Urquhart, JS, Figura, CC, Moore, TJT, Hoare, MG, Lumsden, SL, Mottram, JC, Thompson, MA and Oudmaijer, RD

 The RMS survey: galactic distribution of massive star formation

Article

For more information please contact researchonline@ljmu.ac.uk

The RMS survey: galactic distribution of massive star formation⋆

J. S. Urquhart,[1]† C. C. Figura,[2] T. J. T. Moore,[3] M. G. Hoare,[4] S. L. Lumsden,[4] J. C. Mottram,[5] M. A. Thompson[6] and R. D. Oudmaijer[4]

[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*
[2]*Wartburg College, 100 Wartburg Blvd, Waverly, IA 50677, USA*
[3]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[4]*School of Physics and Astrophysics, University of Leeds, Leeds LS2 9JT, UK*
[5]*Leiden Observatory, Leiden University, NL-2300 RA Leiden, the Netherlands*
[6]*Centre for Astrophysics Research, Science and Technology Research Institute, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK*

ABSTRACT

We have used the well-selected sample of ∼1750 embedded, young, massive stars identified by the Red MSX Source (RMS) survey to investigate the Galactic distribution of recent massive star formation. We present molecular line observations for ∼800 sources without existing radial velocities. We describe the various methods used to assign distances extracted from the literature and solve the distance ambiguities towards approximately 200 sources located within the solar circle using archival H I data. These distances are used to calculate bolometric luminosities and estimate the survey completeness ($\sim 2 \times 10^4$ L$_\odot$). In total, we calculate the distance and luminosity of ∼1650 sources, one third of which are above the survey's completeness threshold. Examination of the sample's longitude, latitude, radial velocities and mid-infrared images has identified ∼120 small groups of sources, many of which are associated with well-known star formation complexes, such as G305, G333, W31, W43, W49 and W51. We compare the positional distribution of the sample with the expected locations of the spiral arms, assuming a model of the Galaxy consisting of four gaseous arms. The distribution of young massive stars in the Milky Way is spatially correlated with the spiral arms, with strong peaks in the source position and luminosity distributions at the arms' Galactocentric radii. The overall source and luminosity surface densities are both well correlated with the surface density of the molecular gas, which suggests that the massive star formation rate per unit molecular mass is approximately constant across the Galaxy. A comparison of the distribution of molecular gas and the young massive stars to that in other nearby spiral galaxies shows similar radial dependences. We estimate the total luminosity of the embedded massive star population to be $\sim 0.76 \times 10^8$ L$_\odot$, 30 per cent of which is associated with the 10 most active star-forming complexes. We measure the scaleheight as a function of the Galactocentric distance and find that it increases only modestly from ∼20–30 pc between 4 and 8 kpc, but much more rapidly at larger distances.

Key words: stars: early-type – stars: formation – ISM: molecules – Galaxy: kinematics and dynamics – Galaxy: structure.

1 INTRODUCTION

Although significantly more luminous than their low-mass counterparts, massive stars ($M_\star > 8$ M$_\odot$) pose a challenging problem for study, particularly in their early formation stages. Massive stars are much less common and have shorter lifetimes than intermediate- and low-mass stars. They are known to form almost exclusively in clusters (de Wit et al. 2004) where source confusion limits the ability to discriminate between individual stars. Because these stars

---

⋆ The full version of Tables 1, 2 and 3 and Figs 1 and 4 is only available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.125.5) or via http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/MNRAS/
† E-mail: jurquhart@mpifr-bonn.mpg.de

form much more quickly than intermediate- and low-mass stars (e.g. Davies et al. 2011; Mottram et al. 2011b), they reach the main sequence while still deeply embedded within their natal clump, so that their formation stages take place beneath hundreds of magnitudes of visual extinction. These observational difficulties complicate the identification of the large statistical samples required to investigate the earliest stages of massive star formation (SF): as a result, our understanding of the initial conditions required or the processes involved in massive SF is much poorer than for lower mass stars.

The motivation for understanding the formation of massive stars is nevertheless strong: these stars are responsible for many of the higher energy events in the Universe and play a significant role in the evolution of their host galaxies (Kennicutt 2005). Throughout their lives they enrich the local chemistry, and inject an enormous amount of radiative and mechanical energy into the interstellar medium (ISM) in the form of UV radiation, stellar winds, jets and outflows, and supernova explosions.

These feedback mechanisms play a role in regulating the SF in their vicinity by disrupting molecular clouds before SF has begun, or by triggering the formation of future generations of stars in the surrounding molecular material. Triggering may occur by sweeping up and compressing the molecular material in surrounding clouds via the collect-and-collapse mechanism (e.g. Whitworth et al. 1994; Deharveng et al. 2003) or through radiatively driven implosion (e.g. Bertoldi 1989; Urquhart et al. 2007a). The processes involved in the formation of massive stars and their subsequent impact on their local environment are a key element to understanding the role massive stars play in shaping the dynamics and structure of their host galaxies.

The Red MSX Source (RMS; Hoare et al. 2005; Urquhart et al. 2008b; Lumsden et al. 2013) survey has been tailored to address many of these outstanding questions by identifying a large well-selected sample of massive young stellar objects (MYSOs) and compact and ultra-compact (UC) H II regions. The RMS survey essentially consists of a suite of follow-up observations and complementary data from other surveys (e.g. 2MASS, UKIDSS, VVV, GLIMPSE, MIPSGAL, ATLASGAL and CORNISH; Skrutskie et al. 2006; Minniti et al. 2010; Lawrence et al. 2007; Benjamin et al. 2003; Carey et al. 2009; Schuller et al. 2009 and Hoare et al. 2012, respectively) of a mid-infrared colour selected sample of ∼5000 MSX sources (Lumsden et al. 2002). This initial sample contains a significant number of dusty objects such as evolved stars that have similar mid-infrared colours as embedded young stars; however, the follow-up observations had been carefully chosen to identify and remove these contaminating sources from the final sample. A data base has been constructed to hold all of these multiwavelength data sets and to compile all of the available data on a source-by-source basis to aid in their classification.[1]

With our programme of follow-up observations and the classification now complete, we have identified approximately 1600 young stellar objects (YSOs) and H II regions located throughout the Galactic plane ($|b| < 5°$). This sample of young embedded massive stars is an order of magnitude larger than that was previously available. The RMS sample provides a sufficient number of sources to allow statistically significant studies of young massive stars as a function of luminosity and environment, while avoiding many of the biases associated with previous surveys of young massive stars. A complete overview of this project, detailed description of the

classification scheme and discussion of the properties of the final embedded catalogue are presented in Lumsden et al. (2013).

We have investigated the Galactic distribution of RMS MYSOs in two previous papers, both of which were focused on smaller subsamples (i.e. Urquhart et al. 2011a, 2012). In this paper, we build on these previous studies and use the full RMS sample to provide a comprehensive picture of the Galactic distribution of massive SF. We will compare this distribution of the now complete RMS sample of massive young stars with the expected positions of the spiral arms and investigate the source and luminosity surface densities as a function of Galactic location.

In the next section, we will summarize the previous molecular line observations and describe additional observations that have been undertaken in an effort to obtain a complete set of radial velocities for this sample. In Section 3, we review the various methods used to determine distances, which are then used to calculate individual source bolometric luminosities and estimate the survey's completeness. We use these distances and luminosities to investigate the Galactic distribution of massive SF and to compare this with the position of the spiral arms and measure the scaleheight of the Galactic disc in Section 4. Finally, in Section 5 we present a summary of our results and highlight our main findings.

2 RADIAL VELOCITY MEASUREMENTS

2.1 Previous observations

Molecular line observations are a crucial part of the RMS follow-up campaign, as they provide radial velocities for the determination of kinematic distances and luminosities. These are used to distinguish nearby low- and intermediate-mass YSOs from the generally more distant MYSOs. They can also be useful in identifying the more evolved stars that contaminate our sample, as these are not generally associated with sufficient cold gas to produce a detectable emission in the lower excitation states (i.e. Loup et al. 1993).

The ^{13}CO (1–0) and (2–1) transitions were chosen as optimal tracers as they have a lower abundance than their ^{12}CO counterparts. ^{12}CO is generally optically thick in these regions, and suffers from saturation, self-absorption and multiple emission components along a given line of sight that can produce complex line profiles, while ^{13}CO is generally free from these problems. Furthermore, ^{13}CO is more abundant than $C^{18}O$, allowing emission to be detected in a modest amount of observing time. We have made ^{13}CO observations towards ∼2000 sources; the results of the majority of these were reported in Urquhart et al. (2007b, 2008c) and the remainder will be discussed in the following subsection.

Although the ^{13}CO transition is much less affected by many of the problems associated with ^{12}CO emission, multiple components are still detected towards ∼60 per cent of the observed sources. We have therefore followed up a large number of these with observations of high-density tracers such as CS (2–1), NH_3 and water maser observations (Urquhart et al. 2009, 2011b). We have complemented these targeted observations with other high-density transitions reported in the literature (e.g. Bronfman, Nyman & May 1996; Schlingman et al. 2011; Wienen et al. 2012).

2.2 Additional observations: ^{13}CO, CS and NH_3

The observations described in this section were made with the 22 m Mopra radio telescope, which is located near Coonabarabran, New

---

[1] http://rms.leeds.ac.uk/cgi-bin/public/RMS_DATABASE.cgi.

South Wales, Australia.[2] The Mopra beam size is approximately 2.5 arcmin for the ammonia observations and 36 and 40 arcsec for the CO and CS observations, respectively.

The CS (2–1) observations were made with Mopra's older cryogenically cooled ($\sim$4 K) superconductor–insulator–superconductor junction mixer, with a frequency range between 85 and 116 GHz. The receiver backend was a digital autocorrelator capable of providing two simultaneous outputs with an instantaneous bandwidth between 4 and 256 MHz. The CS observations were made using a bandwidth of 64 MHz with a 1024-channel digital autocorrelator. This provided a frequency resolution of 62.5 kHz and a velocity resolution of $\sim$0.2 km s^{-1} at the CS transition rest frequency (97.98 GHz).

The UNSW Mopra spectrometer (MOPS)[3] was commissioned in 2005 October and consists of four 2.2 GHz bands that overlap slightly to provide a total of 8 GHz continuous bandwidth. Up to four zoom windows can be placed within each 2.2 GHz band allowing up to 16 spectral windows to be observed simultaneously, each providing a bandwidth of 137 MHz with 4096 channels. MOPS was used for the NH$_3$ (1,1) and (2,2) inversion transitions, and the ^{13}CO (1–0) observations described in the following two subsections.

2.2.1 CS and NH$_3$ observations

NH$_3$ and CS observations were made towards sources previously observed in the ^{13}CO transition as part of our initial programme of follow-up observations. As previously mentioned, multiple velocity components were seen along the lines of sight towards many of these sources, and further observations of higher density tracers were required to identify the velocity component associated with the IR continuum source. The Mopra telescope was used to follow up a large number of sources using CS (2–1) and the two lower excitation inversion transitions of NH$_3$. All of these transitions have high critical densities ($\sim$10^4–10^5 cm^{-3}) and are therefore excellent tracers of high-density molecular gas associated with high-mass star-forming regions.

The CS observations were made towards 127 sources in 2004 September and 2005 August (project reference: M121). The signal-to-noise ratio was improved by tuning both polarizations to the CS frequency. System temperatures were typically $\sim$200 K.

The NH$_3$ observations were made in 2008 April and September towards 499 RMS sources using the 12 mm receiver and MOPS (project reference: M270; Urquhart et al. 2008a). The observations covered a frequency range of 16–27.5 GHz. Each of the 16 zoom windows was centred on the rest frequency of a known transition providing a total velocity coverage of $\sim$2000 km s^{-1} with a resolution of $\sim$0.4 km s^{-1} channel^{-1}. System temperatures were between 65 and 100 K. A detailed description of these observations can be found in Urquhart et al. (2009). Although they included a large number of transitions, we only present the NH$_3$ (1,1) and (2,2) transitions here; however, we have made the data for all of the other transitions available in the RMS data base.

2.2.2 CO observations

The RMS source classification criteria have evolved over time as the data obtained from the various programmes of follow-up observations have been analysed [we refer the reader to Lumsden et al. (2013) for a detailed description of the classification criteria is presented]. This has resulted in a number of sources that were initially rejected being reclassified as either a YSO or H II region. As these reintroduced sources were not included in our initial sample, no molecular line observation had been previously made.

We made observations of the CO (1–0) transitions towards 192 of the reintroduced RMS sources in 2012 March using the MOPS instrument (project reference: M573; Urquhart, Hoare & Morgan 2010a). Three zoom windows were used to cover the ^{12}CO, ^{13}CO and C^{18}O transitions. However, the optically thick ^{12}CO emission generally shows complex emission profiles while the C^{18}O is generally only detected towards a small number of sources. We therefore only present the plots of the ^{13}CO emission profiles and the fit parameters to the emission features seen in these spectra. The average system temperature was $\sim$360 K at 110 GHz.

### 2.3 Observational procedures and data reduction

All molecular line observations were performed in position-switching mode, with typical on-source integration times of $\sim$10, 30 and 40 min for the ^{13}CO, NH$_3$ and CS transitions, respectively. The on-source integration time was split into individual 1 min on- and off-source scans. Reference positions were offset from source positions by 1° in a direction perpendicular to the Galactic plane. These were chosen to avoid contamination by emission at the reference position at a similar velocity.

Weather conditions were stable over the short time periods required to complete the observations of each source, with system temperatures varying by no more than approximately 10 per cent. The telescope pointing was checked every 1–2 h by observing a strong nearby SiO maser (Indermuehle et al. 2013), and the average pointing accuracy was found to be better than 10 arcsec rms. A measurement of an ambient load (assumed to be at 290 K) was made before and after each ^{13}CO and CS observation following the method of Kutner & Ulich (1981) to put the measured antenna temperatures on to the standard T_A^* scale, correcting for atmospheric absorption, ohmic losses and rearward spillover and scattering. This correction is not required for observations below 70 GHz as the atmosphere is significantly less variable, and absorption from water vapour is less of a problem.

The observations of all three transitions were reduced using the ATNF spectral line reducing software (ASAP). Individual on–off scans were processed to remove sky emission and visually inspected to remove poor scans. A low-order polynomial baseline was fitted, after which the individual scans were averaged together to produce a single spectrum for each source.

The CO and CS spectra have been Hanning smoothed to a velocity resolution of $\sim$0.4 km s^{-1} to obtain a final sensitivity of $\sim$70 and 35 mK channel^{-1} beam^{-1}, respectively. The reduced spectra were converted to the telescope-independent main-beam temperature scale (T_{mb}), assuming a main-beam efficiency (η_{mb}) of 0.65 (Urquhart et al. 2010b), 0.60 and 0.55 (Ladd et al. 2005) for the NH$_3$, CS and CO observations, respectively. In Fig. 1, we present an example of each of the emission detected for the three molecules observed.

Gaussian profiles were fitted to the observed emission features in the CO and CS data using the spectral line analysis package xs
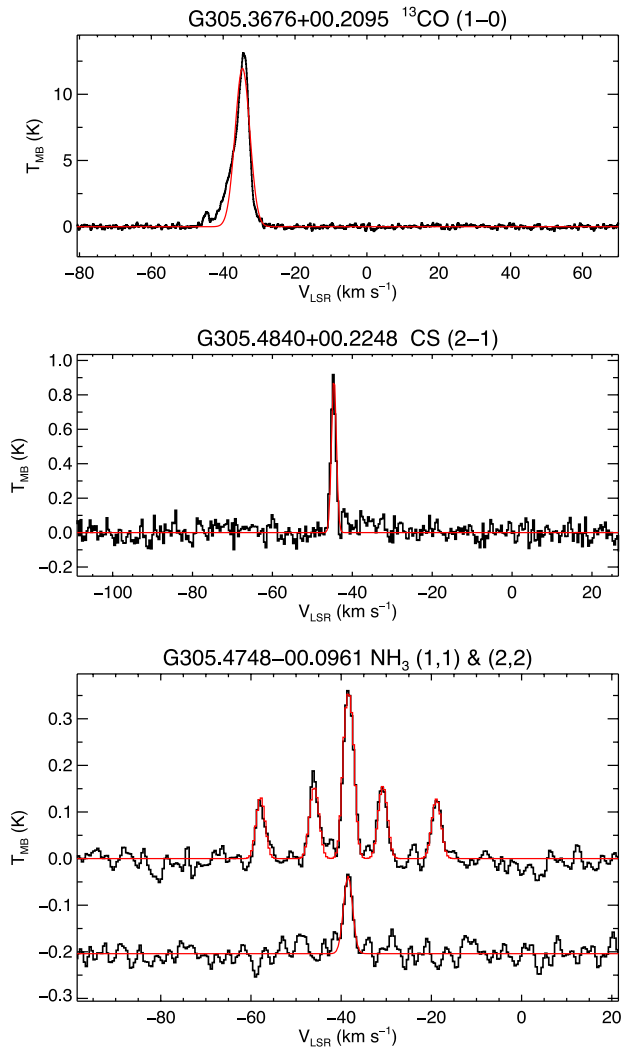
Figure 1. Example spectra obtained from the additional Mopra observations. In the upper and middle panels, we present emission in the CS and ^{13}CO (1–0) transitions towards example RMS sources. The lower panel shows example NH$_3$ (1,1) and (2,2) inversion transitions, with the latter offset by -0.2 K to avoid confusion. The hyperfine structure can be seen clearly in the NH$_3$ (1,1) emission profile. In all of these plots, the black and red lines show the observed spectra and model fits to the data, respectively. All spectra have a velocity resolution of ~ 0.4 km s^{-1}.

written by Per Bergman.[4] Where necessary, a higher order polynomial was fitted to the emission-free parts of the spectrum and subtracted from the baseline before the Gaussian profiles were fitted. As can be seen in the upper and middle panels of Fig. 1, the Gaussian profile provided a good fit to the data in the majority of cases.

The NH$_3$ (1,1) molecule has hyperfine structure and consists of 18 components which need to be fitted simultaneously in order to derive the optical depth and line width (see the lower panel of Fig. 1 for an example). The fitting of these data has been done in the IDL environment using the MPFIT routine.[5] The hyperfine structure is

generally too weak to be observed in the NH$_3$ (2,2) transition, and so these detections have been fitted with Gaussian profiles along with the weaker NH$_3$ (1,1) lines where the hyperfine components are not detected. In all cases, the NH$_3$ (1,1) and (2,2) line widths have been obtained by fitting the hyperfine components to their respective main line emission to remove the effects of line broadening due to optical depth effects.

The radial velocities, main-beam temperatures and full width at half-maximum (FWHM) line widths obtained from these fits are presented for all detected transitions in Table 1.

2.4 Association of sources with known star-forming regions

Examination of *Spitzer* mid-infrared images obtained as part of the GLIMPSE legacy project reveals that many RMS sources are associated with large star-forming complexes. In many cases, a single RMS source is associated with a previously identified complex but, in the majority of cases, we find multiple sources positionally coincident with a given complex. To confirm such associations, we compare their radial velocities and require that these agree within ~ 10 km s^{-1}. Since many complexes are well known (such as W31, M16, M17, etc.) and have been the focus of detailed studies, reliable distances can be readily found in the literature [although care must be exercised regarding the comparison of different distance determinations, for example Westerlund 2 (Dame 2007) and W51 (Clark et al. 2009) – we have only adopted these in the cases where these are broadly consistent].

We have used visual inspection of the mid-infrared images combined with velocity information to identify all of these small groups of sources and, where possible, have associated these groups with a known star-forming complex. This has resulted in ~ 600 sources being associated with approximately 120 known SF regions/complexes and accounts for ~ 40 per cent of the RMS sample of YSOs and H II regions.
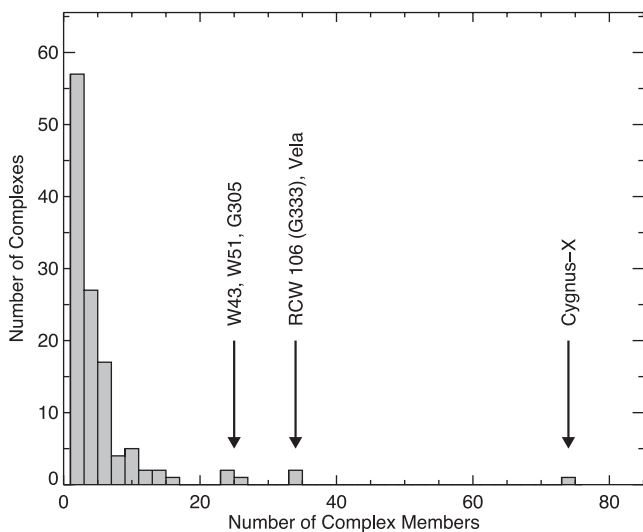
Fig. 2 shows the number of RMS sources associated with each of the 117 complexes identified and Table 2 lists the complex parameters. The vast majority of these complexes are relatively small, consisting of no more than a handful of individual massive star-forming regions. However, given that each one of these regions is probably forming a stellar cluster, these complexes should be viewed as active star-forming regions. If it is typical for a given star-forming complex to have a small number of localized regions of active SF, then complexes with significantly larger numbers of RMS sources stand out as perhaps being outside the norm.

There are only six complexes that have more than 20 members: these are W51, Cygnus-X, W43, G305, RCW 106 and the Vela molecular ridge, which are among the best studied SF regions in the Galaxy (e.g. Parsons et al. 2012; Csengeri et al. 2011; Motte, Schilke & Lis 2003; Hindson et al. 2010; Bains et al. 2006 and Hill et al. 2012, respectively). We note, however, that although Cygnus-X and the Vela molecular ridge are associated with the largest number of RMS sources, these are both relatively nearby (1.4 and 0.7 kpc, respectively) and so the RMS sources are not particularly luminous. Consequently, neither of these star-forming complexes features in our list of the most luminous complexes presented in Table 2. There is therefore a distance bias associated with the number of sources associated with each complex, and care needs to be taken when drawing conclusions from the source counts alone.

Table 1. Fitted molecular line parameters.

Field name	RA (J2000)	Dec. (J2000)	Transition	rms (mK)	V_{LSR} (km s^{-1})	T_R^* (K)	FWHM (km s^{-1})
G305.2017+00.2072	13:11:10.29	−62:34:39.0	NH$_3$ (1,1)	15	−41.38	0.43	12.24
G305.2017+00.2072	13:11:10.29	−62:34:39.0	NH$_3$ (2,2)	16	−41.54	0.18	5.94
G305.2242+00.2028	13:11:22.07	−62:34:48.0	CS ($J = 2$–1)	60	−40.69	0.31	9.65
G305.2242+00.2028	13:11:22.12	−62:34:48.7	NH$_3$ (1,1)	16	−41.45	0.09	10.29
G305.2242+00.2028	13:11:22.12	−62:34:48.7	NH$_3$ (2,2)	14	−41.81	0.08	5.84
G305.2535+00.2412	13:11:35.80	−62:32:22.9	NH$_3$ (1,1)	14	−36.91	0.04	10.84
G305.2535+00.2412	13:11:35.80	−62:32:22.9	NH$_3$ (2,2)	13	−38.55	0.08	3.66
G305.3210+00.0706	13:12:17.96	−62:42:16.3	^{13}CO ($J = 1$–0)	71	−31.71	0.33	2.83
G305.3210+00.0706	13:12:17.96	−62:42:16.3	^{13}CO ($J =$1–0)	71	−38.80	0.74	5.19
G305.3210+00.0706	13:12:17.96	−62:42:16.3	^{13}CO ($J = 1$–0)	71	−45.32	0.59	2.40
G305.3526+00.1945	13:12:29.36	−62:34:41.2	^{13}CO ($J = 1$–0)	67	−37.69	6.20	6.70
G305.3676+00.2095	13:12:36.44	−62:33:43.5	^{13}CO ($J = 1$–0)	65	−34.64	6.58	4.82
G305.3719+00.1837	13:12:39.69	−62:35:15.0	NH$_3$ (1,1)	16	−38.11	0.25	9.03
G305.3719+00.1837	13:12:39.69	−62:35:15.0	NH$_3$ (2,2)	15	−38.48	0.13	5.48
G305.3779+00.2108	13:12:41.56	−62:33:35.4	^{13}CO ($J = 1$–0)	67	−33.57	5.16	3.46
G305.3779+00.2108	13:12:41.56	−62:33:35.4	^{13}CO ($J = 1$–0)	67	−38.53	1.37	5.59
G305.3779+00.2108	13:12:41.56	−62:33:35.4	^{13}CO ($J = 1$–0)	67	−45.27	0.87	0.98
G305.4399+00.2103	13:13:13.84	−62:33:19.0	NH$_3$ (1,1)	15	−	−	−
G305.4399+00.2103	13:13:13.84	−62:33:19.0	NH$_3$ (2,2)	15	−	−	−
G305.4748−00.0961	13:13:45.76	−62:51:27.7	NH$_3$ (1,1)	10	−38.36	0.25	5.17
G305.4748−00.0961	13:13:45.76	−62:51:27.7	NH$_3$ (2,2)	11	−38.48	0.09	2.35
G305.4840+00.2248	13:13:35.99	−62:32:12.8	CS ($J = 2$–1)	46	−44.59	0.55	1.41
G305.4840+00.2248	13:13:36.05	−62:32:13.5	NH$_3$ (1,1)	10	−44.77	0.07	1.92
G305.4840+00.2248	13:13:36.05	−62:32:13.5	NH$_3$ (2,2)	11	−	−	−
G305.5393+00.3394	13:13:59.52	−62:25:05.5	NH$_3$ (1,1)	13	−35.13	0.70	8.40
G305.5393+00.3394	13:13:59.52	−62:25:05.5	NH$_3$ (2,2)	13	−35.17	0.05	2.94
G305.5516+00.0149	13:14:20.94	−62:44:26.4	^{13}CO ($J = 1$–0)	90	−31.06	1.46	1.91
G305.5516+00.0149	13:14:20.94	−62:44:26.4	^{13}CO ($J = 1$–0)	90	−38.88	6.30	3.46
G305.5610+00.0124	13:14:25.82	−62:44:30.8	NH$_3$ (1,1)	14	−39.13	0.09	3.63
G305.5610+00.0124	13:14:25.82	−62:44:30.8	NH$_3$ (2,2)	15	−38.42	0.06	3.11

Notes. A small portion of the data is provided here; the full table is only available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.125.5) or via http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/MNRAS/.

Figure 2. Number of sources associated with each complex. The names of the complexes with the largest number of RMS associations are given.

By grouping sources together and identifying their associated star-forming complexes, we have reduced the number of distances we need to find by ∼500; however, this still leaves some 1100 distances to be determined. The methods used to determine these distances are discussed in Section 3.

2.5 Longitude–velocity distribution

Fig. 3 shows the distribution of the complexes and individual embedded sources as a function of Galactic longitude and radial velocity. The upper and lower panels show the longitudinal distribution of the radial velocity (plotted over the integrated ^{12}CO (1–0) emission mapped by Dame et al. 2001) for the northern and southern Galactic plane, respectively. These longitude–velocity (ℓ–v) plots also show the proposed positions of the four-arm spiral-arm model of Taylor & Cordes (1993) (updated by Cordes 2004), which itself is based on the earlier model of Georgelin & Georgelin (1976). The figure reveals a strong correlation between the distributions of molecular gas and RMS sources within the inner Galaxy (i.e. $|\ell| < 60°$). Outside of this region, the correlation is significantly weaker. The positions of the modelled spiral arms appear to be strongly correlated to the distribution of our sample of YSOs and H ii regions, particularly in the outer Galaxy where the correlation with the molecular gas is significantly poorer. Since the spiral arms in the outer Galaxy are the most distant, the poorer correlation between the RMS sources and the molecular gas there is likely to be due to the sensitivity of the CO survey, with the emission from the more distant molecular clouds being diluted in the 8 arcmin beam used by the Dame et al. survey (cf. Urquhart et al. 2013).

Previous CO, H ii region and mid-infrared surveys have identified the spiral-arm line-of-sight tangents (e.g. Caswell & Haynes 1987; Grabelsky et al. 1987; Bronfman et al. 1988; Alvarez, May & Bronfman 1990; Benjamin 2008). In the northern Galactic plane,

Table 2. Parameters of the 25 most luminous star-forming complexes. The positions and velocities of these complexes have been determined from the mean of all associated RMS sources, and as such are only approximate values. The luminosities of each complex (given in column 9) have been determined from the integrated bolometric luminosities of their embedded YSO and H II region populations. The contribution each complex makes to the total embedded massive star population (column 10) is estimated by dividing the complex's luminosity by the estimate of total Galactic MYSO and embedded H II region luminosity ($L_{Galaxy} = 0.76 \times 10^8$ L$_\odot$; see Section 4.1 for details).

Complex name	ℓ (°)	b (°)	V_{LSR} (km s^{-1})	RMS members (#)	Distance (kpc)	z (pc)	R_{GC} (kpc)	L_{bol} (log[L$_\odot$])	L_{bol}/L_{Galaxy} (per cent)
W51 (A&B)	49.362	−0.330	60.47	24	5.4	−31.1	6.45	6.67	6.82
NGC 3603	291.598	−0.491	12.79	4	7.0	−59.6	8.78	6.61	5.95
W49A	43.145	−0.009	11.14	10	11.1	−1.7	7.61	6.56	5.23
RCW 106 (G333)	333.037	−0.320	−51.65	33	3.6	−20.1	5.54	6.28	2.74
G338.398+00.164	338.404	0.120	−35.74	14	12.8	28.9	5.82	6.21	2.36
GAL331.03−00.15	330.960	−0.185	−91.54	2	9.6	−30.9	4.64	6.04	1.57
G305	305.506	0.085	−36.32	25	4.0	5.8	6.98	5.94	1.26
G282.0−1.2	281.881	−1.605	−6.10	11	7.0	−195.4	9.82	5.93	1.24
W43	30.861	−0.023	93.73	30	5.1	−1.8	4.99	5.93	1.23
AGAL032.797+00.191	32.796	0.198	14.99	2	12.9	44.5	7.37	5.89	1.12
W47	37.601	−0.223	53.48	6	9.9	−38.3	6.07	5.88	1.10
RCW42	273.982	−1.191	37.52	5	5.5	−114.6	9.81	5.86	1.05
AGAL045.121+00.131	45.115	0.131	58.83	2	4.4	10.1	6.23	5.82	0.96
W3	133.797	1.155	−43.42	9	2.0	39.3	9.95	5.80	0.92
AGAL319.399−00.012	319.390	−0.007	−13.50	4	11.6	−1.5	7.57	5.69	0.70
Far 3 kpc arm (south)	348.854	−0.015	13.30	8	11.3	−2.8	3.36	5.67	0.67
Near 3 kpc tangent	336.880	0.011	−123.24	10	7.7	1.6	3.34	5.56	0.53
AGAL020.081−00.136	20.076	−0.139	41.80	2	12.6	−30.5	5.44	5.55	0.51
GAL331.5−00.1	331.553	−0.088	−86.67	9	4.9	−7.7	4.81	5.52	0.48
RCW116B	345.091	1.543	−14.77	14	2.4	64.1	6.23	5.43	0.39
NGC 7538	111.626	0.761	−55.48	9	2.6	35.2	9.79	5.43	0.38
Cygnus-X	79.538	0.942	1.66	74	1.4	23.0	8.36	5.40	0.36
Gum 50	328.567	−0.533	−46.57	1	3.0	−28.3	6.12	5.38	0.35
G010.960+00.017	10.962	0.015	20.29	2	8.1	2.8	5.76	5.36	0.33
GAL336.40−00.23	336.470	−0.232	−86.11	5	10.5	−42.4	4.34	5.30	0.29

Notes. A small portion of the data is provided here; the full table is only available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.125.5) or via http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/MNRAS/.

tangent positions have been reported at $\ell \simeq 30°$ and $47°$, which correspond to the Scutum–Centaurus and Sagittarius spiral arms. Inspection of the distribution of RMS sources in the northern Galactic plane reveals high densities towards the longitudes and velocities of these two tangents and their associations with W43 and W51 star-forming complexes. Similarly, we find high RMS source densities towards the longitudes and velocities of the Sagittarius, Scutum–Centaurus, Norma and near 3 kpc arms ($\sim283°$, $308°$, $328°$ and $337°$; Bronfman et al. 2000). We find G305 and G282.0−0.12 star-forming complexes are associated with the Scutum–Centaurus and Sagittarius tangents, respectively.

The source counts are significantly lower and broadly flat outside the spiral-arm tangents (i.e. $60° < \ell < 280°$) with only two notable exceptions; these are the Cygnus-X region ($\ell \sim 80°$; Reipurth & Schneider 2008) and the Vela molecular ridge ($\ell \sim 268°$; Netterfield et al. 2009). Both of these regions are likely to be associated with the local arm and are therefore relatively nearby (these two regions are labelled in the plots presented in Fig. 3). Incidentally, these two regions are also associated with two highest densities of RMS sources; however, as mentioned in the previous section, most of these are not very luminous.
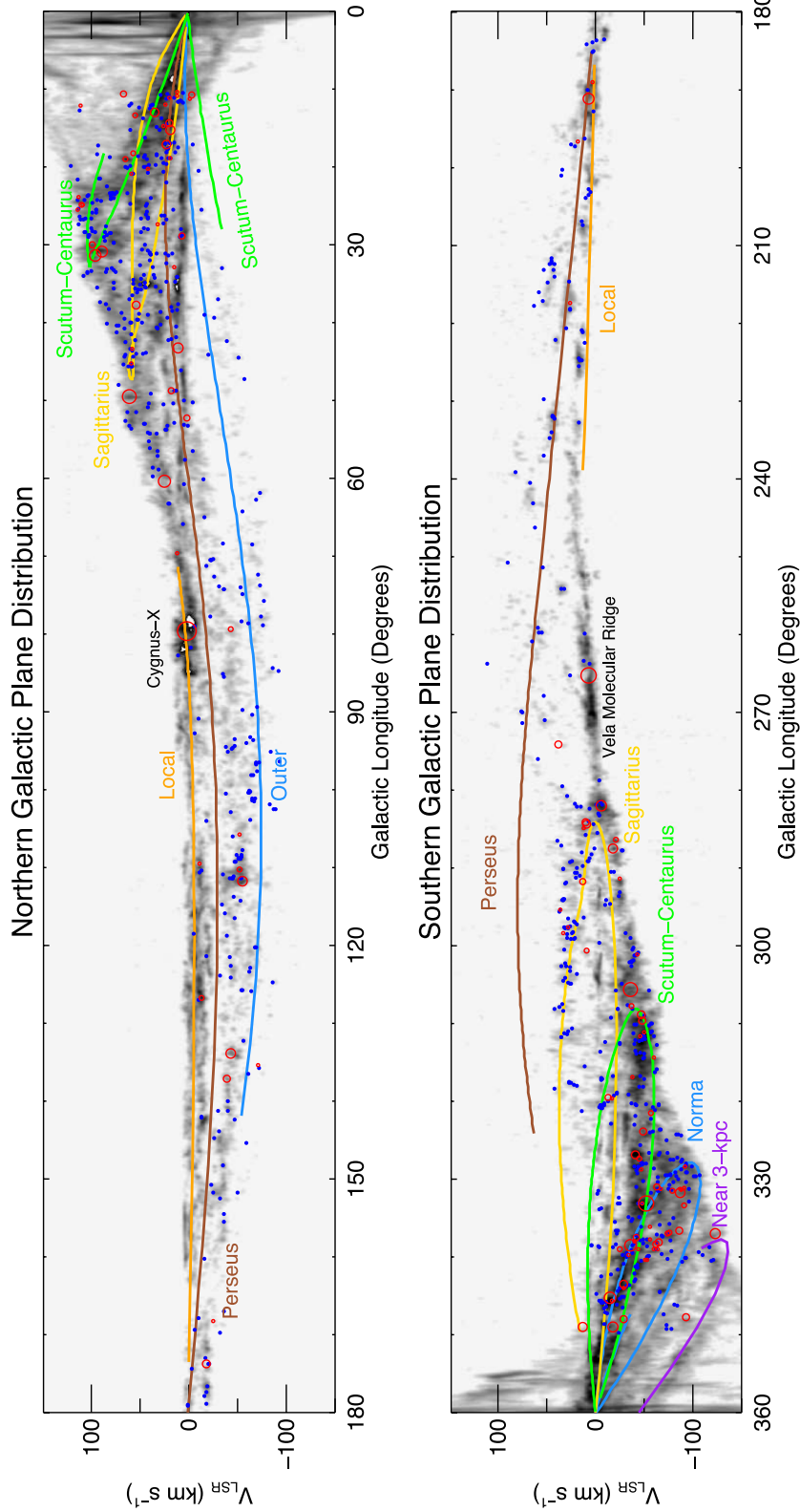
3 DISTANCES AND LUMINOSITIES

3.1 Distance determination

Distances are critical to estimating source luminosities, which allow us to discriminate between genuinely massive stars and nearby lower mass stars and to determine their Galactic distribution. Deriving distances to our sample of MYSOs and H II regions has been an ongoing process, with results having previously been presented for incomplete samples located primarily in the first and fourth quadrants (Urquhart et al. 2011a and Urquhart et al. 2012, respectively).

Maser parallax derived distances are the most reliable methods to determine distances and these are becoming available for an increasing number of star-forming regions (i.e. Reid et al. 2009). At present, these tend to be localized to the relatively nearby parts of the first and second quadrants of the Galaxy, and only provide distances to a small fraction of our sample (no more than a few per cent). Spectrophotometric measurements of H II region exciting stars or associated clusters are the next most reliable method; however, this can produce incorrect distances if the star is incorrectly typed (see Clark et al. 2009 for an example). We have conducted a comprehensive review of the literature and have adopted both parallax and spectrophotometric distances where available, and will continue to update source distances as and when new measurements become available.

For the remaining sources, we have estimated kinematic distances using the Galactic rotation model derived by Reid et al. (2009, circular rotation speed $\theta_0 = 254$ km s^{-1} and distance to the Galactic Centre $R_0 = 8.4$ kpc) and the radial velocities discussed in the previous section. Kinematic distances have associated uncertainties of the order of ±1 kpc (allowing for a ±7 km s^{-1} error in the velocity due to streaming motions). We also note that kinematic distances have been found to deviate significantly from parallax and spectroscopically derived distances (e.g. Moisés et al. 2011).

Figure 3. The upper and lower panels show the longitude and velocity distribution of the RMS sources in the northern and southern Galactic plane, respectively. The complexes are shown as open red circles, the size of which gives an indication of the source density of each complex, while the blue filled circles show the position of the rest of the sample. The grey-scale image shows the distribution of molecular gas as traced by the integrated ^{12}CO emission (Dame, Hartmann & Thaddeus 2001) for comparison. The locations of the spiral and local arms are shown as curved solid lines, coloured to identify the individual arms. The positions of the four main spiral and local arms have been taken from the model by Taylor & Cordes (1993) and updated by Cordes (2004), while the position of the near 3 kpc arm has been taken from Bronfman et al. (2000).

Although these distance anomalies can dramatically affect the derived properties of individual sources, it will not impact the results drawn from the whole sample. Of greater concern is that within the solar circle there are two solutions for each radial velocity (*kinematic distance ambiguity*, or KDA). These distances are equally spaced on either side of the tangent position, and are generally referred to as the *near* and *far* distances. Since the majority of our sources are located within the solar circle, this distance ambiguity affects ∼80 per cent of our sample.

We reduce the number of ambiguities that need to be resolved by applying two initial cuts. First, we place sources with velocities within $10 \, \text{km s}^{-1}$ of the tangent velocity at the tangent distance, since the error in the distance is comparable to the difference between the near/far distance and the tangent distance. Secondly, we place sources at the near distance if a far allocation would result in an unrealistically large displacement from the Galactic mid-plane for an SF region (i.e. $z > 120 \, \text{pc}$; $4 \times$ scaleheight for O-type stars; Reed 2000). These two steps reduce the number of KDAs that need to be resolved by ∼100; however, the vast majority still need to be addressed.

There are a number of ways to resolve these KDAs: H I absorption against a continuum (e.g. Fish et al. 2003; Kolpak et al. 2003; Anderson & Bania 2009; Roman-Duval et al. 2009; Urquhart et al. 2012), H I self-absorption (e.g. Jackson et al. 2002; Roman-Duval et al. 2009; Green & McClure-Griffiths 2011) and matching sources with infrared dark clouds (IRDCs; e.g. Dunham et al. 2011; Ellsworth-Bowers et al. 2013). These techniques are well documented in the literature and so will not be described here in detail. Many of these studies have sources in common with our sample, and we have compared the positions and velocities (i.e. ℓbv parameter space) of these and adopted the given distance solution where a match is found. In the cases where the distance solutions differ, we have made an independent assessment of the available data to determine the most likely distance and have favoured the allocation based on the most reliable data (i.e. the highest resolution and/or data with the highest signal-to-noise ratio). In the cases where there is no clear distinction, we do not resolve the ambiguity (but may do so in the future as more data become available).

Since all of the bright H II regions have been included in previous studies using the H I absorption against a continuum method, the ∼200 sources for which the ambiguity has not been resolved are either radio-quiet or the radio emission is too weak for absorption to be detected in the H I surveys available. We have therefore extracted H I spectra from the Southern Galactic Plane Survey (SGPS; McClure-Griffiths et al. 2005) and the Very Large Array Galactic Plane Survey (Stil et al. 2006) for all these remaining objects. These have been inspected for absorption coincident with the source velocity, indicating the near distance, otherwise the far distance, is considered more likely. Additionally, we inspected the GLIMPSE and MIPSGAL images to correlate sources located at the near distance with IRDCs in an effort to confirm the H I results. If a correlation was found, the near distance was assigned; if no correlation was found, no distance was assigned. Fig. 4 shows some examples of these distance allocations, and in Table 3 we present the distance solutions for all sources examined.

Using a combination of parallax and spectroscopic distances found in the literature, and KDA solutions drawn both from the literature and derived from our own analysis, we are able to assign distances to ∼1650 of the embedded RMS sources. We have been unable to assign a distance to approximately 100 sources; however, this corresponds to ∼7 per cent of the embedded RMS population.
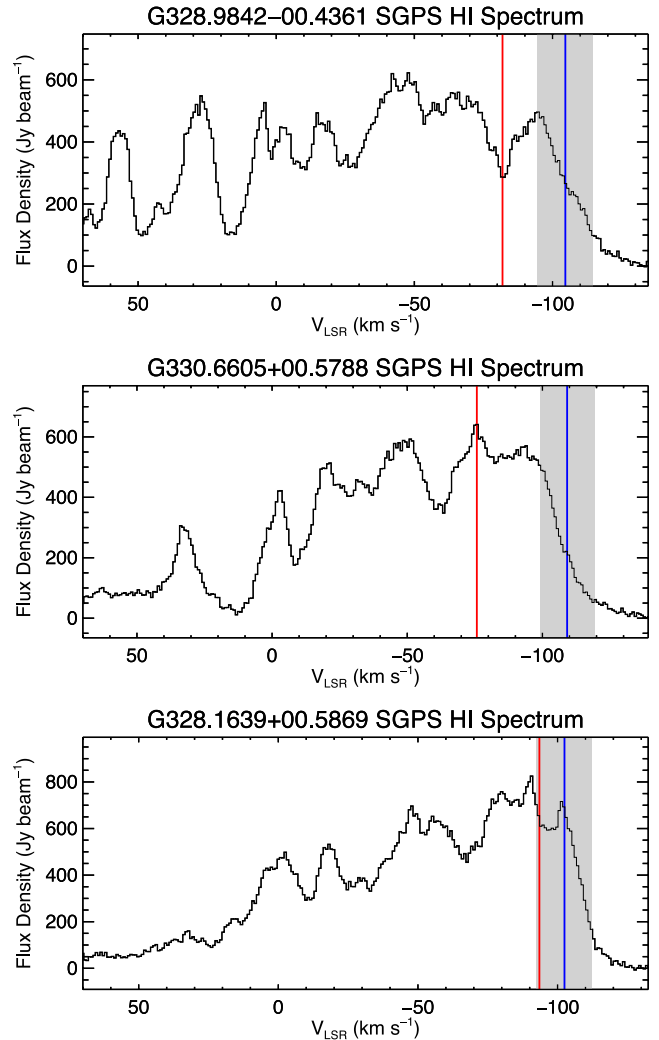
Figure 4. Examples of SGPS H I spectra towards the embedded RMS sources located within the solar circle for which a reliable distance was not available in the literature. The source velocity (v_s) and the velocity of the tangent point (v_t) are shown by the red and blue vertical lines, respectively. The grey vertical band covers the velocity region $10 \, \text{km s}^{-1}$ on either side of the tangent velocity and provides an easy way to identify sources placed at the tangent position. In the upper, middle and lower panels, we present an example of a source located at the near, far and tangent positions.

3.2 Bolometric luminosities and survey completeness

Bolometric luminosities have been estimated for the majority of sources using the distances discussed in the previous section and the integrated fluxes determined from model fits to the spectral energy distributions (SEDs) presented by Mottram et al. (2011b). For very bright sources and sources located in complicated regions, it is not always possible to obtain a sufficient number of photometric points to adequately constrain the SED; this is primarily due to saturation of the mid- and far-infrared images and difficulties subtracting the diffuse background emission (Mottram et al. 2010). In these cases, the luminosities have been estimated by simply scaling the MSX 21 μm flux (see Mottram et al. 2011a for details).

The RMS luminosities have been previously reported by Mottram et al. (2011b), who used a less complete RMS subset to investigate the luminosity functions of MYSOs and H II regions and determine statistical lifetimes of these stages. Rather than reproducing this

Table 3. Results of the H I self-absorption analysis.

MSX name	ℓ (°)	b (°)	V_{LSR} (km s^{-1})	Near (kpc)	Far (kpc)	KDS[a]	Distance (kpc)	R_{GC} (kpc)	z (pc)
G327.9205+00.0921	327.9210	0.0921	−49.5	3.2	11.1	F	11.1	5.9	17.8
G328.5487+00.2717	328.5490	0.2717	−60.5	3.7	10.6	N	3.7	5.6	17.6
G327.8097−00.6339	327.8100	−0.6339	−46.7	3.0	11.2	N	3.0	6.1	−33.6
G328.2275−00.2714	328.2280	−0.2714	−99.5	5.8	8.5	F	7.1	4.4	−33.8
G328.9480+00.5709	328.9480	0.5709	−93.9	5.3	9.0	F	7.2	4.3	71.7
G328.9580+00.5671	328.9580	0.5671	−93.5	5.3	9.1	F	7.2	4.3	71.2
G329.2713+00.1147	329.2710	0.1147	−76.9	4.5	9.9	N	4.5	5.1	9.0
G329.3371+00.1469	329.3370	0.1469	−107.1	6.1	8.3	F	7.2	4.3	18.5
G329.4720+00.2143	329.4720	0.2143	−101.5	5.7	8.7	F	7.2	4.3	27.1
G329.4579+00.1724	329.4580	0.1724	−103.2	5.8	8.6	F	7.2	4.3	21.8
G328.9842−00.4361	328.9840	−0.4361	−81.9	4.7	9.7	N	4.7	5.0	−36.1
G329.6098+00.1139	329.6100	0.1139	−63.9	3.9	10.6	N?	3.9	5.4	7.7
G329.4211−00.1631	329.4210	−0.1631	−76.8	4.5	10.0	N	4.5	5.1	−12.8
G329.8145+00.1411	329.8150	0.1411	−85.0	4.9	9.7	N	4.9	4.8	12.0
G329.3402−00.6436	329.3400	−0.6435	−74.2	4.4	10.1	F	10.1	5.2	−113.2
G330.6605+00.5788	330.6600	0.5788	−75.7	4.4	10.2	F	10.2	5.0	103.2
G330.2923+00.0010	330.2920	0.0010	−64.3	3.9	10.7	N?	3.9	5.4	0.1
G331.0890+00.0163	331.0890	0.0163	−95.6	5.3	9.4	N	5.3	4.5	1.5
G331.0931−00.1303	331.0930	−0.1303	−64.0	3.9	10.8	F?	10.8	5.3	−24.6
G330.9288−00.4070	330.9290	−0.4070	−41.6	2.8	11.9	F	11.9	6.1	−84.2

[a] Kinematic distance solution. 'N' indicates that the source was determined to be at the *near* location, while an 'F' reflects the *far* position. Where a distance solution is considered less reliable, we append '?' to the given allocation.

Notes. A small portion of the data is provided here; the full table is only available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.125.5) or via http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/MNRAS/

work, we aim simply to use the bolometric luminosities to estimate the survey's completeness threshold for the full sample of embedded RMS sources and to investigate the distribution of massive stars across the Galaxy.[6]

In Fig. 5, we show the distribution of YSO and H II region bolometric luminosities as a function of the heliocentric distance. This plot suggests that the sample is complete to young embedded sources with bolometric luminosities over $\sim 2 \times 10^4$ L$_\odot$ to a distance of ~ 18 kpc. While the bolometric luminosities for individual sources can be found in Lumsden et al. (2013), we present the total bolometric luminosities of star-forming complexes in Table 2. These have been estimated by integrating the bolometric luminosities of the embedded massive stellar population associated with each complex.

4 GALACTIC STRUCTURE

With the distances on hand we are in a position to examine the three-dimensional distribution of this sample of young massive stars. In Fig. 6, we present a top-down view of the Milky Way showing the distribution of MYSOs and H II regions with respect to the known large-scale structures of the Galaxy. The background image shown in this figure is an artist's conceptual image of what the Galaxy might look like if viewed from above, and incorporates all that is currently known of the structure of our Galaxy, including the 3.1–3.5 kpc Galactic bar at an angle of 20° with respect to the Galactic Centre–Sun axis (Binney et al. 1991; Blitz & Spergel 1991; Dwek et al. 1995), a second non-axisymmetric structure referred to as the 'long bar' (Hammersley et al. 2000) with a Galactic radius of 4.4

---

[6] The luminosities presented by Mottram et al. (2011b) were estimated using kinematic distances derived from the Brand & Blitz (1993) rotation curve and so have been rescaled to the kinematic distances derived from the Reid et al. (2009) rotation model.
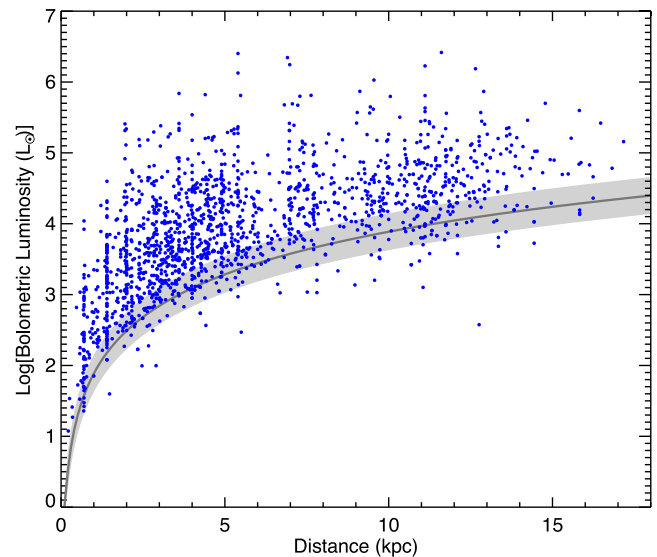
Figure 5. The luminosity distribution as a function of the heliocentric distance. The dark line and light grey shaded region indicate the limiting sensitivity of the MSX 21 μm band and its associated uncertainty. Where multiple sources have been identified within the MSX beam, the luminosity has been apportioned, resulting in some sources being located below the sensitivity limit.

± 0.5 kpc at an angle of 44° ± 10° (Benjamin 2008), the near and far 3 kpc arms, and the four principle arms (i.e. Norma, Sagittarius, Perseus and Scutum–Centaurus arms). The location of the spiral arms is based on the Georgelin & Georgelin (1976) model but has been modified to take account of recent maser parallax distances and updated directions for the spiral-arm tangents (Dame et al. 2001). The RMS survey excluded the innermost 20° of the Galactic longitude (i.e. 350° < ℓ < 10°): we are therefore not sensitive to
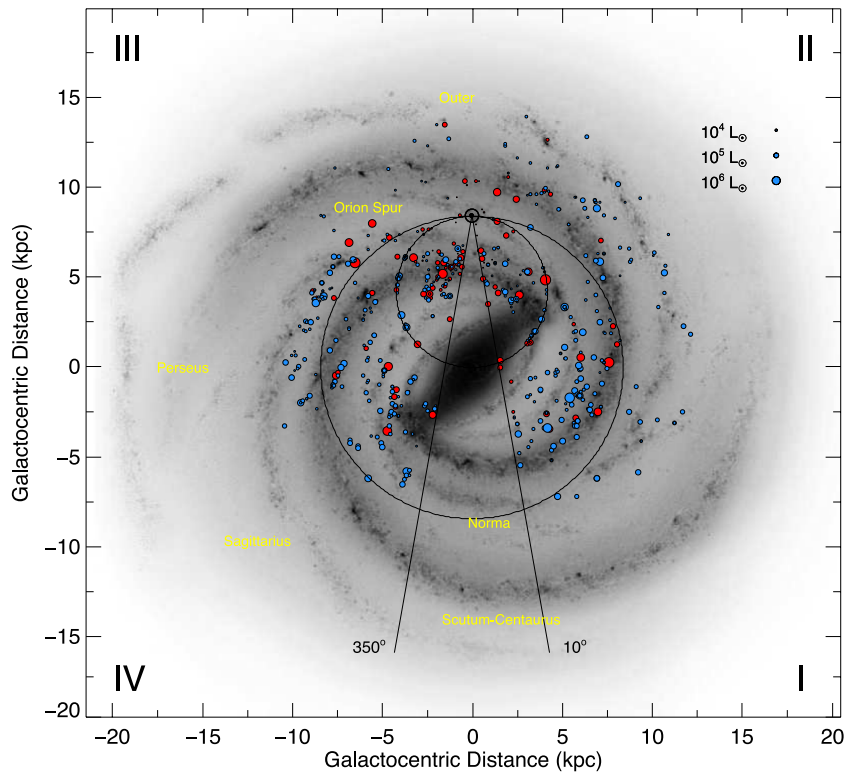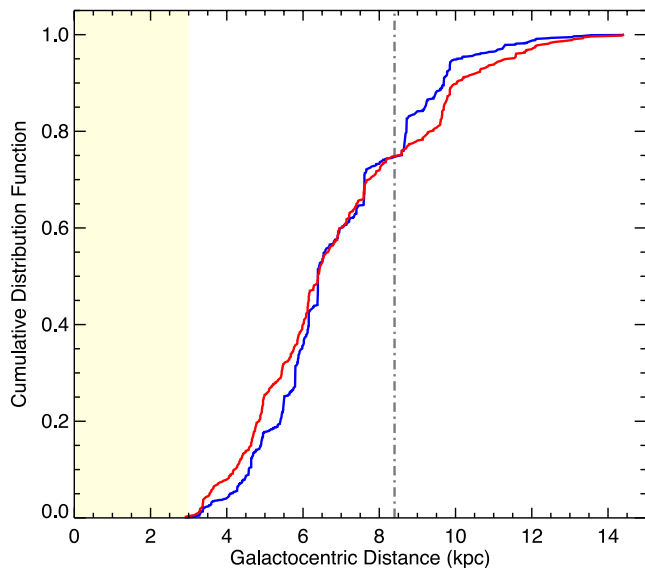
Figure 6. Galactic distribution of all MYSOs and H ɪɪ regions with bolometric luminosities greater than 10^4 $L_\odot$. We show the kinematic positions of the complexes and individual sources as red and blue circles, respectively. The sizes of the markers give an indication of their luminosity, as depicted in the upper-right corner. The background image is a sketch of the Galaxy produced by Robert Hurt of the Spitzer Science Center in consultation with Robert Benjamin. The position of the Sun is shown by the small circle above the Galactic Centre. The two solid black lines enclose the Galactic Centre region that was excluded from the RMS surveys due to problems with source confusion and distance determination. The smaller of the two black dot–dashed circles represents the locus of tangent points, while the larger circle shows the radius of the solar circle.

any SF taking place within ∼3 kpc of the Galactic Centre and any features of Galactic structure within this radius.

The distribution of the young embedded population with respect to the expected positions of the spiral arms given from the models suggests that they are correlated, but quantifying this correlation and estimating its significance is non-trivial. The correlation between the spiral arms and the RMS complexes appears slightly stronger than for the isolated sources in that there are far fewer found to be located in the interarm regions (cf. Stark & Lee 2006). The reason for this is that the more reliable maser parallax and spectrophotometric distances are available for many of the complexes, and for the others, the higher number of associated RMS sources is likely to provide a better estimate of their systemic radial velocities and thus kinematic distances (cf. Russeil 2003). Individual sources are by comparison more poorly constrained. The source density of the inner parts of the spiral arms ($R_{GC} < 10$ kpc) appears to be roughly uniform, which suggests that a similar level of SF is taking place within them. This is in contrast with the result of near- and mid-infrared source counts presented by Benjamin (2008). These authors found enhancements of the source counts towards the Scutum and Centaurus tangents, but not towards the Norma or Sagittarius tangents. This led them to speculate that the Galaxy has two principal arms (i.e. Perseus and Scutum–Centaurus arms) with the Norma or Sagittarius perhaps being optically visible arms that are not associated with any enhancement in the old stellar disc. The emphasized Perseus and Scutum–Centaurus arms portrayed in the background image used in Fig. 6 reflect this (Benjamin 2008; Churchwell et al. 2009).

Comparison of the number of northern and southern RMS sources above the completeness limit shows a similar number of sources (i.e. 47 ± 4.6 and 53 ± 4.0 per cent for the northern and southern Galactic plane, respectively). We also find similar proportions of the northern and southern samples inside and outside the solar circle (∼75 and ∼25 per cent, respectively). The proportional distributions of the RMS sources between the northern and southern Galactic plane and inside and outside the solar circle are similar to that reported by Bronfman et al. (2000) from a study of ∼750 *IRAS* selected candidate UC H ɪɪ regions (67 and 33 per cent inside and outside the solar circle). If we look at the integrated bolometric luminosities of the sources inside and outside the solar circle, we also find the same fractional distribution as the source counts; this is nicely illustrated in the cumulative distribution function of both of these parameters in Fig. 7. The fraction of the total bolometric luminosity of sources within the solar circle is slightly lower than the value of 81 per cent reported by Bronfman et al. (2000); however, it is still consistent within the errors.

The similarity of the distributions of the total luminosity and source counts both inside and outside of the solar circle would suggest that the mean source luminosity is also likely to be broadly similar. Furthermore, given that the proportion of molecular gas inside and outside the solar circle is very similar to the luminosity and source counts, it follows that the average star formation efficiencies (SFE) are also comparable. This is consistent with the results of Snell, Carpenter & Heyer (2002) that the SFE for the population of molecular clouds in the outer Galaxy (as estimated from the clouds' L_{FIR}/M ratio) is similar to that found for the inner

Figure 7. Cumulative distribution function of the RMS source counts (red) and integrated bolometric luminosities (blue) of all MYSOs and H II regions above $2 \times 10^4 \, L_\odot$ as a function of the Galactocentric radius. The shaded region shows the Galactocentric distance region excluded from the RMS survey, while the vertical dashed line indicates the distance of the solar circle.

Galaxy population. The clouds in the outer Galaxy are as active in forming massive stars, despite the fact that the clouds are typically one to two orders of magnitude less massive than those in the inner Galaxy. The difference in star formation rates (SFRs) between inner and outer regions, then, is principally a function of the molecular cloud formation rate, not of SF *within* the molecular clouds. Several factors are thought to contribute to the comparative lack of molecular clouds, including lower metallicity, gas density and ambient pressure (Snell et al. 2002; Leroy et al. 2008; Schruba et al. 2011).

4.1 Surface density of massive SF

In the upper panels of Fig. 8, we present the massive SF and bolometric luminosity surface density distributions as a function of the Galactic radius. These plots include only sources with luminosities above the completeness threshold. However, sources above this limit dominate the total luminosity, and inclusion of the lower luminosity sources has very little impact on the measured values or the overall structure of the distribution. This has been produced in the standard way by dividing the total number of sources in each annulus by its area. We have also applied a heliocentric radius limit of 17 kpc to avoid including areas that are effectively outside the regions the RMS survey is sensitive to and the area of the wedge towards the inner Galaxy excluded from the RMS survey (i.e. $350° < \ell < 10°$). The overall distributions are both highly structured showing numerous peaks, many at similar radii. In the middle and lower panels of this figure, we separate these distributions into the northern and southern populations, respectively, and identify the spiral arms that are coincident with the observed peaks in the first and fourth quadrants.

4.1.1 Source counts

Peaks are seen in the RMS source distribution for the whole sample at ~3, 5, 6 and 10 kpc. The 5 and 6 kpc peaks are coincident with the segments of the Scutum–Centaurus and Sagittarius arms located in the northern Galactic plane, but no peaks are found towards the

Perseus and Norma arms located in this part of the Galaxy. Both of these arms, however, extend into the outer Galaxy and are spread over a larger range of Galactic radii than the other arms (6–10 and 8–13 kpc for the Perseus and Norma arms, respectively), thus smearing their Galactocentric distribution. All four distinct peaks seen in the upper-left panel of Fig. 8 (i.e. ~3, 5, 6 and 10 kpc) are seen in the southern Galactic plane distribution; these can be attributed to the near and far 3 kpc arms, the Norma, Scutum–Centaurus and Sagittarius arms, respectively.

All of the peaks located within the solar circle have been previously reported for RMS subsamples located in the first and fourth quadrants (i.e. Urquhart et al. 2011a and Urquhart et al. 2012, respectively). The overall distribution, as well as that of the northern and southern Galactic plane samples, is also very similar to those presented by Bronfman et al. (2000). Similar features have also been reported in the Galactocentric distribution of methanol masers (Green & McClure-Griffiths 2011), H II regions (Paladini, Davies & De Zotti 2004; Anderson & Bania 2009), thermal dust emission (Dunham et al. 2011) and molecular gas (Rathborne et al. 2009). The correlation of these observed peaks in the Galactocentric distribution, and between the radial velocities and Galactic longitudes of the RMS sources and the spiral-arm models (i.e. Fig. 3), is consistent with four-armed models of the Galaxy (e.g. Georgelin & Georgelin 1976; Nakanishi & Sofue 2006).

We include the H_2 gas surface density derived from CO observations (Nakanishi & Sofue 2006) in Fig. 8 for comparison. The molecular gas surface density peaks at ~5 kpc before decreasing exponentially with a scalelength of ~2.5 kpc (as determined by Nakanishi & Sofue 2006). It is interesting to note that the source densities for the parts of the spiral arms inside the solar circle are very similar (i.e. the peaks have approximately the same height ~5–6 kpc^{-2}), implying that the SFR per unit area is fairly constant. The source distribution is broadly similar to that of the molecular gas, with the peaks in the source distribution identifying localized regions where the SFE and/or SFR have been enhanced. The massive SFR surface density is therefore correlated with the molecular gas surface density. This is consistent with the SFR derived by Misiriotis et al. (2006) from an analysis of *COBE*/Diffuse Infrared Background Experiment (DIRBE; 1.2, 2.2, 60, 100, 140, 240 μm) and *COBE*/Far Infrared Absolute Spectrophotometer (100–1000 μm) maps of the Galactic disc. Observations of nearby spiral galaxies (SINGS; Leroy et al. 2008) found them to have a similar structure to that of the Milky Way with an approximately flat SFE towards their interiors, which decreased exponentially at large radii as the H I/H_2 ratio increases. They infer that the SFRs do not appear to be directly sensitive to environmental conditions within giant molecular clouds (GMCs), but that the formation of GMCs is dependent on local conditions. This would suggest that the spiral structures do not directly influence the SFR but are localized regions where the formation of GMCs is significantly more efficient.

While the correlation between the SFR and molecular gas surface densities is good over the whole range of available Galactocentric distances, the peak in the SFR density at ~10 kpc, associated with the fourth quadrant section of the Sagittarius arm, may be of particular note. The SFR surface density here is perhaps only half that of the segments of the spiral arms located within the solar circle, but is projected against a much lower gas surface density, which may indicate that this spiral arm has a significantly enhanced SFR per unit gas mass beyond what is observed in the other spiral arms.

The final feature to note is the minimum seen in the source density plots at approximately 9 kpc. This minimum in the source counts changes by one bin in the distribution of the northern and
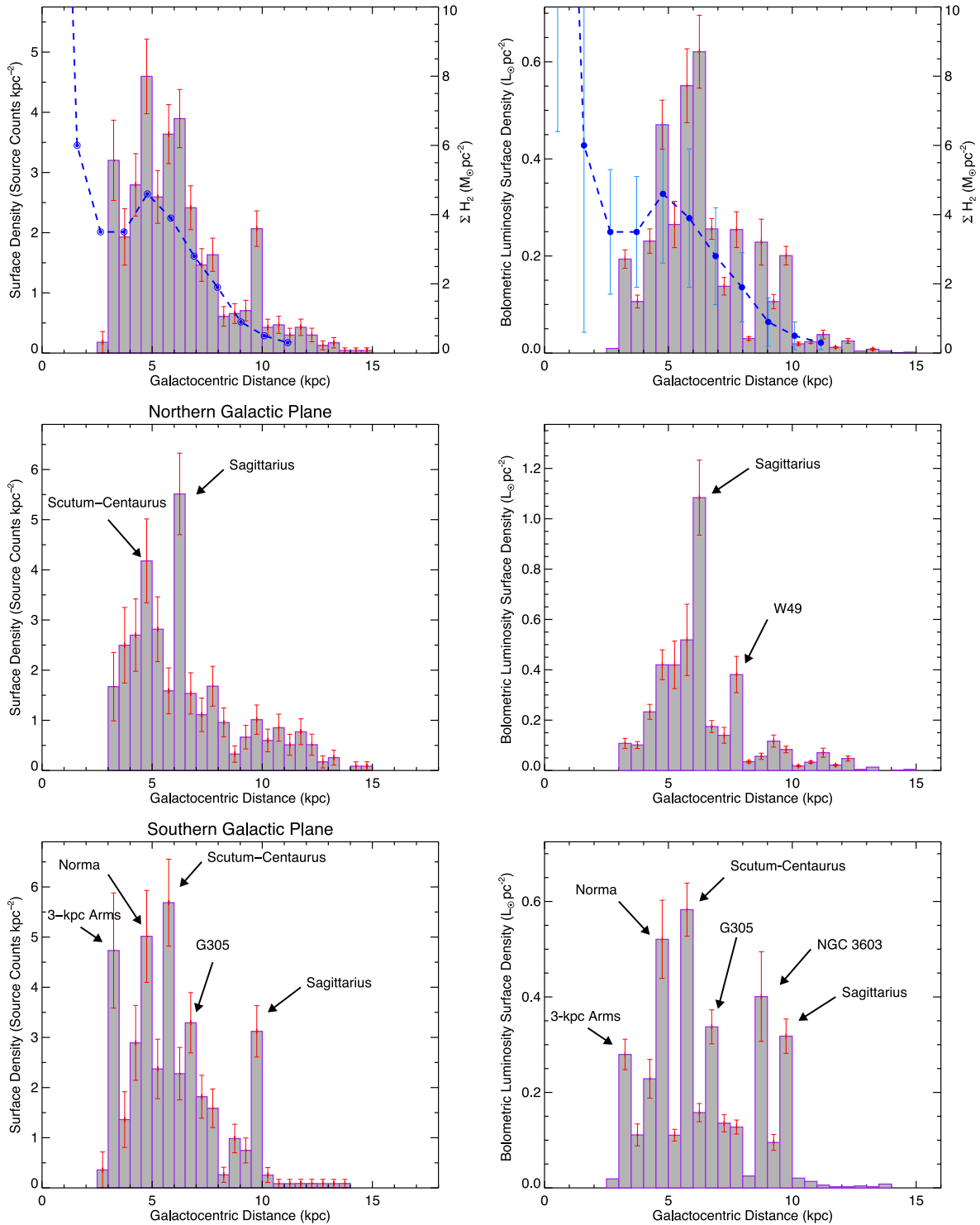
Figure 8. Source and bolometric luminosity surface density distributions of the population of RMS sources above the sample completeness limit ($2 \times 10^4 \, L_\odot$). The upper panel shows the distribution of the whole sample, while the middle and lower panels show the distribution of the northern and southern samples, respectively. The bin size is 0.5 kpc and the uncertainties in the left-hand panels are derived from Poisson statistics, while those in the right-hand panels are estimated assuming a 40 per cent error in the derived source luminosities. The blue circles and connecting dashed blue line indicate the H_2 gas surface density taken from Nakanishi & Sofue (2006, see the right-hand y-axis in the upper panel for values). The light blue vertical lines indicate the errors associated with the H_2 gas surface density points.

southern samples and so is somewhat washed out in the distribution of the whole sample presented in the upper-left panel of Fig. 8. This minimum roughly coincides with a similar feature identified by Lépine et al. (2011) that they associated with a step down in

metallicity and attributed to a kinematic barrier at corotation through which the gas cannot easily pass. Given the difficulties in assigning accurate distances to objects near the solar circle, which have local standard of rest (LSR) velocities near zero, it would be speculative

to draw conclusions regarding this feature at this stage, so we simply note the coincidence.

4.1.2 Bolometric luminosity

There are six peaks present in the luminosity distribution, four of which coincide with the peaks seen in the source count distribution (i.e. ∼3, 5, 6 and 10 kpc). Although the SFR surface densities are similar for the segments of the spiral arms located within the solar circle, we find that the luminosity surface densities increase with increasing distance from the Galactic Centre. As already mentioned, the overall SFR is similar for the most of the arms within the solar circle, so this increase in luminosity per unit area is likely to be due to an increase in the mean luminosity of the embedded stars. This would imply that the SFE is somehow enhanced in the outer parts of the spiral arms or that it is suppressed in the inner regions, perhaps by the interaction of the gas with the Galactic bar.

We estimate the total luminosity of RMS embedded MYSO and H II region populations by multiplying the luminosity in each bin by the area of the bin annulus and summing all of these together to obtain a value of 0.76×10^8 L$_\odot$. In this estimate, we include all RMS sources, including those below the completeness in order to obtain a value that can be compared to that determined by Bronfman et al. (2000) from their *IRAS* selected sample. Our value is approximately half the value determined by Bronfman et al. (2000); however, their fluxes were drawn from the *IRAS* point source catalogue, whose large beam sizes (i.e. ∼5 arcmin at 100 µm) would have likely caused an overestimate of the total luminosity. It is also possible that their sample will include a significant number of more evolved H II regions. The luminosity of the embedded massive star population represents only a few per cent of the total Galactic far-infrared luminosity (i.e. ∼2×10^9 L$_\odot$; Sodroski et al. 1997, *COBE* DIRBE; Bloemen, Deul & Thaddeus 1990, *IRAS*), which would point to a relatively short embedded lifetime.

We use the total Galactic luminosity of the embedded massive star population to estimate the fraction contributed from each of the massive complexes discussed in Section 2.4. This value is given in the last column of Table 2. The full version of this table is only available in electronic form; however, the portion presented in this table lists the 25 most luminous complexes in the Galaxy. It is interesting to note that the 10 most luminous complexes contain only 8 per cent of the embedded MYSOs and H II region population, but contribute approximately 30 per cent of the total luminosity of the Galactic embedded massive star population. A recent study of the *Wilkinson Microwave Anisotropy Probe* free–free foreground emission maps (Murray & Rahman 2010) found that the 18 most luminous Galactic star-forming regions are responsible for the production of over half of the total ionizing luminosity of Galaxy. This would suggest that the formation or early O-type stars is concentrated into a relatively small number of regions in the Galaxy. Furthermore, this suggests that the initial mass function (IMF) is not constant on all scales and supports the cluster-based IMF model (e.g. Kroupa & Weidner 2003).

Some of these complexes are clearly very luminous and contribute significantly to some of the peaks seen in Fig. 8. For example, the increased luminosity surface density seen in the northern distribution at 6 kpc is somewhat due to the presence of the W51 star-forming complex in this bin, which is the most active region in the Galaxy (∼7 per cent of the total RMS luminosity; see Table 2). The additional peaks found at ∼7, 8 and 9 kpc can be attributed to the G305, W49 and NGC 3603 star-forming complexes, respectively, three more of the most active in the Galaxy.

4.1.3 Sagittarius arm

We note that the Sagittarius arm is a prominent feature in all of the plots presented in Fig. 8, much more strongly detected and clearly defined in both hemispheres than, e.g., the Perseus arm. The Sagittarius arm also shows up in the scaleheight distribution presented in the lower panel of Fig. 11. The peak seen at approximately 6.5 kpc is only a feature of the northern Galactic plane sample, and the middle panels of Fig. 8 show that the bin corresponding to this Galactocentric distance is coincident with the Sagittarius arm. This spiral arm therefore has a larger scaleheight than the other segments of the spiral arms located within the solar circle.

This prominence suggests that, as traced by massive SF, the Sagittarius arm is a major feature of the Galaxy, rather than the minor arm implied by the sketch in Fig. 6. This is contrary to the Benjamin (2008) conclusion, although the tracers are quite different as the GLIMPSE data trace an old stellar population so the implication may be that the SF history of the arms is periodic or variable.

4.2 Galactic mid-plane and scaleheight distributions

In Fig. 9, we show the latitude distribution of the whole RMS sample with respect to the Galactic mid-plane ($b = 0$), as a function of the Galactic longitude and radial velocity. This plot reveals that sources occupy a narrow range of distances ($|z| < 100$ pc) within the inner part of the Galactic plane (i.e. $300° < \ell < 60°$) and flare to significantly more positive and negative distances (up to 600 pc in either direction) in the second and third quadrants, respectively. These deviations from the mid-plane follow the structure of the outer Galaxy, which is known to be significantly warped (Oort, Kerr & Westerhout 1958). This is not the first example of infrared sources tracing the Galactic warp, this was previously reported by Wouterloot et al. (1990, see their fig. 3), but provides a nice confirmation of their results.

The velocities of the sources associated with these large excursions from the mid-plane are reasonably coherent and tend towards the more extreme ends of the velocity distributions. Comparison with spiral-arm velocities (cf. Fig. 3) illustrates a correspondence with the Outer arm in the second quadrant and the Perseus arm in the third quadrant. The Perseus arm is also seen in the second quadrant, but lies close to the Galactic mid-plane with perhaps a small excursion to negative distances between $\ell = 90°$ and $120°$. This corresponds to the part of the plane where the largest excursion to positive distances is seen towards the Outer arm. The observed lower source density associated with the third quadrant part of the Perseus arm compared to the Outer arm is likely due to the larger heliocentric distances in this part of the Galaxy.

There are also regions where the velocities are correlated within the inner Galaxy, and these can be matched to the expected spiral-arm tangents: Scutum–Centaurus, Sagittarius, Perseus, Norma and Scutum–Centaurus at $\ell = 20°$, $30°$, $340°$, $330°$ and $300°$, respectively. We also note that nearly all of the complexes identified are located towards the Galactic mid-plane.

In Fig. 10, we show the latitude distribution of the whole sample as well as the subsample of sources with luminosities above 2×10^4 L$_\odot$; scaleheights obtained from fits to both samples agree within the errors with 37.7 ± 0.8 and 39.5 ± 1.7 pc, respectively. This suggests that the SF scaleheight is relatively luminosity independent over the range of luminosities covered by the RMS survey (i.e. ∼10^2–10^6 L$_\odot$). We also find that the scaleheight is similar for the northern and southern Galactic plane subsamples.
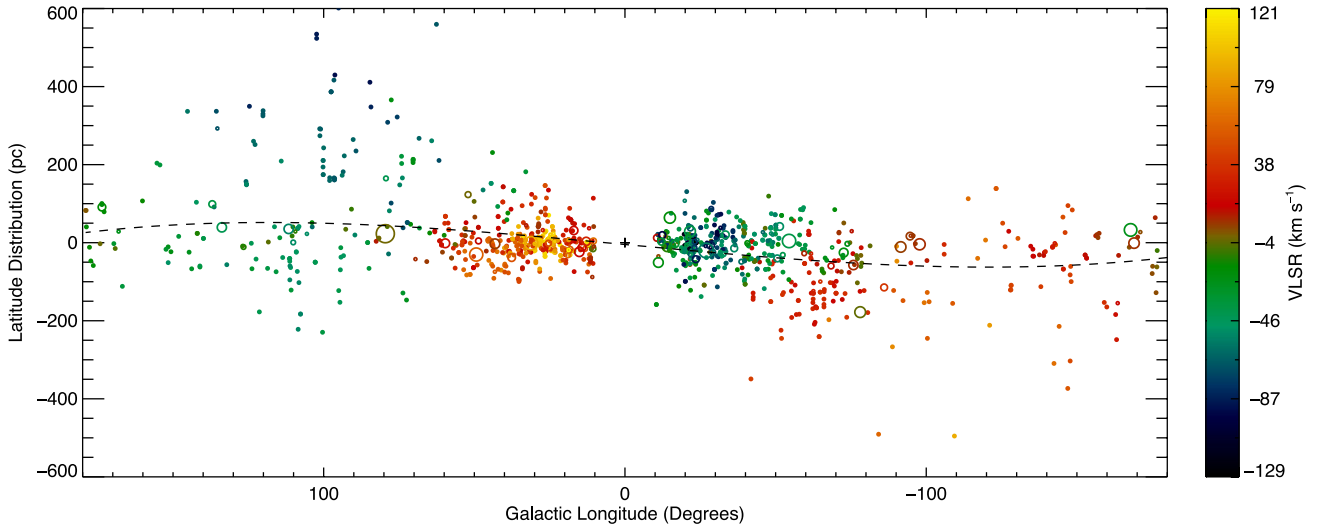
Figure 9. Distance from the Galactic mid-plane (Z_0) for the embedded RMS source population is shown as a function of the Galactic longitude. The complexes are shown as open circles (larger symbol sizes indicate a greater number of members), while individual sources are shown as filled circles. The symbol colours give the source or systematic velocity of the complexes; see the colour bar for values. The black dashed curve shows the result of a third-order polynomial fitted to the source density and thus traces the mean source latitude. This is provided to emphasize the shape of the Galactic warp.

We note that these scaleheights are significantly larger than those derived from the ATLASGAL–CORNISH sample of UC H II regions and their star-forming clumps ($\sim 22.4 \pm 1.9$ and 28.1 ± 2.6 pc, respectively; Urquhart et al. 2013), or methanol masers (27 ± 1 pc; Green & McClure-Griffiths 2011) identified by the Methanol Multibeam (MMB) survey (Caswell 2010) or indeed the subsample of RMS sources located within the Galactic Ring Survey (GRS) region (i.e. 30.2 ± 2.8 pc; Urquhart et al. 2011a). However, most of these surveys have focused on sources located primarily within the inner Galaxy and so the larger scaleheight obtained for the *full* RMS sample is simply a consequence of including the outer Galaxy sources, which have a significantly larger distribution around the mid-plane (e.g. see Fig. 9).
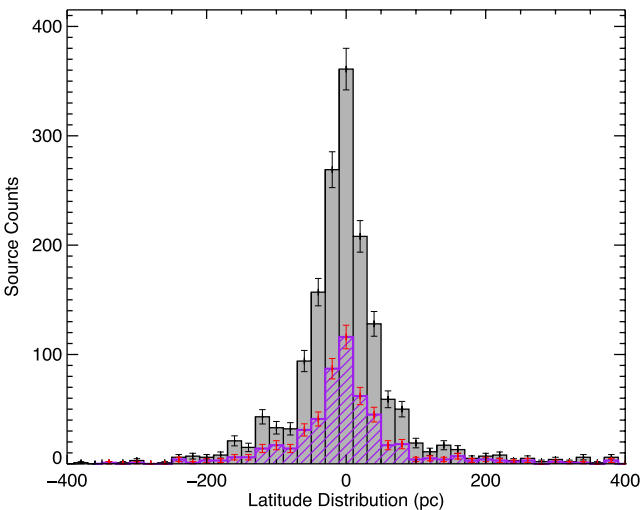

Figure 10. Galactic latitude distribution of the embedded RMS population. The grey histogram shows the distribution of the whole sample while the purple hatched histogram shows the subsample of sources above the completeness limit. The bin size is 20 pc and the errors are determined from Poisson statistics.

In Fig. 11, we present the offset of the Galactic disc from the mid-plane and measurements of the scaleheight as a function of the Galactocentric radius (see Table 4 for values). The upper panel shows the distribution of the mean southern and northern subsample positions about the mid-plane separately to illustrate the effects of the Galactic warp. The lower panel shows the absolute mean height from the mid-plane for the entire sample: the deflection of the southern plane below the mid-plane is similar to that of the northern plane above it. The samples for each of these plots have been divided into eight bins of equal number of sources. The distance from the mid-plane is determined from the average over each bin, and the error is the standard error on the mean. The scaleheight and its associated error are determined from an exponential fit to the histogram of data in each bin. The Galactocentric radius is simply the mean radius of the binned data, and the error is the standard deviation of this value.

These plots show that both the northern and southern samples are tightly associated with the Galactic mid-plane for radii <7 kpc, after which they begin to diverge. The Galactic scaleheight is also seen to increase modestly with increasing Galactocentric distance with a value of ~ 20 pc at a radius of 4 kpc to ~ 30 pc at 8 kpc. At larger distances this slope increases very rapidly, reaching values of almost 200 pc at a Galactocentric distance of ~ 11 kpc. We find no significant difference between the northern or southern samples. The increase in scaleheight with increasing Galactocentric radius is very closely matched by the height of the molecular H_2 layer both in magnitude and slope (e.g. Grabelsky et al. 1987; Bronfman et al. 1988; Wouterloot et al. 1990; Nakanishi & Sofue 2006). Similar increases in scaleheight have been reported from H I studies of the Milky Way (Malhotra 1995) and nearby edge-on spiral galaxies (Rupen 1991), although in these cases the H I discs are significantly thicker, and increase much more quickly (100 to 220 pc between 3 and 8 kpc; Malhotra 1995).

Lin & Shu (1964) predict that the influence of spiral-arm shocks is significantly weaker in the outer Galaxy beyond the corotation radius ($R_{GC} \sim 8$ kpc). This would lead to less confined and smooth spiral features outside this radius, with supernovae probably playing a more important role than the spiral structure in determining
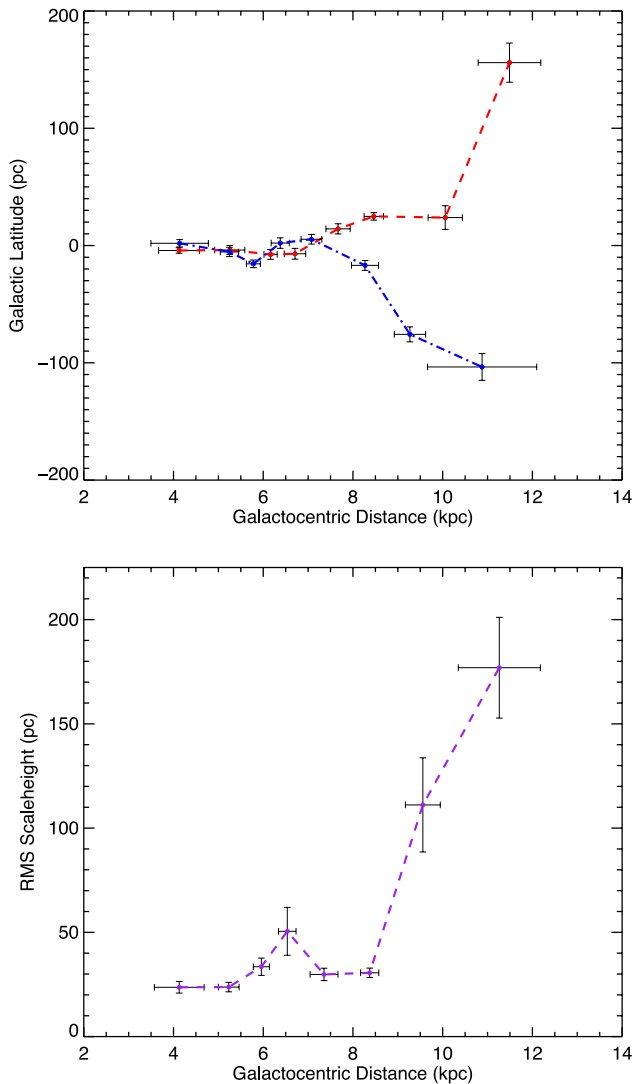
Figure 11. The offset of the mean RMS source position from the mid-plane and the scaleheight as functions of the Galactocentric radius are shown in the upper and lower panels, respectively. The northern and southern Galactic plane subsamples are shown in red and blue in the upper panel, while in the lower panel we show only the distribution of the whole sample in purple.

Table 4. Changes in scaleheight over the disc.

R_{GC} (kpc)	Z_o (pc)	ΔZ_o (pc)	Scaleheight (pc)	ΔScaleheight (pc)
4.15	−1.14	2.04	22.32	1.49
5.27	−7.26	2.63	23.03	2.00
5.99	−11.43	2.97	31.57	3.35
6.55	−1.01	3.07	51.01	12.07
7.37	6.37	3.44	31.93	3.00
8.40	3.53	2.85	31.21	2.99
9.59	−53.40	7.30	113.89	23.79
11.29	55.01	12.71	176.82	24.82

the state of the ISM (Dib et al. 2009). Our RMS sources are young enough that they are still embedded in their natal molecular clouds and have not been dispersed by the Galactic gravitational potential, and are thus a good probe of changes in the molecular ISM. The relatively sharp change in the distribution and increase of the

scaleheight of RMS sources outside the corotation radius would therefore support the idea of the ISM being generally more flocculent, dominated by supernovae and less confined in the outer Galaxy.

5 SUMMARY AND CONCLUSIONS

The RMS survey is a Galaxy-wide mid-infrared selected sample of ∼1750 MYSOs and UC H II regions. In this paper, we summarize the results of our previous molecular line follow-up observations and present additional observations towards ∼800 RMS sources. Observations of the ^{13}CO (1–0) transition were made towards sources not previously observed, and the higher density tracers CS and NH$_3$ were used to identify the correct radial velocity for sources towards which multiple velocity components had been detected.

Combining these results with those from our previous follow-up campaign of observations and with archival line surveys, we have obtained radial velocities to all but a handful of sources. Use of the positional and radial velocities of our sources combined with inspection of mid-infrared images has allowed us to group sources into ∼120 star-forming complexes. Approximately one third of the RMS sample are associated with these complexes.

We have conducted an in-depth literature search to identify sources/complexes for which a reliable distance has previously been determined. Where a distance was not available, we have used the radial velocities and the Galactic rotation curve derived by Reid et al. (2009) to calculate kinematic distances, and have used the H I self-absorption to resolve any outstanding KDAs for sources within the solar circle (∼200 sources). We have obtained distances to ∼1650 H II regions and YSOs, which corresponds to over 90 per cent of the embedded RMS population. These distances are used to estimate bolometric luminosities using model fits to the SEDs previously derived by Mottram et al. (2011a). We find that we are complete above 2×10^4 L$_\odot$ to a heliocentric distance of ∼18 kpc.

We use this sample to investigate the Galactic distribution of massive stars with respect to the position of the spiral arms, the Galactic long bar, the Galactic warp and the flaring of the molecular disc. Finally, we compare the distribution of massive stars in our Galaxy with other nearby spiral galaxies. Our main findings are as follows.

(i) We observe a high surface density of MYSO and H II regions within the solar circle and a decaying exponential at larger distances from the Galactic Centre, consistent with other measurements of the SFR determined from infrared and submillimetre studies. Variation in the SFR and efficiency within the solar circle appear to be attributable to specific large-scale features of Galactic structure, such as the spiral arms and the ends of the bar.

(ii) We have compared the distribution of the RMS sources with the expected position of the spiral arms using longitude–velocity diagrams, in three dimensions and as a function of Galactocentric radii, and have found them to be in good agreement. Our results are therefore consistent with a model of the Galaxy consisting of four major arms.

(iii) The embedded source and H$_2$ gas surface densities have a very similar overall distribution. We find that the source surface densities associated with the segments of the spiral arms located within the solar circle are very similar, which suggests that the massive SFR per unit molecular gas mass is approximately constant for the inner parts of the spiral arms.

(iv) We also find that the luminosity surface density increases with increasing Galactocentric radius for the segments of the spiral

arms located within the solar circle. Given that the source surface densities are approximately the same for the spiral arms in this region, this indicates that the mean source luminosity is increasing with distance from the Galactic Centre. This increase in mean source luminosity can be attributed to a very small number of extremely active star-forming complexes (i.e. W51, W49, W43) where the star and/or clump formation efficiency is significantly enhanced.

(v) We find that the scaleheight of massive SF as measured from the whole sample is 39.5 ± 1.7 pc, which is larger than has been reported previously ($\sim$30 pc). Most earlier studies have concentrated on inner Galaxy samples and exclude the significant flaring of the disc observed in the outer Galaxy. We measure the scaleheight as a function of the Galactocentric distance and find that it increases only modestly between $\sim$4 and 8 kpc (i.e. from $\sim$20 to 30 pc), but much more rapidly at larger distances.

(vi) We estimate the total integrated bolometric luminosity of the embedded MYSO and H II region population to be $\sim 0.76 \times 10^8$ L$_\odot$; however, we find that the 10 most luminous complexes contribute almost 30 per cent of the total integrated RMS bolometric luminosity while comprising only 8 per cent of the sources.

ACKNOWLEDGEMENTS

REFERENCES

Alvarez H., May J., Bronfman L., 1990, ApJ, 348, 495
Anderson L. D., Bania T. M., 2009, ApJ, 690, 706
Bains I. et al., 2006, MNRAS, 367, 1609
Benjamin R. A., 2008, in Beuther H., Linz H., Henning T., eds, ASP Conf. Ser. Vol. 387, Massive Star Formation: Observations Confront Theory. Astron. Soc. Pac., San Francisco, p. 375
Benjamin R. A. et al., 2003, PASP, 115, 953
Bertoldi F., 1989, ApJ, 346, 735
Binney J., Gerhard O. E., Stark A. A., Bally J., Uchida K. I., 1991, MNRAS, 252, 210
Blitz L., Spergel D. N., 1991, ApJ, 379, 631
Bloemen J. B. G. M., Deul E. R., Thaddeus P., 1990, A&A, 233, 437
Brand J., Blitz L., 1993, A&A, 275, 67
Bronfman L., Cohen R. S., Alvarez H., May J., Thaddeus P., 1988, ApJ, 324, 248
Bronfman L., Nyman L.-A., May J., 1996, A&AS, 115, 81
Bronfman L., Casassus S., May J., Nyman L.-Å., 2000, A&A, 358, 521
Carey S. J. et al., 2009, PASP, 121, 76
Caswell J. L. E., 2010, MNRAS, 404, 1029
Caswell J. L., Haynes R. F., 1987, A&A, 171, 261
Churchwell E. et al., 2009, PASP, 121, 213
Clark J. S., Davies B., Najarro F., MacKenty J., Crowther P. A., Messineo M., Thompson M. A., 2009, A&A, 504, 429
Cordes J. M., 2004, in Clemens D., Shah R., Brainerd T., eds, ASP Conf. Ser. Vol. 317, Milky Way Surveys: The Structure and Evolution of Our Galaxy. Astron. Soc. Pac., San Francisco, p. 211
Csengeri T., Bontemps S., Schneider N., Motte F., Dib S., 2011, A&A, 527, A135
Dame T. M., 2007, ApJ, 665, L163
Dame T. M., Hartmann D., Thaddeus P., 2001, ApJ, 547, 792
Davies B., Hoare M. G., Lumsden S. L., Hosokawa T., Oudmaijer R. D., Urquhart J. S., Mottram J. C., Stead J., 2011, MNRAS, 416, 972
de Wit W. J., Testi L., Palla F., Vanzi L., Zinnecker H., 2004, A&A, 425, 937
Deharveng L., Lefloch B., Zavagno A., Caplan J., Whitworth A. P., Nadeau D., Martín S., 2003, A&A, 408, L25
Dib S., Walcher C. J., Heyer M., Audit E., Loinard L., 2009, MNRAS, 398, 1201
Dunham M. K., Rosolowsky E., Evans N. J., II, Cyganowski C., Urquhart J. S., 2011, ApJ, 741, 110
Dwek E. et al., 1995, ApJ, 445, 716
Ellsworth-Bowers T. P. et al., 2013, ApJ, 770, 39
Fish V. L., Reid M. J., Wilner D. J., Churchwell E., 2003, ApJ, 587, 701
Georgelin Y. M., Georgelin Y. P., 1976, A&A, 49, 57
Grabelsky D. A., Cohen R. S., Bronfman L., Thaddeus P., May J., 1987, ApJ, 315, 122
Green J. A., McClure-Griffiths N. M., 2011, MNRAS, 417, 2500
Hammersley P. L., Garzón F., Mahoney T. J., López-Corredoira M., Torres M. A. P., 2000, MNRAS, 317, L45
Hill T. et al., 2012, A&A, 548, L6
Hindson L., Thompson M. A., Urquhart J. S., Clark J. S., Davies B., 2010, MNRAS, 408, 1438
Hoare M. G. et al., 2005, in Cesaroni R., Felli M., Churchwell E., Walmsley M., eds, Proc. IAU Symp. 227, Massive Star Birth: A Crossroads of Astrophysics. Cambridge Univ. Press, Cambridge, p. 370
Hoare M. G. et al., 2012, PASP, 124, 939
Indermuehle B., Edwards P., Brooks K., Urquhart J., 2013, CSIRO Data Archive, p. 1
Jackson J. M., Bania T. M., Simon R., Kolpak M., Clemens D. P., Heyer M., 2002, ApJ, 566, L81
Kennicutt R. C., 2005, in Cesaroni R., Felli M., Churchwell E., Walmsley M., eds, Proc. IAU Symp. 227, Massive Star Birth: A Crossroads of Astrophysics. Cambridge Univ. Press, Cambridge, p. 3
Kolpak M. A., Jackson J. M., Bania T. M., Clemens D. P., Dickey J. M., 2003, ApJ, 582, 756
Kroupa P., Weidner C., 2003, ApJ, 598, 1076
Kutner M. L., Ulich B. L., 1981, ApJ, 250, 341
Ladd N., Purcell C., Wong T., Robertson S., 2005, Publ. Astron. Soc. Aust., 22, 62
Lawrence A. et al., 2007, MNRAS, 379, 1599
Lépine J. R. D. et al., 2011, MNRAS, 417, 698
Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782
Lin C. C., Shu F. H., 1964, ApJ, 140, 646
Loup C., Forveille T., Omont A., Paul J. F., 1993, A&AS, 99, 291
Lumsden S. L., Hoare M. G., Oudmaijer R. D., Richards D., 2002, MNRAS, 336, 621
Lumsden S. L., Hoare M. G., Urquhart J. S., Oudmaijer R. D., Davies B., Mottram J. C., Cooper H. D. B., Moore T. J. T., 2013, ApJS, 208, 11
Malhotra S., 1995, ApJ, 448, 138
McClure-Griffiths N. M., Dickey J. M., Gaensler B. M., Green A. J., Haverkorn M., Strasser S., 2005, ApJS, 158, 178

Minniti D. et al., 2010, New Astron., 15, 433

Misiriotis A., Xilouris E. M., Papamastorakis J., Boumis P., Goudis C. D., 2006, A&A, 459, 113

Moisés A. P., Damineli A., Figuerêdo E., Blum R. D., Conti P. S., Barbosa C. L., 2011, MNRAS, 411, 705

Motte F., Schilke P., Lis D. C., 2003, ApJ, 582, 277

Mottram J. C., Hoare M. G., Lumsden S. L., Oudmaijer R. D., Urquhart J. S., Meade M. R., Moore T. J. T., Stead J. J., 2010, A&A, 510, A89

Mottram J. C. et al., 2011a, A&A, 525, A149

Mottram J. C. et al., 2011b, ApJ, 730, L33

Murray N., Rahman M., 2010, ApJ, 709, 424

Nakanishi H., Sofue Y., 2006, PASJ, 58, 847

Netterfield C. B. et al., 2009, ApJ, 707, 1824

Oort J. H., Kerr F. J., Westerhout G., 1958, MNRAS, 118, 379

Paladini R., Davies R. D., De Zotti G., 2004, MNRAS, 347, 237

Parsons H., Thompson M. A., Clark J. S., Chrysostomou A., 2012, MNRAS, 424, 1658

Rathborne J. M., Johnson A. M., Jackson J. M., Shah R. Y., Simon R., 2009, ApJS, 182, 131

Reed B. C., 2000, AJ, 120, 314

Reid M. J. et al., 2009, ApJ, 700, 137

Reipurth B., Schneider N., 2008, in Reipurth B., ed., Handbook of Star Forming Regions, Vol. I, Star Formation and Young Clusters in Cygnus. Astron. Soc. Pac., San Francisco, p. 36

Roman-Duval J., Jackson J. M., Heyer M., Johnson A., Rathborne J., Shah R., Simon R., 2009, ApJ, 699, 1153

Rupen M. P., 1991, AJ, 102, 48

Russeil D., 2003, A&A, 397, 133

Schlingman W. M. et al., 2011, ApJS, 195, 14

Schruba A. et al., 2011, AJ, 142, 37

Schuller F. et al., 2009, A&A, 504, 415

Skrutskie M. F. et al., 2006, AJ, 131, 1163

Snell R. L., Carpenter J. M., Heyer M. H., 2002, ApJ, 578, 229

Sodroski T. J., Odegard N., Arendt R. G., Dwek E., Weiland J. L., Hauser M. G., Kelsall T., 1997, ApJ, 480, 173

Stark A. A., Lee Y., 2006, ApJ, 641, L113

Stil J. M. et al., 2006, AJ, 132, 1158

Taylor J. H., Cordes J. M., 1993, ApJ, 411, 674

Urquhart J. S., Thompson M. A., Morgan L. K., Pestalozzi M. R., White G. J., Muna D. N., 2007a, A&A, 467, 1125

Urquhart J. S. et al., 2007b, A&A, 474, 891

Urquhart J., Hoare M., Lumsden S., Purcell C., 2008a, ATNF Proposal No. M270, p. 1814

Urquhart J. S., Hoare M. G., Lumsden S. L., Oudmaijer R. D., Moore T. J. T., 2008b, in Beuther H., Linz H., Henning T., eds, ASP Conf. Ser. Vol. 387, Massive Star Formation: Observations Confront Theory. Astron. Soc. Pac., San Francisco, p. 381

Urquhart J. S. et al., 2008c, A&A, 487, 253

Urquhart J. S. et al., 2009, A&A, 507, 795

Urquhart J., Hoare M., Morgan L., 2010a, ATNF Proposal No. 573, p. 3625

Urquhart J. S. et al., 2010b, PASA, 27, 321

Urquhart J. S. et al., 2011a, MNRAS, 410, 1237

Urquhart J. S. et al., 2011b, MNRAS, 418, 1689

Urquhart J. S. et al., 2012, MNRAS, 420, 1656

Urquhart J. S. et al., 2013, MNRAS, 435, 400

Whitworth A. P., Bhattal A. S., Chapman S. J., Disney M. J., Turner J. A., 1994, MNRAS, 268, 291

Wienen M., Wyrowski F., Schuller F., Menten K. M., Walmsley C. M., Bronfman L., Motte F., 2012, A&A, 544, A146

Wouterloot J. G. A., Brand J., Burton W. B., Kwee K. K., 1990, A&A, 230, 21

SUPPORTING INFORMATION

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

Figure 3. The upper and lower panels show the longitude and velocity distribution of the RMS sources in the northern and southern Galactic plane, respectively. The complexes are shown as open red circles, the size of which gives an indication of the source density of each complex, while the blue filled circles show the position of the rest of the sample. The grey-scale image shows the distribution of molecular gas as traced by the integrated ^{12}CO emission (Dame, Hartmann & Thaddeus 2001) for comparison. The locations of the spiral and local arms are shown as curved solid lines, coloured to identify the individual arms. The positions of the four main spiral and local arms have been taken from the model by Taylor & Cordes (1993) and updated by Cordes (2004), while the position of the near 3 kpc arm has been taken from Bronfman et al. (2000). (http://mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stt2006/-/ DC1).

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

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