



**HAL**  
open science

## **GLOSTAR: Radio Source Catalog I. $28^\circ < \ell < 36^\circ$ and $|b| < 1^\circ$**

S.-N. Medina, J. Urquhart, S. Dzib, A. Brunthaler, B. Cotton, K. M. Menten, F. Wyrowski, H. Beuther, S. Billington, C. Carrasco-Gonzalez, et al.

► **To cite this version:**

S.-N. Medina, J. Urquhart, S. Dzib, A. Brunthaler, B. Cotton, et al.. GLOSTAR: Radio Source Catalog I.  $28^\circ < \ell < 36^\circ$  and  $|b| < 1^\circ$ . *Astronomy and Astrophysics - A&A*, 2019, 627, pp.A175. 10.1051/0004-6361/201935249 . hal-03054104

**HAL Id: hal-03054104**

**<https://hal.science/hal-03054104v1>**

Submitted on 3 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# GLOSTAR: Radio Source Catalog I. $28^\circ < \ell < 36^\circ$ and $|b| < 1^\circ$ ★

S.-N. X. Medina<sup>1</sup>, J. S. Urquhart<sup>2</sup>, S. A. Dzib<sup>1</sup>, A. Brunthaler<sup>1</sup>, B. Cotton<sup>3</sup>, K. M. Menten<sup>1</sup>, F. Wyrowski<sup>1</sup>,  
H. Beuther<sup>4</sup>, S. J. Billington<sup>2</sup>, C. Carrasco-Gonzalez<sup>5</sup>, T. Csengeri<sup>1</sup>, Y. Gong<sup>1</sup>, P. Hofner<sup>6,7</sup>, H. Nguyen<sup>1</sup>,  
G. N. Ortiz-León<sup>1</sup>, J. Ott<sup>7</sup>, J. D. Pandian<sup>8</sup>, N. Roy<sup>9</sup>, E. Sarkar<sup>10</sup>, Y. Wang<sup>4</sup>, and B. Winkel<sup>1</sup>

<sup>1</sup> Max-Planck-Institut für Radioastronomie (MPIfR), Auf dem Hügel 69, 53121 Bonn, Germany  
e-mail: [smedina@mpi-fr-bonn.mpg.de](mailto:smedina@mpi-fr-bonn.mpg.de)

<sup>2</sup> School of Physical Sciences, University of Kent, Ingram Building, Canterbury, Kent CT2 7NH, UK

<sup>3</sup> National Radio Astronomy Observatory (NRAO), 520 Edgemont Road, Charlottesville, VA 22903, USA

<sup>4</sup> Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany

<sup>5</sup> Instituto de Radioastronomía y Astrofísica (IRyA), Universidad Nacional Autónoma de México Morelia, 58089, Mexico

<sup>6</sup> Physics Department, New Mexico Tech, 801 Leroy Place, Socorro, NM 87801, USA

<sup>7</sup> National Radio Astronomy Observatory, 1003 Lopezville Road, PO Box O, Socorro, NM 87801, USA

<sup>8</sup> Department of Earth and Space Science, Indian Institute for Space Science and Technology, Trivandrum 695547, India

<sup>9</sup> Department of Physics, Indian Institute of Science, Bangalore 560012, India

<sup>10</sup> Sodankylä Geophysical Observatory, University of Oulu, 99600 Sodankylä, Finland

Received 11 February 2019 / Accepted 20 May 2019

## ABSTRACT

**Context.** Radio continuum surveys of the Galactic plane are an excellent way to identify different source populations such as planetary nebulae, H II regions, and radio stars and characterize their statistical properties. The Global View of Star Formation in the Milky Way (GLOSTAR) survey will study the star formation in the Galactic plane between  $-2^\circ < \ell < 85^\circ$  and  $|b| < 1^\circ$  with unprecedented sensitivity in both flux density ( $\sim 40 \mu\text{Jy beam}^{-1}$ ) and range of angular scales ( $\sim 1''.5$  to the largest radio structures in the Galaxy).

**Aims.** In this paper we present the first results obtained from a radio continuum map of a 16-square-degree-sized region of the Galactic plane centered on  $\ell = 32^\circ$  and  $b = 0^\circ$  ( $28^\circ < \ell < 36^\circ$  and  $|b| < 1^\circ$ ). This map has a resolution of  $18''$  and a sensitivity of  $\sim 60\text{--}150 \mu\text{Jy beam}^{-1}$ .

**Methods.** We present data acquired in 40 h of observations with the VLA in D-configuration. Two 1 GHz wide sub-bands were observed simultaneously and they were centered at 4.7 and 6.9 GHz. These data were calibrated and imaged using the *Obit* software package. The source extraction was performed using the BLOCAT software package and verified through a combination of visual inspection and cross-matching with other radio and mid-infrared surveys.

**Results.** The final catalog consists of 1575 discrete radio sources and 27 large scale structures (including W43 and W44). By cross-matching with other catalogs and calculating the spectral indices ( $S(\nu) \propto \nu^\alpha$ ), we have classified 231 continuum sources as H II regions, 37 as ionization fronts, and 46 as planetary nebulae. The longitude and latitude distribution and negative spectral indices are all consistent with the vast majority of the unclassified sources being extragalactic background sources.

**Conclusions.** We present a catalog of 1575 radio continuum sources and discuss their physical properties, emission nature, and relation to previously reported data. These first GLOSTAR results have increased the number of reliable H II regions in this part of the Galaxy by a factor of four.

**Key words.** surveys – radiation mechanisms: thermal – radio continuum: general – stars: formation – techniques: interferometric – radiation mechanisms: non-thermal

## 1. Introduction

One of the great challenges in modern astronomy is understanding the circumstances of the formation of high-mass stars ( $> 8 M_\odot$ ). Massive stars dominate the energy budget of galaxies, regulate future star formation, and drive their evolution (Kennicutt 2005). Over the past 10–15 years there has been a huge effort to map the Galactic mid-plane at infrared (e.g., GLIMPSE; Churchwell et al. 2009, MIPS GAL; Carey et al. 2009), (sub)millimeter (e.g., HiGAL; Molinari et al. 2010, ATLAS GAL; Schuller et al. 2009, BGPS; Aguirre et al. 2011,

and SEDIGISM; Schuller et al. 2017), and radio wavelengths (e.g., CORNISH; Hoare et al. 2012, MAGPIS; Helfand et al. 2006, THOR; Beuther et al. 2016; Bühr et al. 2015; Wang et al. 2018). These surveys have provided Galaxy-wide and unbiased samples of high-mass star-forming regions (Urquhart et al. 2014a, 2018; Elia et al. 2017) that include all evolutionary stages and, for the first time, allow star formation to be studied in a global context.

The recent upgrade of the *Karl G. Jansky* Very Large Array (VLA) telescope of the NRAO<sup>1</sup> has resulted in a significant increase in frequency coverage and sensitivity (Perley et al.

\* Full version of Table 7 and data from Figs. 7, 9, and 17 are only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](http://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/627/A175>

<sup>1</sup> The National Radio Astronomy Observatory is operated by Associated Universities Inc. under cooperative agreement with the National Science Foundation.

2011). This provides an excellent opportunity to conduct powerful and comprehensive radio-wavelength surveys of both the ionized and the molecular tracers of star formation in the Galactic plane that will complement previous surveys. The GLOSTAR (Global View of Star Formation in the Milky Way) survey<sup>2</sup> is one of these new radio-wavelength surveys. Combined with accurate distances determined by the BeSSeL trigonometric parallax survey (the Bar and Spiral Structure Legacy Survey; Reid et al. 2016), it will provide a more complete view of star formation in the Milky Way. The GLOSTAR-VLA survey (Brunthaler et al., in prep.) uses the wideband (4–8 GHz) C-band receivers of the VLA to conduct an unbiased survey to characterize star-forming regions in the Milky Way. This survey of the Galactic mid-plane allows us to detect tell-tale tracers of early phases of high-mass star formation: compact, ultra-, and hyper-compact H II regions, and 6.7 GHz methanol (CH<sub>3</sub>OH) masers, which trace some of the earliest evolutionary stages in the formation of high-mass stars (Minier et al. 2003) and can be used to pinpoint the positions of very young stellar objects, many of them still deeply embedded in their natal material. The observations also cover emission from the 4.8 GHz formaldehyde (H<sub>2</sub>CO) and multiple radio recombination lines (RRLs) all of which will be presented in future publications. GLOSTAR observations were made with the VLA B- and D-configurations.

In this paper, we present the first source catalog covering the 16 square degrees of the Galactic mid-plane observed with the VLA in its most compact (D) configuration ( $28^\circ < \ell < 36^\circ$  and  $|b| < 1^\circ$ ). We provide a description of the observational and data reduction strategies used in Sect. 2. In Sect. 3 we describe the source extraction and steps used to identify reliable sources and dubious sources that were removed from the final catalog. In the Sect. 4 we compare the detected sources with other Galactic surveys to look for counterparts at other wavelengths, and therefore to verify our catalog. We discuss the physical properties and the resulting statistics of the sources in Sects. 5 and 6 we use the multiwavelength data available to determine their nature (e.g., H II regions, planetary nebulae, extragalactic). In Sect. 7 we focus on the properties and distribution of a sample of H II regions identified from our source classification. Finally, in Sect. 8 we give a summary of our results and highlight our main findings.

## 2. GLOSTAR-VLA survey observations

### 2.1. Observation strategy

The Karl G. Jansky VLA of the national radio astronomy observatory (NRAO) was used in D-configuration to observe the C-band (4–8 GHz) continuum emission. The correlator setup consisted of two 1 GHz wide sub-bands, centered at 4.7 and 6.9 GHz. The primary beam at the full width of the half maximum (FWHMs) are  $\sim 9'$  and  $\sim 6'$  at 4.7 and 6.9 GHz, respectively. Each sub-band was divided into eight spectral windows of 128 MHz, and each spectral window observed 64 channels of 2 MHz. Higher spectral resolution windows were used to cover the most prominent methanol maser emission line at 6.7 GHz, formaldehyde absorption at 4.8 GHz, and seven radio recombination lines. The results of these line observations will be reported in forthcoming papers. The chosen setup avoids strong and persistent radio frequency interference (RFI) seen at 4.1 and 6.3 GHz, and allows a rough estimation of spectral indices.

<sup>2</sup> This VLA survey was part of the proposal for ERC Advanced Investigator Grant (247078) GLOSTAR that partially funded early stages of the research described here.

**Table 1.** Observation epochs.

Region ( $\ell$ )	Observation date (yyyy.mm.dd)
28°–29°	2013.04.09
29°–30°	2013.04.06
30°–31°	2013.04.11
31°–32°	2013.04.15
32°–33°	2013.04.16
33°–34°	2013.04.20
34°–35°	2013.04.29
35°–36°	2013.05.02

The region of the GLOSTAR-VLA survey presented in this paper was observed over a total of eight epochs and a total of 40 h of telescope time (project ID 13A-334). The observations were taken during the spring of 2013 (see Table 1 for details). Each epoch consisted of observations of five hours. The phase calibrator was J1804+0101 and the amplitude calibrators 3C 286 and 3C 48 were observed periodically during each observing block. During each epoch, an area of  $2^\circ \times 1^\circ$  was covered by about 630 pointings. The pointings were on a hexagonal grid with a spacing of  $\theta_{\text{hex}} = 3'.25$ . This corresponds to  $\theta_B/2$ , where  $\theta_B = 6'.5$  is the FWHM width of the primary beam at the central frequency of the higher frequency continuum band (6.9 GHz). Each grid position was observed twice for 11 s which, after considering the slewing time, resulted in a total integration time of 15 s per position. The theoretical noise level from these observations is  $\sim 90 \mu\text{Jy beam}^{-1}$  per position and per sub-band. This noise is improved by a factor of two by combining the two observations of each position and both 1 GHz sub-bands (theoretical noise  $\sim 45 \mu\text{Jy beam}^{-1}$ ).

### 2.2. Calibration, data reduction, and imaging of continuum data

The data were calibrated using the *Obit* software packages (Cotton 2008), which was designed for handling radio astronomy data. *Obit* applications can be accessed through a Python interface (ObitTalk). *Obit* also inter-operates with classic astronomical image processing system (AIPS) software (Greisen 2003) and has access to its tasks.

Observations made with the VLA are well standardized, as is the data calibration, and as such calibration pipelines are now available. These implement the data reduction and calibration using amplitude, phase, and bandpass calibrators (for an example consult the common astronomy software applications (CASA) package pipeline webpage<sup>3</sup>). Following the methodology of these pipelines, we have developed a pipeline in *Obit* to calibrate the GLOSTAR-VLA data with some differences. The details of these differences are described in what follows:

1. One of the first steps that most pipelines apply is to flag<sup>4</sup> the first record of every scan to consider the slew of the antennas. As our target scans last only  $\sim 11$  s this flagging covers an important fraction of our science data. Furthermore, the distance between the target pointings is only a few arcminutes so that the slewing time is minimized. Thus, our pipeline

<sup>3</sup> <https://science.nrao.edu/facilities/vla/data-processing/pipeline>

<sup>4</sup> “Flagging” is a colloquial term used for marking faulty data to exclude them from further use.

only performs the flagging of the first record on scans that observe calibrators.

2. The GLOSTAR observations use a mixed setup of continuum and spectral lines, which are not handled well by pipelines. Thus, the next step of our pipeline is to distinguish between the different observational sets, that is, for continuum, radio recombination line, formaldehyde, and methanol data, and then load solely the continuum data.
3. A standard VLA calibration was applied in the *Obit* package, which alternates editing of data affected by instrumental problems or interfering signals with various calibration steps. External calibration includes VLA system temperature ( $T_{\text{sys}}$ ) calibration, group delay, bandpass, amplitude and phase, and polarization calibration. After a Hanning smoothing of the data, an automated search for outliers in the data is performed by a comparison against running medians in time and frequency for each data product. Calibrator data is checked for adequate signal to noise ratio by an rms/mean test. After the initial pass at data editing, the various parallel hand calibration steps are performed applying further editing based on outliers in the calibration solutions. Following the first pass at calibration, a second set of automated editing steps is done on the calibrated data. The calibration tables are then reset, the flagging table kept, and the calibration is repeated. This assures that the calibration is based on fully edited data. After flagging cross hand outliers, the polarization calibration is performed. Per pointing flagging is also included as part of the imaging. An initial, shallow *clean* is done to determine the approximate flux density in the field. Data for that pointing are then clipped at five times the estimated pointing integrated flux density.
4. The *Obit* task MFImage was used for the continuum imaging (see Cotton et al. 2018 for details). First, we imaged the data after a shallow cleaning to identify strong sources that needed to be imaged in outlier fields in the final imaging. This program uses wideband multifrequency synthesis by dividing the observed bandwidth into frequency bins, that is, sections of the sub-band, with a fractional bandwidth of less than 5%. The number of frequency bins for our observation is eight. These are narrow enough such that the effects of variable spectral index and antenna pattern variations are minor when performing a joint spectral deconvolution.
5. The *cleaning* is driven by a weighted average of the residual images from these frequency bins. Each pointing is imaged separately with a  $0.28 \times 0.28$  square-degrees-sized area, centered on the phase center, using a Briggs' weighting with robustness parameter 0 and using its native restoring beam and with a pixel size of  $2''.5$ . Outlier fields are used if a source with an integrated flux density of more than 10 mJy is known from the NRAO VLA Sky Survey (NVSS) or a previous imaging run from our own data out to a distance of  $0.46^\circ$ . Multiscale clean was used with three different sizes: a delta function, and tapers of 20 and 40 pixels ( $50''$  and  $100''$ ), respectively. Cleaning stopped when 5000 iterations were reached or the maximum in the residual image was below a level 0.25 mJy. Phase only, and amplitude plus phase self-calibration were applied for images with sufficiently bright peaks ( $>10$  mJy for phase only and  $>100$  mJy for amplitude plus phase self-calibrations). The shortest baselines of 30 m allowed imaging of scales up to  $4'$ .
6. After imaging the data from each pointing, the images were merged together in *Obit*. Since all pointings had a slightly different native beam size and orientation due to different  $uv$ -coverage, all pointings were smoothed in a first step to

a circular synthesized beam of  $18''$  in diameter. This beam size corresponds roughly to the major axis of the largest beam size of all the pointings. Primary beam corrections were applied during the merging of the images. After combining all the pointings at each frequency bin, a combined image at the reference frequency was generated.

7. The maximum sensitivity can be obtained by a weighted combination of the overlapping observed pointing images onto a set of images covering the surveyed region. All else being equal, weighting by the square of the antenna power pattern in a given direction gives the optimum sensitivity. The mosaic formation process for each image plane is given by the summation over overlapping pointing images:

$$M(x, y) = \frac{\sum_{i=1}^n A_i(x, y) I_i(x, y)}{\sum_{i=1}^n A_i^2(x, y)}$$

where  $A_i(x, y)$  is the antenna gain of pointing  $i$  in the direction  $(x, y)$  and  $I_i(x, y)$  is the pointing  $i$  pixel value interpolated to the direction  $(x, y)$  and  $M$  is the mosaic image. After combination, the spectrum in each pixel is redetermined. The pointing images have already been corrected by the antenna gain pattern. A more detailed discussion of these image considerations is given by Cotton et al. (2018).

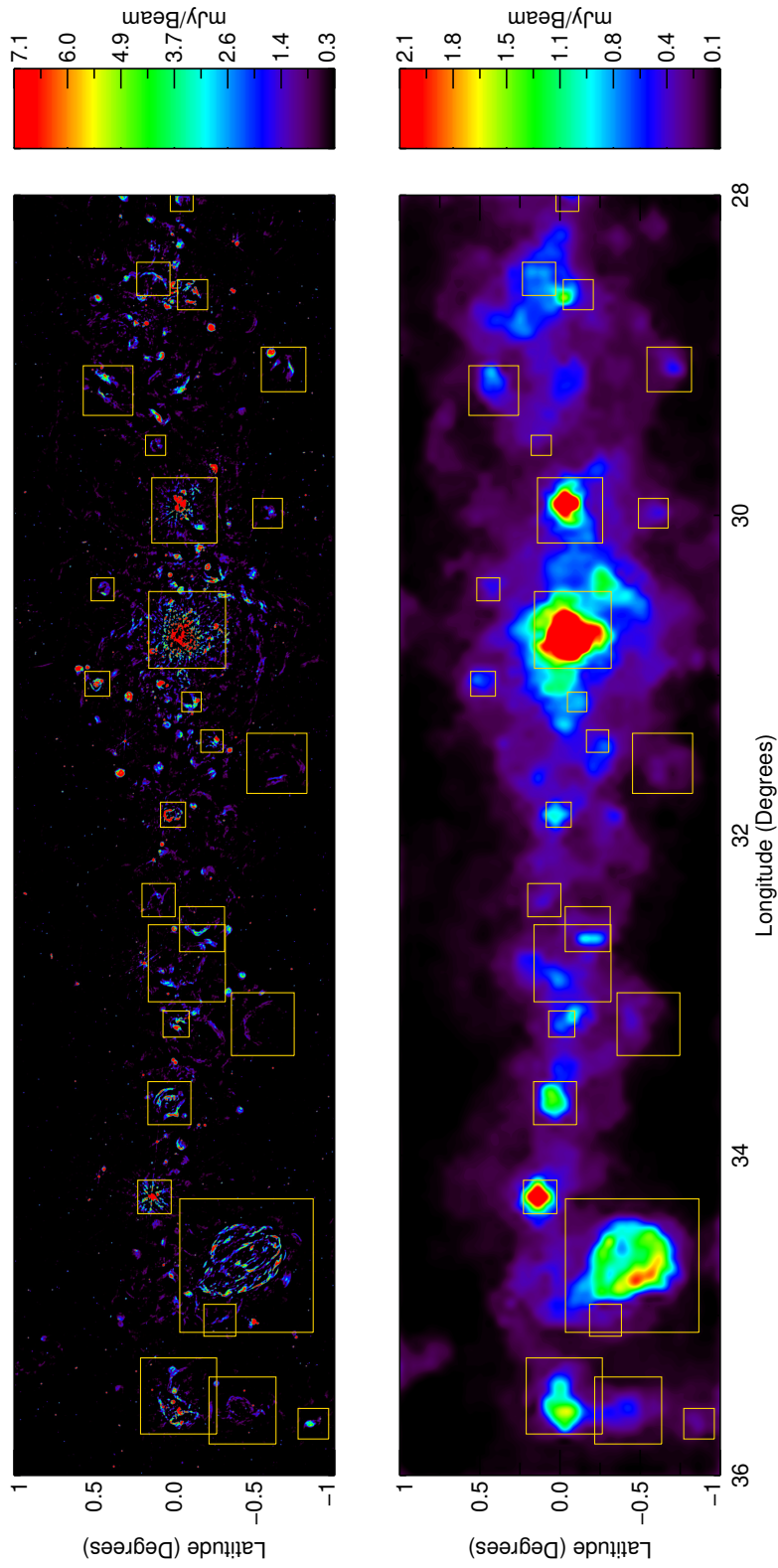
### 3. Analysis of the continuum map

#### 3.1. Radio continuum map

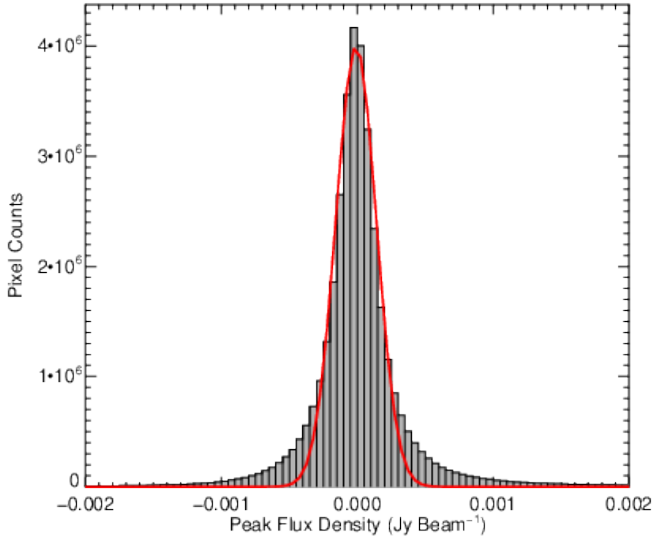
A mosaic of the 16 square degrees that results from the eight sub-images integrated and discussed in this paper is presented in the upper panel of Fig. 1. The effective frequency of the integrated image is 5.8 GHz. Although the observed region extends slightly beyond  $|b| = 1$ , the noise increases significantly in these regions due to poorer  $uv$ -coverage and differences in the beam sizes of the different sub-maps, and therefore these edges have been removed. The final map has a spatial resolution of  $18''$ . In the lower panel of this figure, we present a map of the noise distribution created using SExtractor (Bertin & Arnouts 1996; Holwerda 2005). A full description of the maps can be found in Sect. 3.2. In Fig. 2 we present a histogram of the pixel values from which we estimate the sensitivity of the survey by fitting a Gaussian profile to the pixels clustered around zero. The standard deviation of the distribution of the pixel values is  $\sigma_{\text{rms}} = 150 \mu\text{Jy beam}^{-1}$ . However, as can be seen from the lower panel of Fig. 1 there are strongly localized variations in the noise across the map particularly towards prominent complexes (e.g., W43, W44, G34.26+0.15), most of which are concentrated towards the mid-plane. This is a consequence of the poor  $uv$ -coverage of the observations themselves. These large variations of the noise across the GLOSTAR map produce the non-Gaussian profile of the pixel values in Fig. 2. In regions free of extended emission the noise level is around  $60 \mu\text{Jy beam}^{-1}$ . In Fig. 3 we show the noise as a function of Galactic latitude, which reveals a steep increase in noise closer to the Galactic mid-plane from  $\sim 100 \mu\text{Jy beam}^{-1}$  to  $\sim 450 \mu\text{Jy beam}^{-1}$ .

To measure the accuracy of the astrometry of the GLOSTAR map, we have matched the GLOSTAR unresolved sources ( $Y$ -factor  $< 1.2$ ; this factor is the ratio of the integrated and peak flux densities, and will be fully described in Sect. 5) with more compact radio sources detected by the 5 GHz CORNISH radio continuum survey (Purcell et al. 2013) and measured the positional offsets between them. The offsets in  $\ell$  and  $b$  are plotted in Fig. 4, which are clearly clustered around the center; the mean

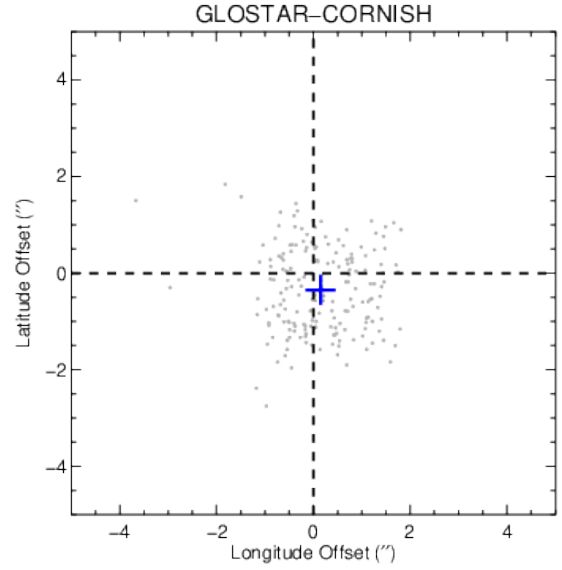




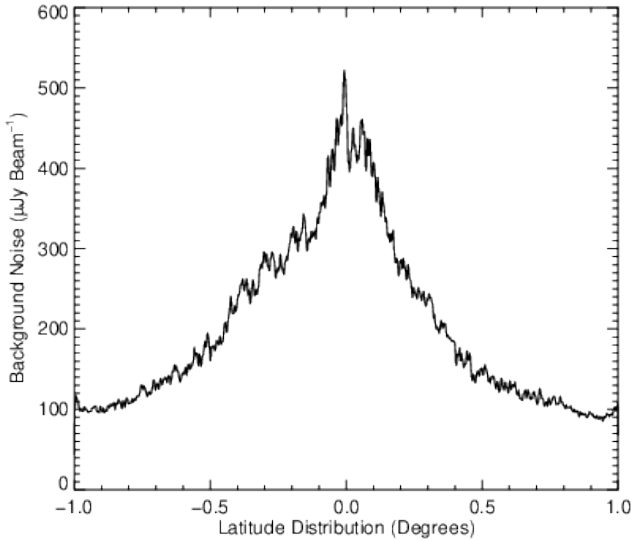
**Fig. 1.** GLOSTAR mosaics and noise map of the present catalog. *Upper panel:* GLOSTAR radio continuum map at 5.8 GHz of 16 sq degrees of the Galactic plane ( $28^\circ < \ell < 36^\circ$  and  $|b| < 1^\circ$ ). *Lower panel:* full resolution noise background map determined by `SEXtractor` using a mesh size of  $80 \times 80$  pixels and threshold of  $5\sigma$ . The yellow boxes correspond to the complexes that have been identified (see Sect. 3.2.3 and Table 2 for more detailed and positions, respectively).



**Fig. 2.** Noise distribution of the map presented in Fig. 1. We have restricted the range of pixel values to between  $-2$  and  $2$  mJy in order to estimate the noise in the map and determine its sensitivity. The red line shows the results of a Gaussian fit to the distribution, which gives the standard deviation of the noise as  $150 \mu\text{Jy}$ . The bin size is  $50 \mu\text{Jy}$ .



**Fig. 4.** Position offsets between GLOSTAR compact sources ( $Y$ -factor  $< 1.2$ ) with their CORNISH counterparts. The blue cross indicates the mean value of the offsets and the dashed lines indicate where the longitude and latitude offsets are equal to zero.



**Fig. 3.** Noise distribution of the map presented in Fig. 1 as a function of Galactic latitude. This has been produced in the same way as Fig. 2 but for each increment in latitude.

values are  $0'.2 \pm 0'.9$  and  $-0'.4 \pm 0'.9$ , respectively, where the uncertainties are the standard deviations in the measurements. We therefore conclude that the GLOSTAR astrometry is reliable. We estimate the combined uncertainty on position, by adding the standard deviations in quadrature, to be  $1'.2$ , which is an order of magnitude lower than our angular resolution ( $18''$ ).

### 3.2. Source extraction

The source extraction has been performed using the BLOBCAT software package developed by Hales et al. (2012). BLOBCAT is a Python script that uses a flood fill algorithm to detect and identify blobs, or islands of pixels representing sources, in two-dimensional astronomical radio-wavelength images. The package has been successfully used to create the source catalog for

The HI/OH/Recombination line survey of the inner Milky Way (THOR; Beuther et al. 2016) survey, and has been adopted by us to facilitate comparison with this complementary VLA continuum and line survey. We have followed the same method used by the THOR team at  $1.4$  GHz, which is outlined in Bihr et al. (2016).

The GLOSTAR radio continuum image has position-dependent noise and so we need to produce an independent noise map of the region to perform the automatic source extraction. We use the rms estimation algorithm within the SExtractor package that has been proven to be trustworthy in creating noise maps from radio data (Bondi et al. 2003; Huynh et al. 2005). This algorithm defines the rms value for each pixel in an image by determining the distribution of pixel values within a local mesh until all values are around a chosen  $\sigma$  value. Most real emission is removed from the noise image and the determined noise map contains the correct noise level. We use a detection and analysis threshold of  $5\sigma$ , a minimum size of five pixels, and a mesh size of  $80 \times 80$  pixels<sup>2</sup> (following the calculation of Hales et al. 2012). The resulting noise map is shown in the lower panel of Fig. 1, its large scale emission is consistent with the large scale emission previously reported by Bihr et al. (2016) at  $1.4$  GHz.

To perform the automatic source extraction, we use the noise map as an input into BLOBCAT along with the FITS<sup>5</sup> image of the region under study. We applied a detection threshold (parameter dSNR in BLOBCAT) of four and a minimum source size of five pixels in diameter ( $\sim 13''$ ). This resulted in the extraction of all sources that have flux above four times the local noise (i.e., the ratio between the input image and the rms map) and have a size comparable to or larger than the beam; these criteria reduce the number of false detections, particularly in complex regions.

These input parameters resulted in the detection of 1975 sources within the  $16$ -square-degree map. However, this sample is likely to contain a significant number of artifacts and emissions from large scale structures that have been over-resolved,

<sup>5</sup> Flexible image transport system.

resulting in the emission being separated into multiple components (e.g., the W44 supernova remnant at  $\ell = 34.6^\circ$ ,  $b = 0.5^\circ$ ), and from the blending of discrete sources towards star formation regions found towards the Galactic mid-plane. In the following subsections, we will describe how we have dealt with some of these issues, while in Sect. 4 we will describe the steps taken to identify and remove artifacts, and check the consistency of our catalog.

### 3.2.1. Superposition of distinct radio sources

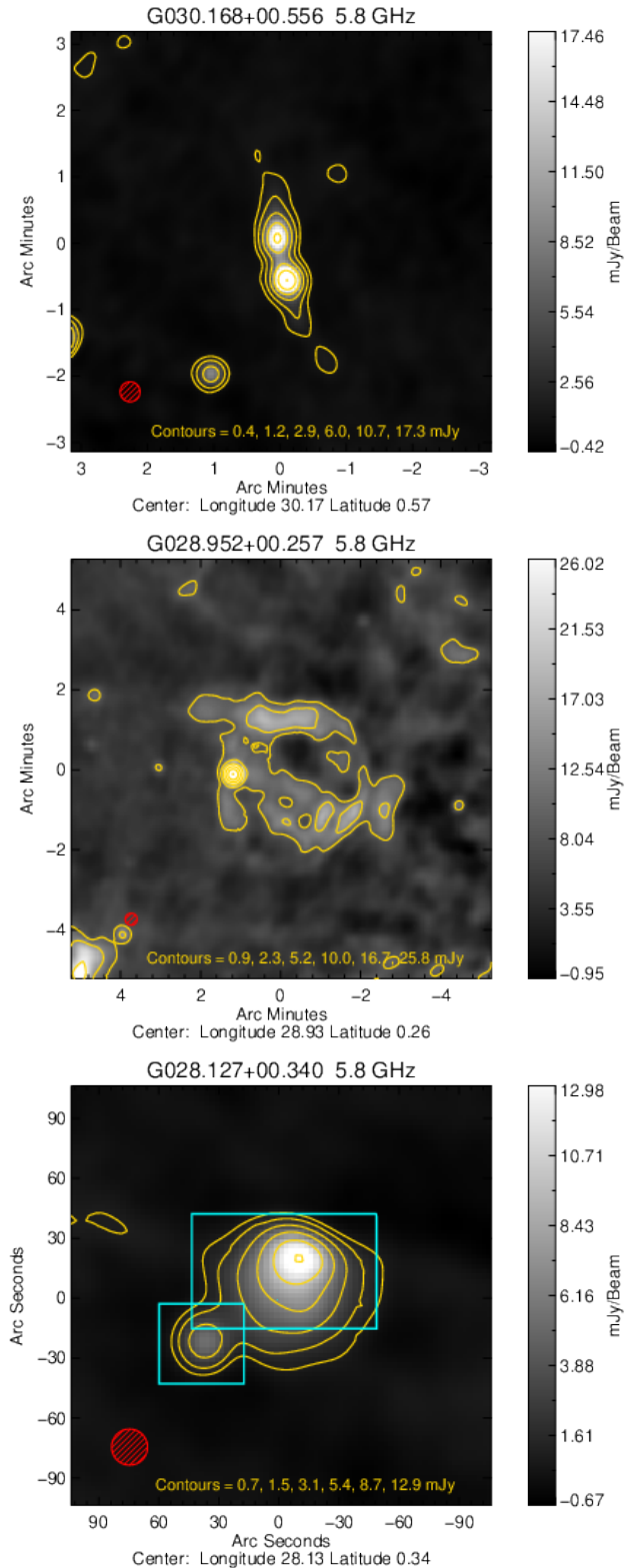
In the visual inspection process, we flagged any sources that appeared to consist of two or more distinct emission peaks. We have grouped these into three categories: (1) elongated “double lobed” sources; these look like extended emission regions powered by twin jets from active galactic nuclei (AGN; Treichel et al. 2001; Malarecki et al. 2015; Neff et al. 2015); (2) superposition of a compact, bright radio source and a larger, more diffuse radio source; (3) superposition of spherical radio sources; these are likely to be H II regions in the same star-forming complexes. An example of each of these three types is shown in the upper, middle, and lower panels of Fig. 5.

We have made no attempt to separate the emission for the double lobed sources as it is likely to be associated with the same object. We have also made no attempt to separate the emission in cases in which a bright compact source is coincident with larger scale structures as there is not enough information to split them into reliable components as any measurement will be highly uncertain. There are a total of 51 such cases. We have, however, separated the emission into two or more sources where the overlap is small and the integrated flux density can be reliably determined (see lower panel of Fig. 5 as an example).

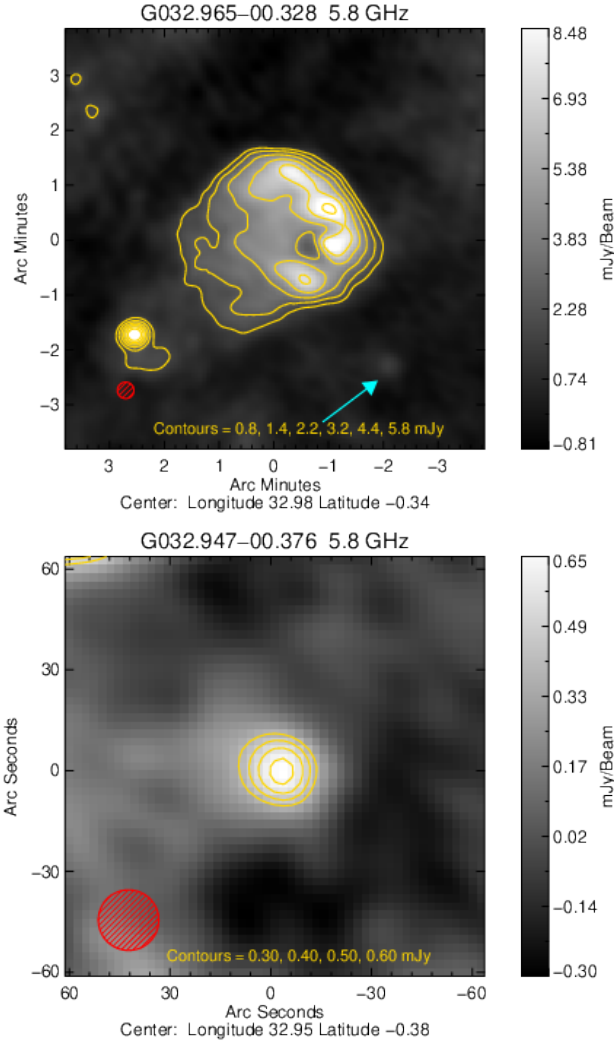
Inspection of the GLIMPSE 4.5, 8.0, and 24  $\mu\text{m}$  mid-infrared images of these blended sources identified 20 cases for which the emission peaks appear to be associated with distinct sources and could be separated. We split them into two or three sources according to the number of infrared counterparts. We use the CASA software to select and split the regions in which emission is above  $4\sigma$  and then calculate their parameters such as flux and size from their emission area. For some sources, for example, uniform and compact radio emission, we make a Gaussian fit using the task `imfit` also from the CASA software suite. An example of these type of sources is shown in the lower panel of Fig. 5. This process resulted in 44 sources being identified from the 20 blended sources examined. These split sources are treated in the source catalog (see Table 7) as individual sources but we also include a flag to the entry so these can be easily identified. The peak and centroid position values are considered the same for these cases.

### 3.2.2. Compact sources in noisy regions

Visual inspection of the extracted sources and comparison with other catalogs (some of which are discussed in Sect. 4.1) revealed a small number of compact (15), low surface brightness sources (with peak flux  $\sim 1 \text{ mJy beam}^{-1}$ ) that are located near regions of bright extended sources. An example is the source in Fig. 6. BLOBCAT does not extract these objects due to the higher localized noise surrounding them and poor dynamic range of snapshot interferometric observations. We have calculated their physical properties again using the CASA task `imfit`. This performs an elliptical Gaussian fit to the region identified around the source. The source parameters for these objects are included in the final catalog presented in Table 7 but, as with the split



**Fig. 5.** Examples of various types of radio sources (see text). The contour levels start at  $3\sigma$  and increase in steps determined using the dynamic range power-law fitting scheme where  $D = N^i + 3$ , where  $D$  is the dynamic range,  $S_{\text{peak}}/\sigma$ ,  $N$  is the number of contours, in this case six, and  $i$  is the power law that determines the separation between consecutive contours; this is a modified version of a scheme developed by Thompson et al. (2006). The red hatched circle shown in the lower left corner of the map indicates the GLOSTAR beam size. The cyan boxes indicate the two sources identified and the region from which the flux has been estimated.



**Fig. 6.** Example of compact source in noisy region. *Upper panel:* example of an instance where a bright extended source has increased the local noise such that a nearby compact  $7\sigma$  source located to the southwest has been missed. The position of this source is indicated by the cyan arrow. *Lower panel:* smaller region centered on the missed compact source.

sources, we also include a flag to the entry so these can be easily identified.

### 3.2.3. Large scale structures

We have identified 27 regions with large scale emission, by comparing the distribution of radio emission to that of the  $8\mu\text{m}$  maps extracted from the GLIMPSE archive. We have drawn a box around the coherent infrared structures that are morphologically correlated with the radio emission, but for which the radio emission has been broken up into two or more components. These 27 regions are indicated by squares in the upper and lower panels of Fig. 1. It is clear that these regions are also correlated with the noisiest regions of the map where much of the large angular scale emission is poorly recovered by the observations. In Table 2 we give the name of each structure, the extent of the boxed region in  $\ell$  and  $b$ , and the integrated flux of the total emission in the box minus the emission from any compact sources that are considered distinct sources. The name of each structure has been constructed from the central coordinate of the enclosing box.

**Table 2.** Large scale structures.

Name	$\ell_{\min}$ ( $^\circ$ )	$\ell_{\max}$ ( $^\circ$ )	$b_{\min}$ ( $^\circ$ )	$b_{\max}$ ( $^\circ$ )	Int flux (Jy)
G028.026-00.045	27.970	28.082	-0.100	0.010	0.32
G028.520+00.132	28.470	28.570	0.049	0.215	0.30
G028.6-00.1 (SNR) <sup>(*)</sup>	28.544	28.694	-0.188	-0.038	0.95
G029.087-00.682	28.976	29.199	-0.756	-0.608	0.22
G029.219+00.415	29.095	29.343	0.344	0.486	0.50
G029.6+00.1 (SNR) <sup>(*)</sup>	29.512	29.612	0.071	0.165	0.13
W43-south center	29.805	30.131	-0.195	0.071	9.08
G029.986-00.582	29.913	30.060	-0.650	-0.513	0.11
G030.462+00.449	30.405	30.519	0.393	0.506	0.16
W43 center	30.525	30.908	-0.238	0.080	13.43
G031.053+00.482	30.992	31.115	0.435	0.529	0.59
G031.166-00.106	31.117	31.214	-0.050	-0.162	0.42
G031.411-00.234	31.355	31.467	-0.286	-0.183	0.60
G031.5-00.6 (SNR) <sup>(*)</sup>	31.400	31.700	-0.790	-0.490	0.14
G031.9+00.0 (SNR) <sup>(*)</sup>	31.821	31.920	-0.053	0.071	1.01
G032.4+00.1 (SNR) <sup>(*)</sup>	32.322	32.488	0.027	0.172	0.10
G032.586-00.172	32.534	32.637	-0.284	-0.059	0.26
G032.8-00.1 (SNR) <sup>(*)</sup>	32.649	32.950	-0.269	0.115	0.67
G033.179-00.010	33.118	33.240	-0.075	0.055	0.66
G033.2-00.6 (SNR) <sup>(*)</sup>	33.033	33.327	-0.708	-0.395	0.09
G033.6+00.1 (SNR) <sup>(*)</sup>	33.567	33.781	-0.067	0.132	0.93
G034.260+00.125	34.177	34.343	0.043	0.207	10.17
W44 - G034.6-00.5 (SNR) <sup>(*)</sup>	34.390	34.930	-0.760	-0.093	2.96
G035.032-00.283	34.994	35.070	-0.363	-0.204	0.04
G035.506-00.026	35.316	35.696	-0.157	0.106	2.36
G035.6-00.4 (SNR) <sup>(*)</sup>	35.502	35.692	-0.593	-0.258	2.36
G035.680-00.868	35.610	35.750	-0.943	-0.792	0.19

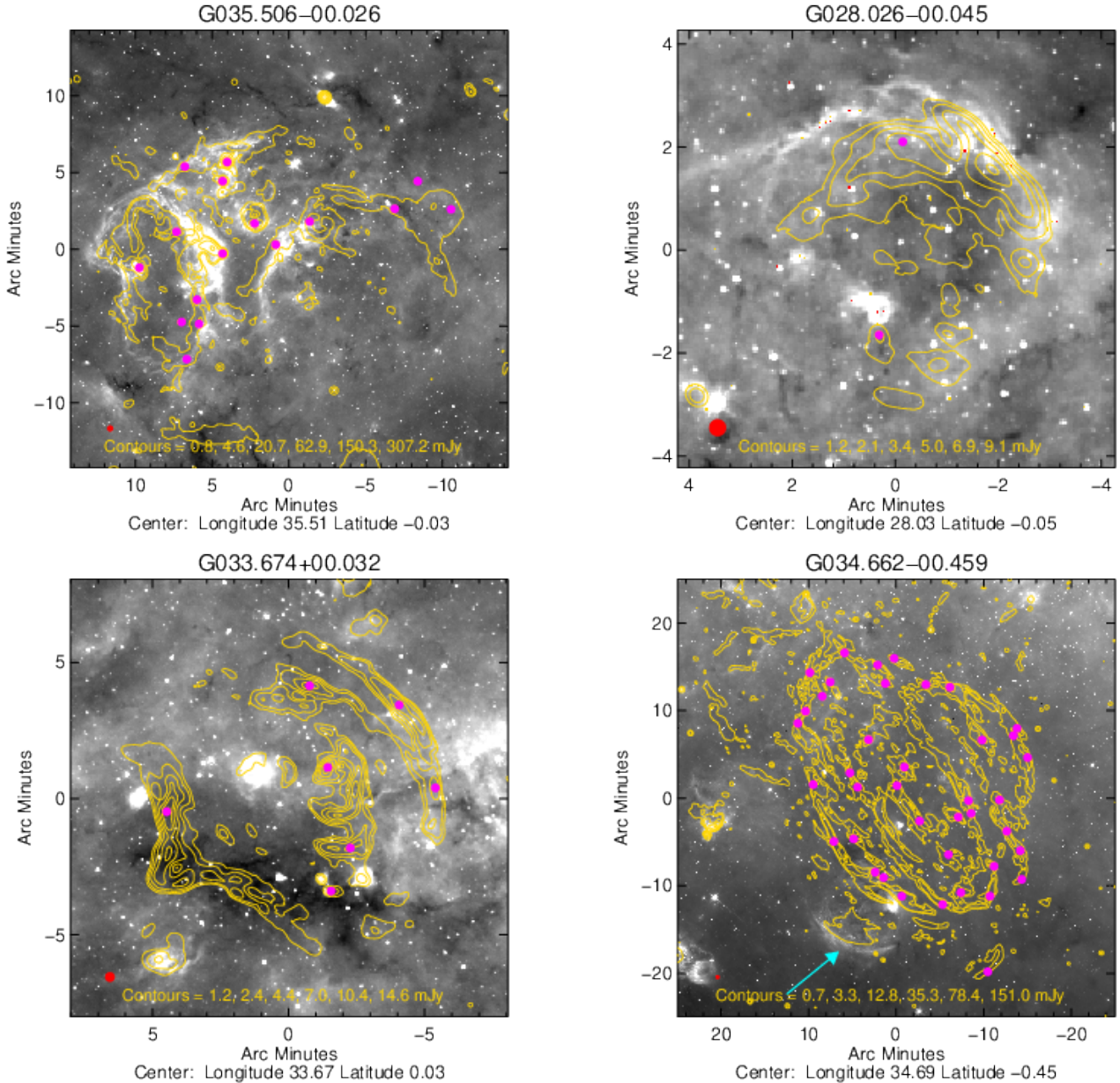
**Notes.** The source name is constructed from the central position of the box used to encapsulate the regions. The sources with <sup>(\*)</sup> are associated with supernova remnants from the catalog of Green (2014).

In Fig. 7 we show some examples of these large scale structures. All extended radio sources in these boxed regions are considered to be part of the large scale structure and have been excluded from the final catalog of radio sources. However, compact radio sources that are considered likely to be associated with discrete objects and not part of the larger scale structure have been retained in the final catalog as distinct sources.

In total, 195 sources are associated with the 27 large scale structures identified. Of the 27 large scale regions identified, two stand out as being particularly significant; these are the W43 star-forming complex ( $29^\circ \leq \ell \leq 32^\circ$  and  $|b| \leq 1^\circ$ ; Blum et al. 1999) and the well-known supernova remnant W44 (Clark & Parkinson 1975). Between them they contain 92 of the 195 individual fragments that have been associated with large scale emission regions. We note that radio recombination lines have been detected towards a radio source (G034.793-00.711) seen towards the southeastern edge of W44 (Lockman 1989; G34.758-0.681). In the lower right panel of Fig. 7 we present a map of the radio emission seen towards W44 and indicate the position of this nearby H II region with an arrow. The association of RRL emission with this radio source provides strong evidence that this source is a H II region that is coincident with the edge of this supernova remnant (G034.8-00.7; see Ortega et al. 2010 for detailed discussion of this H II region) rather than part of W44 itself. Therefore, we treat this as a separate source, which we classify as a H II region.

The W43 region has two prominent components: W43-main and W43-south. W43-main hosts a giant H II region (Blum et al. 1999) and has been classified as a mini star-burst





**Fig. 7.** Example of radio emission associated with large scale structures. The background images are GLIMPSE  $8\ \mu\text{m}$  while the yellow contours are the GLOSTAR 5.8 GHz radio continuum emission. These regions often show coherent infrared structures that are morphologically correlated with most of the radio emission (see examples presented in the *upper panels*), however, the correlation is not always present but the radio emission is clearly correlated (see examples presented in the *lower panels*). These large scale structures have been excluded from the final catalog of radio sources. The contours are determined as described in Fig. 5 and the red filled circle shown in the lower left corners shows the GLOSTAR beam. The magenta circles indicate the peak positions of radio sources associated with these regions. *Lower-right panel:* W44 SNR; the cyan arrow points to the H II region (G034.793-00.711) located close to the edge of this complex (see text for details).

region (Motte et al. 2003). Due to the complexity of the region and our limited  $uv$ -coverage, the central parts of both regions are not well reconstructed in the radio emission map. That could also be a reason for the different integrated flux density measurement of the central part of W43-main between the THOR team, 55.55 Jy at 2 GHz (Bühr et al. 2016), and our value, 13.43 Jy at 6.0 GHz. However, our observations have been able to recover very well much of the emission from the shell structure of W44 (see lower-right panel of Fig. 7). This extended object is  $\sim 30$  pc across with thermal X-ray emission seen towards its center (Clark & Parkinson 1975). We measure an integrated flux density of 2.96 Jy, which is significantly larger than the value of

1.2 Jy determined by Wynn-Williams et al. (1981) from 5 GHz VLA observations; this new integrated flux density estimation is due to our higher sensitivity to more extended emission.

Our sample of large scale emission structures also includes three other well-known regions. The first of these is the very bright ultra-compact H II region G34.3+0.2 (G034.260+00.125 in Table 2). Garay et al. (1986) estimated the integrated flux density at 5 and 8 GHz, to be  $\sim 1.5$  Jy and  $\sim 2$  Jy respectively. We calculate the flux to be significantly higher at  $\sim 10$  Jy, but we are unable to fully resolve the structure of this object with the resolution of the current GLOSTAR data. Our higher integrated flux density determination is likely to be caused by our sensitivity

to larger scale emission that its resolved out in the higher angular resolution data of Garay et al. (1986). The other two regions worth mentioning are G035.506–00.026 and G030.462+00.449, both of which have counterparts in the catalog of star-forming regions reported by Rahman & Murray (2010) (complexes 17 and 23 respectively). The W44 and W43 regions also have counterparts in this catalog. Therefore, we have recovered four of the most luminous free-free sources in the Galaxy.

Searching the literature we find that ten of these large scale structures, including W44, have been previously identified as supernova remnant (SNR) (Green 2014); this corresponds to over a third of this sample. Where a match has been made we have adopted the catalog name from Green (2014) in Table 2.

### 3.3. Catalog completeness and reliability

The BLOBCAT source extraction method was verified by the THOR team (Bihl et al. 2015; Wang et al. 2018). They performed a detailed set of completeness tests. They chose a region of  $0.5^\circ \times 0.5^\circ$  with a constant noise level and added artificial 2D Gaussian sources with the size of the beam and different peak intensities. They measured the fraction of sources recovered as a function of intensity and found that at a threshold of  $7\sigma$  the extraction algorithm was 95% complete. Given that our data are similar to the THOR data (both are VLA continuum surveys and have similar resolutions) we have simply adopted their results.

The CORNISH team used a similar flux threshold ( $7.2\sigma$ ; Purcell et al. 2013) as the minimum required for inclusion in their catalog. This was determined by calculating the number of spurious detections expected as a function of  $\sigma$  and set at a level above which no spurious detections would be expected (see discussion by Medina et al. 2018). The CORNISH survey consisted of  $5.6 \times 10^8$  beams and so a high threshold was required to avoid any false detections. The region presented here consists of approximately 1000 times fewer beams ( $\sim 5.65 \times 10^5$  beams) and so we can set our threshold for reliable sources at a lower level. Based on Gaussian statistics, the number of spurious sources we would expect in our catalog using  $5\sigma$  flux threshold is  $< 1$  and so this is the level above which sources in the catalog are considered to be highly reliable. We find another  $\sim 250$  radio sources between 4 and  $5\sigma$  of which we estimate 7% are likely to be spurious ( $\sim 17$  sources); we include these in the final catalog as the vast majority are likely to be real sources and have counterparts in other surveys, but these are considered to be less reliable.

## 4. Verification of the source catalog

As discussed in the previous section, only a small number of spurious sources are expected in our final catalog. However, due to the nature of these interferometric snapshot observations, the  $uv$ -plane is relatively sparsely sampled, which can result in a significant number of imaging artifacts in the restored maps. As a consequence the number of spurious sources expected from Gaussian statistics is, therefore, a lower limit to the actual number of spurious sources detected.

To improve the reliability of the catalog produced by BLOBCAT we need to identify any spurious noise peaks and imaging artifacts (side lobes and emission from over-resolved structures). We have already taken emission from over-resolved structures into account by identifying the large scale structures as described in Sect. 3.2.3; this reduces the number of sources

in the catalog to 1819. Next we assume that any radio source that has a counterpart in a published radio or mid-infrared catalog is real; this is because it is extremely unlikely that a false detection would randomly appear in the same position in other independent surveys.

### 4.1. Complementary surveys

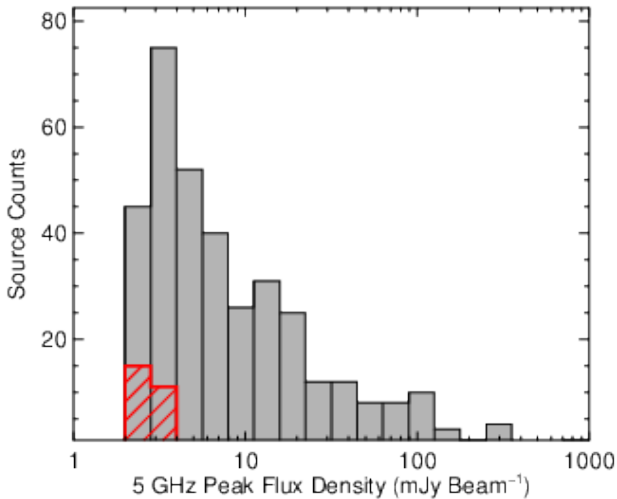
A number of complementary multiwavelength surveys have produced catalogs of sources located in the Galactic plane, which we have used to verify whether the sources recovered from the GLOSTAR map are real and to help determine their nature. Below we provide a brief summary of the surveys used in this paper. In all cases we have used a  $10''$  radius to identify matches with the exception of the ATLASGAL dust emission; this is because the dust emission tends to be more extended. The surveys selected are either well matched in resolution (e.g., WISE, ATLASGAL, MAGPIS, and THOR, with angular resolutions of  $6$ – $19''$ ) or in wavelength (e.g., CORNISH at 5 GHz).

#### 4.1.1. CORNISH

The Co-Ordinated Radio “N” Infrared Survey for High-mass star formation (CORNISH; Hoare et al. 2012) is a sensitive and high-resolution 5 GHz radio continuum survey that focuses on the northern GLIMPSE region ( $10^\circ < \ell < 65^\circ$  and  $|b| < 1^\circ$ ). The survey was conducted using the Very Large Array before its upgrade in B-configuration, which provides an angular resolution of  $1''.5$  and has a  $7.2\sigma$  sensitivity of  $\sim 2.7$  mJy beam $^{-1}$ .

The CORNISH catalog consists of  $\sim 2600$  sources above  $7\sigma$  (Purcell et al. 2013), 353 of which are located in the GLOSTAR region presented in this paper. We find a GLOSTAR counterpart for 327 of these ( $\sim 80\%$ ). Inspection of the emission maps for the brighter CORNISH sources reveals that 20 more are associated with extended or multiplex emission sources found in GLOSTAR where the positions are less well constrained and consequently these sources were not matched. In Fig. 8 we show the peak flux distribution of all 353 CORNISH sources and the 26 sources where we have failed to find a GLOSTAR counterpart using either  $10''$  search radius or through inspection of the emission maps. It is clear from this plot that all of the radio sources that are not picked up by GLOSTAR correspond to the weakest CORNISH sources. We have estimated the noise values from the GLOSTAR maps towards these 26 CORNISH position and find a mean noise value of  $166 \mu\text{Jy beam}^{-1}$  and a standard deviation of  $154 \mu\text{Jy beam}^{-1}$  (minimum and maximum noise values are 47 and  $768 \mu\text{Jy beam}^{-1}$ , respectively), which is significantly lower than the typical CORNISH  $1\sigma$  ( $0.33$  mJy beam $^{-1}$ ) noise values (see Fig. 9).

There are two possible explanations for these missing sources: (1) the flux density is variable and has decreased between the time the CORNISH observations were completed and the GLOSTAR observations were made; and (2) these are spurious CORNISH detections. The variability of radio sources has been investigated by Kalcheva et al. (2018) from an analysis of the CORNISH and MAGPIS integrated flux densities (we will discuss the MAGPIS survey in the next subsection). Kalcheva et al. (2018) estimate that  $\sim 5\%$  of the CORNISH UCH II region sample are potentially variable sources (flux increases of over 50% over a 15 year time period), which is consistent with the number of CORNISH detections that do not have a matching GLOSTAR counterpart ( $\sim 7\%$ ). In Table 5 we provide the source name and integrated and peak flux densities of the unmatched CORNISH sources (taken from



**Fig. 8.** Peak flux distribution of CORNISH sources located in the GLOSTAR field. The gray and red hatched histograms show the distribution of all CORNISH sources and those without a GLOSTAR counterpart, respectively. The bin width is 0.15 dex.

Purcell et al. 2013) along the local rms noise determined from the GLOSTAR maps.

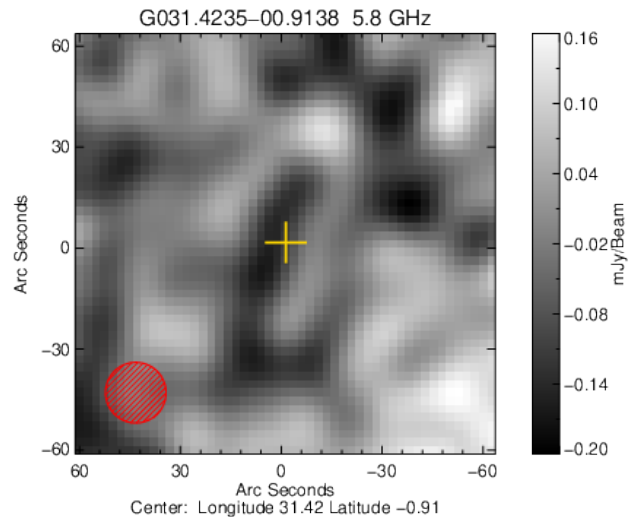
In the case of variable sources, we should expect to find approximately as many sources with significantly larger fluxes in CORNISH than measured in GLOSTAR, and comparing the fluxes we find seven matches for which this is the case (following Kalcheva et al. 2018 we set this criterion to be  $>50\%$  more flux). The radio names of these sources and the flux ratio are given in Table 6. The discrepancy between this number and that of the number of CORNISH sources not detected in GLOSTAR (Table 5) suggests that variability alone cannot fully account for the 26 CORNISH source not detected in GLOSTAR. The highest flux ratio is 3 and so the differences are relatively modest compared to those reported by Kalcheva et al. (2018) (range between 2.4 and 73.9). All of the sources are either marginally detected or unresolved, have integrated fluxes below 10 mJy, and all but one have negative spectral indices, and so the variability appears to be restricted to weak and unresolved extragalactic sources. We note that these radio properties are similar to those reported for T Tauri stars (e.g., Dzib et al. 2013), however, given these radio sources appear to be mid-infrared dark this possibility is considered less likely.

Kalcheva et al. (2018) considered only H II regions in their analysis, whereas none of the variable candidates we have identified have been classified as being H II regions, planetary nebulae (PNe), or radio stars. However, Kalcheva et al. (2018) only found one variable H II region in the region being considered here and, given the shorter temporal baseline ( $\sim 7$  yr), the lack of variable H II regions and the lower levels of variation are not surprising.

#### 4.1.2. MAGPIS

One of the first VLA surveys of a large fraction of the Galactic plane at 5 GHz was conducted in the early 1990s ( $350^\circ < \ell < 40^\circ$  and  $|b| < 0.4^\circ$ ; Becker et al. 1994). This survey is now part of a larger collection of re-reduced radio surveys known as the Multi-Array Galactic Plane Imaging Survey (MAGPIS; Helfand et al. 2006)<sup>6</sup>. This part of the MAGPIS survey has a resolution of

<sup>6</sup> <https://third.ucllnl.org/gps/>



**Fig. 9.** Example of a CORNISH source not detected in GLOSTAR. The grayscale is the GLOSTAR map of a  $2' \times 2'$  region centered on the position of the CORNISH source G031.4235-00.9138 (as indicated by the yellow cross). The noise towards this source is  $63 \mu\text{Jy beam}^{-1}$  and given that the integrated CORNISH flux is 3.47 mJy, it should have been easily detected in the GLOSTAR map. The red hatched circle shown in the lower left corner indicated the size of the GLOSTAR beam.

6–9'' and a  $5\sigma$  sensitivity of  $\sim 2.5 \text{ mJy beam}^{-1}$ . This reprocessing has resulted in the identification of 1272 discrete sources including 218 that are located within the GLOSTAR region presented in this paper. Of these, we have found that 128 are coincident with 125 GLOSTAR sources (the lower resolution of GLOSTAR results in three cases where a GLOSTAR source is associated with two MAGPIS sources). A further 22 MAGPIS sources are coincident with sources associated with large scale structures discussed in Sect. 3.2.3. In total, 150 MAGPIS sources are matched in GLOSTAR, which corresponds to approximately 70% of the MAGPIS 5 GHz sample. We have investigated a few of the MAGPIS sources not matched in GLOSTAR and find many of these to be associated with sources that are over-resolved in MAGPIS. The offsets between the partially recovered fragments and the GLOSTAR source are larger than the  $10''$  used to identify counterparts.

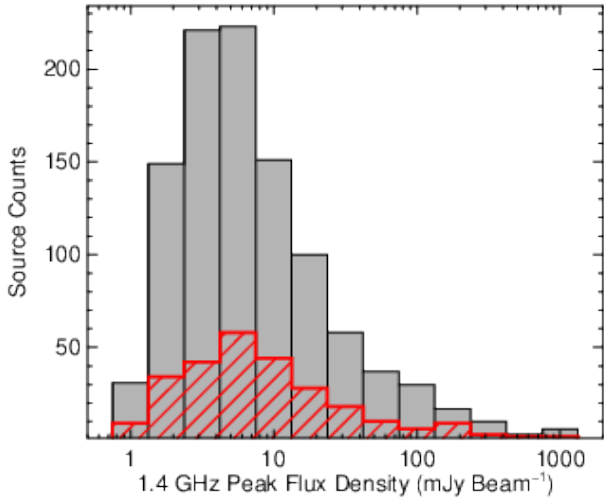
#### 4.1.3. THOR

The H I, OH, Recombination line survey (THOR; Beuther et al. 2016; Bihr et al. 2015; Wang et al. 2018) is a 1.4 GHz radio continuum and line survey of  $\sim 100 \text{ deg}^2$  of the Galactic plane in the first quadrant of the Milky Way. This survey covers the 21 cm H I line, four OH transitions, 19 radio recombination lines, and continuum emission between 1 to 2 GHz. The survey was conducted using the *Karl G. Jansky* VLA in C-array and has a spatial resolution of  $10\text{--}25''$ . The data reduction strategy and source catalog for the first half of the survey ( $\ell = 14^\circ\text{--}37.9^\circ$ ,  $\ell = 47.1^\circ\text{--}51.2^\circ$ , and  $|b| \leq 1.1^\circ$ ) are described in Bihr et al. (2016). The catalog contains a total of  $\sim 4400$  sources, of which  $\sim 1200$  are spatially resolved and  $\sim 1000$  are possible artifacts<sup>7</sup>.

A total of 1036 THOR sources are located in the region of the GLOSTAR survey described here; we are able to find a GLOSTAR counterpart for 711 of these. We have matched a further 61 THOR sources with GLOSTAR sources that have been

<sup>7</sup> These sources have been incorporated into the complete THOR source catalog that consists of  $\sim 10\,000$  sources (Wang et al. 2018).





**Fig. 10.** Peak flux distribution of THOR sources located in the GLOSTAR field. The gray and red-hatched histograms show the distribution of all THOR sources and those without a GLOSTAR counterpart, respectively. The bin width is 0.25 dex.

**Table 3.** Summary of detection categories.

Description	Number of sources
Total number of sources extracted	1975
Number of sources assoc. with large structures	195
Number of split sources	20 (44)
Number of sources recovered near bright sources	15
Number of sources rejected	237
Number of sources in final catalog	1575

identified as being part of large scale structures. In Fig. 10 we show the 1.4 GHz peak flux distribution of all THOR sources located in the GLOSTAR region and the flux distribution of those without a counterpart in GLOSTAR (gray and red-hatched histograms, respectively). It is clear from this plot that a higher proportion of the weaker THOR sources have not been matched. Spectral index measurements are only available in the THOR catalog for the brightest of these unmatched sources (203 of the 264 unmatched) and the majority of these have negative values (71% have  $\alpha < 0$ , with a mean value =  $-0.85$ ), which means they are weaker at 5 GHz than at 1.4 GHz and are hence less likely to be detected in the GLOSTAR map. The smaller proportion of unmatched sources that do have a positive spectral index (29% have  $\alpha > 0$ , with a mean value =  $0.28$ ) are a little harder to explain. However, we note that the THOR team acknowledge that  $\sim 23\%$  of their catalog are possibly artifacts, which corresponds well with the fraction of unmatched sources we have found here.

#### 4.1.4. ATLASGAL

The APEX Telescope Large Area Survey of the Galaxy (ATLASGAL; Schuller et al. 2009; Beuther et al. 2012) is the first high-resolution ( $\approx 20''$  FWHM) ground-based submillimetre ( $870 \mu\text{m}$ ) survey of the thermal dust emission in the inner Galactic plane. It has produced an unbiased catalog of  $\sim 10\,000$  massive pre- and proto-stellar clumps (ATLASGAL Compact Source Catalog (CSC); Contreras et al. 2013;

Urquhart et al. 2014a). Correlating the GLOSTAR and ATLASGAL catalogs is an excellent way to identify embedded or dust-enshrouded objects such as UCH II regions and planetary nebulae (e.g., Irabor et al. 2018; Urquhart et al. 2007, 2009, 2013; Purcell et al. 2013).

As previously mentioned, the dust emission is more extended than the radio sources, the majority of which are unresolved and so rather than using a  $10''$  matching radius we identify matches where the radio source falls within the  $4\sigma$  boundary of the ATLASGAL dust emission. There are 747 ATLASGAL sources located in the GLOSTAR region, of which 132 are positionally coincident with a radio source. Of these, 92 have been previously identified as H II regions (Kalcheva et al. 2018; Anderson et al. 2014; Giveon et al. 2005a), one has been identified as a PN, and one as a radio star by the CORNISH team (Purcell et al. 2013). The relatively low fraction of ATLASGAL matches is to be expected as only clumps hosting H II regions will have a radio counterpart and H II regions are only present in the most massive and evolved clumps.

#### 4.1.5. WISE

The Wide-field Infrared Survey Explorer (WISE) is a NASA infrared-wavelength astronomical space telescope mission that mapped the entire sky in four infrared bands, W1, W2, W3, and W4, centered at 3.4, 4.6, 12, and  $22 \mu\text{m}$  using a 40 cm telescope, which fed arrays with a total of four million pixels; these wavelengths correspond to angular resolutions of 6.1, 6.4, 6.5, and  $12''$ . A full description of the mission can be found in Wright et al. (2010).

The WISE ALL-sky Release Source Catalog contains astrometry and photometry for over half a billion objects. We have searched this data repository to identify and extract the mid-infrared fluxes towards 988 GLOSTAR sources; where two or more WISE sources are found within the search radius, we have picked the infrared counterpart to be the source with the smallest angular offset to the peak radio emission.

The WISE catalog is optimized for compact sources and so may overlook more extended mid-infrared emission associated with more evolved H II regions. We have therefore also cross-matched the GLOSTAR sample with a catalog of H II region candidates identified from a combination of a visual and automatic search of WISE data based on their mid-infrared emission morphology (Anderson et al. 2014). This catalog contains  $\sim 8000$  H II region candidates, of which 667 are located in the GLOSTAR field. We find 130 of these mid-infrared sources to be coincident with a radio source, including 26 not found in the search made of the WISE catalog. Although we have only found a radio counterpart for  $\sim 20\%$  of the H II region candidates, Anderson et al. (2014) estimate that they are sensitive to H II regions with radio fluxes greater than  $0.2 \text{ mJy}$ , which is lower than our  $4\sigma$  sensitivity; this sensitivity can be as poor as  $1.6 \text{ mJy}$  towards the Galactic mid-plane (i.e., see Fig. 3) where most of the H II regions are likely to be located. Also, objects being resolved out ( $>4'$ ) could only explain a small proportion (10%) of unmatched WISE H II regions and so the rest are weak H II regions. In total, we have matched 988 radio sources with a WISE counterpart.

#### 4.2. Correlation with other surveys

We have searched for counterparts in the THOR, CORNISH, and MAGPIS radio continuum surveys and the ATLASGAL dust and



**Table 4.** Statistics of the matches between GLOSTAR and other published surveys.

Survey	Wavelength	Resolution ( $''$ )	rms level $\text{mJy beam}^{-1}$	Num. of sources in region	Num. of matches
CORNISH	6 cm	1.5	0.33	353	327
MAGPIS	6 cm	6–9	0.3	218	128
THOR	20 cm	10–25	0.3–1	1036	772
ATLASGAL	$870 \mu\text{m}$	19.2	50–70	747	132
WISE	3.4–22 $\mu\text{m}$	6–12	...	...	988

**Notes.** A match radius of  $10''$  has been used centered on the peak of the GLOSTAR emission, except for the ATLASGAL dust emission, which is more extended than the radio and mid-infrared catalogs.

**Table 5.** Catalog names and fluxes for CORNISH sources not detected in GLOSTAR.

CORNISH name	$S_{\text{peak}}$ ( $\text{mJy beam}^{-1}$ )	$S_{\text{int}}$ ( $\text{mJy}$ )	GLOSTAR rms ( $\text{mJy beam}^{-1}$ )
G028.1875+00.5047	2.61	2.61	0.083
G028.3660–00.9640	2.92	2.92	0.073
G028.5690+00.0813	2.99	5.34	0.359
G028.9064+00.2548	3.47	3.47	0.768
G029.2555–00.8653	2.37	3.16	0.084
G029.2620+00.2916	2.76	2.76	0.120
G029.2648+00.9527	2.59	2.97	0.060
G029.3096+00.5124	2.84	2.84	0.072
G029.4302–00.9967	2.83	2.83	0.103
G029.4404–00.3199	2.81	2.81	0.160
G029.5184+00.9478	2.43	2.43	0.065
G029.7805–00.2661	3.35	7.53	0.451
G030.1884+00.1110	3.00	4.45	0.223
G030.2193+00.6501	2.80	2.80	0.122
G030.4543+00.3223	2.69	3.74	0.195
G030.6328–00.7232	2.57	2.57	0.163
G031.2859–00.2095	2.43	3.76	0.286
G031.4235–00.9138	2.66	3.47	0.063
G032.2408+00.1667	3.06	3.66	0.218
G032.2783–00.1705	2.84	4.80	0.152
G033.9622–00.4966	2.81	2.81	0.062
G034.1382+00.3805	2.61	3.23	0.139
G034.2171–00.6886	2.16	2.37	0.077
G034.3555–00.0876	2.98	4.61	0.087
G034.3852+00.3526	3.89	4.66	0.105
G035.0605+00.6208	2.76	2.76	0.047

**Notes.** The last column gives the measured rms towards the position of the CORNISH source in the GLOSTAR map. The flux errors are comparable to the CORNISH noise level,  $0.33 \text{ mJy beam}^{-1}$ .

**Table 6.** Variable source candidates where the integrated CORNISH flux is more than 50% higher than the peak GLOSTAR flux.

GLOSTAR name	CORNISH name	$S_{\text{peak,GLOSTAR}}/S_{\text{int,CORNISH}}$
G029.589+00.579	G029.5893+00.5789	2.17
G030.104+00.399 <sup>(*)</sup>	G030.1039+00.3983	1.53
G031.585–00.063	G031.5854–00.0635	1.94
G032.453+00.368	G032.4535+00.3679	2.02
G032.599+00.826	G032.5996+00.8265	1.86
G034.178+00.257	G034.1782+00.2564	1.51
G034.265+00.719	G034.2655+00.7195	3.05

**Notes.** <sup>(\*)</sup>The source was previously reported as a variable radio source by Becker et al. (2010).

WISE mid-infrared surveys. The correlation process with other catalogs allows us to find any large scale structures and all compact sources that have a counterpart. This analysis has identified counterparts for 1315 radio sources and these are therefore considered to be reliable detections.

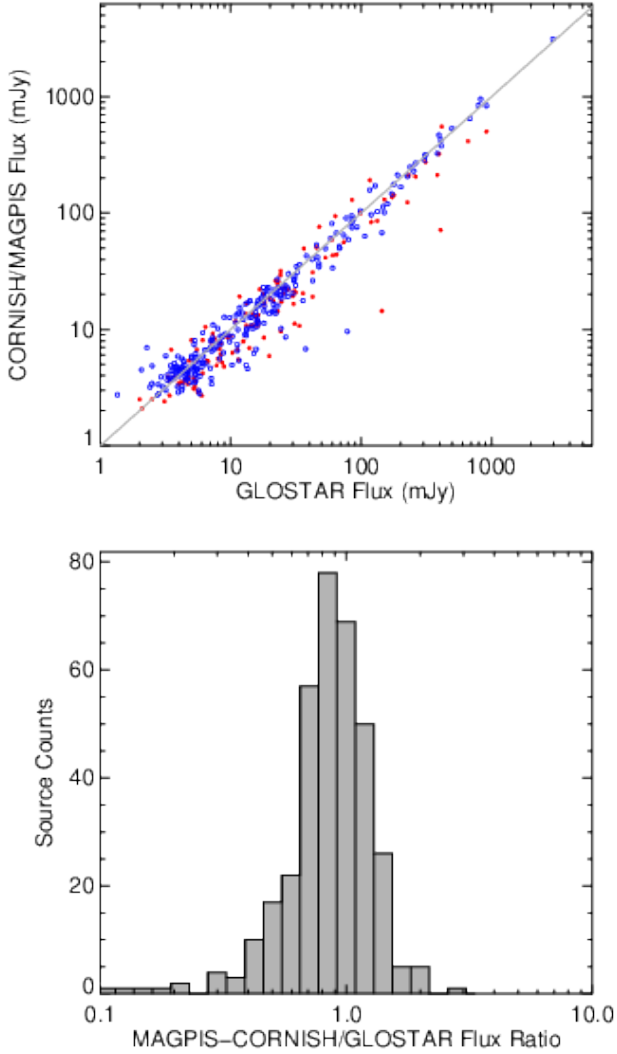
The remaining 465 radio sources are a mix of compact, moderately extended sources, imaging artifacts and spurious detections. The extended sources are also likely to be associated with extended H II regions or supernova remnants. This can be confirmed by comparing the morphology of the radio emission with that at mid-infrared wavelengths. This can also identify weak mid-infrared counterparts to the compact radio sources that were not picked up in the mid-infrared catalogs. The next step is, therefore, to compare the radio and mid-infrared emission maps for these 465 sources; sources for which mid-infrared counterparts can be identified can be considered real. This visual inspection is also useful in identifying and removing imaging artifacts.

We have compared the radio emission with mid-infrared maps extracted from the GLIMPSE 4.5 and  $8.0 \mu\text{m}$  bands, and MIPS GAL  $24 \mu\text{m}$  band archives. We have discarded any compact radio sources that appear to be localized peaks in high-noise regions, sources that are smaller than the beam (these tended to be clustered towards the detection threshold), and extended emission that appears to be associated with large scale structures which, due to poor  $uv$ -coverage, have not been reliably imaged (these have a wider range of fluxes). This process resulted in 237 sources being excluded from the 465 without a counterpart in one of the other surveys.

Combined with the split sources and recovery (by hand) of compact sources in noisy regions, previously discussed in Sect. 3.2, the final radio catalog therefore consists of 1575 sources. In Table 3 we present a summary of the various steps taken in the source verification process and in Table 4 the summary of the search for counterparts with continuum, dust, and mid-infrared surveys.

#### 4.3. Comparison of fluxes with other 5 GHz surveys

As mentioned earlier, we have identified matched counterparts of GLOSTAR sources in the CORNISH and MAGPIS catalogs. Although these surveys were conducted at a higher resolution than the part of the GLOSTAR survey reported here, they were conducted at a similar frequency and can hence provide a useful consistency check on the measured fluxes. Due to the different resolutions, we have compared the GLOSTAR peak fluxes with the integrated fluxes from CORNISH and MAGPIS, except when the sources in those surveys are larger than the GLOSTAR beam in which case we compare the integrated fluxes from both surveys; CORNISH filters out any structures larger than  $\sim 15''$



**Fig. 11.** Comparison of GLOSTAR fluxes with other 5 GHz surveys. *Upper panel:* comparison of the GLOSTAR flux measurements with the Becker et al. (1994) survey (MAGPIS) and CORNISH; data from these are represented as red (filled) and blue (open) circles, respectively. Due to the difference in resolution between the surveys, we compare the integrated CORNISH flux densities with the peak GLOSTAR flux densities, while for the Becker et al. (1994) survey we compare the integrated flux densities with the peak GLOSTAR flux densities, when the Becker et al. (1994) source size is less than  $20''$  and the integrated flux from both surveys if the source is larger than  $20''$ . The gray line is the line of equality. *Lower panel:* flux ratio distribution (as described above), which has a mean of 0.9 and standard deviation of 0.34.

and so this is only a consideration for some of the MAGPIS sources.

In Fig. 11 we show the results of this flux comparison. It is clear that the fluxes from all three surveys are in very good agreement, however, it is noticeable that there is bias towards higher fluxes measured by GLOSTAR. This is to be expected given the lower resolution resulting in more flux being present in the beam and a result of some of the flux being filtered out by the high-resolution surveys due to their poorer sensitivity to larger angular scales. We also note that a handful of CORNISH sources have a significantly higher flux than measured in the GLOSTAR data; these appear as a small cluster of sources above the line of equality towards the lower left corner of the plot and have been previously discussed in Sect. 4.1.1.

## 5. Catalog properties

The final catalog consists of 1575 distinct sources with flux densities above  $4\sigma$ , where  $\sigma$  corresponds to the local noise. We give the names of the sources and the observed parameters in Table 7. The source names are constructed from the peak flux density positions. The position given in Cols. 2 and 3 are the barycentric coordinates determined from the first order moments of the longitude and latitude profiles. The barycentric position generally defines the centroid position of a source. However, if the emission profile of a source is extended and consists of multiple peaks, then this can be offset from the peak position. The mean uncertainties on the peak flux density positions is  $1''.22$  with a standard deviation of  $0''.81$ ; the positional uncertainties associated with individual sources are provided in the full version of Table 7 available at the CDS. Combined with the uncertainty in the map astrometry the mean uncertainty in the catalog positions is  $\sim 2''$ . The signal to noise values are given in Col. 4; these are estimated by dividing the peak flux by the rms at the same position. The flux density values given in Cols. 5–8 have been corrected for clean bias and bandwidth smearing (see Hales et al. 2012 for details) and in Col. 9 we give the  $Y$ -factor, which is the ratio of the integrated and peak flux densities (i.e.,  $S_{\text{int}}/S_{\text{peak}}$ ); this gives an indication of how compact a source is, with unresolved sources having a value of one and very extended sources having values of several 10s. The radius in Col. 10 is determined from the number of pixels associated with the source and is therefore a rather crude estimate of the size of a source particularly for extended structures and should be used with caution. In Cols. 11 and 12 we give the spectral index for compact sources (i.e.,  $Y$ -factor  $< 2$ ) that have a 20 cm counterpart detected by the THOR survey, while in Cols. 13 and 14 are the values calculated from the eight sub-images discussed in Sect. 3.1; this will be described in Sect. 5.2. In Cols. 15 and 16 we give the classification assigned to the source and a reference when available.

### 5.1. Fluxes and angular sizes

In Fig. 12 we show the peak and integrated flux densities distribution of the detected sources. Together with the peak flux density distribution, we also show the integrated flux density distribution of all GLOSTAR sources that are matched to sources previously identified by CORNISH and/or MAGPIS (red histogram overlaid on the upper panel); this shows clearly that the vast majority of 5 GHz sources detected in the GLOSTAR map with a peak flux density above a few  $\text{mJy beam}^{-1}$  have been detected before. The sources not picked up by the previous surveys are all extended and would have been filtered out by CORNISH or in regions not covered by MAGPIS. What is also clear from this plot is that the improved sensitivity of GLOSTAR has resulted in the detections of many previously unknown integrated low-flux density sources.

In Fig. 13 we show the distribution of  $Y$ -factor for the sample; this parameter is the ratio between integrated and peak flux density (i.e.,  $S_{\text{int}}/S_{\text{peak}}$ ) and gives an indication of the source size. While for unresolved sources the  $Y$ -factor should be equal to 1, allowing for measurement uncertainties we classify all sources with a  $Y$ -factor  $< 1.2$  as unresolved (cf. Bühr et al. 2016). We make a further distinction between compact (which includes all of the unresolved sources) and extended sources using a  $Y$ -factor  $= 2$  as dividing line. We note that a significant number of sources have  $Y$ -factors  $< 1$  and this has been discussed in detail by Bühr et al. (2016) who identified three possible explanations for this unexpected result: (1) BLOCAT does not fit enough

**Table 7.** GLOSTAR source catalog.

GLOSTAR name	$\ell$ ( $^{\circ}$ )	$b$ ( $^{\circ}$ )	SNR	$S_{\text{peak}}$ (mJy beam $^{-1}$ )	$\Delta S_{\text{peak}}$	$S_{\text{int}}$ (mJy)	$\Delta S_{\text{int}}$	$Y$ -factor	Radius ( $''$ )	$\alpha$	Spectral index $\Delta\alpha$	$\alpha^{(*)}$	$\Delta\alpha^{(*)}$	Classification Type	Ref. (16)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
G028.091–00.092	28.091	–0.093	11.4	3.12	0.32	4.26	0.35	1.1	16	...	...	–0.07	0.25	...	...
G028.097–00.951	28.097	–0.951	6.1	0.54	0.09	0.45	0.09	0.6	10	...	...	–0.36	0.69	...	...
G028.098–00.781	28.098	–0.781	186.9	17.80	0.96	22.06	1.11	1.2	25	–0.91	0.01	–1.08	0.03	Radio star	8
G028.100+00.644	28.099	0.644	6.5	0.87	0.14	0.76	0.14	0.6	10	...	...	–0.57	0.28	...	...
G028.107–00.301	28.109	–0.300	5.6	0.81	0.15	1.89	0.17	1.3	16	...	...	–0.57	0.62	...	...
G028.109–00.490	28.108	–0.490	20.0	2.84	0.21	2.70	0.20	0.9	15	0.06	0.09	–0.18	0.1	...	...
G028.116–00.368	28.116	–0.368	44.0	5.77	0.34	5.82	0.32	1.0	18	–0.1	0.04	–0.12	0.05	...	...
G028.120+00.419	28.120	0.419	4.9	0.87	0.19	0.64	0.18	0.4	8	...	...	0.15	0.43	...	...
G028.126+00.742	28.126	0.742	33.6	4.26	0.26	3.86	0.23	0.9	17	–0.28	0.04	–0.81	0.1	...	...
G028.127+00.341 <sup>(‡)</sup>	28.127	0.341	58.2	12.98	0.22	51.78	0.45	4.0	...	...	...	...	...	...	...
G028.137–00.194	28.138	–0.194	4.2	0.77	0.19	0.68	0.19	0.4	8	...	...	0.53	0.37	...	...
G028.137–00.567	28.137	–0.566	5.7	1.15	0.21	1.29	0.21	0.7	11	–0.58	0.15	–1.14	0.48	...	...
G028.140+00.329 <sup>(‡)</sup>	28.140	0.329	11.0	4.50	0.43	7.30	1.00	1.6	...	0.23	0.12	–0.52	0.19	HII region	3
G028.144–00.378	28.144	–0.379	8.0	1.08	0.15	1.11	0.15	0.8	12	...	...	–0.46	0.55	...	...
G028.146+00.517	28.145	0.517	8.7	1.21	0.15	1.86	0.17	1.2	16	...	...	–0.66	0.27	...	...
G028.149–00.998	28.149	–0.997	76.9	11.93	0.66	19.71	1.00	1.6	25	...	...	–1.05	0.05	...	...
G028.151+00.164	28.148	0.149	38.6	10.38	0.62	194.53	9.73	18.0	95	...	...	...	...	HII region	1
G028.151–00.958	28.151	–0.957	16.0	2.00	0.16	2.25	0.17	1.0	17	–0.37	0.09	–0.4	0.28	...	...
G028.159–00.024	28.157	–0.020	8.0	2.91	0.40	28.85	1.49	6.6	38	...	...	...	...	Extended/Diffuse	1
G028.159–00.798	28.159	–0.798	29.1	2.97	0.19	2.67	0.17	0.9	16	–0.35	0.05	–0.36	0.23	...	...
G028.160–00.046	28.160	–0.046	4.8	1.68	0.36	1.59	0.36	0.5	9	...	...	–0.15	0.52	HII region	3
G028.163–00.163 <sup>(†)</sup>	28.163	–0.163	7.0	0.50	0.07	2.00	0.34	4.0	...	...	...	...	...	HII region	3
G028.172–00.605	28.173	–0.606	4.9	1.13	0.24	14.55	0.76	5.5	32	...	...	...	...	Ionization Front	1
G028.178+00.074	28.178	0.072	4.3	1.49	0.35	9.56	0.59	2.5	21	...	...	...	...	...	...
G028.187+00.583	28.187	0.583	22.3	3.02	0.21	2.75	0.19	0.9	15	–0.01	0.06	–0.71	0.11	...	...
G028.189–00.744	28.189	–0.744	17.0	1.74	0.14	1.36	0.12	0.7	13	–0.33	0.08	–1.09	0.2	...	...
G028.193–00.785	28.193	–0.785	8.6	0.87	0.11	1.29	0.12	1.1	16	–0.68	0.12	–0.22	0.26	Pulsar	5
G028.197–00.891	28.198	–0.892	4.6	0.48	0.11	0.96	0.11	1.0	13	...	...	–0.94	0.47	...	...
G028.200–00.050	28.191	–0.050	771.1	332.67	17.92	396.10	19.81	1.2	48	...	...	...	...	HII region	3
G028.209+00.226	28.209	0.226	5.8	1.07	0.19	0.95	0.19	0.6	10	...	...	0.22	0.4	...	...

**Notes.** The source names are appended with a <sup>(‡)</sup> to indicate if a source has been split (as described in Sect. 3.2.1) and a <sup>(†)</sup> if the source has been recovered (as described in Sect. 3.2.2). <sup>(\*)</sup>Values calculated using the 8 sub images are described in Sect. 2. Only a small portion of the data is provided here, the full table is available in electronic form at the CDS.

**References.** (1) This work, (2) Kalcheva et al. 2018, (3) Anderson et al. (2014), (4) Giveon et al. (2005a), (5) SIMBAD, (6) Irabor et al. (2018), (7) Ortega et al. (2010), (8) Purcell et al. 2013.

pixels for weak sources, which leads to lower integrated flux densities; (2) unresolved sources are located in slightly negative side lobes from nearby strong sources; (3) weak sources are not cleaned properly. Bühr et al. suggest using the peak flux densities for unresolved sources in any analysis and we have also adopted this recommendation.

BLOCAT gives source size in terms of the number of pixels (parameter  $npix$ ) that are flood-filled in the source extraction process. We convert this into a radial size using

$$R_{\text{source}} = \sqrt{\frac{A}{\pi}}$$

where  $R_{\text{source}}$  is the effective radius, and  $A$  is the area covered by the source calculated by  $npix$  times the pixel size squared, which in this case is  $2''.5 \times 2''.5$ . Figure 14 shows the distribution of the angular radius of the sources.

## 5.2. Spectral index

Astrophysical sources mainly emit two types of radio emission, thermal and non-thermal. The first refers to bremsstrahlung radiation and the latter is also known as synchrotron radiation (see, e.g., Wilson et al. 2012). Thermal bremsstrahlung radiation is produced in H II regions and planetary nebulae and usually has spectral indices between  $-0.1$  and  $2$  depending on whether the emission region is optically thin or thick, respectively. Active galactic nuclei (AGN) galaxies and supernova remnants give rise to synchrotron radiation, which has negative spectral indices, typically between  $-1$  and  $-0.5$  (Condon 1984; Dzib et al. 2013;

Rodríguez et al. 2012 and references therein). The spectral index can, therefore, help to broadly distinguish between different types of radio sources.

As mentioned in Sect. 2.2, the calibration process produced eight sub-images in the frequency range between 4 and 8 GHz, which we used to estimate the spectral index only of the compact sources (i.e.,  $Y$ -factor  $< 2$ ). We used the peak flux density from each sub-image, derived with the CASA task `imfit`, and a  $\chi^2$  linear fit to calculate the spectral index value.

We also estimate the spectral index of the compact sources by combining their peak flux densities with the peak lower frequency flux densities determined by THOR. As previously mentioned, we have matched 711 sources with a THOR source, of which 640 have  $Y$ -factors  $< 2$  and so can be assumed to be relatively compact. The mid-frequencies for THOR and GLOSTAR are 1.5 GHz and 5.8 GHz, respectively. Following Kalcheva et al. (2018) we estimate the spectral index using

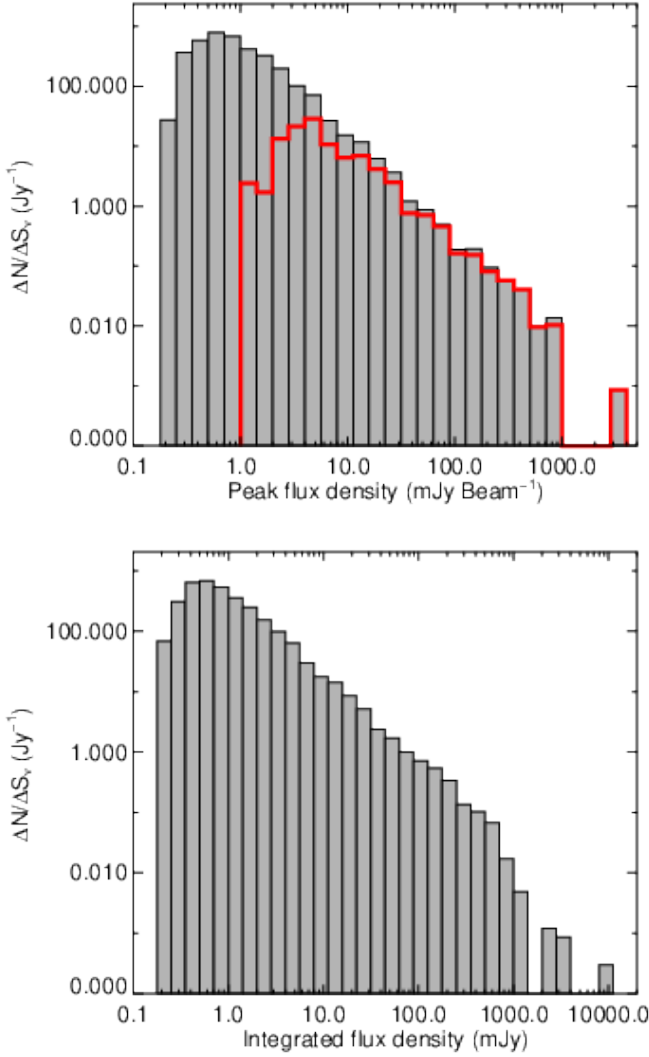
$$\alpha = \frac{\ln(S_{\text{GLOSTAR}}/S_{\text{THOR}})}{\ln(5.8/1.5)},$$

where  $S_{\text{THOR}}$  and  $S_{\text{GLOSTAR}}$  are the peak flux densities at 6 cm and 20 cm, respectively. The uncertainty is calculated using

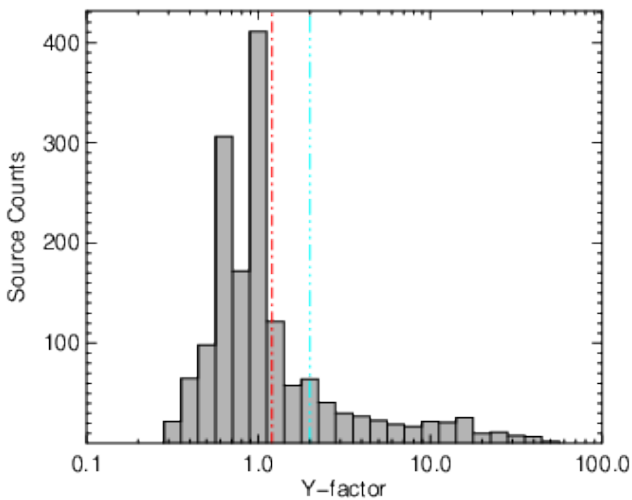
$$\Delta\alpha = \frac{\sqrt{(\sigma_{\text{THOR}}/S_{\text{THOR}})^2 + (\sigma_{\text{GLOSTAR}}/S_{\text{GLOSTAR}})^2}}{\ln(5.8/1.5)}$$

where  $\sigma_{\text{GLOSTAR}}$  and  $\sigma_{\text{THOR}}$  are the uncertainties in the peak flux densities.

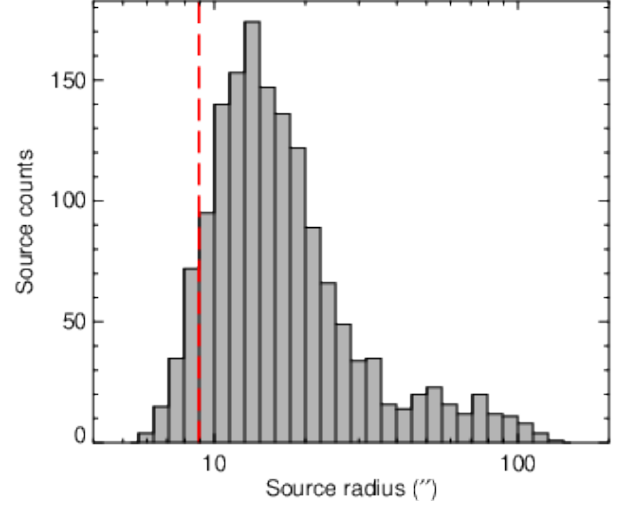
For a comparison of both methods discussed above, we selected sources with reliable GLOSTAR internal spectral



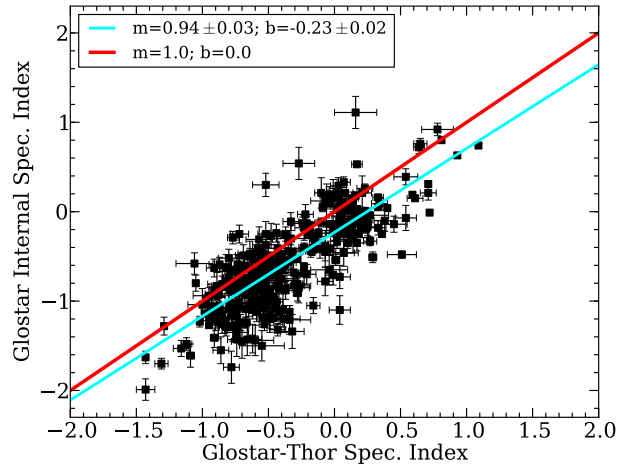
**Fig. 12.** Peak and integrated flux distribution of GLOSTAR sources are shown in the *upper and lower panels*, respectively. The thick red histogram on the *upper panel* flux distribution of sources previously identified by CORNISH and/or MAGPIS. The bin width is 0.15 dex.



**Fig. 13.** Distribution of the Y-factor ( $S_{\text{int}}/S_{\text{peak}}$ ). The red dashed-dotted ( $Y\text{-factor} = 1.2$ ) and cyan dashed-dotted lines ( $Y\text{-factor} = 2$ ) indicate the criterion used to distinguish between unresolved, compact, and extended sources, respectively. The bin size used is 0.1 dex.



**Fig. 14.** Distribution of the source radii. The dashed-dotted red line indicates the resolution of the observations (radio beam). The bin size used is 0.1 dex.

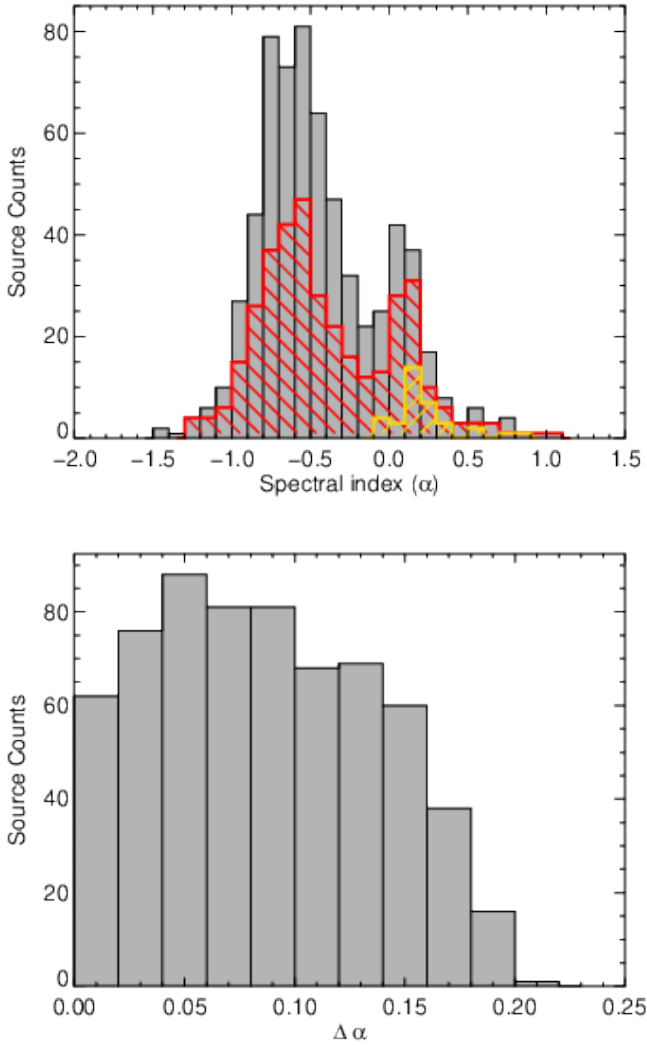


**Fig. 15.** Comparison of spectral index values between only GLOSTAR data with uncertainties less than 0.2, and GLOSTAR combined with THOR data. The  $m$  and  $b$  are the slope and intersection of the linear fit, respectively, and are represented with the blue line. The red line is the line of equality.

indices,  $\sigma_\alpha < 0.2$ , and compared them with their GLOSTAR-THOR spectral indices (see Fig. 15). A linear fit to these values yields that the slope is  $m = 0.94 \pm 0.03$ , close to the slope of the equality line. In conclusion, the spectral index values derived with the two methods are statistically similar.

The spectral index derived from the combination of GLOSTAR and THOR data and their associated uncertainties are given in Cols. 11 and 12 of Table 7. In the upper panel of Fig. 16 we show the distribution of spectral indices; these cover a range of values between  $-1.5$  and  $1$  and have a bimodal distribution with peaks at approximately  $-0.7$  and  $0.1$ . The dip between these two peaks occurs at approximately  $-0.2$ , which we consider the transition from thermal and non-thermal emission. We also show the spectral indices of radio sources that are associated with mid-infrared emission (i.e., have a WISE counterpart) and associated with dust emission (i.e., have an ATLASGAL counterpart); these are shown in red and yellow hatched histograms, respectively. It is clear from a comparison of these distributions that the sources with mid-infrared counterparts span the whole range of spectral indices, while the sources associated with dust





**Fig. 16.** Histograms of the spectral index distribution. *Upper panel:* spectral index distribution of all compact GLOSTAR sources with  $Y$ -factor  $< 2$  and a counterpart detected at 1.5 GHz in the THOR survey. The red and yellow hatched histograms show the distribution of compact radio sources that are associated with mid-infrared sources and dust emission, respectively. *Lower panel:* distribution of the associated uncertainties to the spectral index measurements. The bin sizes used in the upper and lower panels are 0.1 and 0.02, respectively.

are almost exclusively thermal sources. So while sources associated with dust are likely to be a mix of H II regions and PNe, the association of a particular source with mid-infrared emission does not provide any constraints on its nature. We will explore this issue of source types in the next section. In the lower panel of Fig. 16 we show the distribution of the uncertainties associated with the spectral index measurements; these have typical values of  $\pm 0.1$  but can be as large as  $\pm 0.2$ .

We have calculated the spectral indices for 671 GLOSTAR sources without a THOR counterpart. The values and their associated uncertainties are given in Cols. 13 and 14 of Table 7. We note that the limited frequency range or poor signal to noise results in large uncertainties for the majority (74% have  $\Delta\alpha > 0.2$ ). The spectral indices range between  $-1$  and  $1$  for the sources for which they are considered to be reliable. Given that the sample of GLOSTAR-internal spectral index is relatively modest, we decided not to perform a detailed statistical analysis of the spectral index.

## 6. Classification of radio sources

### 6.1. Literature search

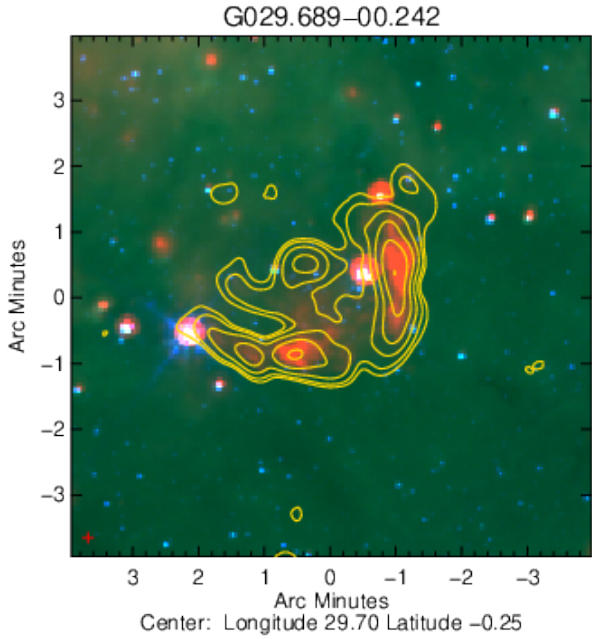
A number of studies have investigated the nature of radio and mid-infrared sources. By correlating our catalog with these we can adopt their classifications. Cross-matching our radio sources with H II region catalogs compiled by Purcell et al. (2013) and Kalcheva et al. (2018) using CORNISH data, Giveon et al. (2005b) using MAGPIS data, and Ortega et al. (2010), we have identified a total of 46 previously known H II regions. As described in Sect. 4.1.5 we also cross-matched our radio sources with a catalog of H II region candidates identified by Anderson et al. (2014) from their analysis of WISE colors. This revealed a match for 130 radio sources including 90 not previously matched to a CORNISH or MAGPIS source. Of these, 43 have spectral indices determined from the THOR and GLOSTAR data; 42 of these have spectral indices between  $-0.17$  and  $0.88$ , which supports their classification as H II regions (Anderson et al. 2014). Combined with the 46 sources identified from the literature, this brings the sample of reliable H II regions in the GLOSTAR region to 136.

A further 28 sources have been identified as PNe by the CORNISH team (Purcell et al. 2013; Irabor et al. 2018), and another 15 as radio stars 12 identified by Purcell et al. (2013) and three found in SIMBAD<sup>8</sup> database. We found, with a cross-match radius of  $10''$ , that three of these candidates have a counterpart in the HASH (Hong Kong/AAO/Strasbourg H $\alpha$ ) planetary nebula database (Parker et al. 2016). Our literature search has also identified another three interesting compact radio sources that are consistent in position with pulsars: G032.763+00.092 with PSR J1850–0006 (Keith et al. 2009), G028.880–00.938 with PSR B1844–04, and G028.193–00.785 with PSR B1842–04 (Hobbs et al. 2004). Radio continuum emission from energetic pulsars is usually associated to the integrated emission of their pulses, but other mechanisms such as a pulsar wind and magnetic activity have been recently discussed by Dzib et al. (2018). For the GLOSTAR data discussed here, the angular resolution is insufficient to decide whether or not the radio emission comes from the pulsar, from the surrounding medium, or from an unassociated object. This will be better constrained with the GLOSTAR VLA B-configuration data when it becomes available.

The importance of pulsars with radio continuum emission is that they may be targeted by very long baseline interferometry (VLBI) observations to determine their astrometric parameters, such as proper motions and trigonometric parallaxes (Briskin et al. 2000). This information may be used to associate them with supernova remnants (SNRs) to better constrain their luminosity, and Galactic electron distribution models (Briskin 2005; Deller et al. 2009).

We also identified two SNRs from a cross-matching with a radius of  $10''$  between the GLOSTAR sources and the THOR  $\sim 20$  SNR candidates identified by Anderson et al. (2017) located in the GLOSTAR field; these are G028.642+00.187 and G028.218–00.080. The low number of matches may be due to the negative spectral index of the SNRs, which are better recovered at lower frequencies, or may be due to the fact that SNRs tend to be extended and so many are likely to be either filtered out by the interferometer entirely or only partially detected and classified as extended emission. To investigate this further we also searched for SNRs using the Green (2014) catalog. This identifies 12 SNRs within the GLOSTAR region, ten of which are matched with the

<sup>8</sup> Set of Indications, Measurements, and Bibliography for Astronomical Data.



**Fig. 17.** Compact SNR G029.689–00.242 identified from the Green (2014) catalog (G029.7–00.3 in their catalog). The background image is a false color composite image produced by combining the GLIMPSE 5.8 and  $8.0\ \mu\text{m}$  bands and the MIPS GAL  $24\ \mu\text{m}$  band; these are shown in blue, green, and red, respectively. The yellow contours show the distribution of the 5.8 GHz radio emission associated with the GLOSTAR source.

large scale emission regions identified in Table 2 and mentioned in Sect. 3.2.3. Of the remaining two SNRs in the field, one is not detected at all (G032.1–00.9), and one is detected as a compact radio source (G029.7–00.3; see Fig. 17 for an image of this compact SNR). The one SNR not recovered in our map happens to be the largest identified by Green (2014) in the GLOSTAR field, with a size of  $40' \times 40'$ , and so its surface brightness may have fallen below the GLOSTAR detection threshold.

## 6.2. Analysis of mid-infrared images

The correlation of compact radio emission with dust and mid-infrared emission is an indicator of the presence of H II regions and there have been a number of studies that have used this correlation to identify samples of H II regions (e.g., Urquhart et al. 2013, 2009, 2007; Thompson et al. 2006). In addition to identifying and removing spurious sources (as discussed in Sect. 4.2) we have, therefore, also used the mid-infrared images to distinguish between the nature of different kinds of radio emission. In the following two subsections we will give a brief description of the rationale used to define these different kinds of sources. To do this, we split the remaining unclassified radio sources into extended sources ( $Y$ -factor  $> 2$ ) and compact sources ( $Y$ -factors  $< 2$ ) because the size and structure of the radio emission itself can provide some useful clues to their nature.

We indicate the emission type attributed to each source and a reference if the identification has come from the literature in the last column of Table 7. In Table 8 we provide a summary of the number of each type identified.

### 6.2.1. Classification of extended radio sources ( $Y$ -factor $> 2$ )

There are 225 radio sources in the catalog that have  $Y$ -factors  $> 2$  that have not been previously classified (see Sect. 6.1 for details).

**Table 8.** Summary of emission types.

Description	Number
Ionization fronts	37
Extended emission sources	62
Planetary nebulae	46
H II regions	231
Radio stars	15
Pulsars	3
SNR	3 (10)

**Notes.** The number in parentheses in the SNR column indicates the number of SNRs identified as large scale structures in Sect. 3.2.3.

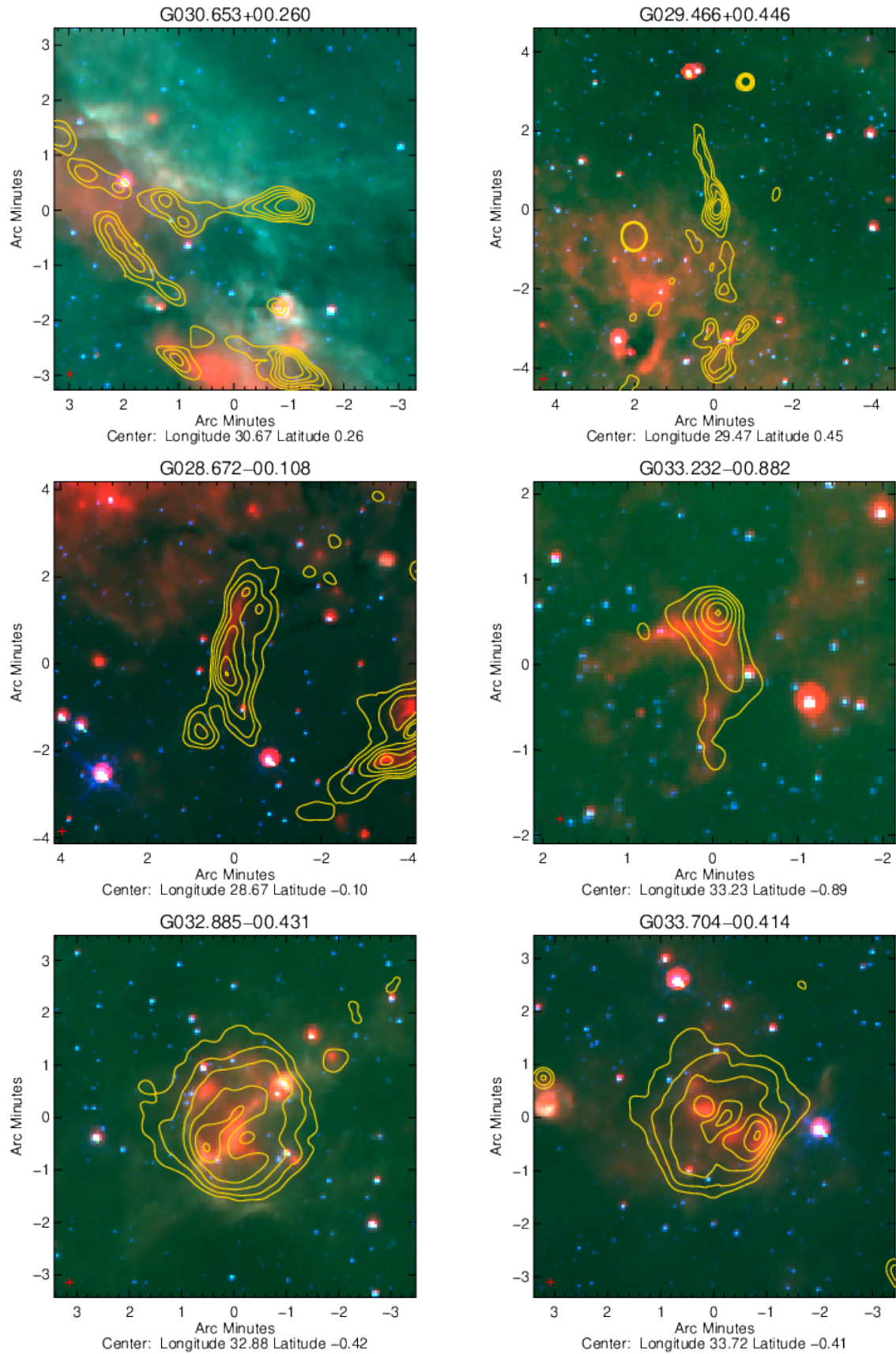
We can assume that any extended radio source is also likely to have thermal bremsstrahlung emission associated with evolved H II regions. In some cases, although the extended radio emission is positionally coincident with the mid-infrared emission, the morphological correlation is poor, and in these cases we classify the emission more loosely as being associated with larger scale extended structures, the nature of which is unclear. In other cases, the radio emission appears to be associated with a section of an ionized rim of an evolved H II region and these are classified as ionization fronts. Although in these cases the radio emission is likely to be associated with an H II region, we are only capturing part of the emission and want to make the distinction between compact H II regions, where we can be confident we are capturing all of the emission, and more extended sources where we are only capturing a less well determined fraction of the total integrated flux density.

We have classified 57 radio sources as being associated with large scale emission and a further 31 sources have been associated with ionization fronts. We have classified 71 of the extended radio sources as H II regions. In Fig. 18 we present three-color composite images of two examples of each of these three morphological types of structures. We classify another two sources as potential PNe due to the compact nature of the mid-infrared emission and relative isolation and lack of any nebulosity typically observed towards star-forming regions. The nature of the remaining 68 extended radio sources that are not classified as either extended emission, ionization front, or H II region is still unclear.

### 6.2.2. Classification of compact radio sources ( $Y$ -factor $< 2$ )

There are 1284 compact radio source in the catalog, of which 133 have been previously classified. In the previous section, we have seen that the spectral index is useful for separating the compact radio sources into thermal and non-thermal bremsstrahlung emission. However, PNe and some extragalactic radio sources, known as high frequency peakers (Dallacasa et al. 2000), can have positive spectral indices and so this alone does not definitively identify H II regions. From an inspection of the mid-infrared and submillimeter dust emission towards all of the remaining compact sources (1151), we have identified another 20 H II region candidates. The mid-infrared emission towards these sources generally reveals a bubble-like or complex morphology and is associated with compact dust emission.

We have identified a further 16 sources as PNe candidates due to the correlation between the compact radio and mid-infrared emission and the lack of any evidence of dust emission from the ATLASGAL maps. This will be further explored via the information provided by the GLOSTAR RRL data. We have also classified six more sources as being associated with



**Fig. 18.** Examples of radio sources classified according to their morphology. *Upper, middle, and lower panels:* examples of sources classified as extended radio sources, ionization fronts, and evolved H II regions, respectively. In all cases the background image is a false color composite image produced by combining the GLIMPSE 5.8 and 8.0  $\mu\text{m}$  bands and the MIPS GAL 24  $\mu\text{m}$  band; these are shown in blue, green, and red, respectively. The yellow contours show the distribution of the GLOSTAR 5.8 GHz radio emission.



an ionization front and four others as being extended emission; although in these cases the radio emission is compact, it is coincident with a strong and/or extended  $8\ \mu\text{m}$  emission. The nature of the remaining 1105 that are not classified is uncertain. The mean value of the spectral index, for those 523 with this parameter, is  $\sim -0.5$  with a standard deviation of 0.3. This suggests that the vast majority of these unclassified sources are likely to be extragalactic background objects.

To calculate the number of extragalactic background objects expected in the surveyed area ( $57\ 600\ \text{arcmin}^2$ ), we have followed Fomalont et al. (1991). These authors observed an area free of stars, and showed that the number of background objects at 5 GHz above a flux  $S$ , is described by

$$\left(\frac{N}{\text{arcmin}^2}\right) = (0.42 \pm 0.05) \left(\frac{S}{30\ \mu\text{Jy}}\right)^{-1.18 \pm 0.19}.$$

Even though our observed frequency is slightly different, this study is the nearest in frequency of this kind to ours. Thus we may use their results for an approximation. Using  $S = 450\ \mu\text{Jy}$ , that is, three times the mean noise level, the number of expected background sources in our map is  $990 \pm 640$ . These numbers also suggest that most of the 1151 unclassified sources are extragalactic radio sources.

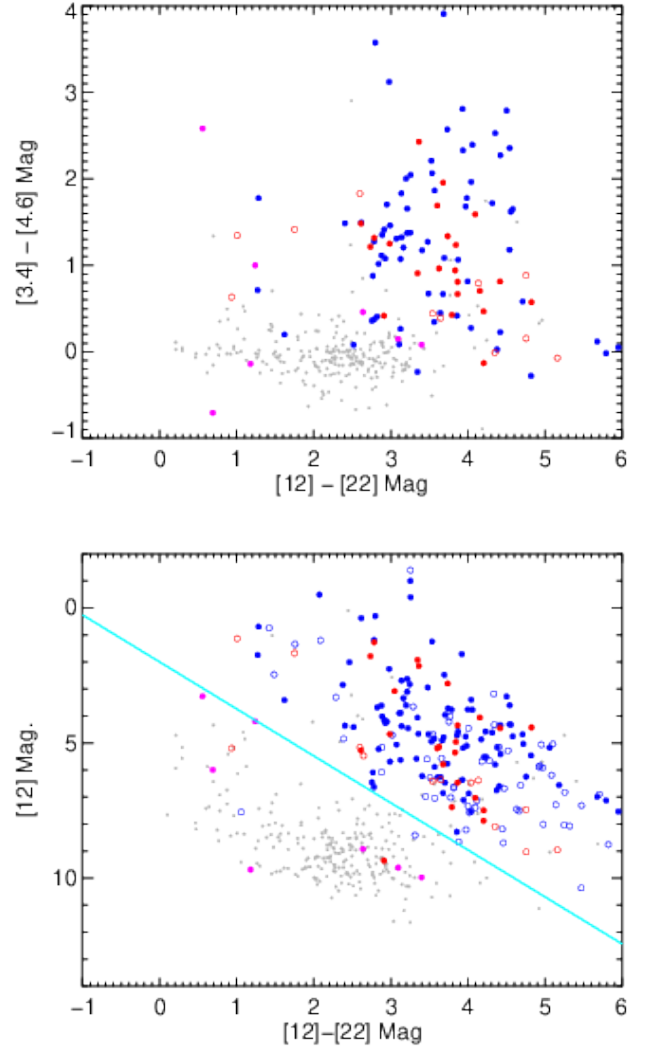
### 6.3. Mid-infrared colors of radio sources

We have identified WISE counterparts for 988 radio sources from the WISE all sky point source catalog and will use these to investigate their mid-infrared color. Of these only 470 have reliable fluxes in all four bands. This increases to 475 when considering the three longest wavelength bands. We will use these to examine the colors of previously identified H II regions in an effort to confirm the classifications made and to look for possible trends that can be used to define a set of criteria to identify new H II region candidates. In Fig. 19 we present color-color and color-magnitude diagrams showing the distribution of all H II regions and PNe identified in the literature (filled blue and red circles, respectively) and from our visual inspection of the mid-infrared images (open blue and red circles, respectively). We also show the distributions of the radio stars (filled magenta circles). The smaller gray circles show the distribution of compact sources with a WISE counterpart that we have been unable to classify.

Inspection of these plots shows that the locations in color-color and color-magnitude diagrams of H II regions and PNe identified in our work agree very well with those published in the literature, providing confidence that our classifications are robust. The unclassified radio sources (shown as gray filled circles) and radio stars almost exclusively occupy the same region of the parameter space as field stars and background objects (cf. Figs. B2 and B3 in Csengeri et al. 2014) while the embedded and dust enshrouded objects (H II regions and PNe) have redder colors and are clearly separated from the field stars. This is particularly clear in the color-magnitude plot shown in the lower panel of Fig. 19, where all of the H II regions and the vast majority of the PNe are located in an area of parameter space defined by

$$[12]\ \text{Mag.} < 1.74 \times ([12] - [22]) + 2. \quad (1)$$

This is also indicated on the lower panel of Fig. 19 with a cyan line. Only two PNe (G029.493+00.931 and G033.120-00.894) and two H II regions (G030.246-00.91 and G033.704-00.414) are found to the left of this color criterion. One of the PNe was identified by the CORNISH team, while the other PN and both H II regions have been classified by us. We note that the latter of the two H II regions is very evolved, with a number of WISE sources being located within its boundary.



**Fig. 19.** Mid-infrared colors of compact GLOSTAR sources associated with a WISE counterpart. Sources classified as H II regions from the literature are shown as filled blue circles, while the H II classified in this study are shown as open blue circles. Similarly, PNe identified by Purcell et al. (2013) are shown as red filled circles, while PNe identified in this paper are shown as open circles. Radio stars identified by Purcell et al. (2013) are indicated by filled magenta circles. The smaller gray circles show the distribution of compact sources with a WISE counterpart that we have been unable to classify.

For the other three sources it is possible that these are the result of chance alignments between radio and mid-infrared sources, but it is also possible that the colors of these objects cover a larger range or that the classifications themselves are wrong. The sample is too small to test any of these possibilities statistically, but we note that the one of the PNe identified by the CORNISH team that falls on the wrong side of this criterion (G033.120-00.894) is positionally coincident with the WISE source J185457.96-001424.3. The positional offset is  $\sim 3''$  and the photometric flux qualities are classified as AABB, and so the correlation and data are reliable<sup>9</sup>.

<sup>9</sup> The quality of the profile-fit photometry measurement is indicated by a four letter code (one for each band [W1/W2/W3/W4]; see Sect. 4.1.5). The letters “A” and “B” correspond to the sources detected in a particular band with a flux signal-to-noise ratio  $> 10$  and between 3 and 10, respectively. Detections with a quality flag of “A” and “B” are considered to be reliable.



**Table 9.** Catalog of new compact H II region candidates.

GLOSTAR name	$S_\nu$ (mJy beam <sup>-1</sup> )	$\Delta S_\nu$	$\alpha$	$\Delta\alpha$	Sub-mm emission	Mid-IR emission	UC H II candidate
G029.367–00.317	2.26	0.24	...	...	✓✓	✓✓	✗
G030.154+00.583	0.55	0.13	...	...	✓	✓	✓
G030.241–00.592	1.73	0.18	–0.12	0.16	✗	✓	✗
G030.371+00.483 <sup>(†)</sup>	3.47	0.33	...	...	✓	✓	✓
G030.678+00.767 <sup>(†)</sup>	3.95	0.26	0.07	0.08	✗	✓	✓
G032.051–00.091	2.12	0.24	–0.34	0.15	✓✓	✓	✗
G032.460+00.386	0.67	0.16	...	...	✓✓	✓✓	✗
G033.714+00.256	1.50	0.22	...	...	✓	✓	✓
G034.274–00.151	1.01	0.16	...	...	✓	✓	✓
G034.338–00.694	0.49	0.12	...	...	✓	✓	✓
G034.440+00.060	1.53	0.35	...	...	✓✓	✓	✓
G035.344+00.348	0.46	0.11	...	...	✓✓	✓	✓

**Notes.** Sources that are bright enough to have been detected by CORNISH are indicated by appending a <sup>(†)</sup> to the source name. The fluxes and spectral indices are drawn from the GLOSTAR catalog (i.e., Table 7). In the last two columns we indicate the presence of compact (✓), extended (✓✓), or absence (✗) of any corresponding ATLASGAL or WISE emission. In the last column we indicate which are still UC H II region candidates after analysis (see text for details).

#### 6.4. New compact H II regions

Our analysis has identified a total of 231 H II regions, of which 46 were previously known (Kalcheva et al. 2018; Ortega et al. 2010; Giveon et al. 2005b) and 90 were proposed as candidates from analysis of the WISE data by Anderson et al. (2014). We have therefore identified 96 new H II regions and confirmed the nature of those proposed by Anderson et al. (2014). The majority of the new H II regions (84) have  $Y$ -factors  $> 1.2$  and are therefore resolved in GLOSTAR, and are unlikely to have been picked up in earlier high-resolution surveys where the extended emission is filtered out (CORNISH and MAGPIS). The remaining 12 H II region candidates are unresolved in GLOSTAR and so could have been picked up in earlier surveys. However, their peak fluxes are lower than the sensitivity limit of CORNISH ( $7\sigma \sim 2.7$  mJy beam<sup>-1</sup>) except in the case of two sources. We checked the CORNISH maps of these two sources but did not find any emission present coincident with the GLOSTAR position (the noise values are 0.34 mJy and 0.32 mJy for G030.371+00.483 and G030.678+00.767, respectively). Both of these sources have mid-infrared counterparts and are hence real. These sources are therefore either both extended and completely filtered out in the CORNISH data or are variable radio sources.

In Table 9 we give the peak flux densities and spectral indices of the new unresolved H II regions; only three sources have a detection in THOR and consequently we only have a measurement for the spectral index for three of these H II regions, all of which are consistent with what is expected for optically thin H II regions (i.e.,  $\alpha = -0.1$ ). This suggests that these are more evolved than the UC H II region stage, which are usually characterized by optically thick emission ( $\alpha \sim 1-2$ ; Kurtz & Hofner 2005). We note that no dust emission is detected towards two of these H II regions (G030.241–00.592 and G030.678+00.767), which is also consistent with them being more evolved having already dissipated their natal clump. We further note that the latter of these is also one of the bright sources we would have expected to have been detected by CORNISH, and so the spectral index and absence of dense envelope are consistent with this source being quite evolved and perhaps explains its non-detection in CORNISH.

In Table 9 we also list the new H II regions association with mid-infrared and submillimeter emission. As can be seen from

this table, all of these sources are associated with mid-infrared emission and the majority are associated with dust emission. The mid-infrared emission looks to be extended in two cases indicating these are likely to be quite evolved H II regions ( $>12''$  in diameter)<sup>10</sup>. We have searched in SIMBAD database towards all of these sources. This revealed that G035.344+00.348 had been previously identified as an H II region (Urquhart et al. 2011; MSX6C G035.3449+00.3474). In the last column of Table 9 we identify the H II regions that are likely to be good UC H II region candidates; this is done by eliminating sources that are not associated with dust, have spectral indices indicating optically thin emission, or are extended in the mid-infrared, all of which point to sources that are more evolved than the UC H II phase.

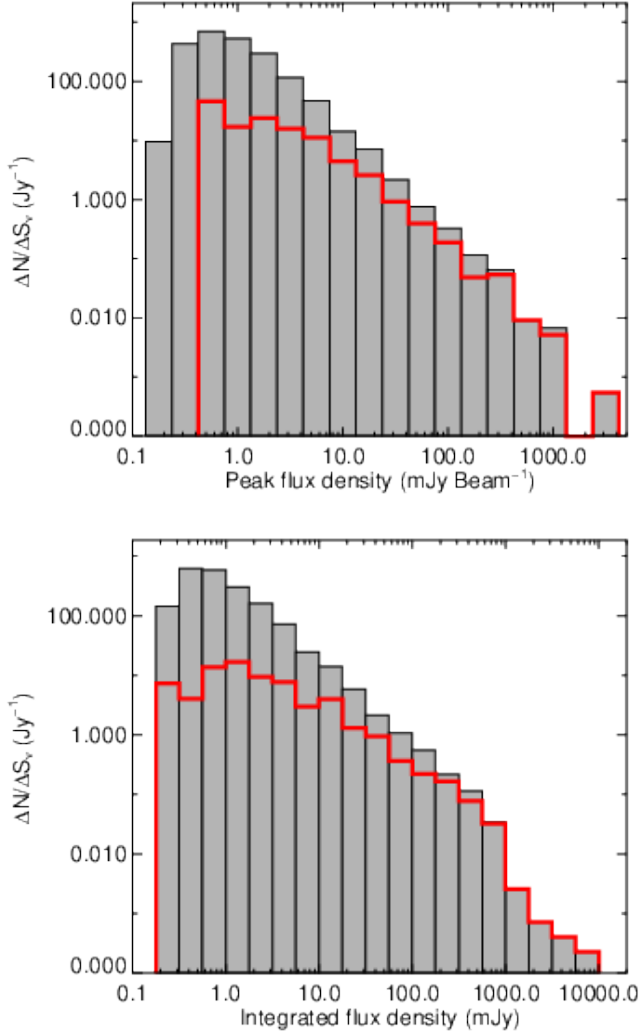
## 7. Properties and Galactic distribution of the H II regions

From a combination of visual inspection of mid-infrared images and cross-matching with other studies, we have identified a total of 231 H II regions in our catalog. This is the most complete sample of H II regions identified in this part of the sky to date, incorporating H II regions with sizes of a few arcseconds to a few arcminutes. In this section we will provide an analysis of their properties and distributions.

### 7.1. Observed and measured properties

In Fig. 20 we present plots of the peak and integrated flux density for the whole GLOSTAR catalog and of the H II regions identified. By comparing these two samples we see that nearly all of the very bright radio sources are associated with H II regions, but that the fraction of the population decreases as the flux decreases and turns over between 0.5 and 2 mJy beam<sup>-1</sup> and between 1–2 mJy for the peak and integrated flux densities, respectively. This indicates that there are relatively fewer weak H II regions in comparison to the total number of radio sources detected at low surface

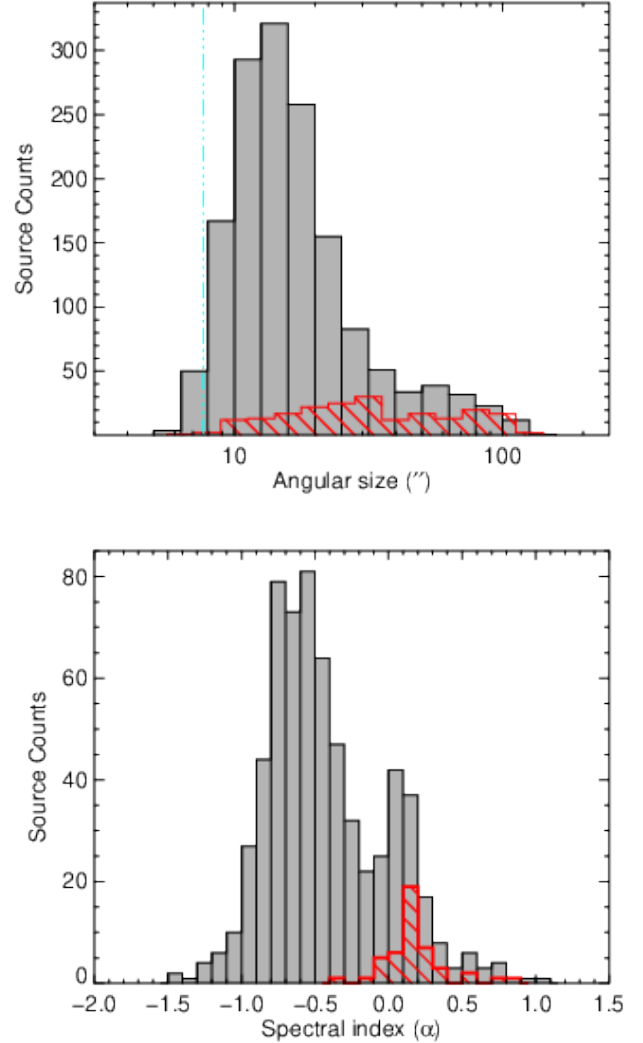
<sup>10</sup> We have made these associations with dust and mid-infrared emission from inspection of the ATLASGAL and WISE maps as not all emission is captured in the catalogs. All but one of these H II regions has a counterpart in WISE and five have a counterpart in the ATLASGAL CSC.



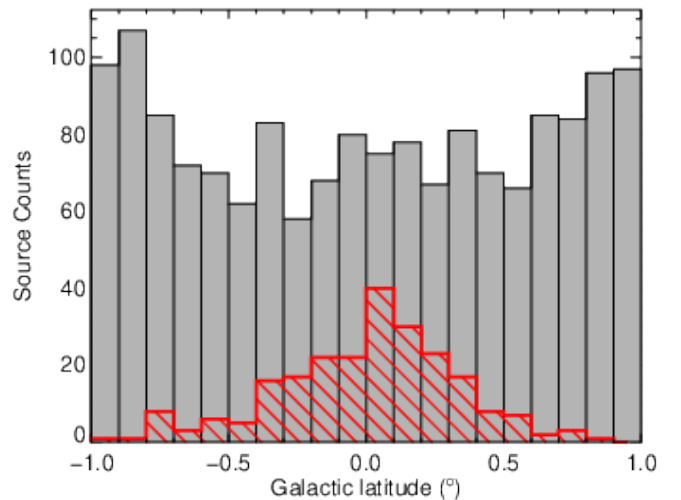
**Fig. 20.** Distribution of H II region properties. *Upper and lower panels:* peak and integrated flux density for the whole GLOSTAR catalog (gray) and those classified as H II regions (red). The bin sizes in both plots is 0.25 dex.

brightnesses. The fact that we find no H II region with fluxes below the turnover in the peak flux density may indicate a minimum flux density threshold for H II regions. However, this corresponds to a flux density of  $\sim 0.4$  mJy beam $^{-1}$ , which is close to the  $3\text{--}4\sigma$  noise found towards the Galactic mid-plane where the vast majority of H II regions are to be found. Radio sources detected below this threshold are primarily extragalactic background sources located away from the mid-plane where the noise level drops off significantly (see lower panel of Fig. 3).

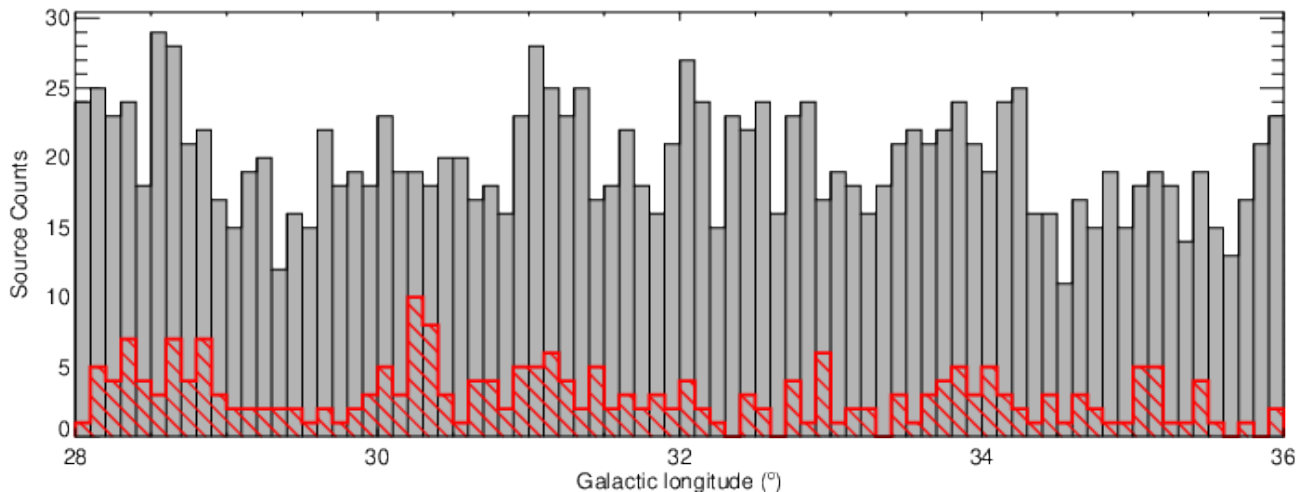
In Fig. 21 we compare the source sizes and spectral indices for the whole GLOSTAR sample and those identified as H II regions. It is clear from the upper panel in this figure that the vast majority of the radio sources are unresolved, with the numbers dropping off steeply for sizes larger than  $\sim 15''$ . The distribution of the H II region sizes appears to be roughly uniform over the whole range of sizes, which is consistent with other studies (e.g., Kurtz et al. 1999; Masqué et al. 2017), and is to be expected given that the sample is likely to cover a larger range of distances and ages. We also note that the H II regions contribute a significant fraction of the radio sources with sizes above  $\sim 30''$ , with the others being classified as ionization fronts or extended emission sources. We find very few radio sources with sizes larger than  $\sim 100''$ , but this is a consequence of the interferometric snapshot



**Fig. 21.** Distribution of H II region properties. *Upper and lower panels:* angular radius and spectral index for the whole GLOSTAR catalog (gray) and those classified as H II regions (red). The dashed-dotted vertical line drawn on the upper panel shows the radius of the beam. The bin sizes used in the upper and lower panels are 0.1 dex and 0.1, respectively.



**Fig. 22.** Distribution of GLOSTAR radio sources as a function of Galactic latitude. The whole catalog is shown as the gray histogram, while the sources identified as H II regions are shown as the hatched red histogram. The bin size used for both distributions is  $0.1^\circ$ .



**Fig. 23.** Distribution of GLOSTAR radio sources as a function of Galactic longitude. The whole catalog is shown as the gray histogram while the sources identified as H II regions are shown as the hatched red histogram. The bin size used for both distributions is  $0.1^\circ$ .

observations themselves as the poor  $uv$ -coverage puts limits on the angular size of structures that can be reliably imaged.

The spectral index distribution of the whole sample is bimodal and peaks at  $\alpha = -0.6$  and  $0.1$ ; the former is likely to be associated with synchrotron emission from the extragalactic background population, while the latter is likely to be associated with thermal emission from Galactic sources such as H II regions and PNe. It is therefore reassuring to see the distribution of H II regions correlating with the thermal emission peak. This is comparable with the spectral index calculation of the THOR team (Wang et al. 2018), but our sources show a higher density of positive index.

## 7.2. Galactic distribution

The majority of the detected radio sources are likely to be extragalactic, and are expected to be smoothly distributed across the whole sky. However, the H II region population should show significant variations in Galactic longitude and latitude, with peaks that correlate with the most active star-forming complexes and the distribution of molecular material, both of which are concentrated in the Galactic mid-plane.

In Fig. 22 we show the Galactic latitude distribution of all the radio sources and the H II region subsample. It is clear from this plot that the H II regions are indeed tightly concentrated around the Galactic mid-plane with a peak at  $b \sim 0^\circ$ , which is consistent with previous studies of H II regions (Urquhart et al. 2013; Kalcheva et al. 2018) and dense, massive star-forming clumps (Urquhart et al. 2014b). The distribution of the full sample is relatively flat for the inner part of the mid-plane but rises close to the edges of the surveyed region ( $|b| \sim 1^\circ$ ), which is probably related to the reduction in the local noise value as one moves away from the Galactic mid-plane where much of the extended emission from extended H II regions and supernova remnants is located (see Figs. 1 and 3).

In Fig. 23 we show the distribution of all GLOSTAR sources and the H II regions as a function of the Galactic longitude. Before we discuss the longitude distribution we should bear in mind that since we have excluded emission associated with the 27 large scale structures, many of which are associated with active H II regions, the peaks seen in the H II region distribution are likely to be somewhat suppressed. All of these large scale structures are Galactic in origin and are unlikely to affect the distribution of the extragalactic population, which is rarely resolved.

The distribution of the whole catalog can be described as being broadly flat. As expected, the distribution of H II regions reveals some significant peaks. We note the strong peak near  $\ell \sim 30^\circ$ , which is related to emission from the massive star-forming regions W43-main and W43-south. This region is located where the Scutum-Centaurus arm interacts with the end of the Galactic long bar (Rodríguez-Fernández & Combes 2008; Nguyen Luong et al. 2011), which is expected to lead to a rise in star formation activity in this area. The latitude and longitude distributions are consistent with the distribution of the high-reliability ( $\sigma \geq 7$ ) CORNISH catalog (Purcell et al. 2013) and the lower frequency THOR catalog (Bühr et al. 2016; Wang et al. 2018), which provides confidence that the catalog we have produced is reliable.

## 8. Summary and conclusions

The GLOSTAR survey is a VLA survey of the Galactic mid-plane first quadrant ( $-10^\circ < \ell < 80^\circ$  and  $|b| < 1^\circ$ ). In this paper we present the radio continuum data for a 16-square-degree part of the survey covering eight degrees in longitude and two degrees in latitude and centered on  $\ell = 32^\circ$ ,  $b = 0^\circ$  (i.e.,  $28^\circ < \ell < 36^\circ$  and  $|b| < 1^\circ$ ). We use the BLOCAT source extraction algorithm to produce a catalog of radio sources and identify a number of different types of morphologies, from unresolved to compact and extended.

We have also identified 27 large scale structures that are over-resolved at the GLOSTAR resolution ( $\sim 18''$ ) and where the emission has been broken up into multiple radio sources. In these cases we have regrouped the various fragments and give the parameters for these objects separately. The individual fragments are removed from the final catalog, which consists of 1575 distinct radio sources. The initial catalog has been cross-matched with other radio and mid-infrared catalogs to verify that the extracted sources are real. We have visually examined mid-infrared and submillimeter images for radio sources without a counterpart to identify and remove artifacts. We discuss the observed properties of this new catalog of sources and compare them to previous studies.

We have used multiwavelength analysis and searched the literature to identify the nature of the radio sources. This has resulted in the identification of 231 H II regions. Of these, only 46 had been previously identified with high-resolution radio

surveys, and a further 90 H II region candidates identified in the mid-infrared by Anderson et al. (2014) have now been confirmed. We have identified a further 96 H II regions not previously known. The majority of the new H II regions are well resolved (71 with  $Y$ -factor  $> 2$ ) and have been missed by previous high-resolution interferometric surveys due to their lack of short baseline data and consequent poor sensitivity to large angular scales. Twelve of the new H II regions identified are considered to be unresolved (i.e.,  $Y$ -factor  $< 1.2$ ), and of these we identify eight UCH II region candidates.

We have investigated the properties of the full sample of H II regions and compared them to the previously identified samples. The improved flux sensitivity ( $\sim 100$ – $400 \mu\text{Jy beam}^{-1}$ ) and sensitivity to larger angular scales ( $\sim 20$ – $200''$ ) of GLOSTAR compared with other surveys has increased the number of H II regions in this part of the Galaxy by a factor of four.

In addition to the H II regions, we have also identified 15 radio stars, 46 PNe, three pulsars, three compact SNR-candidates, and 99 radio sources associated with more extended structures such as the ionization fronts of much more extended H II regions. We are unable to classify 1186 radio sources; these can be broadly categorized as being weak and unresolved, the majority of which are likely to be extragalactic background sources.

Here we present a catalog of 1575 radio sources and we have investigated their nature and physical properties. The data presented here is only a small fraction of the total GLOSTAR coverage, which also includes spectral line and higher resolution data obtained with the VLA in B-array. This will be complemented with observations that are being performed with the 100-m Effelsberg telescope, which will provide sensitivity to scales larger than those seen in the D-configuration observations.

*Acknowledgements.* We would like to thank the referee for constructive comments and suggestions that have helped to improve this work. S.-N.X.M. is a member of the International Max-Planck Research School at the Universities of Bonn and Cologne (IMPRS). N.R. acknowledges support from the Infosys Foundation through the Infosys Young Investigator grant. H.B. acknowledges support from the European Research Council under the Horizon 2020 Framework Program via the ERC Consolidator Grant CSF-648505. This research was partially funded by the ERC Advanced Investigator Grant GLOSTAR (247078). It made use of information from the ATLASGAL database at [http://atlasgal.mpi-fr-bonn.mpg.de/cgi-bin/ATLASGAL\\_DATABASE.cgi](http://atlasgal.mpi-fr-bonn.mpg.de/cgi-bin/ATLASGAL_DATABASE.cgi) supported by the MPIFR, Bonn, as well as information from the CORNISH database at <http://cornish.leeds.ac.uk/public/index.php>, which was constructed with support from the Science and Technology Facilities Council of the UK. This publication also makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. We have used the collaborative tool Overleaf available at: <https://www.overleaf.com/>

## References

Aguirre, J. E., Ginsburg, A. G., Dunham, M. K., et al. 2011, *ApJS*, 192, 4  
 Anderson, L. D., Bania, T. M., Balsler, D. S., et al. 2014, *ApJS*, 212, 1  
 Anderson, L. D., Wang, Y., Bihr, S., et al. 2017, *A&A*, 605, A58  
 Becker, R. H., White, R. L., Helfand, D. J., & Zoonematkermani, S. 1994, *ApJS*, 91, 347  
 Becker, R. H., Helfand, D. J., White, R. L., & Proctor, D. D. 2010, *AJ*, 140, 157  
 Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393  
 Beuther, H., Tackenberg, J., Linz, H., et al. 2012, *ApJ*, 747, 43  
 Beuther, H., Bihr, S., Rugel, M., et al. 2016, *A&A*, 595, A32  
 Bihr, S., Beuther, H., Ott, J., et al. 2015, *A&A*, 580, A112  
 Bihr, S., Johnston, K. G., Beuther, H., et al. 2016, *A&A*, 588, A97

Blum, R. D., Daminieli, A., & Conti, P. S. 1999, *AJ*, 117, 1392  
 Bondi, M., Ciliegi, P., Zamorani, G., et al. 2003, *A&A*, 403, 857  
 Briskeen, W. 2005, *ASP Conf. Ser.*, 340, 489  
 Briskeen, W. F., Benson, J. M., Beasley, A. J., et al. 2000, *ApJ*, 541, 959  
 Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, *PASP*, 121, 76  
 Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, *PASP*, 121, 213  
 Clark, D. H., & Parkinson, J. H. 1975, *Nature*, 258, 408  
 Condon, J. J. 1984, *ApJ*, 287, 461  
 Contreras, Y., Schuller, F., Urquhart, J. S., et al. 2013, *A&A*, 549, A45  
 Cotton, W. D. 2008, *PASP*, 120, 439  
 Cotton, W. D., Condon, J. J., Kellermann, K. I., et al. 2018, *ApJ*, 856, 67  
 Csengeri, T., Urquhart, J. S., Schuller, F., et al. 2014, *A&A*, 565, A75  
 Dallacasa, D., Stanghellini, C., Centonza, M., & Fanti, R. 2000, *A&A*, 363, 887  
 Deller, A. T., Tingay, S. J., & Briskeen, W. 2009, *ApJ*, 690, 198  
 Dzib, S. A., Loinard, L., Mioduszewski, A. J., et al. 2013, *ApJ*, 775, 63  
 Dzib, S. A., Rodríguez, L. F., Karuppusamy, R., Loinard, L., & Medina, S.-N. X. 2018, *ApJ*, 866, 100  
 Elia, D., Molinari, S., Schisano, E., et al. 2017, *MNRAS*, 471, 100  
 Fomalont, E. B., Windhorst, R. A., Kristian, J. A., & Kellerman, K. I. 1991, *AJ*, 102, 1258  
 Garay, G., Rodríguez, L. F., & van Gorkom, J. H. 1986, *ApJ*, 309, 553  
 Giveon, U., Becker, R. H., Helfand, D. J., & White, R. L. 2005a, *AJ*, 129, 348  
 Giveon, U., Becker, R. H., Helfand, D. J., & White, R. L. 2005b, *AJ*, 130, 156  
 Green, D. A. 2014, *Bull. Astron. Soc. India*, 42, 47  
 Greisen, E. W. 2003, *Astrophys. Space Sci. Lib.*, 285, 109  
 Hales, C. A., Murphy, T., Curran, J. R., et al. 2012, *MNRAS*, 425, 979  
 Helfand, D. J., Becker, R. H., White, R. L., Fallon, A., & Tuttle, S. 2006, *AJ*, 131, 2525  
 Hoare, M. G., Purcell, C. R., Churchwell, E. B., et al. 2012, *PASP*, 124, 939  
 Hobbs, G., Lyne, A. G., Kramer, M., Martin, C. E., & Jordan, C. 2004, *MNRAS*, 353, 1311  
 Holwerda, B. W. 2005, ArXiv e-prints [arXiv:astro-ph/0512139]  
 Huynh, M. T., Jackson, C. A., Norris, R. P., & Prandoni, I. 2005, *AJ*, 130, 1373  
 Irlabor, T., Hoare, M. G., Oudmaijer, R. D., et al. 2018, *MNRAS*, 480, 2423  
 Kalcheva, I. E., Hoare, M. G., Urquhart, J. S., et al. 2018, *A&A*, 615, A103  
 Keith, M. J., Eatough, R. P., Lyne, A. G., et al. 2009, *MNRAS*, 395, 837  
 Kennicutt, R. C. 2005, *IAU Symp.*, 227, 3  
 Kurtz, S., & Hofner, P. 2005, *AJ*, 130, 711  
 Kurtz, S. E., Watson, A. M., Hofner, P., & Otte, B. 1999, *ApJ*, 514, 232  
 Lockman, F. J. 1989, *ApJS*, 71, 469  
 Malarecki, J. M., Jones, D. H., Saripalli, L., Staveley-Smith, L., & Subrahmanyan, R. 2015, *MNRAS*, 449, 955  
 Masqué, J. M., Rodríguez, L. F., Trinidad, M. A., et al. 2017, *ApJ*, 836, 96  
 Medina, S.-N. X., Dzib, S. A., Tapia, M., Rodríguez, L. F., & Loinard, L. 2018, *A&A*, 610, A27  
 Minier, V., Ellingsen, S. P., Norris, R. P., & Booth, R. S. 2003, *A&A*, 403, 1095  
 Molinari, S., Swinyard, B., Bally, J., et al. 2010, *A&A*, 518, L100  
 Motte, F., Schilke, P., & Lis, D. C. 2003, *ApJ*, 582, 277  
 Neff, S. G., Eilek, J. A., & Owen, F. N. 2015, *ApJ*, 802, 87  
 Nguyen Luong, Q., Motte, F., Hennemann, M., et al. 2011, *A&A*, 535, A76  
 Ortega, M. E., Paron, S., Cichowolski, S., et al. 2010, *A&A*, 510, A96  
 Parker, Q. A., Bojčić, I. S., & Frew, D. J. 2016, *J. Phys. Conf. Ser.*, 728, 032008  
 Perley, R. A., Chandler, C. J., Butler, B. J., & Wrobel, J. M. 2011, *ApJ*, 739, L1  
 Purcell, C. R., Hoare, M. G., Cotton, W. D., et al. 2013, *ApJS*, 205, 1  
 Rahman, M., & Murray, N. 2010, *ApJ*, 719, 1104  
 Reid, M. J., Dame, T. M., Menten, K. M., & Brunthaler, A. 2016, *ApJ*, 823, 77  
 Rodríguez, L. F., González, R. F., Montes, G., et al. 2012, *ApJ*, 755, 152  
 Rodríguez-Fernández, N. J., & Combes, F. 2008, *A&A*, 489, 115  
 Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, *A&A*, 504, 415  
 Schuller, F., Csengeri, T., Urquhart, J. S., et al. 2017, *A&A*, 601, A124  
 Thompson, M. A., Hatchell, J., Walsh, A. J., MacDonald, G. H., & Millar, T. J. 2006, *A&A*, 453, 1003  
 Treichel, K., Rudnick, L., Hardcastle, M. J., & Leahy, J. P. 2001, *ApJ*, 561, 691  
 Urquhart, J. S., Busfield, A. L., Hoare, M. G., et al. 2007, *A&A*, 461, 11  
 Urquhart, J. S., Hoare, M. G., Purcell, C. R., et al. 2009, *A&A*, 501, 539  
 Urquhart, J. S., Morgan, L. K., Figura, C. C., et al. 2011, *MNRAS*, 418, 1689  
 Urquhart, J. S., Thompson, M. A., Moore, T. J. T., et al. 2013, *MNRAS*, 435, 400  
 Urquhart, J. S., Csengeri, T., Wyrowski, F., et al. 2014a, *A&A*, 568, A41  
 Urquhart, J. S., Moore, T. J. T., Csengeri, T., et al. 2014b, *MNRAS*, 443, 1555  
 Urquhart, J. S., König, C., Giannetti, A., et al. 2018, *MNRAS*, 473, 1059  
 Wang, Y., Bihr, S., Rugel, M., et al. 2018, *A&A*, 619, A124  
 Wilson, T. L., Rohlfs, K., & Huttemeister, S. 2012, *Tools of Radio Astronomy*, 5th edn.  
 Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868  
 Wynn-Williams, C. G., Beichman, C. A., & Downes, D. 1981, *AJ*, 86, 565