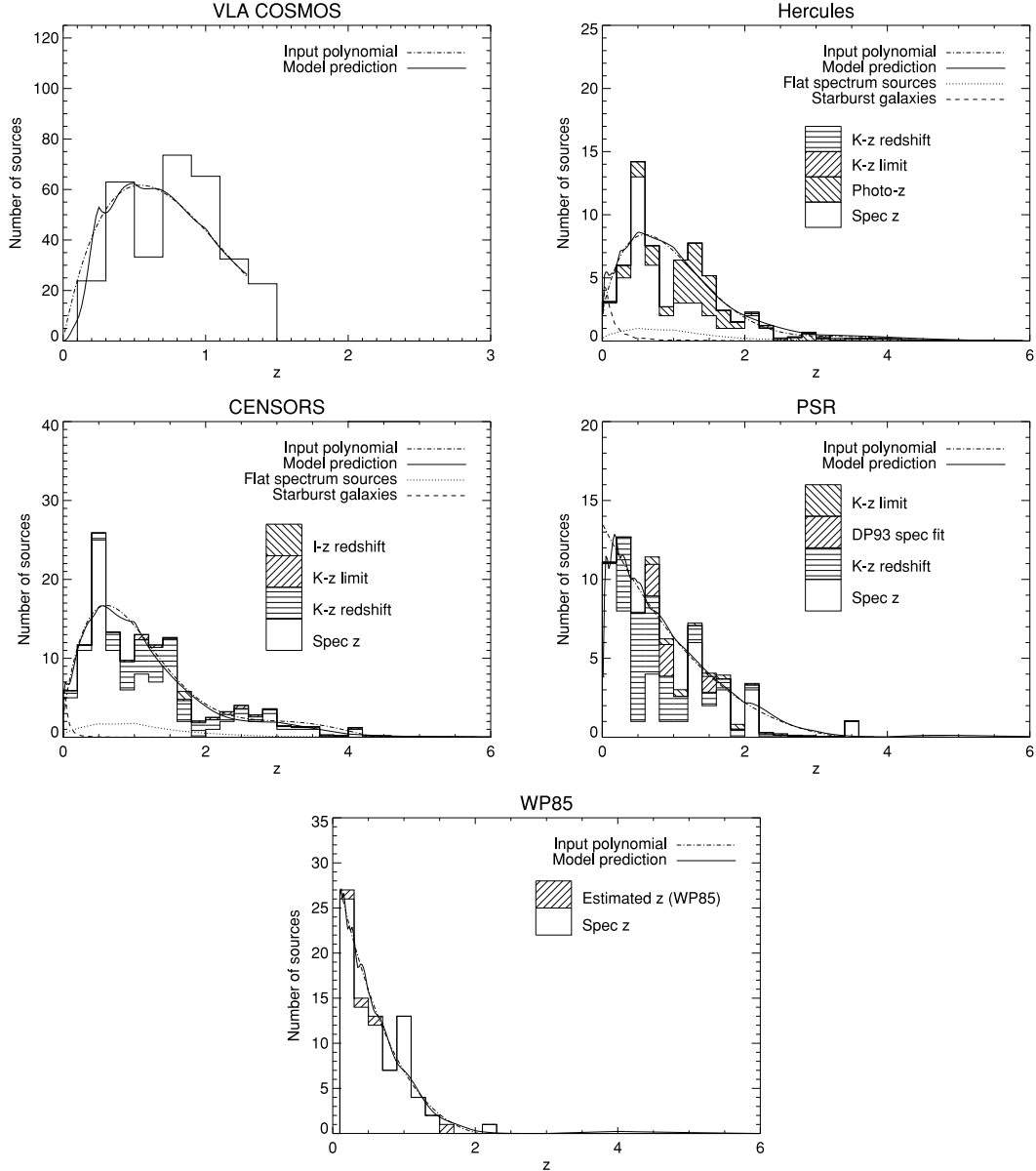
Figure 4. The redshift distributions for the five samples, CENSORS, Hercules, WP85, PSR and VLA-COSMOS, included in the modelling. The histograms are plotted with a bin size of 0.2 in z and incorporate the uncertainties arising from the limits and estimated redshifts by representing them as weighted Gaussian distributions (see Section 3.2 for details). The overplotted solid lines show the model predictions for each sample, and the bin shading shows the contribution from the different redshift determinations. For samples which were compared to some combination of the three input grids, the contributions from the additional grids are also shown separately to illustrate the dominance of the steep-spectrum population at high redshift. See Table 3 for information on which grids were used for each data set.

H, where

$$\sigma^2_{\mathrm{marg},i} = [H]^{-1}_{i,i}, \tag{7}$$

for parameter i, where,

$$H_{i,j} = \frac{\partial^2 \ln L}{\partial p_i \partial p_j}. \tag{8}$$

This is calculated here via the finite difference approximation (note that diagonal terms in the Hessian matrix are simply $-1/\sigma^2_{\mathrm{cond},p}$). It is these marginalized errors that are quoted in the results presented in the remainder of this paper. In general, these are comparable to the conditional errors, but for some grid points they are up to a factor of 2 higher.

4 RESULTS AND DISCUSSION

This section presents the best-fitting steep-spectrum P–z grid produced by running the modelling code described in Section 3 above, and examines how well its predictions agree with the input data sets. The best-fitting P–z grid and its associated error is given in full in Appendix D (with the online version of the article – see Supporting Information).

4.1 Data set comparison

The success of the best-fitting steep spectrum P–z grid, combined where necessary with the unvaried flat-spectrum and starburst grids, at fitting the input data is illustrated in Figs 4–6 for the five sample
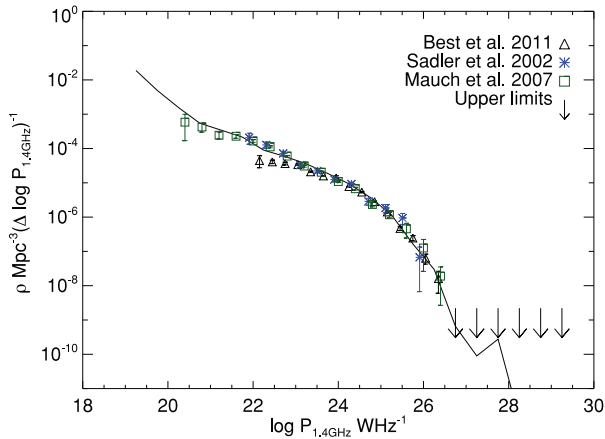
Figure 5. The data for the three input LRLFs used in the modelling process (points) along with the best-fitting model LRLF prediction (solid line) from the $z = 0.1$ column in the combined steep- and flat-spectrum final P–z grids.
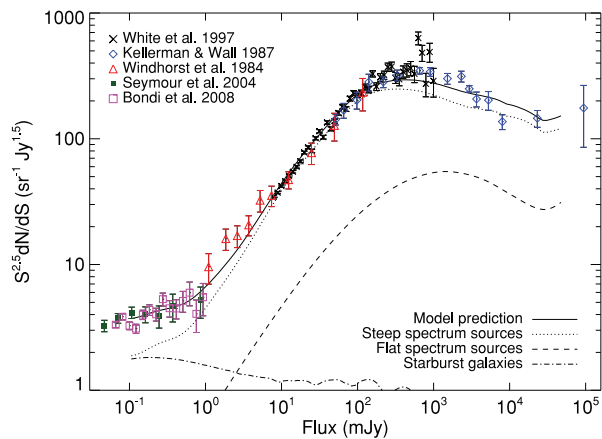

Figure 6. The source count data used in the modelling process (points) along with the best-fitting prediction for the counts from the three final model S–z grids (solid line). Also shown are the contributions to the model counts for the steep-spectrum (dotted line), flat-spectrum (dashed line) and star-forming galaxies (dot–dashed line) separately to illustrate the dominance of the steep population.

redshift distributions, the LRLF and the differential source counts, respectively. The agreement with both the LRLF and the source counts is very good – this is to be expected for two reasons. The constraint provided by the LRLFs is tight and particular to specific parameters, meaning that the model has little freedom to vary it. At the opposite extreme, there are numerous combinations of densities which will satisfy the observed counts, as they depend on the summation carried out across the P–z grid to transform it to S–z. Thus, whilst fitting these data is essential, on their own they do not provide a particularly interesting constraint. The source counts comparison plot (Fig. 6) also breaks down the model prediction into the different contributions from the three populations; this shows that the steep-spectrum sources dominate at the flux densities probed by the current samples, which in turn justifies limiting the fitting to them only at this point.

The redshift distributions also derive from the S–z grid so they prevent non-sensical combinations of densities fitting the source counts, and the stronger constraint that they provide is therefore essential to obtaining information about the evolution of the luminosity function. Again the flat and starburst populations are plotted
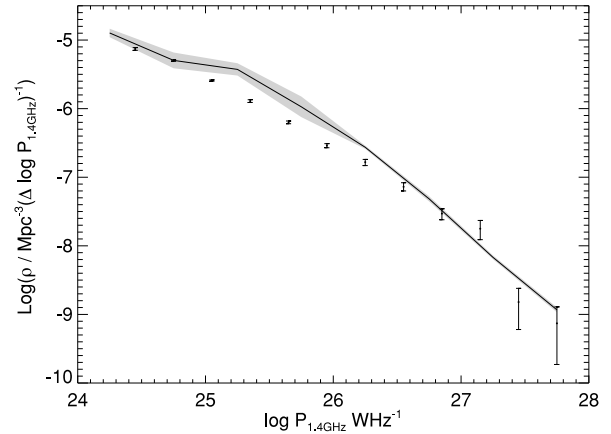

Figure 7. The RLF data points of Donoso, Best & Kauffmann (2009) compared to the best-fitting model prediction at $z = 0.5$ from the steep-spectrum grid (solid line). The shaded region shows the 5σ confidence limits of the model.

separately in Fig. 4 where relevant, showing their small contribution, especially at the high redshifts which are of particular interest here.

The model predictions for all five samples are generally in good agreement with the data across the whole of each redshift range. The total numbers of sources given by the model are 131.1, 67.7, 74.5, 284.8 and 84.5 for CENSORS, Hercules, PSR, COSMOS and WP85, respectively; these are well matched to the actual figures of 135, 64, 74, 314 and 83.

As a further check on the results, the final model grid can also be compared to data sets that were not included in the fitting: for example the $\bar{z} = 0.55$ RLF determined by Donoso et al. (2009) using a sample of $\sim$14000 radio-loud AGN, created by combining the NVSS and FIRST 1.4-GHz radio surveys with the MegaZ-LRG catalogue; this comparison is shown in Fig. 7. The match is reasonable over the luminosity range of the data, and the overprediction seen at $25 < \log P < 26$ is likely to be because Donoso et al. were only considering radio sources associated with luminous red galaxies in their sample, and therefore miss any bluer radio galaxies (this is also why this sample was not included as an explicit constraint).

4.2 Model luminosity functions

Fig. 8 shows the behaviour of the individual best-fitting steep-spectrum luminosity functions with z for $\log P = 25.25$ to $\log P = 27.75$ inclusive, averaged over bins of 0.5 in $\log P$. This luminosity range was chosen as it covers the region of the P–z plane with the most constraints as illustrated in Fig. 2; additionally, high-redshift points are only plotted if they are constrained by both the redshift distributions and the source counts. Also plotted as a comparison is the median of the seven evolutionary models calculated by DP90 (translated to the cosmology used for this work), which all suggest a mild high-redshift decline in the number density of sources of these powers.

The evolving luminosity function (in which the grid densities are plotted against $\log P$ for different redshifts) is shown in Fig. 9. As before, regions of radio power and redshift are excluded if they are not constrained by at least two data sets. Displaying the grid behaviour in this way is a useful alternative to Fig. 8, as it gives an overview of the space density changes with radio power and redshift.
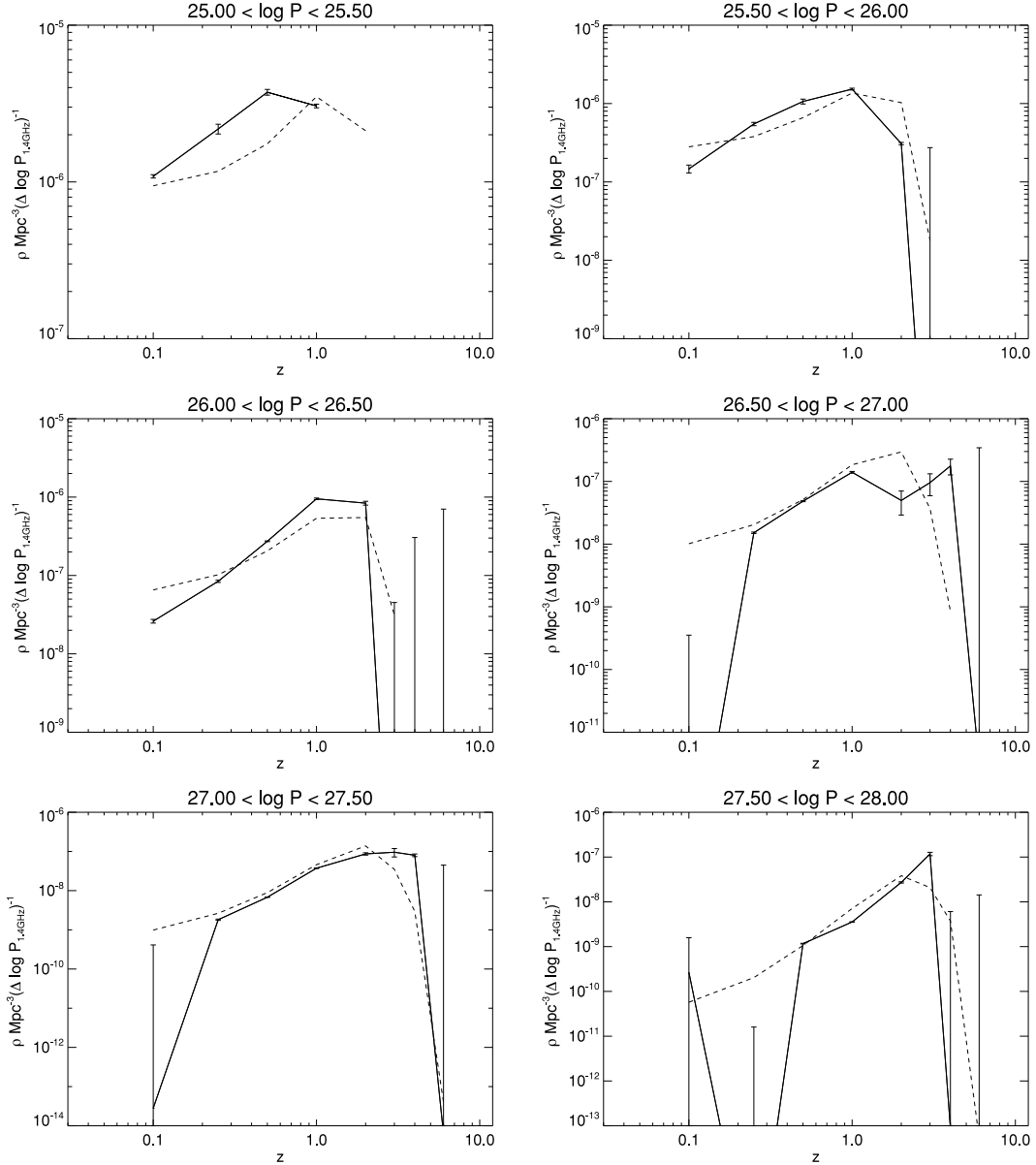
Figure 8. The individual model steep-spectrum RLFs versus redshift from the best-fitting *P–z* grid (solid lines). Points are only plotted here if they are constrained by at least two of the input data sets (see Fig. 2 for details). Also shown are the median of the evolutionary models from DP90 (dashed lines).

The low-*z*, $\log P \geq 26.5$ points are discrepantly low (albeit with large error bars) compared to their neighbours. This is a result of the weaker constraints provided by the upper limits of the LRLF to the minimization process (though this region is also constrained by the source counts and redshift samples). Inspection of Fig. 5 (the combined LRLF from the steep- and flat-spectrum grids) illustrates that the model reaches well below the upper limits of the densities in these regions. Similarly, there are some apparently anomalous high-redshift points (e.g. the 'dip' seen at $z \sim 2$ for $26.5 < \log P < 27$ and the $z = 3$ points at $\log P < 26.5$); Figs 2 and 1 suggest that although these should be constrained by both the redshift distributions and the source counts, the low densities are likely to result from a lack of sources in the radio samples covering this range. However, the 'dip' in particular may not just be the result of low-number statistics as it lies within the 'redshift desert' discussed previously in Section 3.2.

4.3 The high-redshift turnover

Inspection of Figs 8 shows that decreases in space density are present to some degree at $z \gtrsim 0.5$ for all the luminosities considered. At low powers these declines are clear and occur at $z \sim 1$, but at the highest powers the densities remain essentially level with no strong decline, out to $z \sim 4$. The agreement with the DP90 results is reasonable, though visually there seems to be a trend for the low power cut-offs to be at lower redshifts than DP90, and the high power cut-offs to be at higher redshift. However, the range in the DP90 results, and the inability to probe the full distance range for these low powers, means it is difficult to draw firm conclusions about the differences.

The 'strength' of the cut-off, C_σ, between the density at the peak redshift, ρ_{peak}, and that at any of the n subsequent redshift points
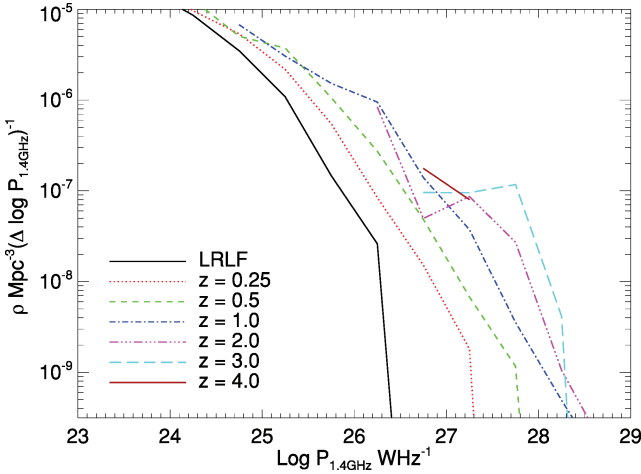
Figure 9. The evolving luminosity function. Points are only plotted here if constrained by more than one data set and errors have been ignored for clarity.

can be quantified as

$$C_\sigma = \frac{\rho_{\rm peak} - \rho_{\rm peak+n}}{\sqrt{(\sigma_{\rm peak}^2 + \sigma_{\rm peak+n}^2)}}, \qquad (9)$$

where $\sigma_{\rm peak}$ and $\sigma_{\rm peak+n}$ are the corresponding errors on the peak and post-peak space densities. However, caution is necessary when calculating C_σ for the results presented here, because of the discrepant, badly constrained, points discussed in the previous section. The quantized nature of the grid also means that it is often relatively flat after the peak, before dropping sharply. To minimize these effects, $\rho_{\rm peak+n}$ is taken as the average of the space densities at redshifts higher than the peak position. This average is weighted by the available volume in each redshift bin, and the $z = 6$ point is ignored in all cases due to the lack of supporting data in the redshift distributions. The error, $\sigma_{\rm peak+n}$, for this average density is calculated by combining the post-peak space density errors in quadrature, taking the volume weighting into account. The results of this calculation are shown in Table 4; they support the previous observation that the declines are strong for the faintest powers, but tend to be weaker at brighter powers.

Fig. 8 and Table 4 both show an apparent luminosity dependence of the peak redshift, $z_{\rm peak}$, but the wide redshift bins in the *P–z* grid means the position of the peak is ambiguous. For a better estimate, polynomials (generally of order 2, but also of order 3 where necessary) were fitted to the model steep-spectrum $\rho - z$ distributions, for various radio power bins, with the aim of roughly parametrizing this $z_{\rm peak}$ versus $\log P$ relation. Fig. 10 shows the results of this for the best-fitting steep-spectrum grid, with error bars showing the uncertainties in the polynomial fit. Also shown is the range in $z_{\rm peak}$ found from repeating this fitting for the different spectral indices and redshift limits discussed in Sections 4.3.1 and 4.3.2. This is not a rigorous analysis but it does illustrate the general increase in $z_{\rm peak}$ over the radio luminosity range probed.

4.3.1 The effect of the spectral index

The creation of the *S–z* grid, used in the modelling for the data from the five samples and source counts, requires a value for the spectral index, α. The choice of α is complicated by the spectral curvature seen in some radio-loud sources (e.g. Laing & Peacock 1980), which can result in an increase in the spectral indices at

Table 4. The redshift at the grid point at which the space density is highest, $z_{\rm grid}$, and the strength of the cut-off, C_σ following that point from the average of the post-peak densities (excluding the $z = 6$ points), using equation (9). Also shown is the peak redshift, $z_{\rm fit}$, determined from the polynomial fitting described in Section 4.3. The different versions illustrate the effects of altering various model parameters; see Sections 4.3.1 and 4.3.2 for full details.

| Radio power range	$z_{\rm grid}$	$z_{\rm fit}$	C_σ
Default: $\alpha = 0.83 + 0.4\log(1+z)$			
25.00–25.50	0.5	0.7	3.8
25.50–26.00	1.0	1.1	9.6
26.00–26.50	1.0	1.4	5.6
26.50–27.00	4.0	2.3	–
27.00–27.50	3.0	2.6	0.7
27.50–28.00	3.0	3.9	10.0
$\alpha = 0.8 + 0.25z$			
25.00–25.50	1.0	0.8	–
25.50–26.00	1.0	1.2	2.4
26.00–26.50	1.0	1.4	4.0
26.50–27.00	2.0	1.6	2.1
27.00–27.50	4.0	2.6	–
27.50–28.00	3.0	3.2	5.9
$\alpha = 1.5$			
25.00–25.50	1.0	0.7	–
25.50–26.00	0.5	0.7	1.5
26.00–26.50	1.0	1.4	4.8
26.50–27.00	1.0	1.4	4.0
27.00–27.50	4.0	2.8	–
27.50–28.00	2.0	2.3	11.4
$\alpha = 0.8$			
25.00–25.50	0.5	0.7	2.4
25.50–26.00	1.0	0.9	7.9
26.00–26.50	1.0	1.4	16.1
26.50–27.00	4.0	2.5	–
27.00–27.50	2.0	1.9	6.7
27.50–28.00	3.0	3.8	6.5
Average strength from 50 random variations of the redshift limits			
25.00–25.50	0.5	0.7	1.8
25.50–26.00	1.0	1.0	8.6
26.00–26.50	1.0	1.4	2.6
26.50–27.00	4.0	2.1	–
27.00–27.50	3.0	2.6	0.2
27.50–28.00	3.0	4.0	9.8
Uncertain z at $+1\sigma$			
25.00–25.50	1.0	0.8	–
25.50–26.00	1.0	1.3	6.5
26.00–26.50	1.0	1.4	6.7
26.50–27.00	2.0	1.8	5.9
27.00–27.50	3.0	2.5	4.4
27.50–28.00	3.0	3.6	15.5
Uncertain z at -1σ			
25.00–25.50	0.5	0.7	3.5
25.50–26.00	1.0	1.2	3.1
26.00–26.50	1.0	1.3	5.0
26.50–27.00	1.0	2.3	2.4
27.00–27.50	3.0	2.1	1.1
27.50–28.00	3.0	3.8	4.6
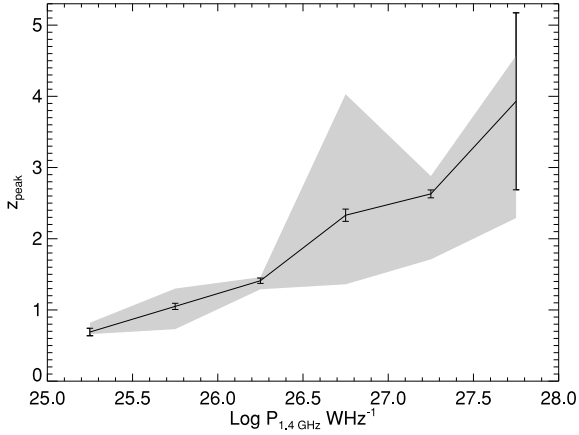
Figure 10. An illustration of the changes in peak redshift with radio power for the best-fitting steep-spectrum grid (solid line). The error bars show the uncertainty in the polynomial fits, and the shaded region represents the range in results which come from repeating this process for the additional grids in Sections 4.3.1 and 4.3.2.

higher redshift (see Jarvis & Rawlings 2000, for a discussion of this effect for flat-spectrum sources); as Fig. 3 shows, this effect is seen in the radio samples used here, albeit with a large scatter. Ignoring this may lead to underestimation of the high-z space density (and increase the significance of any density cut-off) since the steepest spectrum sources are missed. In the modelling results presented in the previous section, attempts were made to take this into account by using the α–z relation from Ubachukwu et al. (1996) in the creation of the S–z grid. This provides a reasonable match to the data (Fig. 3), but it is important to investigate the effect a different choice has on the high-redshift behaviour of the model RLFs. Changing the assumed value of α will move sources into different radio power bins and hence strengthen some turnovers, and weaken others.

Fig. 11 presents the results of using both an extreme spectral index of 1.5, and a stronger increase with redshift (arbitrarily modelled as $\alpha = 0.8 + 0.25z$ to give ultra-steep spectra at $z \gtrsim 2$). Note that these will represent extreme cases since some radio sources with typical $\alpha \sim 0.8$ spectra are found at high redshift (e.g. Jarvis et al. 2009). The cut-off strengths are also given in Table 4. In both cases, the general effect is to increase the densities at $z \gtrsim 0.3$, and weaken the high-redshift cut-off, as sources in the S–z grid have moved into higher radio power bins. Overall, this is a good illustration of how far α needs to be increased to reduce the significance of the cut-offs to the $\sim 3\sigma$ level. Also shown, as a comparison, is the effect of using $\alpha = 0.8$, the low-redshift mean value; this generally decreases the densities but the decline is still present at high significance for all powers.

4.3.2 The effect of the redshift incompleteness

The other uncertainty in the model results comes from the estimated redshifts and redshift limits present to some degree in the five input samples. Attempts are made to take these into account in the modelling process, but this is likely to be less successful for the z-limits as their true value is less constrained. To investigate what effect this has, the model was run 50 times; in each run each limit is assigned a new, higher redshift, drawn randomly from a uniform 10000 element distribution, starting at the given limit, up to a maximum of
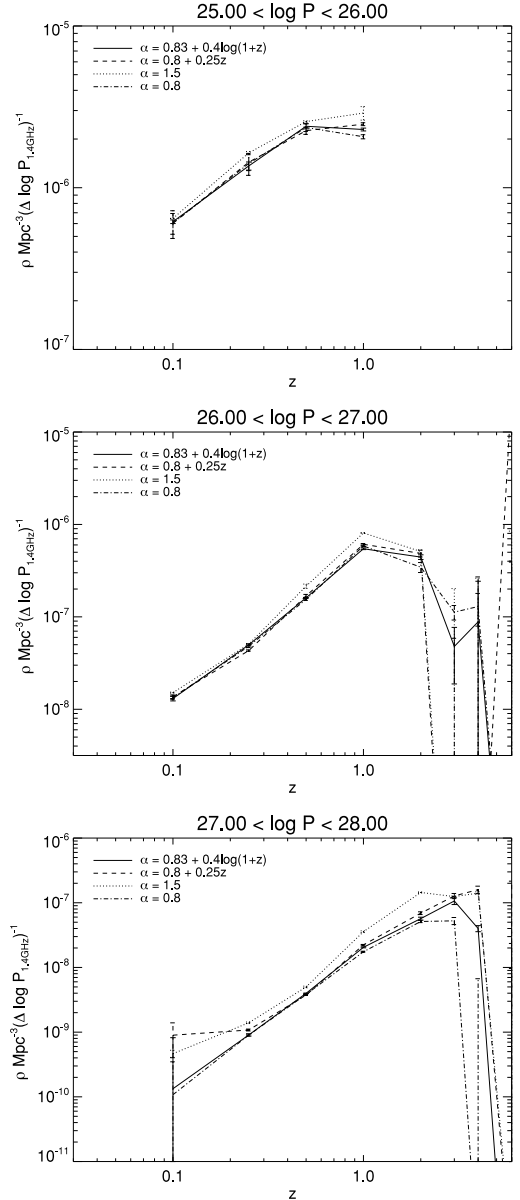

Figure 11. The effect on the steep-spectrum model luminosity functions of changing the value of the spectral index, α, for three luminosity ranges. The '$\alpha = 0.83 + 0.4\log(1 + z)$' line comes from the best-fitting model grid presented earlier. Larger bins of $\Delta \log P = 1$ are shown to better illustrate how the changes affect the results. In all cases, a scatter of $\alpha \pm 0.2$ around the mean value was also included, as previously discussed in Section 3.1.

$z = 6$. As a further check, the modelling was repeated with all the uncertain redshifts, both estimates and limits, moved up and down by 1σ. No other attempt is made to account for the uncertainty in the estimated redshifts and limits in either of these cases.

Figs 12 and 13 show the spread in densities resulting from this, and the cut-off strengths can be found in Table 4. They show that whilst the turnover is preserved in both cases, it is generally strengthened at moderate powers when all the uncertain redshifts are shifted by $\pm 1\sigma$, compared with the original results, because sources are shifted out of these regions to higher power and redshift bins. When the redshift limits are randomly increased the turnover is weakened due to sources being moved into the higher redshift bins.
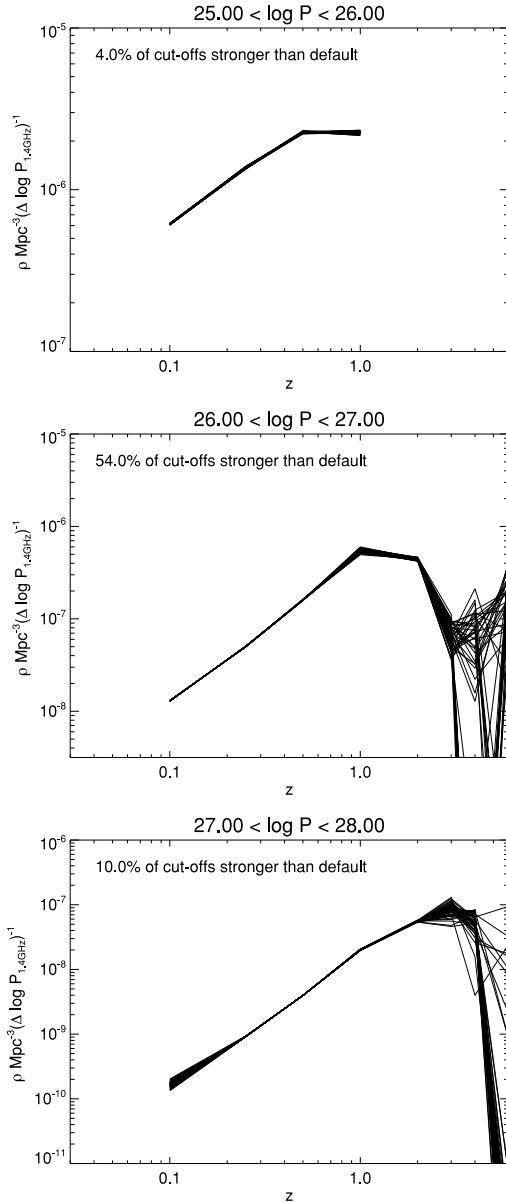
Figure 12. The effect on the steep-spectrum model luminosity functions of repeatedly randomly moving the redshift limits to a new, higher value, up to a maximum of $z = 6$. For this the estimated redshifts were kept at their given values, with no attempt to take their uncertainties into account as done previously. 'Default' in the figure labels refers to the original modelling results presented in Section 4.2. Larger bins of $\Delta \log P = 1$ are shown to better illustrate how the changes affect the results. In all cases an error of $\alpha \pm 0.2$ was also included, as previously discussed in Section 3.1.

Figure 13. The effect on the steep-spectrum model luminosity functions of moving the estimated redshifts and redshift limits up and down by their 1σ values (these were taken as $+0.4$ and -0.1 for the limits); no other attempt is made to take the redshift uncertainties into account. The '$\alpha = 0.83 + 0.4\log(1 + z)$' line comes from the best-fitting model grid presented earlier. Larger bins of $\Delta \log P = 1$ are shown to better illustrate how the changes affect the results. In all cases an error of $\alpha \pm 0.2$ was also included, as previously discussed in Section 3.1.

4.4 Testing the robustness of the redshift turnover

The excellent coverage of the P–z plane in the range $26 < \log P < 28$ to $z \sim 5$, demonstrated in Fig. 1 and 2, allows a further test of how robust the redshift cut-offs seen in Fig. 8 are to possible incompleteness in the radio samples. This is done by determining the number of fake high-redshift radio sources that need to be inserted into this luminosity range to reduce the cut-off strength to $<3\sigma$. In practice, this was split into two parts because of the changing position of the cut-off with luminosity and the range of the different samples. First, different numbers of new Hercules sources, with redshifts
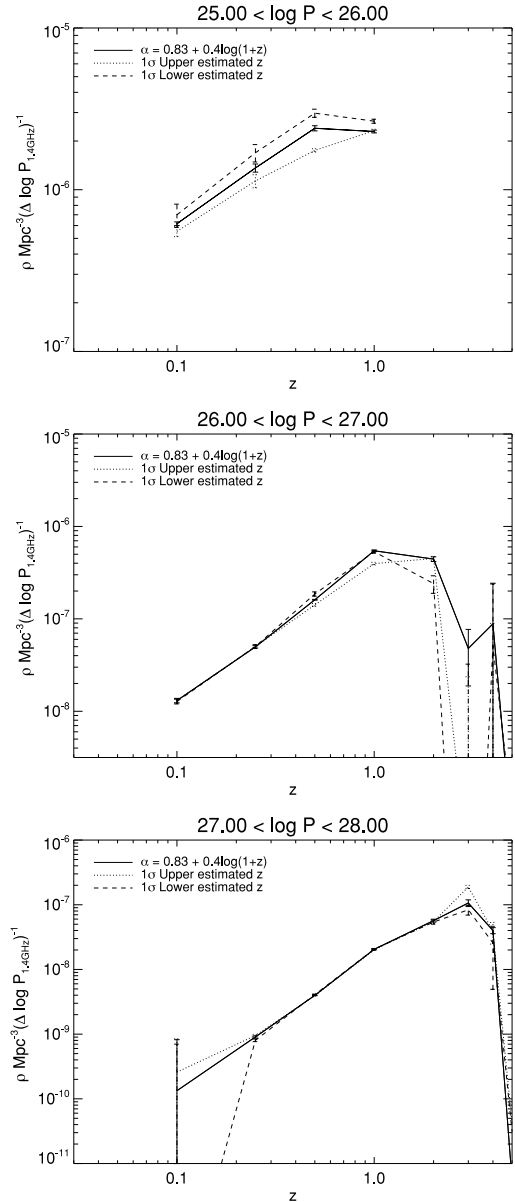
and luminosities randomly selected from $2.5 < z < 4$ and $26 < \log P < 27$, were inserted and the modelling repeated. Next, Hercules was returned to its original composition and the process repeated for CENSORS, but this time with extra sources drawn from $3.5 < z < 6$ and $27 < \log P < 28$. The number of real sources in these two redshift ranges (2.2 in Hercules, 2.6 in CENSORS, incorporating the spread in estimated redshifts as described in Section 3.2) is well reproduced by the polynomial fits used to represent the data, which give 3.0 and 4.2 sources respectively, so adding fake sources in this way should give a good indication of the number needed to significantly affect the turnover.

Table 5. The cut-off strength for the average redshift point following the peak for $26 < \log P < 27$, $2.5 < z < 4$ and $27 < \log P < 28$, $3.5 < z < 6$ for the addition of different numbers of fake sources, $N(\text{fake})$, into the Hercules and CENSORS sample, respectively. The $N(\text{fake}) = 0$ value is for the original version of the redshift samples.

| $N(\text{fake})$	Hercules C_σ	CENSORS C_σ
0	6.5	5.0
1	3.0	3.6
2	2.9	2.5
3	2.3	3.0
4	1.1	2.9
5	0.6	2.2
6	1.0	1.5
7	0.1	1.2

Figure 14. The results of the four smooth fits to the P–z grid for one radio power range. The full set of plots for $\log P = 25.25$ to $\log P = 27.75$ inclusive can also be found in Appendix D.

Table 5 gives the resulting cut-off strengths for the average density following the peak; it indicates that the number of CENSORS or Hercules sources in these ranges has to approximately double to push the significance of the cut-off below 3σ. It should also be noted that these numbers are likely to be a lower limit, as the modelling is likely to overpredict the number of real sources at these redshifts. This is because of the input polynomial used for the fitting, which typically underestimates at lower redshifts, thus leading to overestimation at $z \gtrsim 3$.

Both the CENSORS and Hercules samples contain sources without host galaxy identifications or with no spectroscopic redshifts, so the extra sources required to remove the turnovers could simply be missing. However, this possibility has already been considered in Section 4.3.2, where it was shown that moving all estimated redshifts to their upper limits does not remove the space density declines.

An alternative determination of the number of fake sources required to remove the redshift turnovers in these radio power and redshift ranges can be made by freezing the relevant high-redshift space densities at their peak values, and then calculating the total number of sources that would have been detected in the CENSORS or Hercules samples in the absence of a decline in density. These numbers – 11.7 for Hercules and 12.4 for CENSORS – are comparable to the total number of sources present in these bins with the addition of the fake sources discussed earlier in this section.

4.5 Polynomial approximation of the best-fitting P–z grid

The usefulness of the best-fitting P–z grid (given in Appendix D) to the reader is limited, because it only varies smoothly in regions where it is covered by the available data. To improve this situation the whole grid is fitted four times with a fourth-order polynomial series expansion, similar to the one used for the DP90 models. This provides an easy method to calculate the space density values at any (z, P) point, as well as allowing extrapolations of the data to regions currently not well constrained.

The basic polynomial series used for the fit is

$$\log \rho = \left(\sum_{i=0}^{4} \sum_{j=0}^{4-i} A_{ij} x(P)^i y(z)^j \right) + B x^5 y^5, \tag{10}$$
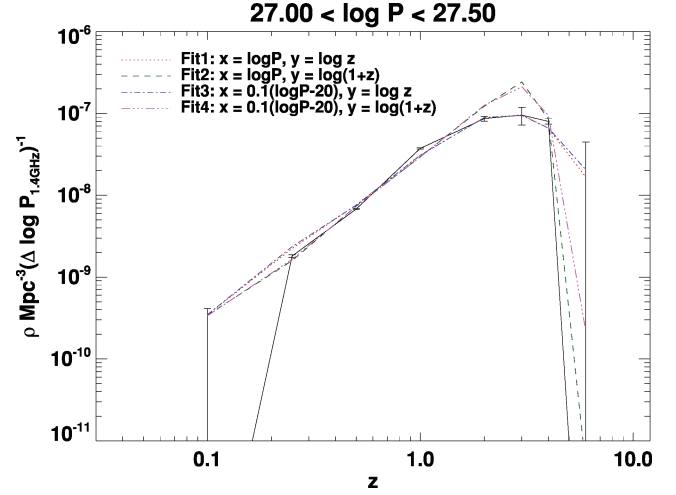
where x and y represent the radio power and redshift axes of the grid, respectively. Only points constrained by two of the input data sets are used in this fitting. The $z = 6$ points are also excluded as the uncertainties in this region are large. The coefficients for the four different versions of this fit are given in Appendix D. These were chosen to give several extensions into the unconstrained region, and the only difference between them is the coordinates used for x and y.

(i) Fit 1: $x = \log P$, $y = \log z$.
(ii) Fit 2: $x = \log P$, $y = \log (1 + z)$.
(iii) Fit 3: $x = 0.1(\log P - 20)$, $y = \log z$.
(iv) Fit 4: $x = 0.1(\log P - 20)$, $y = \log (1 + z)$.

The results of the fits for one radio power range are shown in Fig. 14 (the full set of plots for $\log P = 25.25$ to $\log P = 27.75$ inclusive can also be found in Appendix D) and the agreement with the best-fitting grid is good. However, it should be stressed that this smooth version of the grid is not a perfect representation of the model output and this is why it was not used for the analyses in the previous subsections.

5 SUMMARY AND CONCLUSIONS

The results presented in this paper demonstrate that the method of RLF determination described here works well; it gives an easy means of estimating errors and hence assessing the robustness of any evolutionary behaviour seen, such as the presence of a redshift turnover. Examination of the best-fitting steep-spectrum P–z grid suggests that the turnover in the RLF occurs at $z \simeq 0.7$ for the faintest luminosities considered here, and then moves to $z \gtrsim 2$ for higher powers. These changes are consistent with those seen for steep-spectrum radio sources by Waddington et al. (2001) who found turnovers in redshift at $z \simeq 1$ for low-luminosity sources ($P_{1.4\,\text{GHz}} > 10^{25.1}$ W Hz^{-1}) but at $z \simeq 2$ for the more powerful ($P_{1.4\,\text{GHz}} > 10^{26-27}$ W Hz^{-1}). Similarly a redshift peak at $z \simeq 2$ for the brightest sources is also seen for flat-spectrum quasars, in radio-loud, optical and X-ray selected samples (e.g. Hasinger et al. 2005; Richards et al. 2005; Wall 2008; De Zotti et al. 2010). They are also in broad agreement with the assumption of a low-luminosity peak at $z = 0.706$, and

a high-luminosity peak at $z = 1.91$, in the simulations of Wilman et al. (2008), which make predictions for the next generation of radio telescopes. The results presented here nevertheless suggest that a luminosity-dependent peak, with a high-redshift decline, would be a better representation of the real data than the two population model, with a flat post-peak space density, that these simulations currently adopt.

Physically, the luminosity dependence of the redshift peak in the radio galaxy RLF suggests that the most massive black holes have formed by $z \simeq 4$ and that their lower mass counterparts formed later. This 'cosmic downsizing' may initially appear to be at variance with the hierarchical model of structure formation. However, this discrepancy can be solved if the mode of AGN fuelling changes with cosmic time: in the early Universe, major mergers provide the cold gas to power the accretion at high rates, but at lower redshift it is low-luminosity, radiatively inefficient accretion from hot gas haloes that dominates (Fanidakis et al. 2011).

The data sets available for this work mean that only a narrow range in luminosity is constrained well enough to draw firm conclusions about the luminosity function evolution. Better coverage of the P–z plane will improve this. However, the density turnovers are robust and remain present even in the unlikely scenario that all the estimates are 1σ higher or lower. The turnovers seen in this work are also in good agreement with the work of Wall et al. (2005) who find a cut-off at a significance level $>4\sigma$ for their sample of flat-spectrum quasars.

The agreement with the DP90 results seen at high redshift for the brightest luminosities is not surprising; this region is dominated by CENSORS sources, which have already been shown to be consistent with two of their models (Paper III), both of which include a redshift cut-off.

The modelling method presented here can be easily modified to investigate different populations separately. The modelling results in this paper are limited by the uncertain redshifts present in some samples, and further spectroscopic observations for the CENSORS sample are ongoing to improve this situation. Future increases in sample size would allow independent minimization of all three grids, as well as subdividing the grids further into additional populations, e.g. FR I and FR II galaxies (Fanaroff & Riley 1974), or high- and low-excitation line radio sources (e.g. Hardcastle, Evans & Croston 2006). The upcoming large, deep radio surveys from both the LOw Frequency ARray (LOFAR) and the Australian Square Kilometre Array Pathfinder (ASKAP) will be ideal for this, but complementary redshift data, using the deep multicolour optical photometry from the planned Large Synoptic Survey Telescope (LSST) for example, is essential. Such an extension of this work would yield an invaluable tool for investigating the links between the different AGN subspecies.

ACKNOWLEDGMENTS

REFERENCES

Antonucci R., 1993, ARA&A, 31, 473
Barthel P. D., 1989, ApJ, 336, 606
Best P. N., Longair M. S., Röttgering H. J. A., 1998, MNRAS, 295, 549
Best P. N., Arts J. N., Röttgering H. J. A., Rengelink R., Brookes M. H., Wall J., 2003a, MNRAS, 346, 627 (Paper I)
Best P. N., Peacock J. A., Brookes M. H., Dowsett R. E., Röttgering H. J. A., Dunlop J. S., Lehnert M. D., 2003b, MNRAS, 346, 1021
Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005, MNRAS, 362, 9
Best P. N., Kaiser C. R., Heckman T. M., Kauffmann G., 2006, MNRAS, 368, L67
Best P. N., von der Linden A., Kauffmann G., Heckman T. M., Kaiser C. R., 2007, MNRAS, 379, 894
Blain A. W., Smail I., Ivison R. J., Kneib J.-P., 1999, MNRAS, 302, 632
Bondi M., Ciliegi P., Schinnerer E., Smolčić V., Jahnke K., Carilli C., Zamorani G., 2008, ApJ, 681, 1129
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Boyle B. J., Shanks T., Croom S. M., Smith R. J., Miller L., Loaring N., Heymans C., 2000, MNRAS, 317, 1014
Brookes M. H., Best P. N., Rengelink R., Röttgering H. J. A., 2006, MNRAS, 366, 1265 (Paper II)
Brookes M. H., Best P. N., Peacock J. A., Röttgering H. J. A., Dunlop J. S., 2008, MNRAS, 385, 1297 (Paper III)
Burgess A. M., Hunstead R. W., 2006, AJ, 131, 100
Colless M., 1999, R. Soc. London Philos. Trans. Ser. A, 357, 105
Colless M. et al., 2001, MNRAS, 328, 1039
Croton D. J. et al., 2006, MNRAS, 365, 11
De Zotti G., Massardi M., Negrello M., Wall J., 2010, A&AR, 18, 1
Donoso E., Best P. N., Kauffmann G., 2009, MNRAS, 392, 617
Downes A. J. B., Peacock J. A., Savage A., Carrie D. R., 1986, MNRAS, 218, 31
Dunlop J. S., Peacock J. A., 1990, MNRAS, 247, 19 (DP90)
Dunlop J. S., Peacock J. A., 1993, MNRAS, 263, 936
Dunlop J. S., Peacock J. A., Savage A., Lilly S. J., Heasley J. N., Simon A. J. B., 1989, MNRAS, 238, 1171
Dye S., Eales S., 2010, MNRAS, 405, 2406
Fabian A. C., Sanders J. S., Taylor G. B., Allen S. W., Crawford C. S., Johnstone R. M., Iwasawa K., 2006, MNRAS, 366, 417
Fan X. et al., 2001, AJ, 121, 54
Fan X. et al., 2004, AJ, 128, 515
Fanaroff B. L., Riley J. M., 1974, MNRAS, 167, 31
Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, MNRAS, 410, 53
Grimes J. A., Rawlings S., Willott C. J., 2005, MNRAS, 359, 1345
Hardcastle M. J., Evans D. A., Croston J. H., 2006, MNRAS, 370, 1893
Hasinger G., Miyaji T., Schmidt M., 2005, A&A, 441, 417
Jarvis M. J., Rawlings S., 2000, MNRAS, 319, 121
Jarvis M. J., Rawlings S., Willott C. J., Blundell K. M., Eales S., Lacy M., 2001, MNRAS, 327, 907
Jarvis M. J., Teimourian H., Simpson C., Smith D. J. B., Rawlings S., Bonfield D., 2009, MNRAS, 398, L83
Jones D. H. et al., 2004, MNRAS, 355, 747
Kellermann K. I., Wall J. V., 1987, Obser. Cosmol., 124, 545
Koo D. C., Kron R. G., 1982, A&A, 105, 107
Kron R. G., 1980, ApJS, 43, 305
Labiano A., Barthel P. D., O'Dea C. P., de Vries W. H., Pérez I., Baum S. A., 2007, A&A, 463, 97L
Laing R. A., Peacock J. A., 1980, MNRAS, 190, 903
McCarthy P. J., Kapahi V. K., van Breugel W., Persson S. E., Athreya R., Subrahmanya C. R., 1996, ApJS, 107, 19
Mauch T., Sadler E. M., 2007, MNRAS, 375, 931
Nelder J. A., Mead R., 1965, Comput. J., 7, 308
Osmer P. S., 1982, ApJ, 253, 28
Peacock J. A., 1985, MNRAS, 217, 601
Peacock J. A., Wall J. V., 1981, MNRAS, 194, 331
Richards G. T. et al., 2005, MNRAS, 360, 839
Rigby E. E., Snellen I. A. G., Best P. N., 2007, MNRAS, 380, 1449
Rigby E. E., Best P. N., Snellen I. A. G., 2008, MNRAS, 385, 310
Sadler E. M. et al., 2002, MNRAS, 329, 227

Sandage A., 1972, ApJ, 178, 25

Schmidt M., Schneider D. P., Gunn J. E., 1988, in Osmer P., Phillips M. M., eds, ASP Conf. Ser. Vol. 2, Optical Surveys for Quasars. Astron. Soc. Pac., San Francisco, p. 87

Seymour N., McHardy I. M., Gunn K. F., 2004, MNRAS, 352, 131

Shaver P. A., Wall J. V., Kellermann K. I., Jackson C. A., Hawkins M. R. S., 1996, Nat, 384, 439

Silverman J. D. et al., 2005, ApJ, 624, 630

Smolčić V. et al., 2008, ApJS, 177, 14

Smolčić V. et al., 2009, ApJ, 690, 610

Ubachukwu A. A., Ugwoke A. C., Ogwo J. N., 1996, Ap&SS, 238, 151

Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886

Waddington I., Windhorst R. A., Dunlop J. S., Koo D. C., Peacock J. A., 2000, MNRAS, 317, 801

Waddington I., Dunlop J. S., Peacock J. A., Windhorst R. A., 2001, MNRAS, 328, 882

Wall J., 2008, preprint (arXiv:0807.3792)

Wall J. V., Peacock J. A., 1985, MNRAS, 216, 173 (WP85)

Wall J. V., Peacock J. A., 1999, VizieR Online Data Catalog, 721, 60173

Wall J. V., Cole D. J., Milne D. K., 1968, Proc. Astron. Soc. Australia, 1, 98

Wall J. V., Jackson C. A., Shaver P. A., Hook I. M., Kellermann K. I., 2005, A&A, 434, 133

White R. L., Becker R. H., Helfand D. J., Gregg M. D., 1997, ApJ, 475, 479

Willott C. J., Rawlings S., Blundell K. M., Lacy M., Eales S. A., 2001, MNRAS, 322, 536

Willott C. J., Rawlings S., Jarvis M. J., Blundell K. M., 2003, MNRAS, 339, 173

Wilman R. J. et al., 2008, MNRAS, 388, 1335

Windhorst R. A., Kron R. G., Koo D. C., 1984a, A&AS, 58, 39

Windhorst R. A., van Heerde G. M., Katgert P., 1984b, A&AS, 58, 1

Wolf C., Wisotzki L., Borch A., Dye S., Kleinheinrich M., Meisenheimer K., 2003, A&A, 408, 499

York D. G., Adelman J., Anderson J. E. et al., 2000, AJ, 120, 1579

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Appendix A. Additional spectroscopic observations of Wall & Peacock (1985) radio sources.
Appendix B. The CENSORS data table.
Appendix C. Additional data sets.
Appendix D. Predicted P–z grid.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

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