

SCORPIO – II. Spectral indices of weak Galactic radio sources

F. Cavallaro,^{1,2,3★} C. Trigilio,² G. Umana,² T. M. O. Franzen,⁴ R. P. Norris,^{3,5} P. Leto,² A. Ingallinera,² C. S. Buemi,² J. Marvil,³ C. Agliozzo,^{6,7} F. Bufano,² L. Cerrigone⁸ and S. Riggi²

¹Università degli studi di Catania, I-95123 Catania, Italy

²INAF–Osservatorio Astrofisico di Catania, Via S. Sofia 78, I-95123 Catania, Italy

³CSIRO Astronomy and Space Science, PO Box 76, Epping, NSW 1710, Australia

⁴International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia

⁵Western Sydney University, Locked Bag 1797, Penrith South, NSW 1797, Australia

⁶Millennium Institute of Astrophysics, Santiago 7591245, Chile

⁷Universidad Andrés Bello, Avda. Republica 252, Santiago 8320000, Chile

⁸ASTRON, the Netherlands Institute for Radioastronomy, PO Box 2, NL-7990 AA Dwingeloo, the Netherlands

Accepted 2017 September 20. Received 2017 September 20; in original form 2016 September 13

ABSTRACT

In the next few years the classification of radio sources observed by the large surveys will be a challenging problem and spectral index is a powerful tool for addressing it. Here we present an algorithm to estimate the spectral index of sources from multiwavelength radio images. We have applied our algorithm to SCORPIO, a Galactic plane survey centred around 2.1 GHz carried out with Australian Telescope Compact Array and found we can measure reliable spectral indices only for sources stronger than 40 times the rms noise. Above a threshold of 1 mJy, the source density in SCORPIO is 20 per cent greater than in a typical extragalactic field, like Australia Telescope Large Area Survey because of the presence of Galactic sources. Among this excess population, 16 sources per square degree have a spectral index of about zero suggesting optically thin thermal emission such as H II regions and planetary nebulae, while 12 per square degree present a rising spectrum, suggesting optically thick thermal emission such as stars and UCH II regions.

Key words: techniques: interferometric–Galaxy: stellar content–radio continuum: galaxies–radio continuum: ISM–radio continuum: stars.

1 INTRODUCTION

We are entering a golden age for radio astronomy. There are several new interferometers upcoming, including the square kilometre array (SKA; Carilli & Rawlings 2004) and its precursors, such as the Australian SKA pathfinder (ASKAP, Johnston et al. 2008) and MeerKAT (Jonas 2009) that will deliver an extraordinary number of new discoveries. In particular, we expect a lot of data from the forthcoming radio continuum surveys, both in the Galactic plane (GP) and at high Galactic latitude, like the evolutionary map of the Universe (EMU; Norris et al. 2011) to be carried out with ASKAP, and the MeerKAT international GigaHertz tiered extragalactic exploration survey (MIGHTEE; Jarvis 2015) and the MeerKAT high frequency Galactic plane survey (MeerGAL), both to be carried out with MeerKAT. We will need to classify a large number of sources,

especially in the GP, where the Galactic population adds to the extragalactic population. There are essentially six ways to classify a generic radio source depending on the data available and whether it is resolved or not:

- (i) studying its morphology, which requires that the source is resolved;
- (ii) studying the radio emission mechanism, which requires that the spectral energy distribution (SED) in the radio wavelengths is known;
- (iii) studying the complete SED, which requires the identification of the source counterparts at different wavelengths;
- (iv) studying the polarization, which requires a very high signal-to-noise ratio (S/N); and
- (v) studying the time domain, which requires a sufficient time resolution.

In this paper we will focus on the second method, since it is solely based on radio multiwavelength or wide-band observations and

* E-mail: francesco.cavallaro@oact.inaf.it

no other information is required. The radio SED depends on the continuum emission mechanism, which could mainly be thermal (bremsstrahlung) or non-thermal (synchrotron).

A parameter that characterises the SED at the radio wavelengths is the spectral index α , defined as $S = S_0(\frac{\nu}{\nu_0})^\alpha$, where S is the flux density of the source at a given frequency ν and S_0 is the flux density at the frequency ν_0 . The spectral index can change with the frequency, but both synchrotron and bremsstrahlung lose this dependence in the two optical depth limits $\tau \gg 1$ and $\tau \ll 1$. In particular, in the case of a homogeneous source of bremsstrahlung we have $\alpha = 2$ for $\tau \gg 1$ while, for $\tau \ll 1$, $\alpha = -0.1$ (Burke & Graham-Smith 2009b). In the case of $n \propto r^{-2}$ (like a stellar wind), where n is the electronic density and r is the distance from the centre, $\alpha = 0.6$ (Panagia & Felli 1975; Wright & Barlow 1975). In the case of a homogeneous source of synchrotron $\alpha = 2.5$ for $\tau \gg 1$ while for $\tau \ll 1$, α depends on the spectral index δ of the energy distribution of the relativistic electrons, defined as $\alpha = -\frac{\delta-1}{2}$ so, for typical values of δ , it ranges between -1 and -0.5 (Burke & Graham-Smith 2009a).

The SED can also be modified by synchrotron self-absorption and free-free absorption, causing a low-frequency turnover, and by electron cooling, causing a high-frequency break. Complex models are invoked to reconcile these models with the data, which are beyond the scope of this paper, but they are discussed extensively by Collier et al. (in preparation).

The frequency at which the flux density of a source reaches a maximum is called the turnover frequency. Measuring the turnover frequency can allow us to constrain additional physical parameters, such as magnetic field intensity or electronic density. To derive the spectral index and all the connected informations for a large number of sources, considering that the future surveys will detect millions of sources (Norris et al. 2011), we developed an algorithm to automatically compute the spectral index.

Our algorithm has been tested by using the images of Stellar Continuum Originating from Radio Physics In Our galaxy (SCORPIO, Umana et al. 2015, hereafter Paper I) survey. SCORPIO is a $2 \times 2 \text{ deg}^2$ survey in the GP centred at Galactic coordinates $l = 344^\circ.25$, $b = 0^\circ.66$, carried out with the Australian Telescope Compact Array (ATCA) between 1.1 and 3.1 GHz that acts as a pathfinder for EMU. A pilot experiment took place in 2011 and 2012, observing about a quarter of the selected field. The rms of the map was $\approx 25-30 \mu\text{Jy beam}^{-1}$, with a resolution of 14.0 by 6.5 arcsec (Paper I). Notice that SCORPIO is actually larger than 4 deg^2 because, to have a better coverage of the region, we preferred to observe a larger area. This can also be applied to the pilot, in fact its area is actually 2.333 deg^2 , compared to the declared 1 deg^2 .

In Paper I we produced a catalogue of 614 point sources in the pilot field, listing their flux densities, position and possible infrared counterpart. The data release of the full SCORPIO field, with a catalogue of about 2000 sources, is in progress (Trigilio et al. in preparation). In this paper we present a statistical analysis of the radio SED of the 614 point sources extracted and catalogued in Paper I, to statistically discriminate Galactic from extragalactic point sources. In Section 2 we give a brief overview of the observations and data reduction. In Section 3 we describe our algorithm to automatically reconstruct the source SED and derive the spectral index. In Section 4 we present our results, thoroughly discussing their relation with ATLAS (Australia Telescope Large Area Survey, Norris et al. 2006) ones, highlighting and explaining the differences. Conclusions are reported in Section 5.

Table 1. Sub-band frequency and primary beam radii at 5 per cent of the peak.

| Sub-bands | Central frequency (GHz) | R (arcmin) |
|-----------|----------------------------|-----------------|
| 1 | 1.469 | 30.8 |
| 2 | 1.649 | 27.5 |
| 3 | 1.850 | 24.4 |
| 4 | 2.075 | 23.6 |
| 5 | 2.329 | 20.1 |
| 6 | 2.613 | 18.7 |
| 7 | 2.932 | 16.7 |

2 DATA

The observations were performed in the periods 2011 April 21–24 and 2012 June 3–11, using ATCA in 6A and 6B configurations at 16-cm. The field was observed using a mosaic of 38 pointings. A detailed description of the observations and data reduction is reported in Paper I. Here we briefly describe the main procedures we followed to produce the final images.

2.1 Data cube generation

Flagging was performed in MIRIAD (Sault, Teuben & Wright 1995), using the task MIRFLAG. Radio frequency interference (RFI) was removed by flagging 30 to 40 per cent of the visibilities. The usable range of frequencies covers 1.350 to 3.100 GHz, corresponding to $\Delta\nu/\nu \approx 0.8$. Calibration was performed using the standard MIRIAD tasks and the standard bandpass calibrator, the radio galaxy 1934-638, which had an assumed flux density of 12.31 Jy at 2.1 GHz (Reynolds 1994). We divided our data in seven sub-bands (see Table 1).

The imaging was performed in MIRIAD. The deconvolution was executed with the MFCLEAN task using the Högbom algorithm (Högbom 1974) on every pointing. The cell size was set to 1.5 arcsec and the restored Gaussian beam was set to 14.0 by 6.5 arcsec. The pointings were then joined into one mosaic using the LINMOS task. Finally we obtained seven maps at seven different contiguous frequency bands. This division is very convenient to calculate the SED of the sources, considering that the 16-cm band is often the turnover region for Galactic sources. An all-band map has been obtained using all the UV data to accomplish the best S/N.

The source extraction was conducted on a region of the all-band map as in Franzen et al. (2011). We chose to restrict our studies to the largest region with a uniform and low rms, by taking into account the size of the primary beam at the highest frequency sub-band, given that at this frequency the primary beam is the smallest (see Paper I). The maps of the local noise were used to identify sources on the basis of their S/N. Local maxima above 5σ were identified as sources. A peak position and flux density value were measured by interpolating between pixels. A centroid position, integrated flux density and source area were also calculated by integrating contiguous pixels down to 2.5σ (as in Waldram & Riley 1993), and sources were identified as overlapping if the integration area contained more than one source. Our source extraction process yielded a sample of 614 point sources.

While the majority of the sources extracted by the Franzen et al. (2011) algorithm are point-like, a few of them are slightly resolved. The most extended one is SCORPIO_169a, with an area of 8.3 times the beam area (Paper I). This source has a linear angular scale of ~ 35 arcsec, while our Largest Angular Scale is ~ 2.3 arcmin, so we

Table 2. Product of the primary beam FWHM and frequency at every SCORPIO sub-band, using the more recent measurements of the primary beam for the new 16-cm CABB receiver (d) and the primary beam model by Wierenga & Kesteven (1992) (d').

| ν (GHz) | d (arcmin GHz) | d' (arcmin GHz) |
|----------------|---------------------|----------------------|
| 1.469 | 48.34 | 47.90 |
| 1.649 | 49.34 | 47.90 |
| 1.850 | 50.40 | 47.90 |
| 2.075 | 51.03 | 49.70 |
| 2.329 | 49.91 | 49.70 |
| 2.613 | 49.54 | 49.70 |
| 2.932 | 48.31 | 49.73 |

are sensitive to scales significantly larger than our largest compact source.

The study of SCORPIO extended sources is beyond the scope of this paper, but we are using the extraction algorithm presented in Riggi et al. (2016) and analysing them in Ingallinera et al. (in preparation).

2.2 Primary beam correction

The mosaicking process requires the correction of the telescope primary beam in order to get accurate brightness in the field. A non-accurate knowledge of the primary beam prevents a good flux density measurement, in particular for the sources far from the centre of the individual pointings. As a consequence, the spectral indices of the sources will be affected non-uniformly in the field.

In the last few years measurements of the primary beam shape for the new 16-cm CABB receivers were made across the entire usable frequency range (1.1 to 3.1 GHz). Although an accurate primary beam model is now available, it had not been implemented in MIRIAD when SCORPIO imaging was performed.

Instead, in Paper I, we used the Gaussian primary beam model made by Wierenga & Kesteven (1992) (see Table 2; Paper I). The product of the full width at half-maximum (FWHM) and the frequency d' , is, according to them, constant for the first three sub-bands, then it changes and remains constant for the last four ones because they used different receivers, where the primary beam is assumed to vary as ν^{-1} in a single receiver. New primary beam measurements show that this is not true¹ (in Table 2, the new ones are reported as d values). The primary beam FWHM tends to decrease less rapidly than ν^{-1} below 2.1 GHz. It can be caused by a deterioration in the focus above 1.4 GHz.

In this paper, we use the new version of MIRIAD, in which LINMOS has the right primary beam correction, so we decided to redo the mosaicking, applying the new LINMOS task to the pointings to obtain the corrected maps. The slight differences in the fluxes between the data in this paper and the data in Paper I is due to this correction.

3 IMAGE ANALYSIS

Each of our 614 sources is characterized by a SED that provides us information on the mechanism of the radio emission (thermal or non-thermal), and, if we find the turn-over frequency, some physical

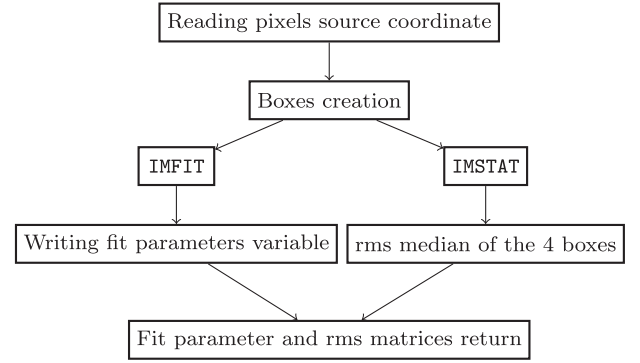


Figure 1. Flowchart for the flux density extraction algorithm.

parameter such as magnetic field or electron density. This information can help us to discriminate between Galactic and extragalactic sources and to characterize the source, giving us hints on its nature.

3.1 Flux density extraction algorithm

In order to automatically estimate the source flux densities at different frequencies, we wrote a PYTHON script that runs in Common Astronomy Software Applications (CASA, McMullin et al. 2007) and uses the IMFIT task (see flowchart in Fig. 1). The script reads the position of each source from the Paper I catalogue and creates a box around it and four other boxes, that are 15 pixels along the x-axis and 10 pixels along the y-axis away from the vertices. The box dimensions are chosen to be three times the beam major and minor axis, respectively (the major axis is oriented along the north–south direction). The script then calls IMFIT in the central box and IMSTAT in the other four ones. IMFIT performs a bidimensional Gaussian fit in the chosen region and returns a PYTHON dictionary *dic* containing all the fit parameters (such as integrated flux density, Gaussian peak position, major and minor axis dimension, etc.). The IMFIT task implements an algorithm that models and subtracts the background from the fit thereby removing the effect of diffuse emission. The script uses IMSTAT to measure the rms and median noise in each of the adjacent boxes, and then calculates the median background noise. At the end, the script returns two matrices with N lines and M columns, where N is the number of sources and M the number of sub-bands. Every element a_{ij} of the first matrix is the fit parameter *dic* of the i^{th} source for the j^{th} sub-band, in a similar way every element b_{ij} of the second one is the local rms of the i^{th} source for the j^{th} sub-band. If the Gaussian fit does not converge for the i^{th} source and the j^{th} sub-band, a_{ij} is simply a 0 (see Fig. 1).

To check the reliability of the algorithm, we applied it to the old all-band map and compared the results with the source flux densities extracted using the Franzen et al. (2011) algorithm. The value of the average ratio $r = S_f/S_c$, where S_f is the flux density measured by the Franzen et al. (2011) algorithm and S_c is the one measured using our algorithm, is $r = 0.97 \pm 0.13$. This value shows that the two methods are consistent.

We also ran our algorithm on the new all-band map, to check for differences with the old map flux densities. We found a ratio of $r = 0.94 \pm 0.44$ between the flux densities measured with our algorithm in the new all-band map and S_c . This higher difference is consistent with the primary beam correction.

¹ http://www.narrabri.atnf.csiro.au/people/ste616/beamshapes/beamshape_16cm.html

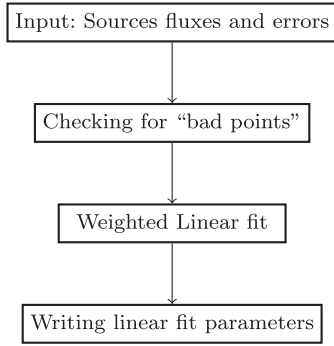


Figure 2. Spectral index extraction algorithm flowchart.

3.2 Spectra construction and control algorithm

For the purposes of this work, we initially assume that all sources have a power-law spectrum. Our algorithm performs a weighted linear fit of $\log S$ as a function of $\log \nu$, where S is the flux density and ν the frequency and the single point error is

$$\Delta S = \sqrt{e_{\text{fit}}^2 + e_{\text{rms}}^2 + e_{\text{cal}}^2}, \quad (1)$$

where e_{fit} is the fit uncertainty, computed by `IMFIT`, e_{rms} is the noise and e_{cal} is the calibration error. To check that the SED has a linear behaviour it fits a second degree polynomial and compares the chi squared (χ^2) of the linear and parabolic fits. In the 30 sources in which the χ^2 of the parabolic fit has a lower value, we ignored the low-frequency turnover by excluding all data at a frequency below that of the peak and redo the fit to get α .

The algorithm performs the following tests for each source in each of the seven sub-bands to flag ‘bad points’ (see Fig. 2 for a simplified flowchart of the procedure):

(i) The flux density S is equal or less than 0. This is the case of non-converged Gaussian fit or Gaussian fit of negative values on the map, for e.g. caused by artefacts:

$$S \leq 0, \quad (2)$$

~36 per cent of the 4298 (614 sources multiplied by the seven sub-bands) flux density measurements were flagged by this criterion.

(ii) The difference between the position of the fitting Gaussian peak and the position given by the source extraction algorithm is greater than half the width of the beam w . This is the case of a Gaussian fit made on a near brighter source with high sidelobes:

$$P_{\text{peak}} - P_{\text{list}} \geq \frac{w}{2}, \quad (3)$$

~6 per cent of the points were flagged by this criterion.

(iii) The difference between the area of the fitting Gaussian at a given frequency and the average of the area of the other frequencies of the same source is greater than 2 times their standard deviation. This is the case of fitting a faint source that can be affected by noise:

$$|A - \bar{A}| \geq 2\sigma, \quad (4)$$

where A and \bar{A} are the area of the fitting Gaussian for a sub-band and the mean area of the other sub-bands, respectively and σ is its standard deviation.

About 1 per cent of the points were flagged by this criterion.

(iv) The flux density is significantly greater than the average \bar{S} .

$$S > n\bar{S}. \quad (5)$$

We tested several values of n between 1 and 2 over ~50 sources. The results have been visually checked and it turns out that $n = 1.8$ was the minimum value above which the fit was unreliable.

About 9 per cent of the points were flagged by this criterion.

If a flux density measurement shows one of those irregularities, it is discarded and does not contribute to the linear fit.

3.3 Grouped components

A sub-sample of 65 of the 614 sources are classified by the source finding algorithm as ‘double sources’ (Franzen et al. 2011). They can be components of larger extended sources or simply be near to each other by projection. To fit these multiple components `IMFIT` needs to read a file that contains a list of estimates of the position of the Gaussians peaks, their dimensions, orientation and brightness. We built a script that automatically writes this file for each group of components, keeping the Gaussian dimensions and orientation fixed. Then the flux extractor algorithm works in the same way it works for the single sources. The algorithm treats each component of a double source as a real component of a larger source, adding the integrated flux densities and using the result as the flux density of the complex source.

3.4 MGPS-2 matches

To check our spectral indices we considered the 43 SCORPIO sources that match with the second epoch Molonglo Galactic Plane Survey (MGPS-2, Murphy et al. 2007) sources (Paper I). MGPS-2 is a survey of the GP carried out with the Molonglo Observatory Synthesis Telescope (MOST) at a frequency of 843 MHz and with a restoring beam of $45 \text{ arcsec} \times 45 \text{ arcsec} \csc \delta$, where δ is the declination so it is suited to expand the SCORPIO sources’ SED and check for spectral index errors. If we exclude the double sources and those embedded in diffuse emission (with confusion because of the large MGPS-2 beam), there are only 16 sources left. We calculated their spectral index, including the 843-MHz MGPS-2 point. In Fig. 3 we show the spectra of the 16 sources. Except for SCORPIO1_178 and SCORPIO1_288 the SEDs are consistent with the 843 MHz flux density. In the case of SCORPIO1_178 the discrepancy is due to the fact that the MGPS-2 centroid is not centred on the SCORPIO source and its beam encompasses another brighter source leading to a higher flux density, while in the case of SCORPIO1_288 it can be due to a turn-over effect. In Fig. 4 we show the effect of the large beam size of MGPS-2 in the case of an isolated source (left-hand panel) and SCORPIO1_178 (right-hand panel).

4 SPECTRAL INDICES ANALYSIS

Since the sub-band maps have a lower S/N than the all-band map, some of the faint sources are detected in less than two sub-bands and hence no spectral index can be derived. In the end we extracted the SED for 510 sources of the 614 detected, of which 74 have no rejected frequency points. However not all of the 510 spectral indices are accurate due to the errors on the flux density extraction in the sub-bands. To define the criteria that only select sources have an accurate spectral index, we ran a simulation. We used the `UVGEN` task in `MIRIAD` (Sault et al. 1995) to simulate 10 ATCA fields with a total of 301 point sources with $\alpha = 0$, and a flux density almost uniformly distributed between 500 μJy and 200 mJy at the same frequency and with the same bandwidth of SCORPIO.

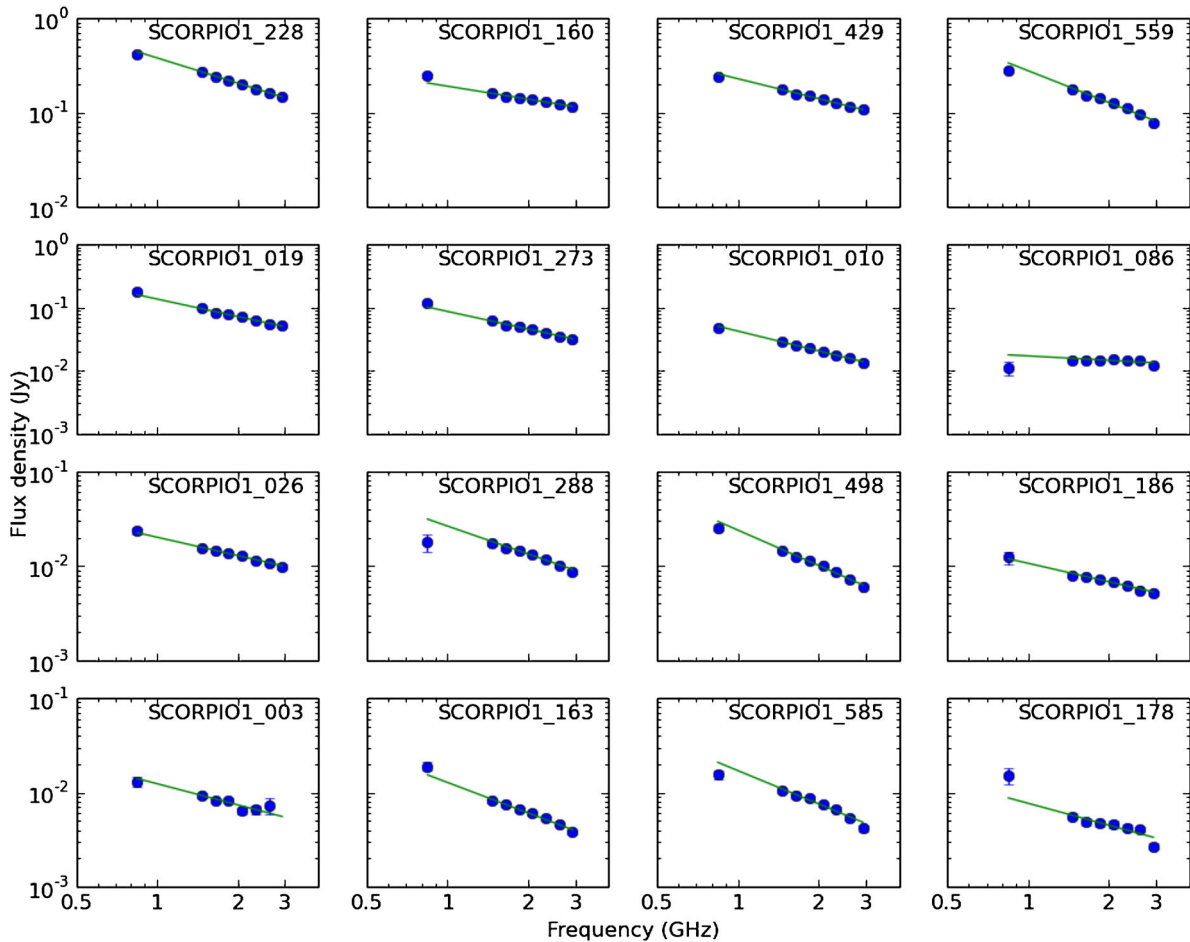


Figure 3. Spectra of the 16 sources selected among the 43 sources that match with the MGPS-2.

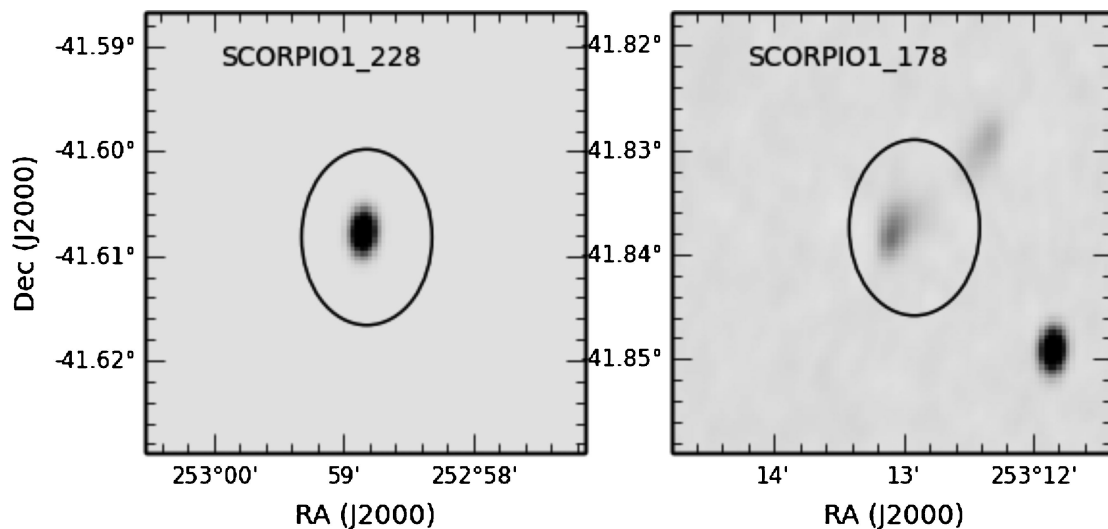


Figure 4. Example of SCORPIO sources with a MGPS-2 match. The Molonglo beam is reported as the black ellipse to stress the effects of a large beam size in flux density measurements.

We mapped the whole field and derived the spectral index by following the same method used for the SCORPIO data. As shown in Fig. 5, simulated sources with a S/N smaller than ~ 40 can have a spectral index error greater than 0.5. Therefore in the following analysis we will only refer to the sources brighter than 40σ .

We also tried to divide the band in three sub-bands instead of seven. We found out that the standard deviation of the spectral indices of sources with a S/N greater than 40 and smaller than 100 is 0.2 in the case of three sub-bands, 0.03 in the case of seven sub-bands. This is due to the smaller number of points leading to

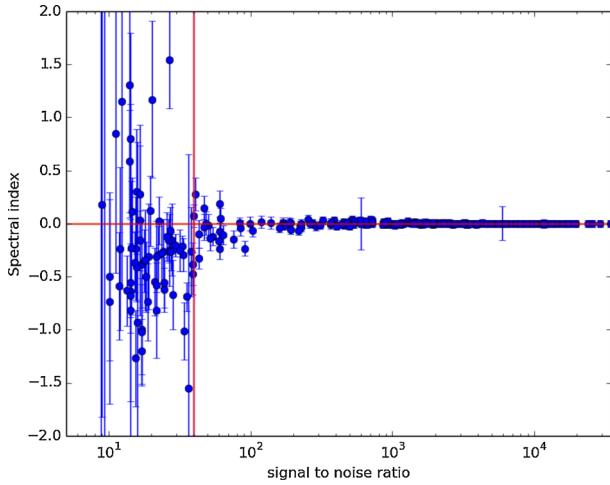


Figure 5. Spectral indices against S/N in the simulation. All the sources have been built to have a flat SED, but the spectral indices are not accurate enough below 40σ (indicated by the red vertical line).

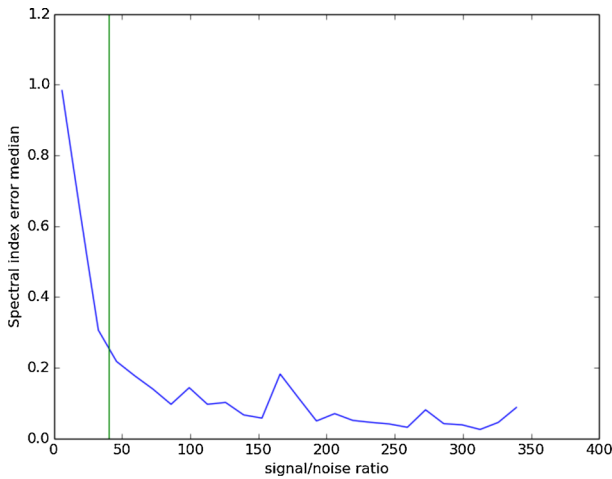


Figure 6. The median on the spectral indices error against S/N in SCORPIO. The green vertical line indicates the 40σ limit.

a worse linear fit. In the three sub-bands case we should choose a 100σ threshold. Therefore we did not use the three sub-bands.

This limiting S/N is also illustrated by Fig. 6. The median of the spectral index errors depends on the S/N as an hyperbola and beyond $\sim 40\sigma$ it decreases rapidly. At 40σ the median on the spectral index errors is ~ 0.2 , thus it is assumed as the minimal value necessary for a good separation. Our threshold of 40σ agrees with results by Rau, Bhatnagar & Owen (2016) that suggest a value between 16σ and 100σ as the limit of reliability for spectral indices. Finally, considering that in the pilot field the rms is $\sim 25\text{--}30 \mu\text{Jy beam}^{-1}$ (Paper I), we adopted a lower limit of 1 mJy for the brightness of the sources. Imposing this limit fixes the number of our sources to 306.

In the Paper I catalogue we list only 260 sources with flux densities greater than 1 mJy. This discrepancy has two causes. The first is the difference in the flux densities due to the updated primary beam model. The second is the different method used, in the catalogue and in this paper, to extract the flux density. To have a consistent way to extract it in the SCORPIO–ATLAS comparison, we decided to use, as our measured flux densities, the value of the linear fit of the SEDs at 2.1 GHz. As noticed in Section 3.1, the differences in methods and primary beam correction can lead to a slightly higher

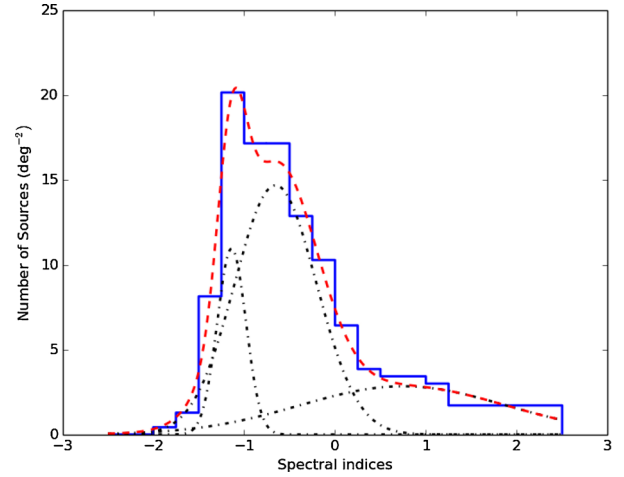


Figure 7. Spectral index distribution of the SCORPIO survey using a 0.25 wide bin in blue straight line, modelled using three Gaussians (red dashed line). In black dash-point line the three Gaussian components of the final model.

Table 3. Criteria to exclude models based on their higher BIC in respect with other models as in Kass & Raftery (1995).

| ΔBIC | Evidence against higher BIC |
|-----------------------------|-----------------------------|
| $0 < \Delta\text{BIC} < 2$ | Marginal |
| $2 < \Delta\text{BIC} < 6$ | Positive |
| $6 < \Delta\text{BIC} < 10$ | Strong |
| $\Delta\text{BIC} > 10$ | Very strong |

flux density in our measures compared to the Franzen et al. (2011) ones.

4.1 Results

In Fig. 7 we show the source density distribution of all the 306 sources brighter than 1 mJy as a function of their spectral indices. The distribution is asymmetric, with a main peak at $\alpha \sim -0.9$ and a long tail at $\alpha \gtrsim 0$.

We modelled the spectral index distribution as the sum of several populations, both Galactic and extragalactic, and fit it with a variable number of Gaussians. We used the Bayesian Information Criterion (BIC; Schwarz 1978) to select the best-fitting model, which will have the smallest BIC. We used the BIC formula as defined in Wit et al. (2012)

$$\text{BIC} = -2 \cdot \ln \hat{L} + k \cdot \ln n, \quad (6)$$

where \hat{L} is the maximized value of the likelihood function, n is the number of data points and k is the number of free parameters of the model. We used the criteria described in Kass & Raftery (1995) to exclude models (see Table 3). The BIC parameter penalises fits with a larger number of parameters so we can be confident in choosing models with a high number of parameters. Table 4 shows that a model consisting of three Gaussians (see Fig. 7) has a better BIC than the other considered models, such as combinations of any number of Gaussians from one to five and a skewed Gaussian. The latter is a function similar to an asymmetric Gaussian due to a ‘skew’ parameter defined as $f(t) = A\phi(t)\Phi(kt)$, where A is the amplitude parameter, k is the skew parameter, $\phi(t)$ is the normal distribution and $\Phi(t) = \int_{-\infty}^t \phi(x) dx$ is the cumulative distribution function.

Table 4. BIC for every model considered for the SCORPIO survey. Every Gaussian needs three parameters, while the skewed Gaussians need four.

| Model | BIC |
|--|------|
| Gaussian | 146 |
| Two Gaussians | 96.7 |
| Three Gaussians | 71.7 |
| Four Gaussians | 80.7 |
| Five Gaussians | 83.3 |
| One Gaussian plus one skewed Gaussian | 159 |
| Two Gaussians plus one skewed Gaussian | 112 |

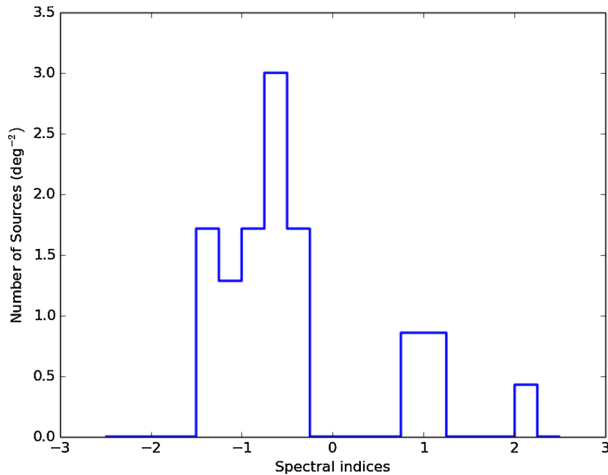


Figure 8. Spectral index distribution of the SCORPIO group sources using a 0.25 wide bin.

As reported in Section 3.3 we detected 65 group of components that we treated as components of 31 sources. With the same method used for single sources, we checked if they have a different spectral index distribution with respect to the former (Fig. 8). We have a total of 11.5 group of components per square degree with a flux density greater than 1 mJy. Among them, $\sim 9.5 \text{ deg}^{-2}$ have a negative spectral index, 1.6 deg^{-2} have a spectral index around 1 and 0.4 deg^{-2} around 2. These numbers are compatible with the all-source results but our sample is too small to make a strong statement.

In the following we compare the statistical results of the spectral index analysis on the SCORPIO data with those of ATLAS, a survey performed at high Galactic latitude. ATLAS can be used as a template for the extragalactic population of SCORPIO, as the Galaxy is optically thin at radio wavelengths and the cosmological distribution of galaxies is isotropic at these scales.

4.2 Comparison with ATLAS

ATLAS (Norris et al. 2006) is a survey conducted on two fields, the *Chandra* Deep Field-South (CDF-S) and the European Large Area ISO Survey-South 1 (ELAIS-S1) at 1.4 GHz with ATCA. The third ATLAS data release (Franzen et al. 2015) includes observations taken using the CABB receiver with a 500 MHz bandwidth covering 1.3–1.8 GHz. Most importantly, the fields are far away from the GP, thus we expect to find only an extragalactic population.

In Franzen et al. (2015), the ATLAS DR3 CABB data were divided into two sub-bands and two separate mosaics of each field were created, one using the lower sub-band centred at 1.40 GHz and the other one using the higher sub-band centred at 1.71 GHz. For

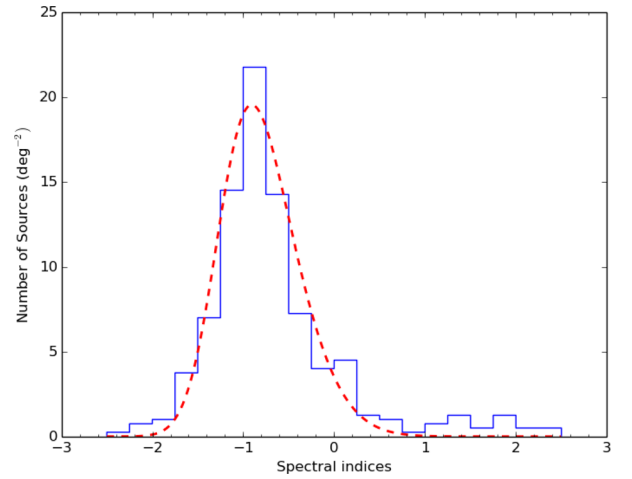


Figure 9. Number of sources per square degree in a 0.25 wide bin histogram for the ATLAS survey modelled as a skewed Gaussian (red dashed line).

consistency with SCORPIO, we extracted the spectral indices of the ATLAS sources running the algorithm described in Section 3 on the two mosaic maps. Then we normalized the number of sources per square degree for both the surveys and only considered the sources brighter than 1 mJy at 2.1 GHz (extrapolated using the spectral index). As shown in Fig. 9, the synchrotron peak is the predominant component in ATLAS. There is a minor peak around 0, which can also be explained by extragalactic sources, the Gigahertz Peaked Spectrum (GPS) sources (O’Dea, Baum & Stanghellini 1991), that can have the turnover frequency between 1.4 and 1.71 GHz. We tried to fit this distribution with different models, as in Section 4.1, and we found that the Gaussian model has a BIC value of 85.4, higher than the skewed Gaussian one (81.7). Consequently the latter is the one we selected (see Fig. 9 for the model). Assuming that virtually all the ATLAS sources are extragalactic, we used the ATLAS model as a template for the extragalactic population of SCORPIO.

We found 20 per cent more sources per square degree in SCORPIO than in ATLAS (120 versus 100) and we found that a model using the ATLAS skewed Gaussian plus another skewed Gaussian gives a BIC of 76.8, better than almost all the models shown in Table 4 (with the exception of the one that uses three Gaussians). Nevertheless the model including the ATLAS extragalactic template has, as anticipated before, a physical motivation making it the more reliable one. Fig. 10 shows that the ATLAS Gaussian fits quite well the peak at $\alpha = -0.9$ and, most importantly, another skewed Gaussian fits the other peaks. Maybe larger statistics could help us in concluding that they are part of more than one Galactic population, as expected, but, with our data, we can only say that they are part of a generic Galactic population.

In Fig. 11 the difference between the SCORPIO and the ATLAS population per spectral index bin is shown. The error e_i associated with the i^{th} bin is

$$e_i = \sqrt{N_{Si}} + \sqrt{N_{Ai}}, \quad (7)$$

where N_{Si} and N_{Ai} are, respectively, the number of SCORPIO sources and ATLAS sources in the i^{th} bin. We can recognize three regions of the spectral index α , which are given below.

(i) $-2.5 < \alpha < -0.5$: the number of excess sources is not significantly different from zero, implying that all SCORPIO sources in this spectral index range may be extragalactic. In this spectral

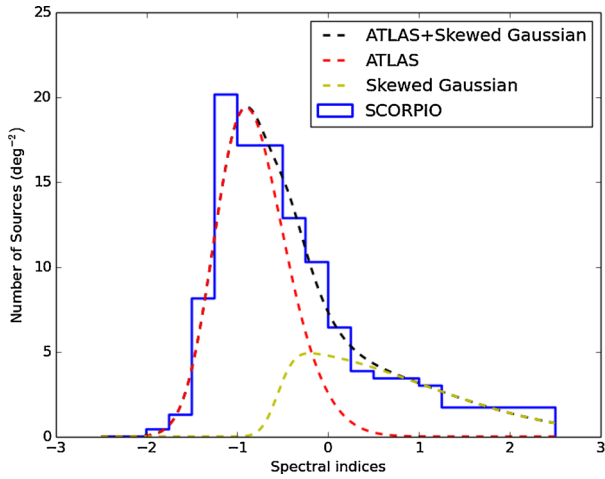


Figure 10. SCORPIO spectral index distribution modelled as a sum of the ATLAS model and a skewed Gaussian.

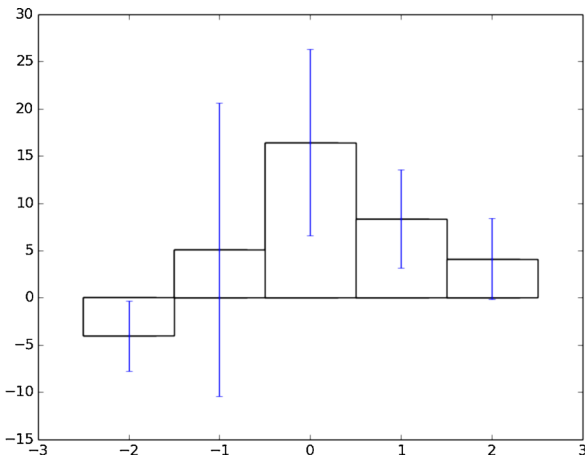


Figure 11. The difference between the SCORPIO and the ATLAS spectral indices histogram.

index range there can also be unresolved or almost unresolved young supernovae remnants (YSNRs). To know what size we should expect from different age SNRs, we consider the size of the SN1987A remnant. The SN1987A remnant is located in the Large Magellanic Cloud at ~ 51.4 kpc from the Sun. It expanded from 0.21 to 0.39 pc between 1995 and 2010 with an almost linear expansion rate of about 0.01 pc/yr $^{-1}$ (Chiad 2012). Therefore a 100 yr old similar supernovae would have a linear dimension of ~ 1.1 pc and an angular dimension of ~ 11 arcsec at a 20 kpc distance from the Sun. Given the 14.0 arcsec \times 6.5 arcsec resolution of SCORPIO, we may not be able to resolve it.

We now estimate how many YSNRs are expected. We assume that in the last 1000 yr there were seven supernovae in our Galaxy within 5 kpc from the Earth in the GP. Given the position of the Sun in the GP, from the point of view of the Galactic Centre this corresponds to a ~ 50 deg circular sector. Assuming that the distribution is uniform in the GP, we estimate about one supernova (SN) every ~ 20 yr in our Galaxy. Considering that ~ 66 per cent of the OB stars lies in the direction of the Galactic Center (Robin et al. 2003) in a 120 deg 2 area, we can assume a supernova (SN) every ~ 40 square degree younger than 100 yr in the direction of the SCORPIO field.

Moreover we have to consider that SN discovery is typically performed at optical wavelength. Thus, even if the remnants of these

Table 5. SCORPIO ID, flux density, spectral index and catalogue tag on Anderson et al. (2014) of the seven cross-matched sources. The catalogue tags stand for radio Quiet and for Candidate, indicating, respectively, a H II region detected in the infrared but without a counterpart in the radio catalogue considered in Anderson et al. (2014), the SGPS and a bubble candidate to be a H II region.

| SCORPIO ID | Flux density (mJy) | Spectral index | Catalogue tag |
|---------------|--------------------|------------------|---------------|
| SCORPIO1_123b | 40.5 ± 1.3 | 1.1 ± 0.1 | Q |
| SCORPIO1_136 | 0.7 ± 0.1 | -0.7 ± 0.6 | Q |
| SCORPIO1_169 | 24.5 ± 1.3 | -0.34 ± 0.13 | Q |
| SCORPIO1_218 | 182 ± 5.5 | -0.09 ± 0.05 | C |
| SCORPIO1_219 | 9.2 ± 0.7 | -1.0 ± 0.4 | C |
| SCORPIO1_459 | 4.7 ± 0.2 | -1.5 ± 0.4 | Q |
| SCORPIO1_483 | 21.5 ± 0.7 | -0.12 ± 0.04 | C |

supernovae are bright in radio (at peak we expect almost 1000 Jy in L band at 20 kpc, Bufano et al. 2014), they could be very faint in the optical due to the GP absorption: considering that $m = A + M + \mu$, assuming a SN with a peak absolute magnitude of $M = -17$ mag, we would have an apparent magnitude as high as $m \sim 26$ mag, having in the direction of SCORPIO an extinction $A = 26.7$ (Schlafly & Finkbeiner 2011) and a distance module, corresponding to 20 kpc, of $\mu \sim 16.5$. This means that it is possible for a supernova in our Galaxy to happen undiscovered, but to leave a remnant bright enough to be detected at radio wavelength. Even with this premise, considering a density of 0.025 supernova younger than 100 yr per square degree in the direction of SCORPIO, we would not be able to detect more than one YSNR in the relatively small SCORPIO field of view (~ 5 square degree for the whole field), confirming the extragalactic origin of the sources in this α range.

(ii) $-0.5 < \alpha < 0.5$: the excess here is $\sim 16 \pm 10$ per square degree. These are typical spectral indices of Galactic source that show thermal emission in an optically thin environment, e.g. H II regions, Planetary Nebulae (PNe) and Luminous Blue Variable (LBV) (e.g. Cerrigone et al. 2008; Agliozzo et al. 2012, 2014; Ingallinera et al. 2014; Paladini et al. 2015).

(iii) $0.5 < \alpha < 2.5$: there is a significant excess of about 12 ± 8 sources per square degree with a spectral index in this range, suggesting stellar winds, interacting stellar winds and thermal emission in an optically thick environment, e.g. Wolf-Rayet, LBV, OB stars and compact and ultracompact H II regions (e.g. Leto et al. 2009; Scuderi et al. 1996; Umana et al. 2005).

Obviously we expect some overlap between spectral index regions. Furthermore, the different frequency range of the ATLAS survey makes it likely to have a surplus of extragalactic flat sources due to possible turnover positions.

Note that the difference between the total sources surplus in SCORPIO (20 per cent) and the one in the interval $-2.5 < \alpha < 2.5$ is due to the smaller bandwidth in ATLAS, resulting in a larger number of spurious spectral indices.

We cross-matched our point sources with the Anderson et al. (2014) H II regions catalogue, extracted from infrared colours, and we found seven matching sources, reported in Table 5; four of them do not have a counterpart in the Southern Galactic Plane Survey (SGPS, McClure-Griffiths et al. 2005), suggesting that many of the ‘radio quiet’ H II regions emit in radio being actually ‘radio weak’; the remaining three are H II region candidates. The spectral indices of SCORPIO1_123b, SCORPIO1_169, SCORPIO1_218 and SCORPIO1_483 are consistent with thermal free-free emission in an optically thick medium in the first case, and in an optically thin

medium in the latter ones. The other three sources present a much less reliable spectral index because they are resolved and embedded in diffuse emission.

5 SUMMARY AND CONCLUSIONS

We have described:

(i) An algorithm based on the *CASA* task *IMFIT* to automatically extract the flux density of a large number of point-like sources by fitting Gaussians and to calculate their spectral indices.

(ii) The application of this algorithm to the SCORPIO data to derive the SCORPIO sources spectral index distribution, along with the discussion about the different source populations that contribute to the distribution.

(iii) The comparison between SCORPIO and an extragalactic survey, ATLAS.

We have found that, with a bandpass like the SCORPIO one, the S/N for the object in the field has to be at least 40 in order to rely on the spectral index α with an uncertainty $\lesssim 0.2$.

We have found that, in the Galactic Plane, the source count with flux densities greater than 1 mJy is about 20 per cent higher than the source count at high Galactic latitude. We found 16 Galactic sources per square degree with a spectral index of about $\alpha = 0$, suggesting optically thin thermal emission such as H II regions and planetary nebulae, while the remaining 12 sources per square degree present a spectral index $0.5 < \alpha < 2.5$, pointing to an optically thick thermal emission such as stars and compact H II regions.

These results are very important for planning forthcoming radio surveys. This work will be eventually extended to the whole SCORPIO field. Considering that the survey will be $\sim 5 \text{ deg}^2$ and Galactic source density above 1 mJy of 20 deg^{-2} as reported in this paper, we expect about 100 Galactic sources brighter than 1 mJy. Thanks to the larger sample that we will consider, we will be able to confirm or not the results of this paper, validating their applications to the future surveys.

REFERENCES

- Aglizzo C., Umama G., Trigilio C., Buemi C., Leto P., Ingallinera A., Franzen T., Noriega-Crespo A., 2012, *MNRAS*, 426, 181
- Aglizzo C., Noriega-Crespo A., Umama G., Flagey N., Buemi C., Ingallinera A., Trigilio C., Leto P., 2014, *MNRAS*, 440, 1391
- AMI Consortium: Franzen et al., 2011, *MNRAS*, 415, 2699
- Anderson L. D., Bania T. M., Balser D. S., Cunningham V., Wenger T. V., Johnstone B. M., Armentrout W. P., 2014, *ApJS*, 212, 1
- Bufano F. et al., 2014, *MNRAS*, 439, 1807
- Burke B. F., Graham-Smith F., 2009a, *An Introduction to Radio Astronomy*. Cambridge Univ. Press, New York, p. 131
- Burke B. F., Graham-Smith F., 2009b, *An Introduction to Radio Astronomy*. Cambridge Univ. Press, New York, p. 234
- Carilli C. L., Rawlings S., 2004, *New Astron. Rev.*, 48, 979
- Cerrigone L., Umama G., Trigilio C., Leto P., Buemi C. S., Hora J. L., 2008, *MNRAS*, 390, 363
- Chiad B. T., 2012, *Int. J. Astron. Astrophys.*, 2, 199
- Franzen T. M. O. et al., 2015, *MNRAS*, 453, 4020
- Högbom J. A., 1974, *A&AS*, 15, 417
- Ingallinera A. et al., 2014, *MNRAS*, 437, 3626
- Jarvis M., 2015, *Proc. Sci.*, The Many Facets of Extragalactic Radio Surveys: towards new scientific challenges, SISSA, Trieste, PoS#7
- Johnston S. et al., 2008, *Exp. Astron.*, 22, 151
- Jonas J. L., 2009, *Proc. IEEE*, 97, 1522
- Kass R. E., Raftery A. E., 1995, *J. Am. Stat. Assoc.* 90 (430), 773

- Leto P., Umama G., Trigilio C., Buemi C. S., Dolei S., Manzitto P., Cerrigone L., Siringo C., 2009, *A&A*, 507, 1467
- McClure-Griffiths N. M., Dickey J. M., Gaensler B. M., Green A. J., Haverkorn M., Strasser S., 2005, *ApJS*, 158, 178
- McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, *ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI*. Astron. Soc. Pac., San Francisco, p. 127
- Murphy T., Mauch T., Green A., Hunstead R. W., Piestrzynska B., Kels A. P., Sztajer P., 2007, *MNRAS*, 382, 382
- Norris R. P. et al., 2006, *AJ*, 132, 2409N
- Norris R. P. et al., 2011, *PASA*, 28, 215
- O’Dea C. P., Baum S. A., Stanghellini C., 1991, *ApJ*, 380, 66
- Paladini R., Ingallinera A., Aglizzo C., Tibbs C. T., Noriega-Crespo A., Umama G., Dickinson C., Trigilio C., 2015, *ApJ*, 813, 24
- Panagia N., Felli M., 1975, *A&A*, 39, 1
- Rau U., Bhatnagar S., Owen F. N., 2016, *AJ*, 152, 124
- Reynolds J. E., 1994, Technical Report No. 39.3/040, A Revised Flux Scale for the AT Compact Array. Australia Telescope National Facility
- Riggi S. et al., 2016, *MNRAS*, 460, 1486
- Robin A. C., Reylé C., Derrière S., Picaud S., 2003, *A&A*, 409, 523
- Sault R. J., Teuben P. J., Wright M. C. H., 1995, in Shaw R., Payne H. E., Hayes J. J. E., eds, *ASP Conf. Ser. Vol. 77, Astronomical Data Analysis Software and Systems IV*. Astron. Soc. Pac. p. 433
- Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, 737, 103
- Schwarz G. E., 1978, *Ann. Stat.*, 461, 464
- Scuderi S., Panagia N., Stanghellini C., Umama G., Trigilio C., 1996, in A. R. Taylor A. R., Paredes J. M., eds, *ASP Conf. Ser. Vol. 93, Radio Emission from the Stars and the Sun*. Astron. Soc. Pac., San Francisco, p. 53
- Umama G., Buemi C. S., Trigilio C., Leto P., 2005, *A&A*, 437, L1
- Umama G. et al., 2015, *MNRAS*, 454, 902
- Waldram E. M., Riley J. M., 1993, *MNRAS*, 265, 853
- Wieringa M. H., Kesteven M. J., 1992, *Measurements of the ATCA Primary Beam*, AT Memo 39.3/024
- Wit E., Van Den Heuvel E., Jan-Willem R., 2012, *Stat. Neerlandica*, 217, 236
- Wright A. E., Barlow M. J., 1975, *MNRAS*, 170, 41

SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](https://www.mnras.org/) online.

Table A1

Please note: Oxford University Press is 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.

APPENDIX A: CATALOGUE

We created a new catalogue of the spectral indices of the 510 sources listed in [Paper I](#) for which we could calculate it (see [Table A1](#) for the first 10 entries). The columns are

- (i) ID of the source in the SCORPIO catalogue;
- (ii) Galactic longitude;
- (iii) Galactic latitude;
- (iv) flux density at 2.1 GHz;
- (v) error on the flux density (see details on how we calculated that on [Paper I](#));
- (vi) spectral index; and
- (vii) error on the spectral index.

Table A1. First 10 entries of the spectral indices catalogue. The first column is the ID of the source in the SCORPIO catalogue (Paper I), the second and the third columns are respectively the Galactic longitude and latitude of the centroid of the source, the fourth column is the flux density at 2.1 GHz, the fifth is the error on the flux density and the sixth and the seventh columns are respectively the spectral index calculated as explained in this paper and its error.

| SCORPIO ID | l (deg) | b (deg) | S (mJy) | ΔS (mJy) | α | $\Delta\alpha$ |
|--------------|------------|------------|------------|---------------------|----------|----------------|
| SCORPIO_001 | 343.0025 | 1.7604 | 33.33 | 1.03 | -0.39 | 0.07 |
| SCORPIO_002 | 343.0051 | 0.2234 | 0.49 | 0.09 | 6.50 | 3.14 |
| SCORPIO_003 | 343.0134 | 0.6086 | 7.74 | 0.29 | -0.74 | 0.18 |
| SCORPIO_004 | 343.0138 | 1.7208 | 1.60 | 0.17 | -2.80 | 1.45 |
| SCORPIO_006 | 343.0152 | 0.1166 | 1.27 | 0.16 | 0.02 | 0.85 |
| SCORPIO_007 | 343.0157 | -0.1830 | 1.51 | 0.12 | -0.32 | 0.39 |
| SCORPIO_005b | 343.0186 | 1.1577 | 106.56 | 3.20 | -0.40 | 0.10 |
| SCORPIO_010 | 343.0203 | 0.3109 | 20.97 | 0.64 | -1.04 | 0.03 |
| SCORPIO_011 | 343.0216 | 0.7407 | 1.12 | 0.10 | -1.50 | 0.73 |
| SCORPIO_012 | 343.0277 | 0.8702 | 1.17 | 0.09 | 0.65 | 0.35 |

This paper has been typeset from a \LaTeX file prepared by the author.