



Publication Year	2021
Acceptance in OA	2022-06-10T14:02:16Z
Title	Extremely deep 150 MHz source counts from the LoTSS Deep Fields
Authors	Mandal, S., PRANDONI, ISABELLA, Hardcastle, M. J., Shimwell, T. W., Intema, H. T., Tasse, C., van Weeren, R. J., Algera, H., Emig, K. L., Röttgering, H. J. A., Schwarz, D. J., Siewert, T. M., Best, P. N., BONATO, MATTEO, BONDI, MARCO, Jarvis, M. J., Kondapally, R., Leslie, S. K., Mahatma, V. H., Sabater, J., Retana-Montenegro, E., Williams, W. L.
Publisher's version (DOI)	10.1051/0004-6361/202039998
Handle	http://hdl.handle.net/20.500.12386/32271
Journal	ASTRONOMY & ASTROPHYSICS
Volume	648

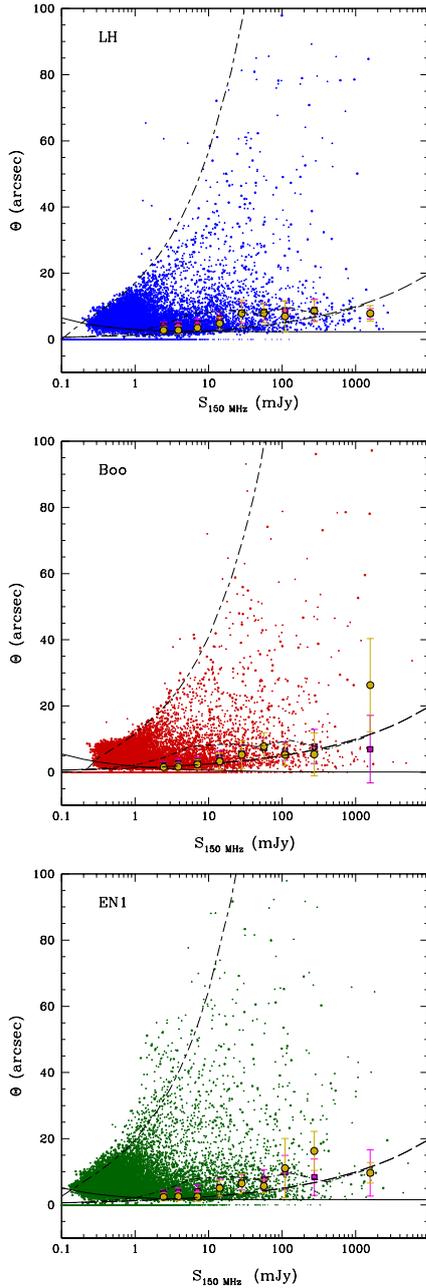


Fig. 4. Source intrinsic (deconvolved) angular sizes as a function of the measured 150 MHz total flux densities. Deconvolved sizes are defined as the geometric mean of the major and minor FWHM axes, except for well resolved radiogalaxies, which are better described by their major axis. Deconvolved sizes of point sources are set to zero. Raw (•) and final (▲) catalogs of the LH, Boo and EN1 fields are shown in the *top*, *middle* and *bottom* panels, respectively. The short-long-dashed lines in the three panels define the maximum size (Θ_{\max}) a source can have for a given measured total flux density before dropping below the detection threshold. The solid lines give the minimum size (Θ_{\min}) below which deconvolution is not considered reliable. Both lines have been drawn assuming the median noise in the masked area (see last column of Table 1). The long-dashed lines indicate the Windhorst et al. (1990) median size – flux density relation, converted to 150 MHz, while the dot-dashed lines indicate the median size – 150 MHz flux density relation derived from the simulated T-RECS catalogs directly at 150 MHz (Bonaldi et al. 2019). The filled black-bordered magenta squares and golden circles with error bars represent the median source sizes for the raw and final catalogs respectively. Medians are computed only for those flux density bins where unresolved sources represent less than 50% of the total number of sources.

of $\alpha_{150\text{MHz}}^{1.4\text{GHz}} = -0.7^{+0.02}_{-0.04}$ at $S_{150\text{MHz}} \sim 1\text{--}2$ mJy (LH; Mahony et al. 2016).

As shown in Fig. 4 the median sizes of both raw and final catalogs (respectively indicated by filled black-bordered magenta squares and golden circles with error bars) are compared with the Windhorst et al. (1990) size – flux density relation converted to 150 MHz (long-dashed line). We see a discrepancy at intermediate flux densities (10–100 mJy), where the measured sizes appear in slight excess to what was predicted by Windhorst et al. (1990). We therefore decided to consider also the median size – flux density relation derived from the Tiered Radio Extragalactic Continuum Simulation (T-RECS) catalogs at 150 MHz (Bonaldi et al. 2019, dot-dashed line), which implement different size – flux density scaling relations for star-forming galaxies and AGN. This seems to better reproduce our measured sizes at flux densities $S_{150\text{MHz}} \sim 10\text{--}100$ mJy, where extended radio galaxies (with typical sizes of hundreds of kpc) are expected to provide a significant contribution to the total radio source population. We note, however, that the afore-mentioned analysis is limited to flux densities $S_{150\text{MHz}} \gtrsim 2$ mJy, while the large majority of the sources in the LoTSS Deep Fields are fainter. Most of these sources cannot be reliably deconvolved, implying that no direct information on their size distribution can be obtained. Several attempts have been made to estimate the intrinsic source sizes at sub-mJy flux densities, based on deep samples carried out over a wide range of observing frequencies (from 330 MHz to 10 GHz). Some of these works have proposed a steepening of the Windhorst et al. (1990) median size – flux density relation at sub-mJy fluxes, with $m = 0.4\text{--}0.5$ in the range 0.1–1 mJy (Richards 2000; Bondi et al. 2003, 2008; Smolčić et al. 2017). A smooth transition from a flatter to a steeper relation at sub-mJy flux densities could again be justified by a smooth transition from a flux regime dominated by extended radio galaxies ($S \gg 1$ mJy) to a flux regime dominated by radio sources triggered by star-formation (or by composite SF/AGN emission), confined within the host galaxy.

In order to establish which size–flux density relation would best quantify the incompleteness of our catalogs we have decided to include in our analysis the results from other deep surveys. Figure 5 shows the existing measurements of (median) source sizes in various flux density bins for a number of surveys (different colors/symbols refer to different observing frequencies). Also shown are the median sizes derived by combining together the three LoTSS fields (raw and final catalogs, respectively indicated by filled black-bordered magenta squares and golden circles). To make the comparison meaningful, all flux densities referring to a different observing frequency have been converted to 1.4 GHz, assuming $\alpha = -0.7$. Also shown are various size – flux density relations: the ones proposed by Bonaldi et al. (2019, converted to 1.4 GHz) and Windhorst et al. (1990) (dot-dashed and long-dashed lines respectively), and some modifications of the latter. The short dashed lines show the ones obtained by rescaling the Windhorst et al. (1990) relation by $1.5\times$ and $2\times$ (i.e., assuming $k=3$ and $k=4$ in Eq. (7)), while the dotted line assumes a smooth transition between $m=0.3$ and $m=0.5$ going from mJy to sub-mJy flux densities, i.e.,:

$$m = m(S) = 0.3 + 0.2 \times \exp(-S_{1.4\text{GHz}}^2) \quad (8)$$

with $S_{1.4\text{GHz}}$ expressed in mJy. Focusing on the sub-mJy regime, it is clear that both the Windhorst et al. (1990) and the steeper $m(S)$ relations are consistent with the observed sizes, especially when considering only the 1.4 GHz surveys (black filled triangles). Surveys undertaken at higher frequencies seem to point toward the steeper relation, but these samples may be

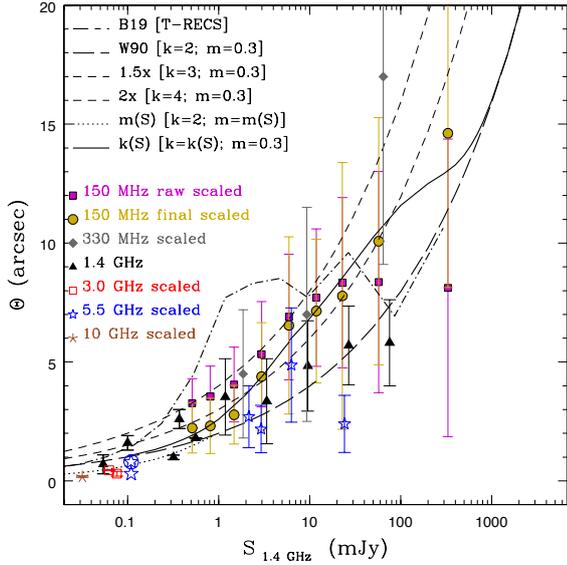


Fig. 5. Source median angular size vs. 1.4 GHz total flux density, as estimated in some of the deepest radio samples available so far. Different colors/symbols correspond to different observing frequencies: 330 MHz (gray filled diamonds – Owen et al. 2009); 1.4 GHz (black filled triangles – Richards 2000; Bondi et al. 2003, 2008; Muxlow et al. 2005; Prandoni et al. 2018); 3 GHz (red empty squares – Bondi et al. 2018; Cotton et al. 2018); 5.5 GHz (blue stars – Prandoni et al. 2006; Guidetti et al. 2017); 10 GHz (brown asterisks – Murphy et al. 2017). Also shown are the median sizes measured in our raw and final catalogs (150 MHz), combined together (filled black-bordered magenta squares and golden circles). We note that Guidetti et al. (2017) gives different median sizes for the AGN and star-forming galaxy subpopulations. The latter population is indicated as a circled blue star in the figure. All flux densities have been converted to 1.4 GHz, assuming a spectral index $\alpha = -0.7$. Various median size – flux density relations are shown for comparison: the ones proposed by Bonaldi et al. (2019) and Windhorst et al. (1990) (dot-dashed and long-dashed lines respectively), and some revised versions of the latter. The short dashed lines show the relations obtained by rescaling the Windhorst et al. (1990) relation by $1.5\times$ and $2\times$ (i.e., assuming $k = 3$ and $k = 4$ in Eq. (7)); the dotted line assumes a smooth transition between $m = 0.3$ and $m = 0.5$ going from mJy to sub-mJy flux densities, as described by Eq. (8); the solid line assumes a value of k varying with flux density according to Eq. (9) (see text for more details).

biased toward a flatter spectrum population, resulting in an over-estimation of the flux densities once converted to 1.4 GHz assuming a too steep spectral index. We also caution that higher frequency surveys more easily miss extended flux, and resolution bias issues can indeed mimic a steepening of source median sizes getting close to the flux limit of a radio survey. At larger flux densities ($S_{1.4\text{GHz}} \gtrsim 1$ mJy) the median sizes are observed to lie between the Windhorst et al. (1990) relations described by $k = 2$ and $k = 4$, with a tendency for larger sizes going to lower frequency. Indeed some analyses of source counts with shallower LOFAR surveys in the LH and Boo fields claimed in the past a better consistency with a $k = 4$ Windhorst et al. (1990) median size – flux scaling relation (Mahony et al. 2016; Retana-Montenegro et al. 2018). It is interesting to note, however, that the LoTSS final catalogs are characterized by smaller median sizes than the raw catalogs at their faint end ($S_{1.4\text{GHz}} \lesssim 5$ mJy), indicating that confusion significantly affects the measured sizes of the faintest sources, and that a significant number of faint sources were deblended. On the other hand, the final

catalogs tend to be characterized by larger sizes at the bright end ($S_{1.4\text{GHz}} \gtrsim 100$ mJy), likely as a consequence of the association of multiple components into single sources, after visual inspection of the radio/optical images (see Sect 3). After accounting for these effects, LoTSS median sizes (golden filled circles) are consistent with the Windhorst et al. (1990) $k = 2$ size – flux relation up to $S_{1.4\text{GHz}} \sim 2$ mJy. Then they smoothly increase and become consistent with the Bonaldi et al. (2019) relation at $S_{1.4\text{GHz}} \gtrsim 10$ mJy. At $S_{1.4\text{GHz}} \gtrsim 100$ mJy the LoTSS source median sizes show large uncertainties. At these large flux densities also the Bonaldi et al. (2019) relation is poorly determined, being based on a simulated catalog covering a similar area to the one covered by the LOFAR deep fields (25 deg^2). It is interesting to note, however, that both are consistent with the Windhorst et al. (1990) relation. Based on all the above considerations, a good description of the observed median sizes can be obtained by the following analytical form, which assumes the Windhorst et al. (1990) relation with a varying $k = k(S)$, i.e.,:

$$k = \begin{cases} 3.5 - 1.5 \times \exp(-S_{1.4\text{GHz}}/2) & S_{1.4\text{GHz}} < 4.5 \\ 2 + 1.5 \times \exp(-S_{1.4\text{GHz}}/200) & S_{1.4\text{GHz}} \geq 4.5 \end{cases} \quad (9)$$

where $S_{1.4\text{GHz}}$ is expressed in mJy (see solid line in Fig. 5).

Another important consistency check regards the angular size distribution of the sources. Figure 6 shows the cumulative size distributions of the final catalogs combined together, in four flux density bins (yellow solid lines). Such distributions can be considered reliable only down to a flux-dependent minimum intrinsic size (see vertical gray lines), below which most of the sources cannot be reliably deconvolved and they are conventionally assigned $\Theta = 0$. The observed distributions are compared with various realizations of the cumulative distribution function described by Eq. (6), obtained by varying either the function exponent q (left and right columns respectively) or the assumed median size – flux relations (see various black lines). The original function proposed by Windhorst et al. (1990) (Eq. (6) with $q = 0.62$, see left column) does provide a good approximation of the observed distributions, when assuming the original $\Theta_{\text{med}} - S$ relation described by Eq. (7), only at flux densities $S_{150\text{MHz}} \gtrsim 10$ mJy (see long-dashed lines). This is perhaps not surprising considering that this relation was calibrated at 1.4 GHz down to a few mJy fluxes. At the lowest flux densities ($S_{150\text{MHz}} \lesssim 1$ mJy) we need to assume a steepening of the parameter m (see Eq. (8)), to get a good match with observations (dotted line in the top left panel). This is consistent with what proposed for higher frequency deep surveys (as discussed earlier in this section). At intermediate fluxes ($S_{150\text{MHz}} \sim 1-10$ mJy), on the other hand, none of the discussed median size – flux relations can reproduce the observed size distribution (see second-row panel on the left). It is interesting to note, however, that if we assume a steeper exponent for the distribution function described by Eq. (7) (i.e., $q = 0.80$), we get a very good match with observations at all fluxes, when assuming a flux-dependent scaling factor ($k = k(S)$; see Eq. (9)) for the Windhorst et al. (1990) median size – flux relation (black solid lines on the right). The median sizes derived from the T-RECS simulated catalogs (Bonaldi et al. 2019) also provide good results for $q = 0.80$ (dot-dashed lines on the right), except again at intermediate fluxes ($S_{150\text{MHz}} \sim 1-10$), where they show strong discrepancies with observations also in Fig. 5. This seems to indicate that the number density of extended radio galaxies in this flux density range is over-estimated in the T-RECS simulated catalogs.

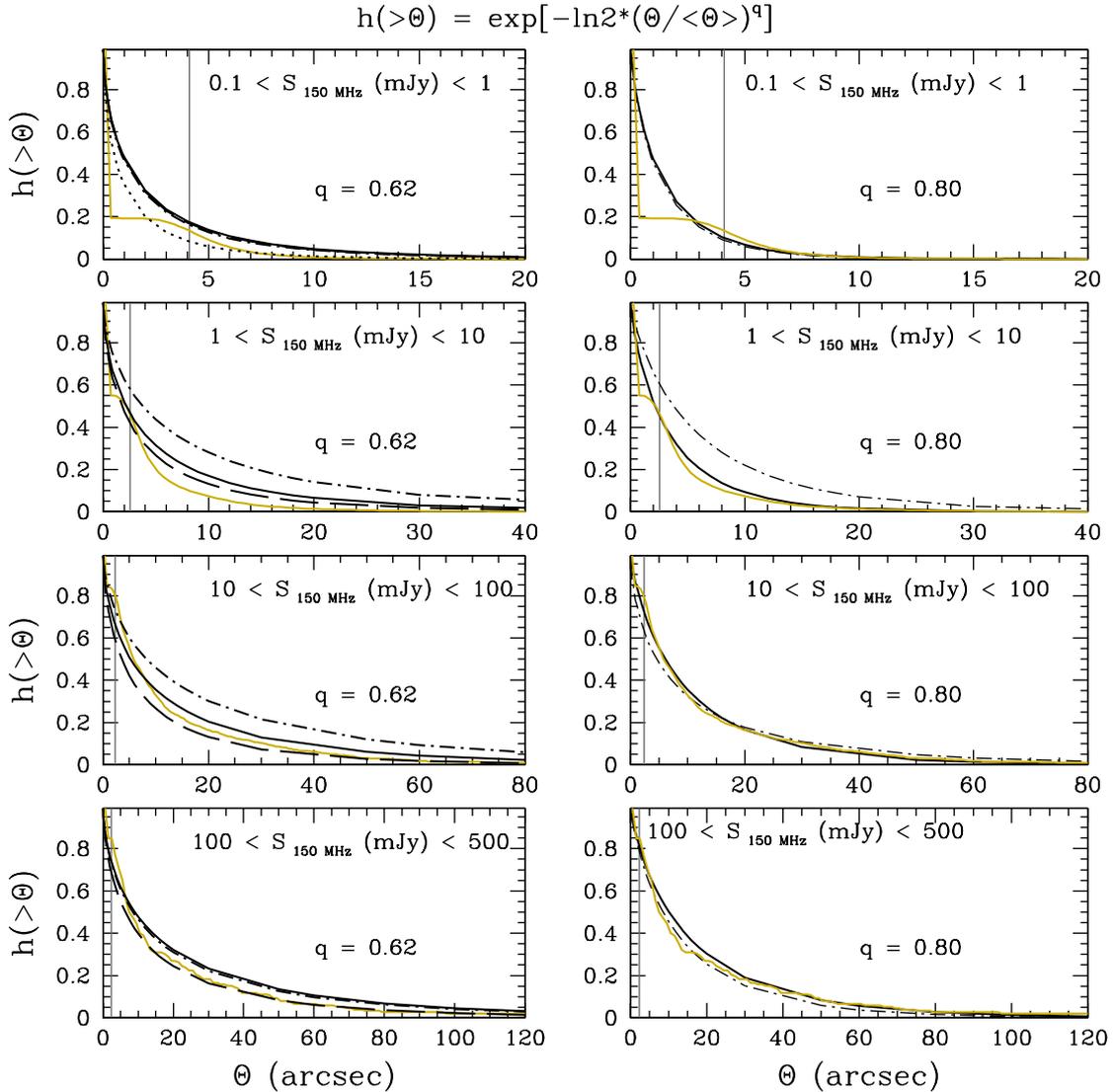


Fig. 6. Source size cumulative distribution of final catalogs (yellow solid line) in four 150 MHz flux density bins (*top to bottom*). The vertical gray lines in all panels provides an approximate indication of the minimum intrinsic angular size to which the observed distributions can be considered reliable (most of the sources below this line cannot be reliably deconvolved and they are conventionally assigned $\Theta = 0$). Also shown for comparison are various realizations of the cumulative distribution function described by Eq. (6). The two columns correspond to two different values for the function exponent q : the original one proposed by Windhorst et al. (1990) ($q = 0.62$) on the *left*, and a steeper one ($q = 0.80$) on the *right*. In addition we also vary the median size – flux relation. In particular we assume the original Windhorst et al. (1990) relation (black long-dashed line), the revised versions with flux-dependent m and k parameters, as described by Eqs. (8) and (9) (black dotted and solid lines respectively) and the one describing the T-RECS catalogs (Bonaldi et al. 2019; black dot-dashed line). All such realizations are shown on the *left*; on the *right* we only show the realizations obtained using the Bonaldi et al. (2019) and the revised Windhorst et al. (1990) $k = k(S)$ relations.

Correction for resolution bias

The correction factor c that needs to be applied to the source counts to account for the resolution bias can be defined as (Prandoni et al. 2001):

$$c = 1/[1 - h(> \Theta_{\text{lim}})] \quad (10)$$

where $h(> \Theta_{\text{lim}})$ takes the form of the integral of the angular size distribution proposed by Windhorst et al. (1990, see Eq. (6)), and Θ_{lim} is the limiting angular size above which the catalogs are expected to be incomplete. Following Prandoni et al. (2001), this is defined as:

$$\Theta_{\text{lim}} = \max[\Theta_{\text{min}}, \Theta_{\text{max}}] \quad (11)$$

where Θ_{max} and Θ_{min} are as defined in Eqs. (4) and (5) respectively. We notice that Θ_{lim} is always equal to Θ_{max} , except for the lowest flux bins, where Θ_{max} becomes unphysical (i.e., tends to zero). Θ_{min} accounts for the effect of having a finite restoring beam size (that is $\Theta_{\text{lim}} > 0$ at the survey limit) and a deconvolution efficiency which varies with the source peak flux density (see Prandoni et al. 2001 for more details).

Figure 7 (left panel) shows the correction factor derived assuming the median size – flux relations discussed above, combined with appropriate values of the q exponent in Eq. (6), based on our analysis of the source size distribution (see Fig. 6 and related discussion). A caveat to keep in mind is that the resolution bias correction does depend on both the source flux density and the noise value at the source position (and/or the source signal-to-noise ratio; see Eqs. (4) and (5)). The corrections

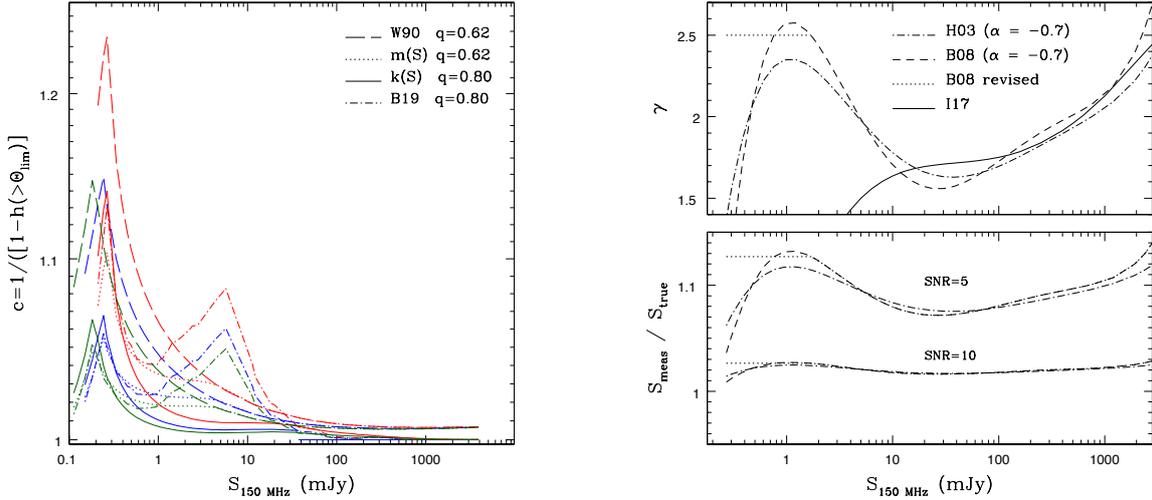


Fig. 7. *Left:* flux-dependent correction to be applied to source counts to account for incompleteness due to resolution bias, for four median size - flux relations: the one derived from the simulated T-RECS catalogs (Bonaldi et al. 2019; dot-dashed lines), the one proposed by Windhorst et al. (1990, long-dashed lines), the revised version with $m = m(S)$, which better describe source sizes at 1.4 GHz sub-mJy fluxes (see Eq. (8), dotted lines), and the revised version with $k = k(S)$ proposed by us (Eq. (9), solid lines). We also vary the q exponent of the integral distribution function. Based on our analysis of the source size distribution (see Fig. 6 and related discussion), we assume the original value proposed by Windhorst et al. (1990, $q = 0.62$) for the Windhorst et al. (1990) median size - flux relation and for the revised version with $m = m(S)$. We assume a steeper $q = 0.80$ for the revised version with $k = k(S)$ and for the Bonaldi et al. (2019) relation (see legenda). Different colors refer to different fields: LH (blue), Boo (red), EN1 (green). The corrections account for noise variations in the masked images through an empirical relation between source flux and source signal-to-noise ratio, calibrated for each field (we assume here the median noise of the masked images; see last column of Table 1). *Right:* Eddington bias for different underlying number-count distributions, as illustrated in the top panel: source counts' slope (γ ; $dN/dS \sim S^{-\gamma}$) derived from the sixth-order polynomial fit proposed at 1.4 GHz by (a) Hopkins et al. (2003; dot-dashed line) and (b) Bondi et al. (2008; dashed line), both converted to 150 MHz assuming a spectral index $\alpha = -0.7$; we also show a revised version of the Bondi et al. (2008) fit, which assumes a constant Euclidean slope ($\gamma = 2.5$) from 2 mJy all the way down to 0.1 mJy (dotted line). The polynomial fit proposed by Intema et al. (2017) at 150 MHz and valid only for the bright end of the counts is also shown for reference (solid line). The flux boosting ($S_{\text{meas}}/S_{\text{true}}$) corresponding to the three cases illustrated above is shown in the bottom panel for two different source signal-to-noise ratios: $S/N = 5$ and $S/N = 10$.

presented in Fig. 7 (left panel) account for local and radial variations of the noise through empirical relations between source flux and local noise or signal-to-noise ratio, specifically derived for each field. Such relations describe average trends only, and hence the corrections presented here should be considered as indicative. The corrections effectively applied to the counts are based on the actual source flux density, noise and signal-to-noise ratio distributions. It is interesting to note, however, that, as a consequence of radially-increasing noise (and/or limited dynamic range around bright radio sources), the correction factor c does not necessarily converge to 1 at large flux densities. As shown in Fig. 7, in the masked regions of our fields this only happens when assuming the shallower integral distribution function ($q = 0.62$). For the steeper one ($q = 0.80$), the expected number density of very extended sources is small, and resolution bias effects become negligible at $S_{150\text{MHz}} \gtrsim 500$ mJy.

5. Eddington bias

While correcting for resolution bias is important to account for missed resolved sources, Eddington bias (Eddington 1913, 1940) should be taken into account to get an unbiased census of unresolved sources. Due to random measurement errors the measured peak flux densities will be redistributed around their true value. In presence of a source population which follows a nonuniform flux distribution, this will result in a redistribution of sources between number-count flux density bins. The way the sources are redistributed depends on the slope of source counts. If the source number density increases with decreasing flux, the flux densities tend to be boosted and the probability to detect a source

below the detection threshold is higher than the probability to miss a source above the threshold, artificially boosting the detection fraction. As a consequence the catalog incompleteness at the detection threshold is also biased.

There are two main approaches to correct for Eddington bias, both requiring an assumption about the true underlying source counts distribution (see Hales et al. 2014a for a full discussion): one can build the source counts using the boosted flux densities and then apply a correction to each flux density bin, or one can correct the source flux densities, before deriving the counts. As demonstrated by Hales et al. (2014b) the two approaches give very similar and consistent results, and we decided to follow the latter approach. A maximum likelihood solution for the true source flux density can be defined as follows (see Hales et al. 2014a and references therein):

$$S_{\text{true}} = \frac{S_{\text{meas}}}{2} \cdot \left(1 + \sqrt{1 - \frac{4\gamma}{(S/N)^2}} \right) \quad (12)$$

where $\gamma = \gamma(S)$ is the slope of the counts at the given flux density ($dN/dS \sim S^{-\gamma}$), and S/N is the source signal-to-noise ratio. The slope of the counts can be modeled from empirical polynomial fits of the observed counts:

$$\log \left(\frac{dN(S)}{dS} S^{2.5} \right) = \sum_{i=0}^n a_i (\log S)^i. \quad (13)$$

It is then easy to demonstrate that:

$$\gamma = 2.5 - \sum_{i=0}^n i \cdot a_i (\log S)^{i-1}. \quad (14)$$

In order to derive γ we can use one of the several counts' fits available in the literature. [Intema et al. \(2017\)](#) derived a fifth-order polynomial fit which describes the 150 MHz normalized counts of the TIFR GMRT Sky Survey (TGSS-ADR1), but this fit is only valid down to a flux limit of 5 mJy. The deepest fits available in the literature have been obtained at 1.4 GHz. We start by exploring the sixth-order ($n=6$) polynomial fits obtained by [Hopkins et al. \(2003\)](#) and [Bondi et al. \(2008\)](#) for 1.4 GHz normalized source counts (converted to 150 MHz using $\alpha = -0.7^8$). The two cases are illustrated in the top right panel of Fig. 7 (dot-dashed and dashed lines respectively), where the derived counts' slope is shown. Both cases are consistent with the [Intema et al. \(2017\)](#) 150 MHz fit (indicated by the solid line) at bright flux densities ($S > 100$ mJy), while significant discrepancies are observed at fainter fluxes, where the deeper 1.4 GHz fits better describe the well-known flattening of the normalized counts. Both the 1.4 GHz fits show an increasing slope below 10 mJy, reaching a maximum around 1 mJy. This maximum is more pronounced in the case of [Bondi et al. \(2008\)](#), and is consistent with an Euclidean slope of $\gamma \sim 2.5$. At $S < 1$ mJy both slopes show a rapid drop. The reality and strength of this drop is unclear, as this is the flux regime where the fits are less reliably constrained. We then explore a third case, i.e., a modification of the [Bondi et al. \(2008\)](#) fit, which assumes a constant Euclidean slope at flux densities $S < 2$ mJy. This represents an extreme scenario, which however might be favored by the recent 150 MHz source counts modeling proposed by [Bonaldi et al. \(2019\)](#), that indicates a flatter slope in the flux range 0.1–1 mJy. This last case is illustrated by the dotted line in Fig. 7 (top right panel). The flux density boosting expected for the three aforementioned scenarios is illustrated in the bottom right panel of Fig. 7 for two signal-to-noise ratio values, $S/N = 5$ and $S/N = 10$.

Once the point source flux densities are corrected for Eddington bias, we can obtain an estimate of the catalog incompleteness at the detection threshold through the use of Gaussian Error Functions (ERF).

6. Differential source counts

The differential source counts, normalized to a non-evolving Euclidean model, obtained from the LoTSS Deep Fields are shown in Fig. 8, together with other count determinations obtained in the same fields from previous low-frequency surveys (see legend). The top and bottom panels refer to counts derived from the raw and final catalogs, respectively (see filled boxes). The source counts obtained from the final catalogs are reported in tabular form in the appendix (Tables A.1–A.3, for LH, Boo and EN1 fields respectively), and are also shown in Fig. 10, where they are compared to counts extrapolated from higher frequencies. In deriving the counts we applied a 'fiducial' model for the systematic corrections described in Sects. 4 and 5. Specifically we assumed the [Windhorst et al. \(1990\)](#) size – flux relation with $k = k(S)$ (Eq. (9)), in combination with a 'steep' ($q = 0.80$ in Eq. (6)) integral size distribution, to estimate the resolution bias, and we assumed the [Bondi et al. \(2008\)](#) source counts best fit to

⁸ We note that assuming a single spectral index is a crude approximation. In principle we should account for the intrinsic scatter in the spectral index distribution of the sources, as well as for possible deviations of the mean spectral index with flux density, due to the varying relative contribution of the individual source populations. Such an approximation is however acceptable, since the largest uncertainties in this analysis come from the assumptions on the counts' slope at the faintest fluxes, which is very poorly known.

estimate the Eddington bias. The uncertainties associated with such assumptions are factored into systematic error terms (see Sys^- and Sys^+ columns in the counts' tables), that are defined as the maximum discrepancy between the 'fiducial' counts and those obtained assuming the other discussed models (shown in Fig. 7). Also shown in Fig. 8 are the counts obtained from the three LoTSS fields before applying the corrections for resolution and Eddington bias (empty boxes). As expected such corrections are only relevant for the lowest flux density bins. We cut the source counts at a threshold of $\sim 7\sigma_{\text{med}}$, where systematic errors dominate over Poissonian (calculated following [Gehrels 1986](#)) by factors ~ 5 –10.

The normalized 150 MHz source counts derived from the three LoTSS Deep Fields are in broad agreement with each other, and show the well known flattening at $S \lesssim \text{few mJy}$. The observed field-to-field variations for the counts derived from the final catalogs are typically on the order of a few percent at sub-mJy fluxes, and of 5–10% at flux densities 1–10 mJy, with EN1 showing somewhat larger deviations (up to 15–20% in some flux density bins). This is consistent with expectations from sample variance for surveys covering areas of 5–10 deg² ([Heywood et al. 2013](#); [Prandoni et al. 2018](#)). We notice that all of the other 150 MHz counts shown in Fig. 8 come from previous shallower LOFAR observations, but the catalogs did not go through the source association and deblending post-processing described in Sect. 3. This explains why determination of previous source counts are more in line with those we obtained from raw catalogs (see top panel of Fig. 8). The largest discrepancies are observed in the Boo field with the counts by [Retana-Montenegro et al. \(2018\)](#). The systematically higher counts that we obtain at sub-mJy flux densities may be the result of the new calibration and imaging pipeline, that produces higher fidelity, higher dynamic range images (Paper I). Another source of systematic differences may be a residual offset in the absolute flux scale.

When comparing the source counts derived from raw and final catalogs (top and bottom panels of Fig. 8), we notice a very interesting feature: the latter show a much more pronounced drop at flux densities around a few mJy, which results in a more prominent 'bump' in the sub-mJy regime. For a more quantitative analysis of this feature we have combined the sources in the three fields and produced a combined version of the source counts. This allows us to smooth out the aforementioned field-to-field variations, as well as reduce the scatter at bright flux densities, where Poissonian errors dominate. The combined counts derived from raw and final catalogs are shown in Fig. 9 and listed in Table 3. We notice that in this case we included LH and Boo sources down to a 5σ flux density limit, to increase the statistics available in the first two flux density bins. From the comparison of the raw and final counts we see that the latter are systematically lower by a factor 5–15% in the range $S \sim 1$ –10 mJy. This deficiency appears to be counterbalanced by an 8–16% excess at flux densities 0.2–0.6 mJy. This form of compensation is likely the result of source deblending (see Sect. 3 for more details). Indeed, for each field we have ~ 1000 s deblended sources (see Paper II), and deblending is mostly effective at the faint end of the counts, where it contributes to the increase of the number of the sources, through the splitting of confused (brighter) sources into several fainter ones. Artifact removal, on the other hand, has probably a limited effect. Artifacts mostly affect the flux density range 2–20 mJy⁹, but only < 100 sources per field are confirmed

⁹ This was verified by running PyBDSF on the negative raw map (i.e., the raw map multiplied by -1), with the same input parameters used to extract the source catalog (see e.g., [Prandoni et al. 2018](#)).

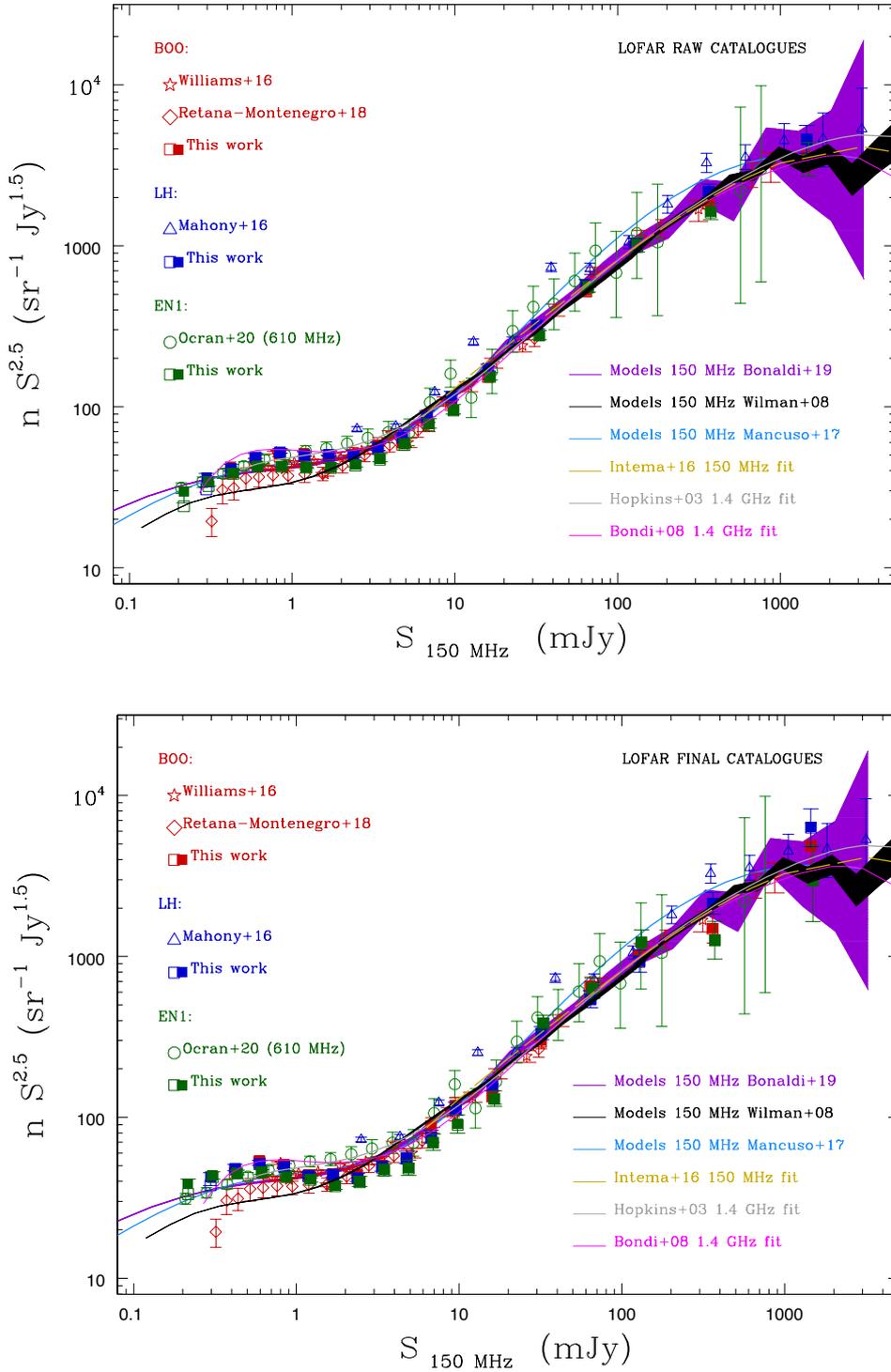


Fig. 8. Normalized 150 MHz differential source counts in the three LoTSS Deep Fields, as derived from raw (*top*) and final (*bottom*) catalogues (filled squares). Error bars correspond to the quadratic sum of Poisson and systematic errors. Also shown are the counts obtained without applying the corrections discussed in Sects. 4 and 5 (empty squares). The counts are derived by using total flux densities for both point and extended sources. In both figures, Wilman et al. (2008), Bonaldi et al. (2019) and Mancuso et al. (2017) 150 MHz models are shown for comparison, as well as other existing 150 MHz counts’ determinations in the same fields (see legend). Since published 150 MHz counts are missing for EN1, we show a recent determination obtained at 610 MHz (Ocran et al. 2020) and rescaled to 150 MHz, assuming $\alpha = -0.7$. Also shown are the counts’ best fits discussed in Sect. 5.

as such (see Paper II). The fact that source cataloguing processes can significantly affect the derived number counts is well known. Significant discrepancies are observed, for instance, when comparing counts derived from source catalogs with those directly obtained from Gaussian components (see e.g., Hopkins et al. 2003). While PyBDSF creates a source catalog, our association and deblending procedure aims at further combining Gaussian components together (association) or splitting them into several independent sources (deblending), hence modifying the source statistics and flux distribution of the original PyBDSF catalog.

As a final remark we note that none of the best fits from previous source counts (shown in Fig. 8) provide a good description of

the pronounced ‘drop and bump’ feature that we observe at 150 MHz. We have therefore derived a new fit that better matches the faint end of the 150 MHz counts derived from our final catalogs. The slope of the counts is modeled by a 7th order polynomial function defined in the log-log space, according to Eq. (13) (see Sect. 5 for more details). To better constrain the bright end of the counts, where the LoTSS Deep Fields provide poor statistics, we have included the counts derived from the TGSS-ADR1 (Intema et al. 2017), up to a flux density of 10 Jy (beyond that the TGSS-ADR1 itself is affected by poor statistics). The resulting coefficient values and their uncertainties are listed in Table 4; the fitted curve is shown in Fig. 9.

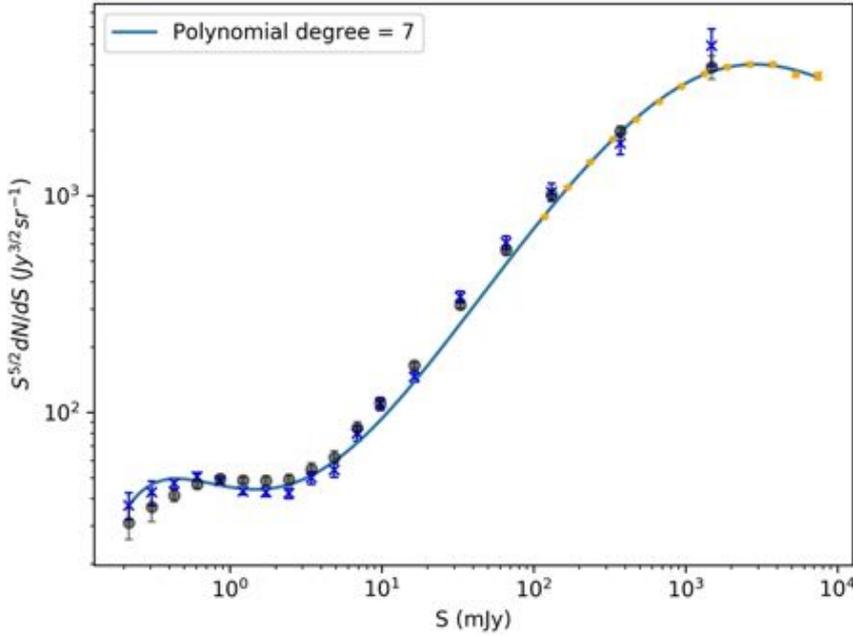


Fig. 9. 150 MHz Euclidean normalized differential source counts as derived from the LoTSS Deep Fields: raw catalog are indicated by transparent black circles and final catalog by blue crosses. Also shown are the counts obtained from the TGSS-ADR1 (Intema et al. 2017, orange filled circles), which better describe the counts’ bright end. Over-plotted is the best fit obtained by modeling the counts in the log–log space with a 7th order polynomial function, according to Eq. (13) (see Table 4 for the values of the best-fit coefficients and associated errors).

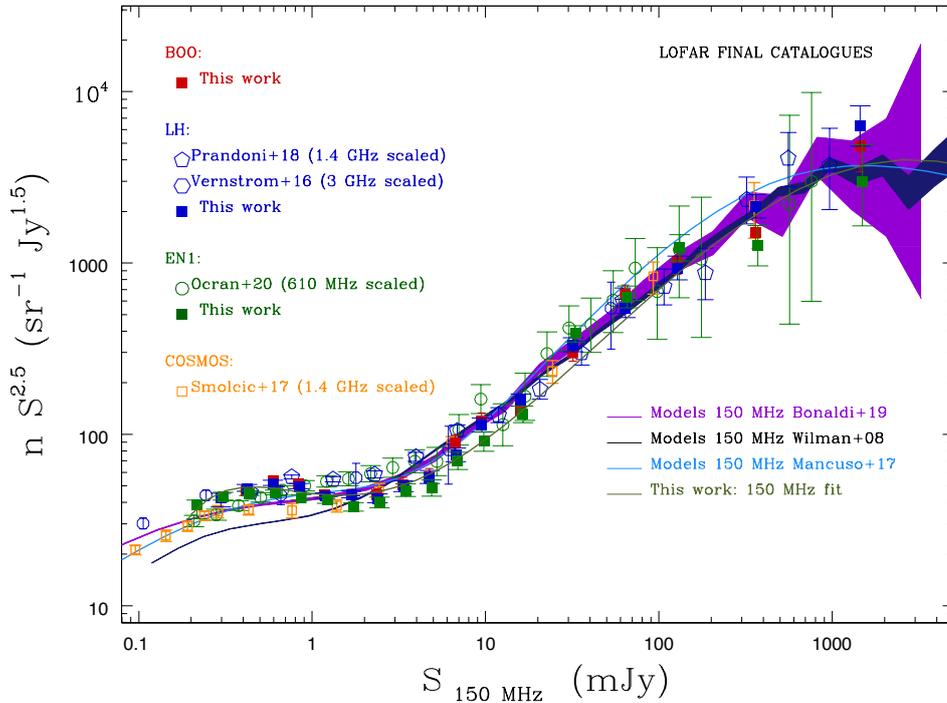


Fig. 10. Normalized 150 MHz differential source counts in the three LoTSS Deep Fields, as derived from final catalogs (filled squares), together with their best fit (gray solid line). They are compared to counts derived from higher frequencies surveys (see legend), and rescaled to 150 MHz by assuming $\alpha = -0.7$. Error bars correspond to the quadratic sum of Poisson and systematic errors. Also shown are the Wilman et al. (2008), Bonaldi et al. (2019) and Mancuso et al. (2017) 150 MHz models. The counts are derived by using total flux densities for both point and extended sources.

6.1. Qualitative comparison with higher frequency counts

Our source counts are the deepest available at 150 MHz. Their faint end can therefore only be probed against counts obtained from higher frequency surveys of similar depth. In Fig. 10 we show the normalized differential counts in the three LoTSS Deep Fields, as derived from final catalogs, together with counts derived from higher frequencies surveys over the same regions of the sky, and rescaled to 150 MHz by assuming $\alpha = -0.7$. We also add the COSMOS 3GHz Large project (Smolčić et al. 2017), which provides the deepest counts to date, over a degree-scale field. None of the higher frequency counts show the pronounced ‘drop’ at mJy flux densities, that we observe at 150 MHz. The only exception is COSMOS, which is in agreement with our 150 MHz counts at $S \gtrsim 1$ mJy, but falls below them at lower flux

densities. We note that sample variance can in principle explain variations $\gtrsim 15\text{--}25\%$ in the range 1–10 mJy for surveys covering areas $\lesssim 2$ deg² (Heywood et al. 2013), like COSMOS and the ones by Ocran et al. (2020) and Vernstrom et al. (2016) in EN1 and LH respectively, but can hardly justify the observed discrepancies with Prandoni et al. (2018) sample, which covers more than half (~ 6.6 deg²) of the LoTSS LH field. In addition, sample variance is expected to be on the order of 5–10% at sub-mJy fluxes for a sample like COSMOS, i.e., smaller than what observed. It is then likely that survey systematics additionally contribute to the observed field-to-field variations. In this respect we note that the excess at mJy flux densities is significantly mitigated when comparing the higher frequency counts to the LoTSS ones derived from raw catalogs (see Ocran et al. 2020 counts in top panel of Fig. 8), indicating that the catalog

Table 3. 150 MHz normalized differential radio-source counts as derived from combining the raw and final catalogs of the three LoTSS Deep Fields.

$\langle S \rangle$	$N(\text{raw})_{-\sigma_{\text{tot}}}^{+\sigma_{\text{tot}}}$	$N(\text{final})_{-\sigma_{\text{tot}}}^{+\sigma_{\text{tot}}}$
0.22	$30.92_{-6.65}^{+2.20}$	$37.15_{-7.92}^{+2.59}$
0.31	$36.57_{-7.04}^{+3.13}$	$42.65_{-7.57}^{+3.40}$
0.43	$41.29_{-2.17}^{+3.13}$	$47.08_{-1.86}^{+3.40}$
0.61	$46.42_{-0.52}^{+3.16}$	$50.55_{-0.62}^{+3.40}$
0.86	$49.39_{-0.42}^{+3.15}$	$47.97_{-0.68}^{+2.85}$
1.22	$48.48_{-0.53}^{+3.04}$	$43.17_{-0.83}^{+2.23}$
1.73	$48.24_{-0.67}^{+4.13}$	$42.77_{-1.07}^{+2.77}$
2.45	$48.73_{-0.87}^{+4.76}$	$42.30_{-1.37}^{+3.30}$
3.46	$54.60_{-1.19}^{+6.20}$	$49.81_{-1.93}^{+4.45}$
4.89	$61.95_{-1.64}^{+7.08}$	$54.21_{-2.61}^{+5.71}$
6.92	$84.37_{-2.47}^{+9.42}$	$79.79_{-4.10}^{+8.09}$
9.79	$110.2_{-3.7}^{+8.2}$	$109.6_{-6.2}^{+8.8}$
16.5	$164.5_{-4.6}^{+6.5}$	$145.9_{-7.5}^{+8.3}$
32.9	$312.8_{-10.7}^{+11.2}$	$341.5_{-19.2}^{+20.3}$
65.8	$560.6_{-24.1}^{+25.2}$	$607.1_{-42.9}^{+46.1}$
132	1012_{-55}^{+58}	1041_{-95}^{+103}
372	1977_{-115}^{+123}	1739_{-183}^{+203}
1489	3932_{-458}^{+515}	4921_{-870}^{+1027}

Notes. $\langle S \rangle$ is the geometric mean of the respective flux density bin, expressed in mJy; $N(\text{raw})$ and $N(\text{final})$ indicates the normalized source counts obtained from the raw and final catalogs respectively (in $\text{Jy}^{1.5} \text{sr}^{-1}$); $\pm \sigma_{\text{tot}}$ are the total errors on the counts, estimated as the quadratic sum of Poissonian and systematic errors. We note that only EN1 and LH sources contribute to the first flux density bin.

Table 4. Coefficients for 7th order polynomial function defined by Eq. (13), which best-fit the LoTSS and TGSS-ADR1 150 MHz normalized source counts in the flux density range: 0.2 mJy–10 Jy.

Coefficient	Value	Error (\pm)
a_0	1.655	0.017
a_1	-0.1150	0.0654
a_2	0.2272	0.097
a_3	0.51788	0.18616
a_4	-0.449661	0.185032
a_5	0.160265	0.084983
a_6	-0.028541	0.018280
a_7	0.002041	0.001497

Notes. The polynomial fit is shown in Fig. 9.

post-processing undertaken for the LoTSS deep fields did play a role here. Indeed we could count on very complete identification rates ($\geq 97\%$ of radio sources have an optical/IR counterpart), enabling a very accurate process of association, deblending and artifact removal. This is not true for other samples which went through similar procedures, like the Prandoni et al. (2018) sample (identification rate $\sim 80\%$ down to $S_{1.4\text{GHz}} = 0.12$ mJy) or the Ocran et al. (2020) sample (identification rate $\sim 90\%$). On the

other hand, COSMOS counts may suffer from flux losses at their faintest end, being derived from a higher frequency (3 GHz), higher (sub-arcsec) resolution survey. Another source of systematics may be the extrapolation of the higher frequency counts to 150 MHz, which is sensitive to the assumed spectral index. In particular, adopting the same spectral index value for all the sources may not be appropriate. There are indeed indications that the spectral index may flatten at mJy/sub-mJy regimes, due to the emergence of a population of low power, core dominated AGN (e.g., Prandoni et al. 2006; Whittam et al. 2013) and re-steepen again at μJy levels, where star-forming galaxies become dominant (e.g., Owen et al. 2009).

6.2. Qualitative comparison with models

The source counts derived from the LoTSS Deep Fields provide unprecedented observational constraints to the shape of the source counts at 150 MHz sub-mJy flux densities. As such they can be compared with counts predictions based on existing evolutionary models of radio source populations. A comprehensive comparison with models is beyond the scope of this paper, and will be the subject of forthcoming papers, where counts and luminosity functions will be presented and discussed for various radio source populations. Here we only provide some first qualitative considerations.

In Figs. 8 and 10 we compare the LoTSS source counts to the 150 MHz determinations derived from the Wilman et al. (2008) and Bonaldi et al. (2019) simulated catalogs¹⁰ (black and dark violet shaded curves, respectively), as well as from Mancuso et al. (2017) models (light blue curve). We notice that the Bonaldi et al. (2019) and Wilman et al. (2008) source counts are very similar at the bright end and better reproduce the observations than Mancuso et al. (2017). On the other hand, Bonaldi et al. (2019) and Mancuso et al. (2017) counts are very similar at the faint end, and in better agreement with the observations than Wilman et al. (2008). Nevertheless, neither Bonaldi et al. (2019) nor Mancuso et al. (2017) can reproduce the pronounced bump at sub-mJy flux densities, observed in the counts derived from final catalogs (see Fig. 8, bottom panel, or Fig. 10). In addition all models appear to over-estimate the counts derived from our final catalogs at intermediate flux densities ($S \sim 2\text{--}20$ mJy). An over-prediction of the observed counts over T-RECS and Wilman et al. (2008) simulations in the flux density range 3–12 mJy was also noticed by Siewert et al. (2020) in their analysis of the HETDEX field of LoTSS-DR1 (Shimwell et al. 2017, 2019).

In an attempt to better understand where the evolutionary models may fail, we compare the observed source redshift distribution (using redshifts from Paper IV) with those of the Bonaldi et al. (2019) simulated catalog. We restrict this comparison to the EN1 field, as it is the deepest and has the most complete optical coverage among the three LoTSS Deep Fields. We limit our analysis to the flux density range $0.25 \text{ mJy} \leq S_{150} \leq 20$ mJy. At lower flux densities the effects of the visibility function and of incompleteness cannot be neglected anymore; at larger flux densities the available source statistics is very sparse. In the flux density regime under consideration, we have 14,916 sources; 14 432 of them (97%) have measured redshifts (91%

¹⁰ Bonaldi et al. (2019) present three simulated catalogs, each covering a different area of the sky. The one used here is the so-called medium tier, which covers a 25 deg^2 field of view, providing a very good match to the LoTSS Deep Fields. We use a new version of the catalogs originally presented in Bonaldi et al. (2019), which better reproduce the observations at the bottom and top of the covered frequency range (Bonaldi, priv. comm.).