Mon. Not. R. Astron. Soc. **384,** 711–718 (2008) doi:10.1111/j.1365-2966.2007.12733.x

Statistical properties of extragalactic sources in the New Extragalactic *WMAP* Point Source (NEWPS) catalogue

J. González-Nuevo,[1⋆] M. Massardi,[1,2] F. Argüeso,[3] D. Herranz,[4] L. Toffolatti,[5] J. L. Sanz,[4] M. López-Caniego[4,6] and G. De Zotti[1,7]

[1]*SISSA-ISAS, via Beirut 4, I-34014 Trieste, Italy*
[2]*Australia Telescope National Facility, CSIRO, PO Box 76 Epping, NSW 1710, Australia*
[3]*Departamento de Matemáticas, Universidad de Oviedo, Avda. Calvo Sotelo s/n, 33007 Oviedo, Spain*
[4]*Instituto de Física de Cantabria (CSIC-UC), Avda. los Castros s/n, 39005 Santander, Spain*
[5]*Departamento de Física, Universidad de Oviedo, Avda. Calvo Sotelo s/n, 33007 Oviedo, Spain*
[6]*Astrophysics Group, Cavendish Laboratory, J.J. Thompson Avenue, CB3 0E1 Cambridge*
[7]*INAF-Ossevatorio Astronomico di Padova, vicolo dell'Osservatorio 5, I-35122 Padova, Italy*

Accepted 2007 November 15. Received 2007 November 13; in original form 2007 September 10

ABSTRACT

We present results on spectral index distributions, number counts, redshift distribution and other general statistical properties of extragalactic point sources in the New Extragalactic *WMAP* Point Source$_{5\sigma}$ (NEWPS) sample. The flux calibrations at all the *Wilkinson Microwave Anisotropy Probe* channels have been reassessed both by comparison with ground-based observations and through estimates of the effective beam areas. The two methods yield consistent statistical correction factors. A search of the NASA Extragalactic Data base (NED) has yielded optical identifications for ∼89 per cent of sources in the complete subsample of 252 sources with signal-to-noise ratio (S/N) $\geqslant 5$ and $S \geqslant 1.1$ Jy at 23 GHz; five sources turned out to be Galactic and were removed. The NED also yielded redshifts for ≃92 per cent of the extragalactic sources at $|b| > 10^\circ$. Their distribution was compared with model predictions; the agreement is generally good but a possible discrepancy is noted. Using the 5 GHz fluxes from the GB6 or PMN surveys, we find that ∼76 per cent of the 191 extragalactic sources with $S_{23\,\mathrm{GHz}} > 1.3$ Jy can be classified as flat-spectrum sources between 5 and 23 GHz. A spectral steepening is observed at higher frequencies: only 59 per cent of our sources are still flat-spectrum sources between 23 and 61 GHz, and the average spectral indexes steepen from $\langle\alpha_5^{23}\rangle = 0.01 \pm 0.03$ to $\langle\alpha_{41}^{61}\rangle = 0.37 \pm 0.03$. We think, however, that the difference may be due to a selection effect. The source number counts have a close to Euclidean slope and are in good agreement with the predictions of the cosmological evolution model by De Zotti et al. The observed spectral index distributions were exploited to get model-independent extrapolations of counts to higher frequencies. The risks of such operations are discussed and reasons of discrepancies with other recent estimates are clarified.

Key words: surveys – galaxies: active – cosmic microwave background – radio continuum: galaxies – radio continuum: general.

1 INTRODUCTION

The statistical properties of extragalactic sources above ∼10 GHz are still largely unknown. This is due to the great difficulty of carrying out extensive surveys at high radio frequencies, because of the very small fields of view of ground-based radio telescopes and of the need of relatively long on-source integrations. Up to the end of the last century, the highest radio frequency at which large-area sky surveys had been performed was 5 GHz (GB6, Gregory et al. 1996; PMN, Griffith et al. 1994; Wright et al. 1994; Griffith et al. 1995; Wright et al. 1996). Recently, the situation has considerably improved thanks to the Pilot ATCA survey that covered the declination region between -60° and -70° at 18 GHz down to 100 mJy (Ricci et al. 2004; Sadler et al. 2006), to the 9C survey covering 520 deg^2 at 15.2 GHz to a flux limit of 25 mJy but going down to 10 mJy on small areas (Taylor et al. 2001; Waldram et al. 2003), and to data from the Very Small Array in the declination range $-5^\circ < \delta < 60^\circ$ at 33 GHz (Cleary et al. 2005). The on-going Australia Telescope 20 GHz (AT20G) Survey (Ekers

⋆E-mail: gnuevo@sissa.it

2006), once completed, will cover the whole Southern hemisphere down to 50 mJy.

Moreover, the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) mission has produced the first all-sky surveys of extragalactic sources at 23, 33, 41, 61 and 94 GHz (Bennett et al. 2003; Hinshaw et al. 2007). From the analysis of the first three years survey data, the *WMAP* team have obtained a catalogue of 323 extragalactic point sources (EPS; Hinshaw et al. 2007), substantially enlarging the first-year one that included 208 EPS detected above a flux limit of ∼0.8–1 Jy (Bennett et al. 2003), with an estimated completeness limit of ∼1.2 Jy. The average estimated spectral index for detected sources is ≃0.0, thus confirming that the source class which dominates the counts at bright fluxes is essentially made of flat-spectrum quasi-stellar objects (QSOs) and BL Lacs, with minor contributions coming from steep- or inverted-spectrum sources (Toffolatti et al. 1998; De Zotti et al. 1999; Toffolatti et al. 1999; Bennett et al. 2003).

In order to take full advantage of all the information content of the *WMAP* three-year survey data, (López-Caniego et al., hereafter LC07 2007) have used the MHW2 filter (González-Nuevo et al. 2006) to obtain estimates of (or upper limits on) the flux densities at the *WMAP* frequencies of a complete all-sky sample of 2491 sources at $|b| > 5^\circ$, brighter than 500 mJy at 5 GHz, or at 1.4 or 0.84 GHz in regions not covered by 5 GHz surveys but covered by either the National Radio Astronomy Observatory/Very Large Array Sky Survey (NVSS) (Condon et al. 1998) or the Sydney University Molonglo Sky Survey (Mauch et al. 2003). This work yielded 5σ detections of 368 extragalactic sources, including 98 sources not present in the *WMAP* three-year catalogue. The results were organized in a catalogue dubbed New Extragalactic *WMAP* Point Source$_{5\sigma}$ (NEWPS) Catalogue.

As pointed out by LC07, the new flux estimates showed small systematic differences, increasing with frequency, with those obtained by the *WMAP* team. These differences were attributed to different approximations of the beam shapes. In fact, the beams are complex and asymmetrical. Taking also into account that sources are observed with different beam orientations, it is clear that the effective beam areas and, therefore, the flux calibrations, are uncertain.

To check the calibration of LC07 fluxes, we have looked for ground-based measurements of NEWPS$_{5\sigma}$ sources at frequencies close to the *WMAP* ones, finding small but appreciable systematic differences. This prompted us to investigate in more detail the calibration problem. The adopted methods and the results are presented in Section 2. In Section 3, we present the optical identifications and the redshift distribution of our sources, based on literature data. After having recalibrated the NEWPS$_{5\sigma}$ flux densities, we have investigated the spectral index distributions in different frequency intervals (Section 4) and used them to obtain model-independent estimates of the mm-wavesource counts, beyond the flux density intervals where they are directly determined (Section 5). Finally, in Section 6, we summarize our main conclusions.

Throughout this paper, we have used the Bayes-corrected version of the NEWPS Catalogue to take into account the well-known Eddington bias (see Section 2.1 or LC07 for more details Eddington 1940).[1] The spectral index, α, is defined as $\alpha \equiv -\log(S_2/S_1)/\log(\nu_2/\nu_1)$, where S_1 and S_2 are the fluxes at the frequencies ν_1 and ν_2, respectively (i.e. $S_\nu \propto \nu^{-\alpha}$).

[1] The Bayes-corrected version of the NEWPS Catalogue can be found at http://max.ifca.unican.es/caniego/NEWPS/.

2 STATISTICAL FLUX CALIBRATION

We have studied in detail the calibration of the estimated NEWPS fluxes, after the application of the Bayesian correction (see next section), by adopting two independent approaches: (i) comparison with ground-based measurements at the closest frequencies and (ii) direct estimates of the effective beam areas for the brightest sources.

2.1 The Eddington bias and the Bayesian correction

It has long been realized (Eddington 1940) that fluxes of sources detected above a certain S/N threshold in noisy fields are, on average, overestimated since there are similar numbers of positive and negative fluctuations but faint sources are more numerous than the bright ones. Thus, the number of faint sources exceeding the threshold because they are on top of positive fluctuations is larger than that of brighter sources falling below the threshold because they happen to be on top of negative fluctuations. This effect is known as the Eddington bias.

The normalized distribution of true fluxes, S, of extragalactic sources, is usually well described by a power law

$$P(S|q) = kS^{-(1+q)}, \quad S \geqslant S_{\rm m}. \tag{1}$$

If the slope q is known, the maximum likelihood estimator of the true fluxes of the sources is easily calculated (Hogg & Turner 1998) from the observed ones. Unfortunately, in many cases and, in particular, in the case of the *WMAP* survey, this condition is not satisfied. For this reason, Herranz et al. (2006) and LC07 have gone beyond the results of the previous work by solving simultaneously for q and S. We remind here only the asymptotic limits of the estimators, valid in the high signal-to-noise ratio (S/N) regime, referring the reader to LC07 for a detailed description of the calculations:

$$q \simeq \left[\frac{1}{N}\sum_{i=1}^{N} \ln\left(\frac{S_i^o}{S_{\rm m}^o}\right)\right]^{-1}, \tag{2}$$

$$S_i \simeq S_i^o\left(1 - \frac{1+q}{r_i^2}\right), \tag{3}$$

where $S^o_{\rm m}$ is the minimum observed flux and $r_i = S_i^o/\sigma_i$ is the S/N of the source. First, we estimate the slope, q, from our observed fluxes (equation 2) and then we use it to calculate the correction for each source (equation 3).

2.2 Ground-based measurements

Many *WMAP* sources are calibrators for the Australia Telescope Compact Array (ATCA) and/or for the very large array and, as such, are continuously monitored at several frequencies. A list of 599 sources brighter than 500 mJy at 12 mm (25 GHz), some of which have also data at 7 and 3 mm (43 and 100 GHz), is available on the ATCA website.[2] We have cross-correlated this list with the complete NEWPS$_{5\sigma}$ subsample flux with $S_{23\,\rm GHz} \geqslant 2$ Jy. Within a search radius of 0.348° (the beam size of the *WMAP K* band), we have found 49 associations, 47 of which have flux measurements also at 7 mm and 49 at 3 mm. The numbers of NEWPS$_{5\sigma}$ detections for these 49 sources are: 48, 48, 43 and 8 at 33, 41, 61 and 94 GHz, respectively. The median flux density ratio for pairs of *WMAP* channels was obtained using the Kaplan & Meier (1958) estimator, taking into

[2] http://www.narrabri.atnf.csiro.au/calibrators/

Table 1. Results of the self-calibration process (effective area/symmetrized area; see Section 2.3) and of the source calibration by comparison with ground-based measurements (ground-based fluxes/NEWPS fluxes or average ratio of the corrected flux to the uncorrected one; see Section 2.2). The correction factors for the *W* band are in parenthesis because they are very uncertain (see the text).

WMAP band & central frequency (GHz)	Average effective area (sr)	Symmetrized area (sr)	Self-calibration correction average value + error	Calibrators correction average value + error
K (23)	$(2.595 \pm 0.014) \times 10^{-4}$	2.46×10^{-4}	1.050 ± 0.006	0.99 ± 0.06
Ka (33)	$(1.569 \pm 0.012) \times 10^{-4}$	1.44×10^{-4}	1.086 ± 0.008	1.12 ± 0.07
Q (41)	$(1.016 \pm 0.007) \times 10^{-4}$	8.94×10^{-5}	1.136 ± 0.008	1.15 ± 0.08
V (61)	$(4.81 \pm 0.07) \times 10^{-5}$	4.19×10^{-5}	1.15 ± 0.02	1.32 ± 0.09
W (94)	$(3.36 \pm 0.17) \times 10^{-5}$	2.07×10^{-5}	1.62 ± 0.08	1.41 ± 0.10

account also the 5σ upper limits. The comparison with calibrators was performed in the space of *WMAP* and calibrator frequencies, defining a 5×3 matrix of flux density ratios $R(\nu_W, \nu_c)$, with $\nu_W = [23, 33, 41, 61, 94]$ GHz and $\nu_c = [25, 43, 100]$ GHz.

Since the number of 94 GHz detections is too small, the application of the Kaplan–Meier estimator to data at this frequency yielded unreliable results. We therefore decided not to take into account the $R(94\,\mathrm{GHz}, \nu_c)$ values and to replace them with extrapolations from the other elements of the matrix. In practice, for each of the calibrators frequencies we have linearly fitted the R [23, 33, 41, 61] GHz, $\nu_c(i)$ data (where $i = 1, 2, 3$) and estimated the $R[94\,\mathrm{GHz}, \nu_c(i)]$ elements by a direct extrapolation of the fits. Next, we estimated the 5×5 matrix $R(\nu_W, \nu_W)$ by interpolating (we have used a two-dimensional spline interpolation) the estimated $R(\nu_W, \nu_c)$ matrix. The calibration factors for each *WMAP* channel were finally obtained as the diagonal elements of the $R(\nu_W, \nu_W)$ matrix. The results and their errors are given in the last column of Table 1. Due to the uncertainty of the correction factor for the *W* band, hereinafter we will use only the corrected fluxes relative to the *K, Ka, Q* and *V WMAP* bands for our statistical analysis.

2.3 Self-calibration: methodology

As stated in Section 1, the uncertainty on flux calibration stems from our imperfect knowledge of the asymmetric beam response functions. *WMAP* calibration is based on the brightness temperature of the CMB dipole; in order to go from temperatures to fluxes it is necessary to multiply the temperature maps by a numerical factor that depends on the frequency of observation (thus transforming the maps into units of Jy/sr, e.g.) and then to multiply again by the effective area (solid angle) subtended by the point source. Therefore, a good knowledge of the effective area of the beams is required. In LC07, the effective areas of the different *WMAP* beams were calculated using the symmetrized beam profiles given by the *WMAP* team.[3] Obviously, this is only a first-order approximation. In particular, for the *W* and *V* bands the beams are known to be highly non-Gaussian and non-symmetric. Thus, accurate flux determinations require the knowledge of *the real, non-symmetric effective beam areas* at each source position. The *WMAP* team provided real beam maps, but it must be noted that the *effective* beam area at a given position in the sky is the result of a complicated composition of the nominal beam over many pointings with different orientations at different times.

Fortunately, there is a simpler way to calculate the effective beam area, only using *WMAP* data. In the absence of noise of any kind (i.e. if only the point source signal is present), it would be sufficient to use known point sources to measure directly the beam area: the effective beam area would simply be the ratio of the total observed source flux to its central value (i.e. the value corresponding to the central pixel). The total NEWPS fluxes have been calculated using the formula:

$$F_i = A_b I_i(0), \tag{4}$$

where A_b is the effective beam area and $I_i(0)$ is the peak intensity for the *i*th source, directly calculated from the filtered image.

The SExtractor package (Bertin & Arnouts 1996) is a well-known software specifically designed for obtaining accurate photometry of compact astronomical objects on images. This software allows us to estimate directly the total flux of each source, F_i, without any assumption about the beam shape or size.

To minimize the errors on the estimated fluxes, we need to select only *bright* sources, i.e. sources with a high S/N. Therefore, for each *WMAP* band we have selected the $\mathrm{NEWPS}_{5\sigma}$ sources with $S \geqslant 1$ Jy and local S/N $\geqslant 5$ (in the real space) at that frequency; their numbers are 123, 75, 67, 28 and 3 at 23, 33, 41, 61 and 94 GHz, respectively. Then, for each NEWPS source satisfying this selection criterion, we have projected a flat patch around its position. Subsequently, we have run SExtractor on the resulting patches to obtain a direct estimation of the total flux and of its photometric error. Besides, for each detected source the peak intensity and the background rms at its position has been obtained. Then, we have calculated the effective area of the beam at the position of the source as $A_{b_i} = F_i / I_i(0)$. Finally, we have computed the average of all the individual effective areas, weighted by the corresponding errors, for each *WMAP* channel.

The results are presented in Table 1. In the first column, we have the *WMAP* frequency band. The second column shows the average effective beam areas calculated with the SExtractor method, and their errors, while the third column contains the corresponding beam area calculated using the symmetrized beam profiles (this is the beam area used in NEWPS for unit conversion). If the estimated average effective areas apply to the whole sky, the correction factor that should be applied to NEWPS fluxes is $A_b, \mathrm{SEx}/A_{b,\mathrm{sym}}$. This factor is given in the fourth column of Table 1. A comparison with the correction factors estimated from calibrators (last column) shows a very encouraging agreement. Since for the *W* band, we have only three sources that satisfy the selection criterion, the correction factor is much more uncertain than suggested by the formal error, that only includes the dispersion of the estimated values.

For the following analysis, we will *correct the NEWPS fluxes in LC07 multiplying them by the self-calibration factors* listed in Table 1, except for the *W* band, that will not be considered further

[3] http://lambda.gsfc.nasa.gov/product/map/dr2/beam_profiles_ get.cfm

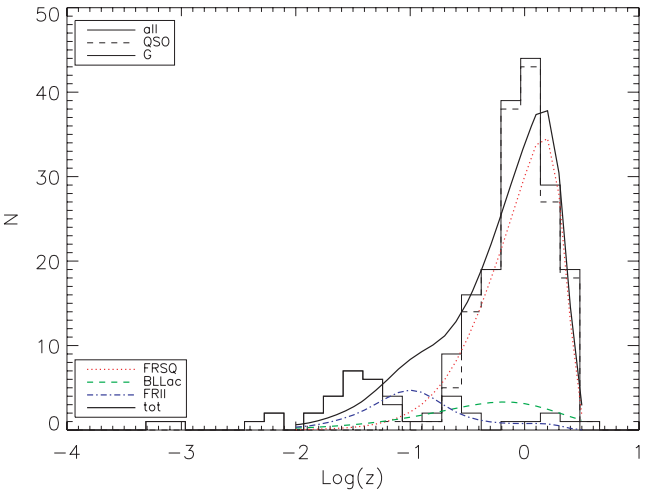

Figure 1. Redshift distributions of the full $|b| > 10^\circ$ sample, QSOs and galaxies (solid, dashed and dot–dashed histograms, respectively). The dotted, dashed and dot–dashed curves display, for comparison, the predictions of the model by De Zotti et al. (2005) for Flat-Spectrum Radio QSOs, BL Lacs, FR II sources, while the solid line shows the total.

because of the lack of a reliable calibration factor, the relatively high number of upper limits and its uncertain completeness level.

3 OPTICAL IDENTIFICATIONS AND REDSHIFT DISTRIBUTION

We have searched the NASA Extragalactic Data base (NED[4]) for optical identifications of the 252 sources at $|b| > 5^\circ$ with $\geqslant 5\sigma$ detections at 23 GHz and flux density above the completeness $S_{23\,\mathrm{GHz}} = 1.1$ Jy. The majority of sources (181, i.e. 72 per cent) are classified as QSOs, 44 (17 per cent) as galaxies, five are Galactic objects [two signal-to-noise ratio (SNR), two H II regions, 1 PN], and the remaining ones (22, i.e. 9 per cent) are unclassified. The galactic objects were obviously dropped from the sample.

The NED data base also provided the redshifts of 215 sources [87 per cent of the extragalactic sample; 95 per cent (42) of galaxies and 96 per cent (173) of QSOs]. Although the redshift completeness is already rather high, it can be increased restricting the sample to $|b| > 10^\circ$. In fact, in the Galactic latitude range $5^\circ < |b| < 10^\circ$ there are 22 extragalactic objects, only eight of which have redshift measurements, not surprisingly because of the substantial Galactic extinction. Our extragalactic sample comprises 225 objects at $|b| > 10^\circ$ and 207 (92 per cent) of them have redshifts (167 QSOs and 40 galaxies). Optical (b_J) magnitudes of these sources were recovered from the SuperCOSMOS Sky Survey archives.[5] The median b_J is 16.7 for galaxies and 18.2 for QSOs.

The redshift distribution for the latter subsample is shown in Fig. 1, where the dashed, dot–dashed and solid histograms refer to QSOs, galaxies and to the total, respectively. The median redshift of the sample is 0.86 (0.052 for Galaxies only, 0.994 for QSOs). We have also plotted, for comparison, the redshift distributions predicted by the De Zotti et al. (2005) model for different source populations. The agreement is generally good, except for the dip in the data around $\log(z) = -1$, where the model predicts a little bump due to Fanaroff–Riley type I (FR II) sources. The reason of this discrepancy is unclear. The most obvious option is that FR II sources are not correctly modelled, but other possibilities, such as a large-scale inhomogeneity in the distribution of bright radio sources, cannot be ruled out.

4 SPECTRAL INDEX DISTRIBUTIONS AND SPECTRAL CLASSES

To study the spectral properties of NEWPS sources, we defined a complete subsample comprising all the 191 NEWPS sources with $S_{23} \geqslant 1.3$ Jy and S/N $\geqslant 5$, with counterparts in the 5 GHz surveys (GB6 or PMN). We have confined ourselves to fluxes somewhat above the 23 GHz completeness limit ($\sim$1.1 Jy) in order to have more than 50 per cent 5σ detections at 61 GHz (we have a detection rate of 52 per cent); otherwise, the uncertainties on the spectral index distribution obtained from Survival Analysis become unacceptably large.

In addition to fluxes in the *WMAP* channels, we have included in our analysis 5 GHz measurements. Since the angular resolution of the 5 GHz catalogues is much higher than that of any of the *WMAP* channels, we have degraded the 5 GHz catalogue to the *WMAP* resolution. In practice, whenever more than one 5 GHz sources are present within a *WMAP* resolution element, we have summed up their 5 GHz fluxes, weighted with the *WMAP* beam centred at the position of the NEWPS source. To correct for the contribution of the background of faint 5 GHz sources, we have selected a set of 4141 fields (control fields) at 23 GHz with area equal to that of a *WMAP* resolution element at the considered frequency and devoid of NEWPS sources. The mean flux (120 mJy) of 5 GHz sources found in control fields, weighted with the *WMAP* beam with axis towards the field centre, has been subtracted from the sum of 5 GHz fluxes (or the flux of the only source) in the fields of NEWPS sources.

4.1 Spectral index distributions

The spectral index distributions for the frequency intervals 5–23 and 23–61 GHz are shown in Fig. 2. The latter distribution has been obtained by the Kaplan–Meier (1958) estimator, taking into account also the upper limits on the fluxes, yielding lower limits on spectral indices. Note that, after application of the Kaplan–Meier estimator it is no longer possible to distinguish between data and limits. Interestingly, the distribution of α_{23}^{61} shows a hint of a second peak in the range 0.4–0.8, not seen in the distribution of α_5^{23}, consistent with the transition of a subset of sources from flat to steep spectra. We have applied the non-parametric two sample tests implemented in the public ASURV code[6] to see whether the two distributions are consistent with being drawn from the same parent distribution. This hypothesis is rejected with a high significance level (more than $\sim 8\sigma$). As noted below, however, the difference may be due, to a large extent, to a selection effect.

4.2 Source spectral classes

Radio sources show a variety of spectral behaviours in this frequency range (cf. Sadler et al. 2006). Fig. 3 shows some examples: a flat-spectrum (top left-hand panel), a steep spectrum (bottom left-hand panel), a spectrum flattening at $\sim$10 GHz (top right-hand

[4] http://nedwww.ipac.caltech.edu/
[5] http://www-wfau.roe.ac.uk/sss/
[6] http://www.astrostatistics.psu.edu/statcodes/sc_censor.html

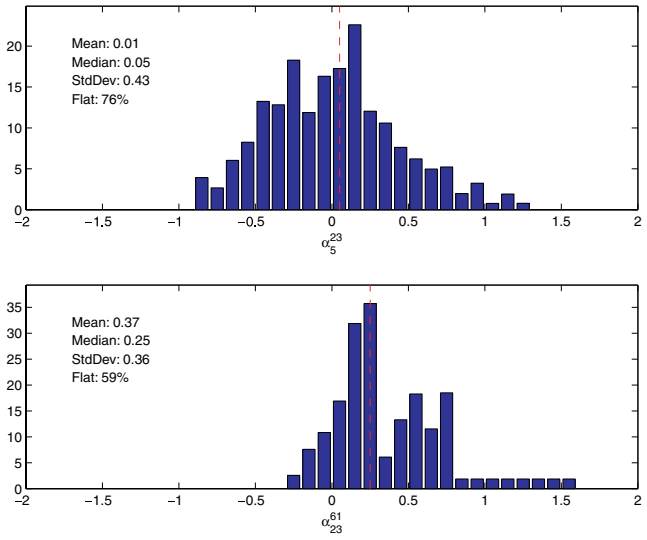

Figure 2. Spectral index distributions for two different frequency ranges: 5–23 GHz (upper panel) and 23–61 GHz (lower panel). For both distributions the mean, median, standard deviation and fraction of flat-spectrum sources (see Section 4.2) are shown. The vertical red dashed lines correspond to the median values of the distributions. See Section 4 for more details on the calculations.

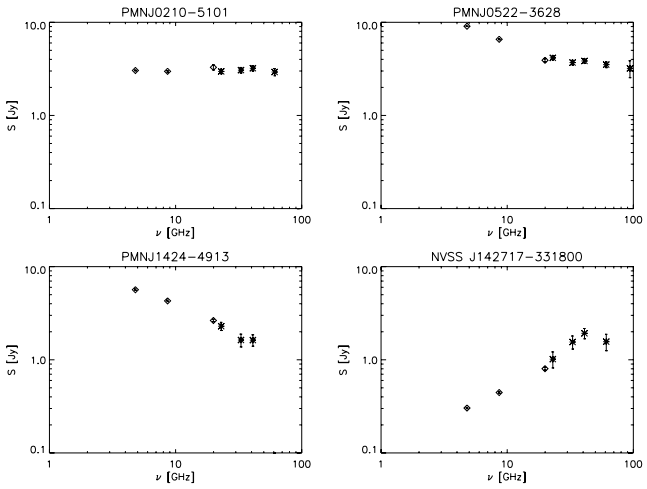

Figure 3. Examples of typical radiosource spectra at mm-wavelengths: a flat-spectrum source (top left-hand panel); a steep-spectrum source (bottom left-hand panel); a source whose spectrum flattens at $\nu \sim 10$ GHz (top right-hand panel); a High Frequency Peaker (HFP) source (bottom right-hand panel). Data from the NEWPS Catalogue (asterisks) and from the AT20G Survey [courtesy of the AT20G group, Massardi et al. (2007)] (diamonds).

panel) and a spectrum peaking at high radio frequency (bottom right-hand panel). The data are a combination of NEWPS fluxes with the AT20G measurements (Massardi et al. 2007).

We have classified our sources into three main spectral classes: flat-spectrum sources ($-0.5 \leqslant \alpha \leqslant 0.5$), steep-spectrum sources ($\alpha > 0.5$) and inverted-spectrum sources ($\alpha < -0.5$). Between 5 and 23 GHz flat-spectrum sources dominate the sample: they amount to 76 ± 7 per cent (see Table 2), in agreement with conclusions of previous works (Bennett et al. 2003; Ricci et al. 2004; Sadler et al. 2006). However, at higher frequencies, we find a trend towards a steepening of the spectral indices. The 5–23 GHz spectral index distribution has a median (mean) value $\alpha = 0.05 \pm 0.04 (0.01 \pm 0.03)$, with a

Table 2. Distribution of the 191 sources in our sample in three spectral classes for the 5–23 and 23–61 GHz frequency intervals: flat spectrum ($-0.5 \leqslant \alpha \leqslant 0.5$), steep spectrum ($\alpha > 0.5$) and inverted spectrum ($\alpha < -0.5$). See Section 4.2 for more details.

Class (per cent)	5–23 GHz	23–61 GHz
Flat	76 ± 6	59 ± 6
Steep	17 ± 1	41 ± 4
Inverted	7 ± 1	0

dispersion $\sigma = 0.43$, in full agreement with the results by Ricci et al. (2004) who used the multifrequency ATCA survey data. The median (mean) spectral index between 23 and 61 GHz for the same sample, selected at 23 GHz, is $\alpha = 0.25 \pm 0.05$ (0.37 ± 0.04) (see Fig. 2); correspondingly, the fraction of the steep-spectrum sources increases from 17 ± 1 to 41 ± 4 per cent.

The quoted errors are Poissonian in the case of α_5^{23}. As for α_{23}^{61}, we must take into account that its distribution includes a substantial fraction of lower limits, redistributed with the Kaplan–Meier (1958) method. Clearly, limits cannot be given the same weight as actual measurements in the error estimate. As discussed by Cantor (2001), in this case a simple and conservative estimate of the standard error on the redistribution of limits is provided by the Peto formula

$$\sigma_{23}^{61} = P \left(\frac{1-P}{N_{\rm u}} \right)^{1/2}, \qquad (5)$$

where P is the probability that α_{23}^{61} is larger than the mean (or median) value and $N_{\rm u} = 91$ is the number of lower limits. The probability P is approximated by the fraction of sources with α_{23}^{61} larger than the mean or the median ($P = 0.44$ or 0.50, respectively). The global error is computed as the quadratic sum of the Poisson error for the total number of data points (measurements plus limits) plus the contribution given by equation (5). In the same way, we have estimated the errors on the fraction of sources of steep- and flat-spectrum sources by identifying the probability P with the fraction of sources with spectral index above the boundary for their spectral class.

The 23 GHz selection obviously biases the spectral index distribution in favour of flat/inverted values, which yield brighter fluxes at higher frequencies. In the case of a Gaussian distribution, a dispersion σ_α translates into an apparent flattening of spectral indices by (Danese & de Zotti 1984; Condon 1984)

$$\Delta\alpha \simeq (1-\gamma)\sigma_\alpha^2 \ln(23/5), \qquad (6)$$

γ being the slope of differential counts. For an Euclidean slope ($\gamma = 2.5$) and $\sigma_\alpha = 0.3$ (note that the distribution in Fig. 2 is broadened by errors on flux measurements Ricci et al. 2004), $\Delta\alpha \simeq -0.2$. Thus, the observed steepening above 23 GHz may be largely a selection effect.

On the other hand, a real steepening is expected in this frequency range since the absorption optical depth decreases with increasing frequency at least as ν^{-2} so that an increasing fraction of the emitting regions becomes optically thin, i.e. take on spectral indices $\alpha \simeq 0.7$–0.8 or steeper, because of electron ageing effects.

5 NUMBER COUNTS

In Fig. 4, we have compared the differential source counts of NEWPS$_{5\sigma}$ sources at 23, 33, 41 and 61 GHz, normalized to the

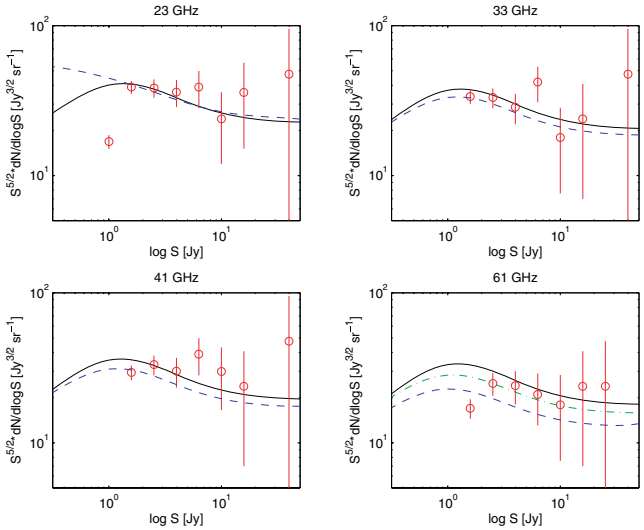

Figure 4. Differential number counts, in bins of $\Delta \log(S) = 0.2$, multiplied by $S^{\frac{5}{2}}$, of the NEWPS$_{5\sigma}$ sources (circles with Poisson error bars). The solid curve shows, for comparison, the total counts predicted by the De Zotti et al. (2005) model. The dashed lines correspond to the counts extrapolated from 5 GHz, for the 23 GHz case, and from 23 GHz for the other frequencies, using the spectral index distributions of Section 4.1. The extrapolated counts are too low at 61 GHz, suggesting that the true steepening of spectral indices above 23 GHz is less pronounced than implied by the distribution yielded by the Kaplan–Meier estimator. A good fit is obtained damping down the secondary peak of the distribution of α_{23}^{61} (Fig. 2) by a factor of 0.2 (dot–dashed line).

Table 3. Parameters of the power-law approximation to the differential source number counts calculated from our sample. The columns contain: (1) the frequency, (2) the lower flux limits, (3) the normalization, A, and (4) the slope, γ, of the counts with their errors and (5) the probability of exceeding the corresponding χ^2 value if the model is correct. Only the data points on-source counts plotted in Fig. 5 and above the lower flux limit, $S_{\rm lim}$, of column 2 are used in this calculation.

Frequency (GHz)	$S_{\rm lim}$ (Jy)	A (σ_A) (Jy^{-1} sr^{-1})	γ (σ_γ)	Probability
23	1.1	34.4 (3.5)	2.36 (0.10)	0.84
33	1.1	27.8 (3.2)	2.25 (0.11)	0.08
	1.4	34.6 (5.0)	2.40 (0.13)	0.30
41	1.1	30.9 (3.4)	2.41 (0.11)	0.99
61	1.1	18.5 (2.8)	2.24 (0.14)	0.07
	1.4	25.8 (4.5)	2.47 (0.15)	0.94

Euclidean slope and in bins of $\Delta \log(S) = 0.2$, with the predictions of the cosmological evolution model by De Zotti et al. (2005). The agreement is generally good above the completeness limits of the sample. However, the flux density range probed with sufficient statistics is quite limited.

Over this range, the differential counts can be described by the usual power-law approximation:

$$\frac{{\rm d}n}{{\rm d}S} = AS^{-\gamma}. \tag{7}$$

The best-fitting values of the normalization, A, and of the slope, γ, are given, with their errors, in Table 3, where the goodness of the fit is quantified by the probability of exceeding the corresponding χ^2 value assuming that the model is correct. The normalization, A, is almost constant up to 41 GHz, in keeping with the fact that the sample is dominated by sources with mean spectral index close to 0, but slightly decreases at 61 GHz, as the effect of the high-frequency steepening.

Above 30 GHz, the values of A are significantly higher than the estimates by Waldram et al. (2007), based on extrapolations of the 9C counts at 15 GHz. As discussed below, the discrepancy arises because the spectral index distribution of the relatively faint sources of their sample is not appropriate at the high flux densities of interest here.

5.1 Number counts estimates by extrapolation from lower frequencies

The spectral index distributions of Section 4.1, combined with the observational determinations of 5 GHz counts, extending over a very broad flux density interval, can be exploited to get model-independent estimates of the high-frequency counts. Alternatively, the comparison of the extrapolated counts with the observed ones provides a test of the Kaplan–Meier estimates of the spectral index distributions.

Under the assumption that the spectral index distribution is independent of flux density (only valid for limited flux density intervals, see below), the source number counts can be extrapolated from the frequency ν_1 to the frequency ν_2, convolving the counts at frequency ν_1 with the probability that a source with $S_1(\nu_1)$ has $S_2(\nu_2)$:

$$\frac{{\rm d}N_2[S_2(\nu_2)]}{{\rm d}\log S_2(\nu_2)}\,{\rm d}\log S_2(\nu_2) = $$
$$= \int \frac{{\rm d}N_1[S_1(\nu_1)]}{{\rm d}\log S_1(\nu_1)}\,P[S_1(\nu_1)|S_2(\nu_2)]\,{\rm d}\log S_1(\nu_1). \tag{8}$$

The probability distribution $P[S_1(\nu_1)|S_2(\nu_2)]$ is straightforwardly obtained from the corresponding spectral index distribution.

However, to extrapolate the 5 GHz counts we need such distribution for sources selected at the same frequency, while that of Section 4.1 refers to the 23 GHz selection. In practice, to derive the 23 GHz counts knowing those at 5 GHz and our distribution of α_5^{23} we need to invert equation (8) with $\nu_1 = 23$ and $\nu_2 = 5$ GHz. This is a classical inversion problem (Fredholm equation of the first kind) and we have used the Lucy's iterative method (Lucy 1974) to solve it:

$$\frac{{\rm d}N[S_{23}(j)]^{r+1}}{{\rm d}N[S_{23}(j)]^{r}} = \sum_i \frac{{\rm d}N^0[S_5(i)]}{{\rm d}N^r[S_5(i)]}\,P[S_{23}(j)|S_5(i)], \tag{9}$$

where for the 5 GHz counts, ${\rm d}N^0[S_5(i)]$, we have used the description provided by the De Zotti et al. (2005) model that accurately reproduces the data. We have initialized the method (i.e. for $r = 0$) adopting Euclidean counts, ${\rm d}N/{\rm d}S \propto S^{-5/2}$, but we have also checked that our results are insensitive to the function used to initialize the process. Additionally, we have improved the stability of the method by smoothing the ${\rm d}N[S_{23}(j)]^{r+1}$ in each iteration.

The results are shown by the dashed line in the upper left-hand panel of Fig. 4: the extrapolated counts are in good agreement with the direct estimate and with the De Zotti et al. (2005) model for $S > 1$ Jy, but are well above the model prediction at fainter fluxes. Since the downturn of the counts below $\simeq$1 Jy is born out by the 9C counts at 15 GHz (Waldram et al. 2003) and by the ATCA pilot survey counts at 18 GHz (Ricci et al. 2004), the discrepancy must be attributed to a failure of the basic assumption underlying the extrapolation. This is confirmed by the direct evidence of a change with decreasing flux density of the proportion of flat- and steep-spectrum sources contributing to the corresponding 5 GHz counts.

This illustrates the risks of extrapolations of counts with incomplete information about the statistical properties of source populations.

Additional evidence of substantial changes of the spectral index distribution with decreasing flux density at ∼20 GHz is provided by Waldram et al. (2007), who carried out multifrequency follow-up of a sample complete to 25 mJy at 15 GHz of extragalactic sources from the 9C survey. The median spectral index between 15 and 43 GHz was found to be 0.89 for these sources, much fainter than ours. They also reported a spectral steepening with increasing frequency.

The observed distribution of α_{23}^{61} can be directly exploited to extrapolate the 23 GHz counts to higher frequencies. Since the De Zotti et al. (2005) model provides a good description of the counts (as determined by the 9C and ATCA surveys) also below the limit of *WMAP* ones at this frequency, we have used it to carry out the extrapolations. The dashed lines in Fig. 4 show that the extrapolated counts tend to be lower than those directly observed, particularly at 61 GHz. This is unexpected, since the observed spectral index distributions are, if anything, broadened by the contributions of measurements errors on fluxes and this leads to a flattening of the effective spectral index and, correspondingly, to an overestimate of the high-frequency counts, contrary to what is found here.

The likely explanation is that the Kaplan–Meier estimator overpopulates the steep portion of the distribution of α_{23}^{61}. The observed counts are recovered if we scale down by a factor of 0.2 the distribution at $\alpha_{23}^{61} > 0.5$ (dot–dashed line in Fig. 4).

6 CONCLUSIONS

We have presented and discussed a variety of statistical properties that characterize the extragalactic sources in the NEWPS Catalogue (LC07). The flux calibration has been investigated with two different methods and correction factors have been derived, except for the highest frequency *WMAP* channel, for which the available information is insufficient.

A search of the NED data base has yielded optical identifications for 89 per cent of the 252 sources in the complete NEWPS sample with $S_{23\,\mathrm{GHz}} \geqslant 1.1$ Jy; five sources turned out to be Galactic and were dropped. At high Galactic latitudes ($|b| > 10^\circ$), we obtain a subsample of 207 sources (92 per cent) with spectroscopic redshifts (again from the NED), for a total number of extragalactic sources of 225. The corresponding redshift distribution is in generally good agreement with the predictions of the model by De Zotti et al. (2005), except possibly around $\log(z) \simeq -1$ where the model predicts a significant contribution from FR II sources that is not observed.

The distribution of spectral indices between 23 and 61 GHz, α_{23}^{61}, was obtained with the Kaplan–Meier estimator, taking into account the upper limits on 61 GHz fluxes. It is shifted towards steeper values compared with the distribution of α_5^{23}, obtained using the GB6 or PMN 5 GHz fluxes. Although the steepening can be just a selection effect, hints of a real high-frequency steepening are noted.

The source number counts obtained from our sample have close to Euclidean slope and are in good agreement with the prediction of the cosmological evolution model by De Zotti et al. (2005).

We have shown that the use of our spectral index distribution to extrapolate the 5 GHz counts to high frequencies leads to inconsistent results below $S_{23\,\mathrm{GHz}} \simeq 1$ Jy because of a substantial change in the mixture of flat- and steep-spectrum sources. The change clearly visible comparing our distribution of spectral indices with that obtained by Waldram et al. (2007) for a much fainter sample at the nearby frequency of 15 GHz, accounts for the difference between our high-frequency counts and extrapolations from 9C counts. This is an example of the risks inherent in extrapolations with incomplete information of source properties.

Extrapolations to higher frequencies of the 23 GHz counts using the estimated distribution of α_{23}^{61} tend to underpredict the observed counts, especially at 61 GHz, implying that the Kaplan–Meyer estimator probably overpopulates the steep portion of the spectral index distribution. The observed counts are recovered if the distribution is decreased by a factor of 0.2 for $\alpha_{23}^{61} > 0.5$.

ACKNOWLEDGMENTS

We acknowledge partial financial support from the Spanish Ministry of Education (MEC) under project ESP2004–07067–C03–01 and from the Italian ASI (contract Planck LFI Activity of Phase E2) and MUR. JGN acknowledges a postdoctoral position at the SISSA-ISAS (Trieste). Thanks are due to the Australia Telescope 20-GHz group for the useful exchange of information and for the data used in Fig. 3. This research has made use of the NASA/IPAC Extragalactic Data base (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. It has also made use of data obtained from the SuperCOSMOS Science Archive, prepared and hosted by the Wide-Field Astronomy Unit, Institute for Astronomy, University of Edinburgh, which is funded by the UK Particle Physics and Astronomy Research Council.

REFERENCES

Bennett C. L. et al., 2003, ApJS, 148, 97
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Cantor A. B., 2001, Stat. Med., 20, 2091
Cleary K. A. et al., 2005, MNRAS, 360, 340
Condon J. J., 1984, ApJ, 287, 461
Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693
Danese L., de Zotti G., 1984, A&A, 131, L1
De Zotti G., Toffolatti L., Argüeso F., Davies R. D., Mazzotta P., Partridge R. B., Smoot G. F., Vittorio N., 1999, in Maiani L., Melchiorri F., Vittorio N., eds, Proc. AIP Vol. 476, 3K Cosmology. Am. Inst. Phys., New York, p. 204
De Zotti G., Ricci R., Mesa D., Silva L., Mazzotta P., Toffolatti L., González-Nuevo J., 2005, A&A, 431, 893
Eddington A. S., 1940, MNRAS, 100, 354
Ekers R., 2006, in De Zotti G. et al., eds, Proc. Int Conf. Ser., CMB and Physics of the Early Universe. Ischia, Italy, p. 15
González-Nuevo J., Argüeso F., López-Caniego M., Toffolatti L., Sanz J. L., Vielva P., Herranz D., 2006, MNRAS, 369, 1603
Gregory P. C., Scott W. K., Douglas K., Condon J. J., 1996, ApJS, 103, 427
Griffith M. R., Wright A. E., Burke B. F., Ekers R. D., 1994, ApJS, 90, 179
Griffith M. R., Wright A. E., Burke B. F., Ekers R. D., 1995, ApJS, 97, 347
Herranz D., Sanz J. L., López-Caniego M., González-Nuevo J., 2006, Proc. IEEE Symp., Signal Processing and Information Technology. IEEE, New York, p. 541
Hinshaw G. et al., 2007, ApJS, 170, 288
Hogg D. W., Turner E. L., 1998, PASP, 110, 727
Kaplan E. L., Meier P., 1958, J. Am. Stat. Assoc., 53, 457–481
López-Caniego M., Herranz D., González-Nuevo J., Sanz J. L., Barreiro R. B., Vielva P., Argüeso F., Toffolatti L., 2006, MNRAS, 370, 2047
López-Caniego M., González-Nuevo J., Herranz D., Massardi M., Sanz J. L., De Zotti G., Toffolatti L., Argüeso F., 2007, ApJS, 170, 108 (LC07)
Lucy L. B., 1974, AJ, 79, 745
Mauch T., Murphy T., Buttery H. J., Curran J., Hunstead R. W., Piestrzynski B., Ropbertson J. G., Sadler E. M., 2003, MNRAS, 342, 1117
Massardi M. et al., 2007, MNRAS, in press (astro-ph/0709.3485)

Planck Collaboration, 2005, Planck: The Scientific Programme, ESA-SCI, 1 (astro-ph/0604069)
Ricci R. et al., 2004, MNRAS, 354, 305
Ricci R., Prandoni I., Gruppioni C., Sault R. J., de Zotti G., 2006, A&A, 445, 465
Sadler E. M. et al., 2006, MNRAS, 371, 898
Taylor A. C., Grainge K., Jones M. E., Pooley G. G., Saunders R. D. E., Waldram E. M., 2001, MNRAS, 327, L1
Toffolatti L., Argüeso Gómez F., De Zotti G., Mazzei P., Franceschini A., Danese L., Burigana C., 1998, MNRAS, 297, 117
Toffolatti L., Argüeso F., De Zotti G., Burigana C., 1999, in de Oliveira-Costa A., Tegmark M., eds, ASP Conf. Ser. Vol. 181, Microwave Foregrounds. Astron. Soc. Pac., San Francisco, p. 153
Waldram E. M., Pooley G. G., Grainge K., Jones M. E., Saunders R. D. E., Scott P. F., Taylor A. C., 2003, MNRAS, 342, 915
Waldram E. M., Bolton R. C., Pooley G. G., Riley J. M., 2007, MNRAS, 379, 1442
Wright A. E., Griffith M. R., Burke B. F., Ekers R. D., 1994, ApJS, 91, 111
Wright A. E., Griffith M. R., Hunt A. J., Troup E., Burke B. F., Ekers R. D., 1996, ApJS, 103, 145

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