Rubies in the Dust: Tracing High Mass Star Formation Throughout the Milky Way

Mark John Gallaway
University of Hertfordshire

June 2011

A thesis submitted in partial fulfilment of the requirements of the

University of Hertfordshire
for the degree of Doctor of Philosophy

June 2011

Rubies in the Dust: Tracing High Mass Star Formation Throughout the Milky Way

Mark John Gallaway
June 2011

University of Hertfordshire
Submitted for the degree of Doctor of Philosophy

Abstract

Over the last decade a number of potential tracers of massive star ($M > 8\mathrm{M}_{\odot}$) formation have been put forward. In this thesis I attempt to understand how these tracers relate to one another and attempt to identify the most suitable tracer for future surveys for massive star formation sites.

In this thesis we examine a number of these tracers; the Methanol Maser Multi- Beam Survey (MMB), the Red MSX Survey (RMS), the Boston University Five Colleges Radio Astronomical Observatory (BU-FCRAO) Galactic Ring Survey (GRS), the BOLOCAM Galactic Plane Survey (BGPS) and the Perretto & Fuller (P&F) Infrared Dark Cloud (IRDC) Catalogue, in addition to the Cyganowski Extended Green Objects Catalogue. This work employs a bespoke non-circular aperture photometry technique, K=1 Nearest Neighbour Analysis and Minimum Spanning Trees (MSTs) in multi-dimensional parameter space with oversampling, edge weighing, mean edge fracturing and convex hull fitting. Additional, new ^{13}CO observations were made of the young infrared cluster BDS[2003] 107 (Bica 107) and its environs.

We see that despite not being contained within the GLIMPSE Point Source Archive the bulk of masers have an infrared bright counterpart. Photometry of the counterparts shows that they occupy the same colour spaces as that previously determined in Ellingsen (2006); [3.6]-[4.5]>1 and [8.0]<1. We show that the bulk of RMS MYSOs do not exhibit masing and that a significant fraction of MYSOs are not found within the RMS . Additionally, we see that the EGO−RMS association rate is higher than expected.

The BGPS, GRS and P&F IRDC exhibit clustering and elongating, with a common characteristic clustering scale of the order of $6 - 8$ pc. We see that the BGPS is more strongly associated with massive star formation than the GRS. Additionally, we see that although in general all three hull types occupy similar co-located spatial positions they also appear as isolated hulls.

The analysis of Bica 107 shows that it is part of a larger star forming region containing Bica 108 and the ultra compact HII region, G5.89. The maser associated with Bica 107

appears to lie on the edge of the cluster's expanding CO shell.

The observation that the IRAC colour-magnitude occupied by the masers from the Ellingsen sample is consistent with the MMB, sample suggest that these objects have broadly consistent colours during their masing phase. This can be attributed to the dust and gas envelope being radiatively dominant.

The cross matching results indicate that the majority of MYSOs do not exhibit masing. The RMS appears to be missing MYSOs due to missing sources in the MSX catalogue and a photospheric bluing due to MSX large beam width, moving candidates outside the RMS colour cut. The RMS EGO relationship appears to be inconsistent with observed MYSO evolution and may be indicative of multiple EGO generation mechanism as suggested by De Buizer and Vacca (2010).

The BPGS and GRS objects and IRDCs do not appear to form a star formation sequence and their existence is not necessarily an indicator of on-going star formation; rather they are an indication of the potential for star formation. All three species types showing signs of clustering and elongation. The shared characteristic scale is suggestive that there may be a processes acting below the scale of the GMC but above that of a single star forming region.

The maser associated with Bica 107 appears to be either an example of triggered star formation or late onset star formation within the region and is not an example of continuing star formation within Bica 107.

We conclude that a GLIMPSE based colour-selected survey, with follow-up observation to reduce contamination, would be the most appropriate method for identifying MYSOs, given the reliability of the tracers examined in this thesis.

Acknowledgements

I wish to thank the University of Hertfordshire, the Science and Technology Facilities Council, ESO and IAU for their support.

This research has made use of NASA's Astrophysics Data System Abstract Service; the NASA / IPAC Infrared Science Archive (which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration); SAOImage DS9, developed by Smithsonian Astrophysical Observatory; FunTools developed by High Energy Astrophysics Division of the Smithsonian Astrophysical Observatory and data products from the GLIMPSE survey, which is a legacy science program of the Spitzer Space Telescope, funded by the National Aeronautics and Space Administration.

The James Clerk Maxwell Telescope is operated by The Joint Astronomy Centre on behalf of the Science and Technology Facilities Council of the United Kingdom, the Netherlands Organisation for Scientific Research, and the National Research Council of Canada. The UKIDSS project is defined in Lawrence et al 2007. UKIDSS uses the UKIRT Wide Field Camera (WFCAM; Casali et al 2007) and a photometric system described in Hewett et al 2006. The pipeline processing and science archive are described in Irwin et al (2008) and Hambly et al (2008).

The code used within this thesis was written using the Endthought distribution with the Matplotlib, Pyfits and Apypl modules.

I would like to thank the the members of the MMB survey, especially James Green, Shari Breen and Simon Ellingsen for all their help and positive input.

I wish to thank my supervisors Dr Phil Lucas, Dr Mark Thompson and Professor Janet Drew for their time, input and support although, I still haven't forgiven Mark for Sweden. In addition, I would like to thank Jane Thompson and Kirsty Lucas for the loan of the husbands for observing trips and conferences.

A special thank you to Dr Avril Day-Jones for all her advice and support on the good days, the bad days and the days I could have walked away. Thanks to Dr Jim Collett

for his support with the trickier pieces of maths, Dr Martin Hardcastle for his support on the cluster and Pam Green for her help on cracking the orphan removal tool algorithm. In addition, I would like to thank Prof. Melvin Hoare for his advice during my first year progression and Prof. James Hough for being a supportive and understanding head of department.

Extra special thanks to Gillian Stansfield and Dr Andrew Norton for putting me on the PhD path and to David Murray for being an excellent office mate (Mollusc).

At the risk of running out of superlatives I wish to thank Dr Stuart Alder for his help, support and friendship. Nobody could wish for a better friend. It has been a journey we have made together although perhaps your journey has been weirder than mine.

This thesis is dedicated to my wife Sarah (a.k.a Mrs G) who has supported me emotionally and financially through the last seven years and has been a partner beyond praise for the last twenty three years.

Sine te nihil sum

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

Twinkle twinkle star
How I wonder what you are
from spectrography I have gleaned
That you are made of Hydrogen.
Unknown

Motivation

Massive stars, stars with masses $> 8M_{\odot}$, provide the bulk of the ionising radiation in the Galactic Plane (Zinnecker and Yorke 2007; Urquhart *et al.* 2008). They enhance the (ISM) by ejecting metals via outflows in the early part of their lives and by strong stellar winds and supernovae at the end (Chu and Gruendl 2011; Zinnecker and Yorke 2007; Deharveng and Zavagno 2008). They are the primary producer of the heavy elements without which complex life, as we know it, would be impossible (Broadley *et al.* 2007; Zinnecker and Yorke 2007). Furthermore, massive stars appear to be important in the production of further star formation which they trigger by the blowing of ionized hydrogen (HII) bubbles (Zinnecker and Yorke 2007; Deharveng and Zavagno 2008). It has also been suggested that it is supernovae that inject the turbulence that create the large scale star forming structures known as Giant Molecular Clouds (GMC) (McKee and Ostriker 2007; Vazquez-Semadeni 2010). At larger scales massive stars delineate the Galaxy's spiral structure, given that the spiral arms are in the regions in which star formation is ongoing.

Massive stars have comparatively short main sequence lives (of the order of $\sim 10^7$ yrs) when compared to the lives of low mass stars (10^{10} yrs) and reach the end of their lives before they are able to migrate away from their formation site. Hence, with the exception of isolated runaway stars, they trace sites of star formation. However, despite this the structure of the Milky Way is poorly understood and there are many questions yet to be resolved.

We see therefore that massive stars play an important role as tracers of structure and the evolution of the Galaxy both directly by the alteration of the Galactic metallicity but also in the production of new generations of stars. However, the processes that lead to high mass star formation, unlike low mass stars, are relatively unknown. This is due to their rapid evolution onto the main sequence resulting in nuclear burning commencing whilst deeply embedded. Additionally, high mass stars are very rare within the Galaxy, consequently they have high average distances when compared to low mass stars. This is further complicated by massive stars being exclusively found in OB associations [1] which often suffer from confusion and high levels of extinction.

In order to further our understanding of high mass star formation and the structure of the Galaxy we need to move away from the small sample of currently identified young massive stellar objects. There are a number of known and proposed tracers of massive star formation, for example colour selected infrared objects, Class II Methanol Masers, Infrared Dark Clouds (IRDC) and Extended Green Objects (EGO). However, the relationship between these tracers and the nature of their association with massive star formation is not well understood.

In this thesis I investigate the relationship between a number of these tracers in order to further the detection of sites of massive star formation and to answer the question, "What techniques and indicators should be used to identify Massive Star Formation?"

1.1 Star Formation

1.1.1

The construction of the star formation sequence is one of the major achievements in astrophysics in the last half century, although there are still major questions to be answered such as: the origin of the Initial Mass Function (IMF), the nature of Brown Dwarf formation and the formation of massive stars.

In broad terms star formation can be considered to consist of a series of density enhancements moving from low density to high density, large objects to small scale objects, Giant Molecular Clouds to stars.

The sequence commences with a Giant Molecular Cloud , the largest star formation structure within the Galaxy and the principal host of observed star formation (Shu *et al.* 1987). Processes within the GMCs cause the formation of local density environments, IRDC which appear as dark silhouettes against the Galactic mid-infrared background, although the nature of IRDCs within the star formation sequence, if they are indeed part of the sequence, is yet to be fully clarified.

Within the GMC more local density enhancements result in regions known as clumps, with masses of the order of a Galactic star cluster inside which we find yet denser matter

[1] With the exception of runaway stars.

(cores) with masses of a single star or binary.

The formation of a high mass star results in the creation of an ionizing front which evolves through a series of phases, Hyper Compact, Ultra Compact and Compact characterised by increasing size. These ionised bubbles may go on to trigger further star formation and lead to the destruction of the GMC. They may also contribute to the formation of the next generation of GMCs.

Figure 1.1 illustrates this sequence for high mass stars. It should be noted that there are substantial differences between high and low mass star formation beyond the core phase. Furthermore, these phases are not exclusive within a GMC, which can demonstrate several, if not all, of these phases simultaneously. Nor should it be concluded that any individual phase before the warm core phase (as discussed in Chapter 6) is indicative of star formation. Rather it is an indication of *potential* star formation (Shu *et al.* (1987); Chrysostomou and Lucas (2005); McKee and Ostriker (2007), Stahler and Palla (2004) Chapter 3).

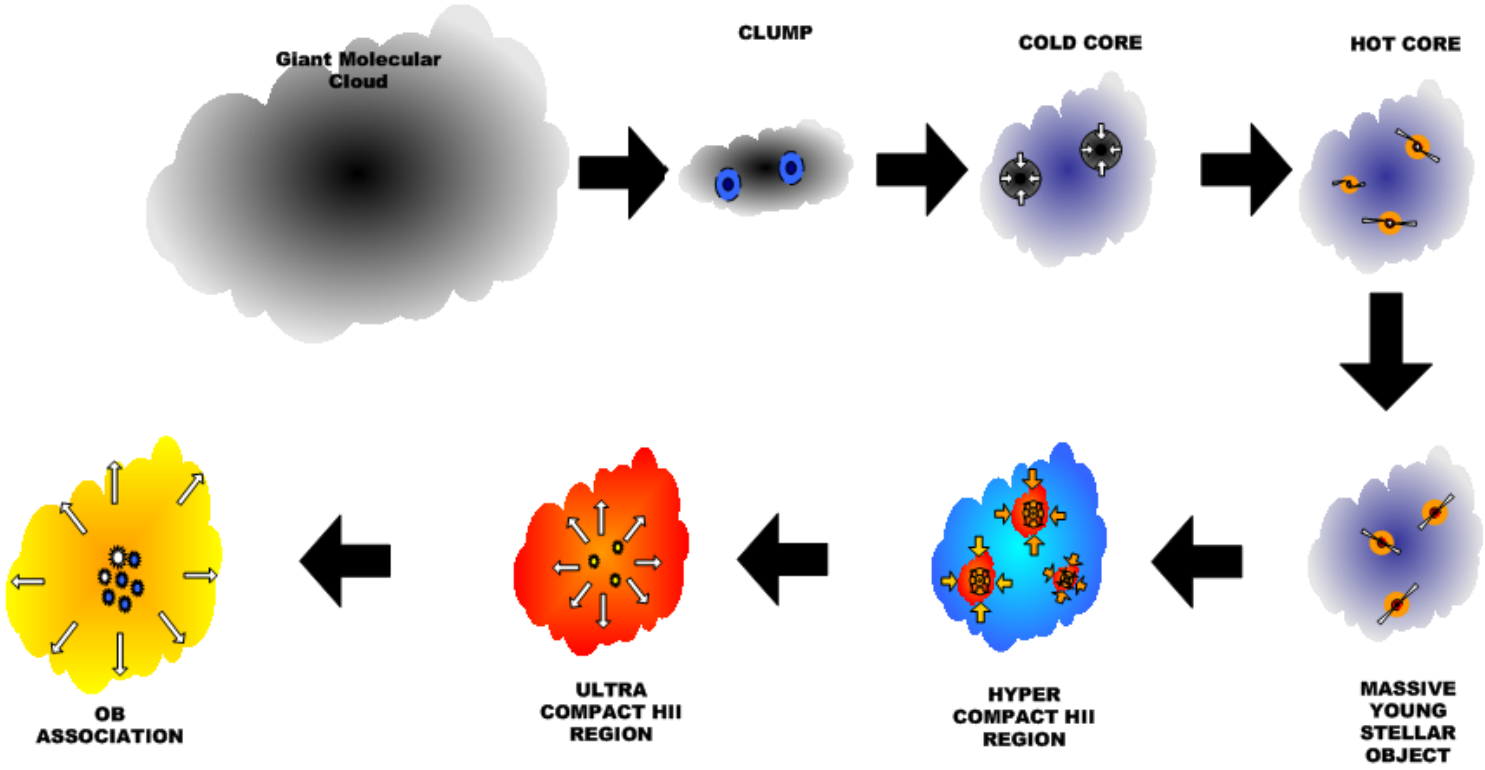


Figure 1.1: Massive Star Formation

Illustration of the current picture of massive star formation showing the progression from GMC to OB association Note that low mass star formation differs somewhat after the formation of the core.

Giant Molecular Clouds

GMCs are the first visible indication of the *potential* onset of star formation. Yet the processes that leads to the formation of a GMC are still mostly unknown with self gravitation of the large scale atomic medium or turbulent interaction between large scale Galactic structures being the most likely formation mechanisms (McKee and Ostriker 2007; Vazquez-Semadeni 2010). Physically GMCs have column densities of the order of 10^{20} cm^{-2} and diameters tens of parsecs (McKee and Ostriker (2007) and Stahler and Palla (2004) Chapter 3). Chemically they consist largely of molecular hydrogen with a thin sheath of atomic hydrogen which protects the molecular hydrogen from cosmic rays and ultraviolet radiation, thereby allowing internal cooling (Solomon and Rivolo 1989; McKee and Ostriker 2007).

The predominance of H_2 poses a problem to observation. The interior temperature of GMCs are often of the order of $10 - 30$ K (Purcell 2006; Krumholz 2011) but the $J = 1$ transitional state has an excitation energy $100 - 200$ K above ground state (Krumholz 2011). This is due to the molecular hydrogen's low mass and moment of inertia with the energy levels within a rotator scaling with the root of the moment of inertia. Consequently we do not observe H_2 rotational lines within the GMC and an alternative tracer molecule is required.

Fortunately CO is a common molecular species within GMC and has a J=1 excitation energy of 5.5 K thus we observe bright CO emission lines within GMCs. At high densities the most common CO isotope, ^{12}CO, becomes optically thick and we need to turn to the optically thin ^{13}CO for these regions. In locations where the temperature is very low, < 10 K, CO starts to condense onto dust grains resulting in depletion. In such cases alternative molecules such as ammonia may be used.

An alternative to the usage of molecular lines is to observe the dust within the GMC. Although dust is only a minor contributor to the overall mass of the GMC, the dust grains found within GMCs are typically infrared absorbers and sub-millimetre emitters. GMCs are optically thin in the sub-millimetre and as such dust provides an alternative method for observing GMCs.

In the seminal paper (Larson 1981) the characterisation of three GMC properties was determined from CO observations. These became known as Larson's Laws. Larson (1981) states that all molecular clouds are bound and that

$$\alpha_{vir} \equiv \frac{5\sigma^2 R}{GM} \tag{1.1}$$

where α_{vir} is a normalised numerical coefficient such that $\alpha_{vir} = 1$ represents non-magnetic gas in virial equilibrium and σ is the mean velocity as determined by the line-width, R is the radius of the cloud, G is the universal gravitation constant and M is the mass of the cloud (McKee and Ostriker 2007; Bertoldi and McKee 1992).

It is also stated that there is a line width-radius relationship within molecular clouds

determined to be approximately

$$\sigma = \sigma_{pc} R_{pc}^{0.5} \tag{1.2}$$

where σ_{pc} is the mean velocity per parsec and R_{pc} is the cloud radius in parsecs.
by Solomon and Rivolo (1989). By substitution we come to the third conclusion in Larson (1981) that the surface densities of molecular clouds is a constant.

$$\alpha_{vir} = \left(\frac{5}{\pi}\right) \frac{\sigma_{pc}^2}{G\Sigma} \tag{1.3}$$

where Σ is the surface density.

The implication of Larson's Laws is that GMCs are in virial equilibrium, that they are confined gravitationally rather than being pressure confined and that GMCs have similar column densities. Additionally it suggests that for sub-structures within the GMC that are smaller than the sonic scale (approximately 0.1 pc) are subject to sub-sonic turbulent (Ballesteros-Paredes *et al.* 2007).

Recent examination of the BU-FCRAO Galactic Ring Survey by Heyer *et al.* (2011) indicates that Larson's laws may not hold up to examination at greater angular resolution and greater depth and that magnetic fields may be the principal support against collapses and the turbulence within the clouds maybe associated with Alfvén waves.

The investigation into exact processes that lead to star formation within GMCs and the time scales over which those processes occur is a subject of intense research. There are two primary theories put forward. Slow star formation, where the GMCs have long lives and are supported against collapse by, for example, the GMC's magnetic field which is slowly overcome by ambipolar diffusion as discussed in chapter 10 of Stahler and Palla (2004). Alternatively, it is thought that star formation may be relatively rapid, with cloud lives times of the order of a supersonic crossing time. Fast star formation requires the formation of density enhancements via gravoturbulence. Turbulent shocks sweep up material causing density enhancements the at the stagnation points of convergent flows. Regions where the Jeans length is greater than the size of the fluctuation scale of the turbulence are unable to collapse before they are dispersed by next shock. Those smaller than the Jeans length and are magnetically supercritical may collapses to form stars (Klessen *et al.* 2005). The more massive stars interact with the GMC via massive outflows and the blowing of HII bubbles. This interaction has two opposing effects. Locally we see the triggering of star formation via Radiative Driven Implosion and via Collect and Collapse (Deharveng *et al.* 2005). However, it is equally true that this interaction may also be preventing further star formation by, for example, the dispersion of quiescent cores. At the level of the GMC the massive stars contribute to the wide scale turbulence resulting in the cloud being dispersed within a few dynamic crossing times ie in the order of 10Myr (Ward-Thompson and Whitmorth 2011; Ballesteros-Paredes *et al.* 2007; McKee and Ostriker 2007) .

Currently, the observational and theoretical models seem to support the fast star formation theories. The observation (as per Larson's laws) that velocity dispersion scale with size

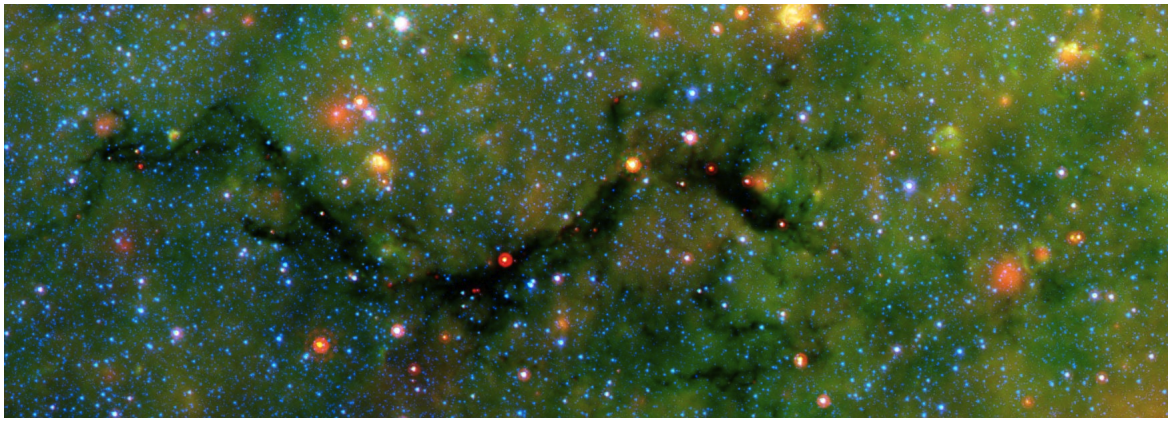


Figure 1.2: The Snake
IRAC / MiPSGAL composite image of the large infrared dark cloud known as the "Snake". The Snake is seen as the dark band against the bright 8.0 μm *and* 24 μm *background. In this image blue is* 3.6 μm*, green* 8.0 μm *and red* 24 μm*. Image courtesy of the NASA Spitzer image archive.*

as $\sigma = \sigma_{pc} R_{pc}^{0.5}$ is agreement for the power spectrum expected for supersonic turbulence. The Fast Star Formation model is further supported supported by the observation that few GMC are starless. Additionally, it is difficult to see how the initial cores would form if star formation is slow. Initial concerns regards how molecules such as H_2 and CO can form in the relatively short time available in the fast model have been addressed by the observation of such molecular in the warm ISM, suggesting that GMC may form from warm molecular material. (Ballesteros-Paredes *et al.* 2007).

1.1.2 Infrared Dark Clouds

Observations conducted by the Infrared Space Observatory (ISO) and Midcourse Space Experiment (MSX) space telescopes resulted in the discovery of ribbons of material with high levels of extinction ($A_v > 2$ where A_v is the level of visual extinction in magnitudes) in silhouette against the Galactic mid-infrared background, which became known as infrared dark clouds (IRDCs) (Perault *et al.* 1996; Carey *et al.* 1998). Figure 1.2 shows an example of an IRDC (known as the Snake) observed using the Infrared Array Camera (IRAC) and Multiband Imaging Photometer for Spitzer (MiPS) cameras aboard Spitzer.

In the sub-millimetre IRDCs appear as bright extended objects which trace the mid-infrared silhouette indicating that these are cold, dense objects that have been suggested to be a precursor to star formation (Peretto *et al.* 2008). Herschel Photodetecting Array Camera and Spectrometer (PACS) and Spectral and Photometric Imaging Receiver (SPIRE) observations of the Snake show numerous point sources along its length with dust temperatures of 20 K and solar like masses. Hence the Snake does not appear to be, as yet, forming high mass stars; however other such observations show IRDCs forming both high and low mass stars (Beuther 2011).

Simon *et al.* (2006a) undertook a survey of IRDCs using the 8.3 μm MSX band as the background for silhouetting. This resulted in the identification of 10,931 IRDCs with angular diameters greater than 20 arcseconds containing 12,774 cores. However, given that these IRDCs are detected by the omission of 8.3 μm background emission it is possible that some of the IRDCs are just holes within the background (Simon *et al.* 2006a).

The work of Simon *et al.* (2006a) was followed up by the inspection of the CO J=1-0 Boston University Five College Radio Astronomy Observatory Galactic Ring Survey (BU-FCRAO GRS) data (Simon *et al.* 2006a). This allowed the characterisation of IRDCs as having typical sizes of 5pc, masses of $\sim 5 \times 10^3 M_\odot$ and H_2 densities of $\sim 2 \times 10^3 cm^{-3}$ as implied by their high level of extinction. This is similar to to the densities found within molecular clumps (Simon *et al.* 2006a).

Given that IRDCs are detected against the infrared background it is clear that many IRDCs are not detected as they are obscured by foreground emission. Simon *et al.* (2006a) estimates that the 10,931 IRDCs listed in Simon *et al.* (2006a) represents 33% of the Galactic population. This being the case and assuming a typical mass, IRDCs could contain 5% of the molecular gas within the Galaxy or $10^8 M_\odot$ (Simon *et al.* 2006a). This would correspond to a star formation rate within IRDCs of $\sim 2 M_\odot yr^{-1}$ and implies that *all* star formation occurs within IRDCs.

The IRDC MSX survey of Simon *et al.* (2006a) was followed by the IRAC / MiPS survey of Peretto and Fuller (2009). The greater sensitivity and angular resolution of Spitzer over MSX allowed for the detection of 11,303 IRDCs within $10^o < |l| < 65^o$ and $|b| < 1^o$. Within the overlap between the two surveys, there is approximately double the number of IRDCs in Peretto and Fuller (2009) over Simon *et al.* (2006a). Peretto and Fuller (2009) noted that 20% to 68% of IRDCs showed signs of ongoing star formation. Further analysis of the Peretto and Fuller (P&F) catalogue by Kauffman and Pillani (2010) indicates that most IRDCs are not dense enough to form high mass stars but may go on to form low or intermediate mass stars. This would appear to be consistent with the increased number of IRDCs observed by Peretto and Fuller (2009) which suggests that not all IRDCs can be star forming. Recent observations by Jiménez-Serra *et al.* (2010) suggested that IRDCs are formed by turbulence within the host GMC rather than by gravitational collapse.

1.1.3 Clumps and Cores

Molecular clouds are far from internally uniform in density. Instead they consist of a series of over densities embedded into the large scale lower density structure (Blitz and Williams 1999). Structures with enough mass to form clusters I will term *clumps* and those that will go on to form individual stars (or binaries) I will term *cores*. However, we should be aware that there is not a distinct boundary in l-b-velocity (*lbv*) space from where a cloud ends and a clump starts. Rather it is the identification of an increase in density by, for example a density contour map. Consequently, surveys that identify cores and clumps, such as those discussed in section 1.3, have results influenced by both the physical limitations

of the survey and by the extraction algorithm used (Schneider and Brooks 2004; Pineda *et al.* 2009). Table 1.1 indicates the physical scales used to distinguish between clouds, clumps and cores.

	Size (pc)	Mass ($M_\odot$)	Density (cm^{-3}
Cloud	3-20	$10^3 - 10^4$	$10^3 - 10^4$
Clump	0.5-3	$10 - 10^3$	$10^4 - 10^6$
Core	<0.5	$10^3 - 10^6$	$> 10^6$

Table 1.1: Table shows size, mass and density classifications of clouds, clumps and cores. Taken from Purcell (2006) and references therein. We see that the effective boundaries are defined by diameter and density, as opposed to mass.

Physically some clumps appear to be gravitationally bound whilst others are not, in which case it is assumed that they are pressure confined (Blitz and Williams 1999). Cores on the other hand appear to be similar to hydrostatic Bonnor-Ebert spheres (Stutz *et al.* 2008). Cores are characterised by two phases; cold and hot. When the core is cold (20 K) it is quiescent with the low temperatures and high densities plus the shielding provided by the molecular cloud from ionizing radiation allowing the formation of complex molecules such as methanol and ammonia. In this condition the core is visible at millimetre and sub-millimetre frequencies via molecular rotational transitions and vibrational emission. Clearly the core cannot stay hydrostatic forever. Some will eventually disperse but others begin to collapse causing a rise in temperature and liberating ices frozen on the surface of dust grains. The core makes the transition from cold core to hot core (< 300 K) and becomes visible as a deeply embedded infrared source. Consequently, the hot core is the first sign of *active* star formation; at this stage the core could be considered to be analogous to Class 0 Young Stellar Object (YSO) (see Section 1.1.5).

1.1.4 The Formation of Clusters

It is clear that most stars form in clusters (Lada and Lada 2003; McKee and Ostriker 2007) and that most massive stars form in OB associations (Zinnecker and Yorke 2007). What is not clear is why the IMF appears universal (if indeed it is), nor why more massive stars appear to be found close to the centre of clusters (Mass Segregation). Any model for the production of stellar clusters must, to some extent, address these questions.

There are currently two rival theories as to the processes that occur once the core starts to collapse; Monolithic Collapse (McKee and Tan 2003) and Competitive Accretion (Bonnell 2000).

In Monolithic Collapse it is assumed that each core will go on to form a single star (or a multiple system) and that the only material available for accretion is the material contained in the core. Hence, the core mass sets the final stellar mass. Monolithic Collapse therefore suggests that the IMF is set by the Core Mass Function (CMF) which is supported by recent core observations (Chabrier and Hennebelle 2010). Mass Segregation in this case could be produced by either the most massive stars forming in the densest regions of the clump, which may tend towards the centre of the core, or by dynamic processes which result in more massive objects sinking to the centre of the cluster as the cluster relaxes.

In Competitive Accretion the core forms a cluster. Inside the core a number of overdensities form which become gravitational sinks for the formation of stars. However, unlike Monolithic Collapse these sinks share a common source of material on which to feed. Hence the larger sinks[2] will drag in more material and form massive stars thereby starving the lower mass sinks which will go on to produce low mass stars. Supporters of Competitive Accretion argue that the Competitive Accretion models produce mass-segregated clusters and reproduce a Salpter IMF (Bonnell 2000). However, it does not explain why we see a Salpeter CMF nor does it counter the arguments for mass segregation above.

1.1.5 Towards the Onset of Hydrogen Burning and Beyond

In this section I begin to separate star formation into low and high mass. The rate at which a core collapses to form a star is constrained by the free-fall time t_{ff}, the rate at which a body collapses under gravity alone, and the Kelvin Helmholtz time t_{KH}, the time taken for a body to radiate away a significant portion of its gravitational energy. At $8M_{\odot}$ $t_{KH} < t_{ff}$ indicating that stars over this mass evolve straight onto the main sequence without the observed pre-main sequence phases seen with low mass star formation (Zhang 2005) (see also Section 1.1.5). The consequence of this is that massive stars enter the main sequence whilst deeply embedded and hence are unobservable. Attempts to scale up the low mass star formation models below have failed to produce massive stars which has led to a number of alternative star formation methods being proposed.

Low Mass Star Formation

The low mass star formation sequence outlined here was first adopted by C.J. Lada in 1987 and, except for the addition of Class 0 YSO, has changed little. The Lada schema classifies YSO from Class 0 to III based on their spectral index (α_{IR}) traditionally taken between 2.2 and 10 μm.

$$\alpha_{IR} = \frac{d\log(\lambda F_{\lambda})}{d\log\lambda} \tag{1.4}$$

where d is distance, λ is wavelength and F_{λ} is the flux at wavelength λ.

[2]Which will tend to be near the centre of the gravitational well.

Objects undetectable at less than 20μm are classified as Class 0, those with $\alpha_{IR} > 0$ Class I. Objects with $-1.5 < \alpha_{IR} < 0$ are said to be Class II whilst those meeting $\alpha_{IR} < -1.5$ are Class III (Stahler and Palla (2004) section 4.1.2). The changes in class maps onto physical changes as the YSO evolves, as I will show.

The star formation sequence outlined here is derived from McKee and Ostriker (2007); as well as Chrysostomou and Lucas (2005) and the references therein. The sequence commences with the spherical collapse of a core. Initially the core is optically thin and can radiate out the released gravitational energy. However, as the collapse continues the density increases and as a consequence of the increased density the core makes the transition from optically thin to optically thick, which results in the temperature rising. Once the temperature reaches 2000 K molecular hydrogen starts to disassociate and the temperature starts levelling off with the disassociation of H_2 acting as a sink. Material continues to rain down onto the core, increasing the mass and disassociating further material until the system becomes unstable and collapses, resulting in a sharp rise in temperature to 10^5 K and densities to 10^{-2}g cm^{-3} at which point we can consider this a class 0 protostar. Radiatively, the protostar is dominated by its envelope and appears observationally as a grey body peaking at ~ 100 μm. The material within the envelope closest to the protostar begins falling onto the core supersonically resulting in an inside-out collapse. Angular momentum within the protostellar envelope prevents any further spherical collapse and the envelope starts flattening to form an accretion disc. Angular momentum may be being removed from the disk via interaction with the local magnetic field and the disc to create a disk wind driven jet (Pudritz and Banerjee 2005). Alternatively, the interaction could be occurring at the co-rotation radius in which case the angular momentum could be transferred to the jet via an x-wind (Shu 2001). In either case we see the launching of a collimated jet.

As the envelope becomes depleted emission from the disc becomes dominant and the object enters the Class I YSO phase. Eventually the rate at which material is passing across the disc drops off and radiation becomes photospherically dominant and we enter the Class II T-Tauri phase. Eventually accretion falls off and the protostar becomes a Class III YSO (a weak lined T-Tauri. Although there maybe a tenuous disc and envelope may exist, the loss of the jet results in the spinning up of the YSO as angular momentum, which normally be removed by the jet, is dumped onto the core. The photosphere, which has been puffed-up by the accretion energy being dumped on it, begins to contract under gravity. Convective mixing ensures that this contraction occurs without a significant increase in temperature and the YSO moves vertically down the HertzsprungRussel (HR) diagram following the Hayashi track until the opacity is reduced to the point where radiation starts to transfer energy. At this point the temperature rises causing the YSO to track horizontally across the HR diagram following a Henyey track. Eventually the core reaches the densities and temperatures required for hydrogen fusion (lithium and deuterium will occur briefly before) and the star enters hydrostatic equilibrium and thereby the Main Sequence.

Figure 1.3: Lada YSO Sequence

Pictorial representation of the Lada YSO sequence, showing the relationship between disc, envelope and protostar during the YSO's evolution onto the main sequence. When the object is radiatively dominated by the envelope it is known as a Class 0 YSO. As the ratio between photospheric and accretion emission increases as photospheric emission increases and accretion emission decreases the YSO passes thought the Class 0,I,II & III stages before entering the main sequence.

High Mass Star Formation

I now move on to massive stars. Scaling up low mass star formation results in two problems; how to transport sufficient material onto the core in such a short period and how to overcome the radiation pressure from the hot core, especially when fusion sets in at $\sim$ 10 $M_{\odot}$ (Yorke 2004; Zinnecker and Yorke 2007). For accretion to take place gravity must exceed the radiative pressure. Hence

$$\frac{\kappa_{eff} L}{4\pi r^2 c} < \frac{GM_*}{r^2} \tag{1.5}$$

where κ_{eff} is the effective opacity of the material being accreted, L is the combined luminosity of the central source and the disc and M_* is the mass of the central source, r is the radius and c is the speed of light in a vacuum. Given a central source of $5M_{\odot}$ we find that for dust with characteristics similar to dust found in the ISM ($\kappa_{\rm eff} \sim 100\ {\rm cm^2 g^{-1}}$), the radiative pressure exceeds the gravitational force and, in theory at least, accretion will cease (assuming that the accretion is Spherical Accretion). This is the radiation problem of massive stars.

The obvious solution is to reduce κ_{eff}. This can be achieved by increasing the effective grain size or by destroying the grains completely. Alternatively, the accreting material could form "blobs" with $\kappa_{\rm eff} = \pi {\rm R}^2_{\rm blob}/{\rm M}_{\rm blob}$, where R_{blob} is the radius of the blob and M_{blob} is its mass (Yorke 2004). If we scale up the blob idea we can see that high mass stars may form by collisions between lower mass stars as proposed by Bonnell *et al.* (1998). However, recent N-body simulations by Baumgardt and Klessen (2011) suggest that the build up of massive stars by mergers is an unlikely mechanism for the formation of high mass stars. Likewise models of grain growth within accretion discs suggest that radiation pressure is not unduly affected by grain growth (Zinnecker and Yorke 2007). Krumholz *et al.* (2005) have suggested that Rayleigh-Taylor instabilities within the disc creates "fingers" of material that can overcome the radiation pressure and thereby accrete.

An alternative approach to addressing the radiation problem of massive stars is to effectively reduce L. It is highly probable that the collapse of the molecular core will result in the formation of a disc, a conclusion supported by the large amount of indirect evidence for the existence of discs surrounding massive star forming objects (Kraus *et al.* 2010; Pestalozzi *et al.* 2004). The effect of the disc is to move the radiation field from spherical to a "flashlight" morphology, where the bolometric luminosity may be an order of magnitude (or greater) between polar and disc positions (Yorke and Bodenheimer 1999; Zinnecker and Yorke 2007). Hence, the effect of the disc is to shield the accreting material from the radiation pressure of the central source. However, the disc will not extend all the way to the surface of the central object so the material will at some point be exposed to the full radiation field. For these dust grains to accrete they must have either grown to the point where the gaseous component of the inflow is the dominant opacity component or alternatively has been completely destroyed (Zinnecker and Yorke 2007).

Massive YSO and the Formation of HII Regions

Let us first consider a model HII region. We assume we have an ionizing source which is producing N_i ionizing photons a second. Their photons will ionize their environment supersonically driving out a spherical ionization front (a Strömren sphere) until the point is reached where the mass of the ionized material enclosed by the Strömren sphere is such that the recombination rate is approximately equal to the ionization rate. As this point is approached, the rate of expansion becomes subsonic and the pressure difference between the inside and outside of the sphere becomes dominant. As the HII region expands the pressure within the HII region decreases until the system slowly reaches equilibrium (Stahler and Palla (2004) section 15.1) .

Returning to massive star formation. As material is increasingly dumped onto the hot core the core temperature and pressure increases until at 10 $M_\odot$ fusion becomes inevitable and the object enters the main sequence as a Massive YSO (MYSO) [3]. However, the outer regions of the MYSO are effectively isolated from the fusing core and hence accretion continues. Eventually enough Ultraviolet (UV) photons are being produced in the core for this separation to end and for the ionisation of the disc to start.

We have seen above that an HII region is driven out supersonically. This being the case it would appear that the formation of the HII region would terminate accretion. However, this is clearly not the case. It is possible that the rate of inflowing material is such that the formation of an HII region is suppressed. It has been suggested by Keto (2007) that the observation of Hypercompact HII regions, (regions with sizes 0.01 pc) are photoevaporated discs trapped by the gravity of the MYSO and by the pressure of inflowing material. Hence the final mass accreted by the MYSO is this trapped ionized material (or part thereof). Eventually a combination of increasing radiation and thermal pressure, together with the reduction of the ram pressure from the inflow, results in the HII region expanding into an Ultracompact HII region (UCHII) and the start of interaction with the remnants of the molecular clump and/or molecular cloud (Zinnecker and Yorke 2007). At this stage it is often impossible to distinguish if the UCHII region is being powered by a single star or an OB association. The UCHII will continue to expand, until as stated above, the internal pressure matches the ambient pressure.

The formation of the HII region marks the beginning of the dissipation of the IRDC and ultimately the GMC. However, new generations of star formation may be formed as the HII regions sweep up material along its edge (collapse and collect (Deharveng *et al.* 2005)) or by causing the collapse of quiescent cores or clumps as they are overrun via radiation driven implosion (Sandford *et al.* 1982) by the ionization front.

So we see that star formation is a progression of evolutionary states from GMC to YSO (or MYSO and HII region) via a series of density enhancements. It is tempting to consider this as a "Russian Doll" with denser more evolved objects enclosed with less dense less evolved objects. However this, as we will see, is not necessarily the case.

[3] This assumes that a single core yields a single star which might not be the case. In the case of a core forming multiple stars this process still holds although at smaller scales.

1.2 Methanol Masers

Since the suggestion of their existence and subsequent discovery in the 1960s, astronomical masers have become an important tool in furthering our understanding of both star formation and Galactic structure. The processes involved in the creation of astronomical masers are both complicated and in many regards, unknown. Here I outline the general physics that underpin their formation. For those who wish a more detailed understanding I suggest the excellent Astronomical Masers by Elitzur (1992).

A maser is the amplification of microwave radiation caused by stimulated emission, MASER being an acronym of **M**icrowave **A**mplification by **S**timulated **E**mission of **R**adiation.
We start with the standard equation of radiative transfer...

$$\frac{dI_\nu}{ds} = j_\nu - \kappa_\nu I_\nu \tag{1.6}$$

as per equation 7.2 of Burke and Graham-Smith (2009). Where I_ν is the intensity at frequency ν, s is the path length, κ_ν is the linear absorption coefficient at frequency ν and J_ν is the specific emissivity at at frequency ν

Let us consider a simple system with two energy levels with an energy difference of $h\upsilon_{12}$ and the number density of each is N_1 and N_2. In this situation the equation of radiative transfer, 1.6 can be written as...

$$\frac{dI}{ds} = h\nu_{12}[B(N_2 - N_1)I + A] \tag{1.7}$$

As per Equation 7.47 Burke and Graham-Smith (2009)
where B is the Einstein coefficient for a stimulated transition between level 1 and 2 and A is the Einstein coefficient for the spontaneous transition from level 2 to level 1. I is intensity, and s is the coherence path length. Traditionally equation 1.7 is used in absorption however in this discussion we are dealing with an emission process so the sign on the right hand side of the Equation of Radiative Transfer has been changed to a positive.
As in general A is very small for masing species we can rewrite equation 1.7 as ...

$$\frac{dI}{ds} = h\nu_{12}[B(N_2 - N_1)I] \tag{1.8}$$

In local thermal equilibrium we find that the relationship between the number density of the two energy levels is...

$$N_1(BI + C) \approx N_2(BI + C) \tag{1.9}$$

where C is the collisional rate.

In a system in thermal equilibrium we find that $N_1 \gg N_2$. In order for masing to occur we must produce an inverted population, a situation where $N_2 \gg N_1$. This is done by the addition of a pumping mechanism where the rate of the pump is R. Hence Equation 1.9 becomes...

$$N_1(BI + C) \approx N_2(BI + C) - R \tag{1.10}$$

As per Equation 7.48 Burke and Graham-Smith (2009)
By rearranging 1.10 so that $N_2 - N_1$ is isolated on the left hand side and substituting into Equation 1.7 we get...

$$\frac{dI}{ds} = h\nu_{12}\frac{BRI}{C + BI} \tag{1.11}$$

As per Equation 7.49 Burke and Graham-Smith (2009)
We can safely assume that the collisional rate will be much larger then the B Einstein coefficient so $C \gg BI$. Hence Equation 1.11 can be written as..

$$\frac{dI}{ds} = \alpha_{12}I \tag{1.12}$$

where $\alpha_{12} = h\nu_{12}\frac{BR}{C}$.
Equation 1.12 can now be solved using the integrating factor method to give...

$$I = I_0 e^{\alpha_{12}s} \tag{1.13}$$

As per Equation 7.50 Burke and Graham-Smith (2009)
Now let us consider the scenario where $C >> BR$. In this state collisional transitions are dominant and the population ratio N_2/N_1 is small. In this situation the distribution of the two energy levels follows a Boltzmann distribution and the system is in local thermodynamic equilibrium (LTE), α is small and consequently the gain, the measure of I/I_0, is also small.

Now let us consider a scenario where the radiative pump R is dominant over the collisional rate. Such a situation, where $N_2 > N_1$, is known as an *inverted* population. In this case we see that α may be large and that consequently the gain will be exponential. This clearly cannot continue and once the pumping rate matches the stimulant emission rate, a situation known as *saturation* arises. At this point α becomes effectively a constant and the gain becomes linear.

A more detailed discussion of this process can be found in Stahler and Palla (2004), pages 497-503.

There are a number of further effects which arise as a consequence of the exponential gain. The gain is dependent on the stimulated emission rate which is in turn dependent on the line shape which is described as a Lorentzian. Consequently, the amplification is peaked sharply at the central frequency and the line is narrowed when compared to the Gaussian distribution we would expect from a thermally broadened line.

Now let us consider our masing medium on a large scale. It is unlikely to be uniform in density, velocity, temperature and inversion state and as a consequence certain pathways through the medium will be coherent while others will not. Additionally some pathways will have longer coherence lengths than others and consequently will provide a high level

of gain. To an external observer these pathways appear as tight, compact maser *spots*. So we see that as well as amplifying either background or spontaneous internal emission, the line width of the emission is narrowed and the radiation beamed as a result of the masing process. It was these characteristics which indicated that the extreme bright radio frequency emission coming from astronomical sources were indeed masers.

I now turn to a specific maser species, the 6.67 GHz (4.49 cm) maser emission linked to the $5_1 \longrightarrow 6_0 A$ toroidal transition of methanol which was first detected in May 1991 (Menten 1991).

Methanol masers are broken into two classes, Class I and Class II (Sobolev *et al.* 2005). Class I methanol masers are collisionally pumped and are often associated with outflows (Sobolev *et al.* 2005) whilst Class II methanol masers are radiatively pumped and seem to be uniquely associated with high mass star formation (Minier *et al.* 2003) although the reasons for this are still largely unknown, as is the exact nature of the environment that leads to masing.

Sobolev *et al.* (1997) have expanded their models for the 12 GHz$2_0 \longrightarrow 3_1$E transition to encompass the Class II 6.67 GHz maser line. Both maser species have been observed to be coincident (although not exclusively so) and hence the environment required to produce both species can be determined by adjusting the free parameters to fit the brightness ratio of these two species.

Sobolev *et al.* (1997) assumes that the pumping mechanism is far infrared emission from a dust cocoon shrouding the masing methanol. This cocoon is re-radiating radiation from the central source, presumably a High Mass YSO or warm core. It is further assumed that the source that is being amplified is a background or associated HII region. With these assumptions Sobolev *et al.* (1997) concludes that a CH_3OH fractional density of 10^{-7} is required, in addition to gas densities of 3×10^3 cm$^{-3} < \mathrm{n_H} < 10^8$ cm^{-3} and dust temperatures of > 150 K, where n_H is the hydrogen number density.

Evolutionary, 6.67 GHz methanol masers are thought to trace the period covered by the onset of the warm core and hence the sublimation of methanol off the surface of dust grains, to the destruction of the masing environment caused by an expanding HII region. This is derived from their association with the early stage tracers Class I Methanol Masers and EGO and the late stage tracers, UCHII regions and OH masers (*Breen et al 2011 in prep.*). It would therefore appear that, initially at least, the source of the seed photons for the maser is not a HII region but rather spontaneous emission from within the cloud.

Observationally Class II methanol masers are some of the brightest radio sources in the sky with a number of masers having fluxes over a kilojansky. This together with their compactness, the relative transparency of the Galaxy at 6.67 GHz and their association with star formation make them ideal tracers of star formation and Galactic structure.

There are a number of major unknowns connected to the relationship between Class II methanol masers and high mass star formation. Firstly, the number of "potential" masing

sources that are at the correct evolutionary stage for masering to occur, but for some reason do not, is not known. Nor is the "filling factor", the chance lying on a beam line, which is related to morphology of the masing environment. If, as suggested by Pestalozzi *et al.* (2004), the masers are forming in discs, then it is unlikely that the masers are being beamed edge on as there is likely to be a loss of coherence due to the Keplarian rotation of the disc. There is also a further issue concerning infrared luminosity and maser luminosity. If Class II masers are indeed pumped by infrared radiation then we may expect that there would be a luminosity relationship. However, no such relationship is observed (see Chapter 2).

The variable nature of Class II methanol masers is also poorly understood. Observationally they seem to vary with periods less than a year, in some cases periodically but the reasoning behind this variability is not known. Changes in the luminosity of the pump or the masing environment seem likely explanations (Goedhart *et al.* 2009).

It is clear that there is yet much work to be done, both observationally and theoretically, before we can fully understand the location of masers within their environments and why Class II methanol masers are solely associated with mass star formation.

1.3 Surveys

1.3.1 The Methanol Multi-Beam Survey

The Methanol Multi-Beam (MMB) Survey is the first targeted survey of the Galactic Plane for methanol masers. The survey uses the Parkes radio telescope for the initial detection with follow-up using the Australian Telescope Compact Array (ATCA) with a 6km East - West baseline. The survey utilises a bespoke seven beam receiver covering $6-7$ GHz, hence the 6.035 GHz OH maser is also covered as part of this survey. Targets too far north to be observed by ACTA were observed with the Multi-Element Radio Linked Interferometer Network (eMERLIN). The plan to place the receiver onto the Lovell Telescope for the Northern Hemisphere run is currently on hold. The target sensitivity was ≤ 0.2 Jy at 1σ Green *et al.* (2009a) with a resolution of 32 arcseconds for the single dish observations and sub-arcsecond for the interferometric follow-up. The survey, when complete, will cover $-170^{o} < l < 168^{o} |b| < 2^{o}$ plus the Large Magellanic Cloud (Green *et al.* 2009b). As of March 2010, the survey has detected 1001 masers, of which over 350 are new detections. Of these 1001 detections, 61 are masers detected by the splitting of single maser spots by the high resolution interferometric observations. This raises the question as to at what scale do we see a change from one maser representing one object to multiple masers per object. There is also the issue that although the single dish component of the MMB is untargeted the high resolution follow up is not and we should be aware that there are two spatial resolution limits within the catalogue. Notwithstanding this, the MMB is the deepest and most complete untargeted survey of Class II methanol masers to date.

1.3.2 The Boston University Five Colleges Astronomical Observatory Galactic Ring Survey

One of the surprising outcomes from past Galactic ^{12}CO surveys is that 70% of molecular gas inside the solar orbit is found within a ring located 5 kpc from the Galactic centre (the so called 5kpc ring) (Clemens *et al.* 1988; Jackson *et al.* 2006). The Boston University Five Colleges Astronomical Observatory Galactic Ring Survey (BU-FCRAO GRS, or just, GRS) is a ^{13}CO J= 1 $\rightarrow$ 0 survey of the 5kpc ring utilising the Five Colleges Astronomical Observatory 14m dish. The lower abundances of ^{13}CO over ^{12}CO means that ^{13}CO is a deeper molecular gas tracer than the more plentiful ^{12}CO which becomes optically thick when dealing with long sight lines. The survey covers the region $18^{o} < l < 55.7^{o}$ and $1^{o} < |b|$ and a velocity range of -5 to 135 kms^{-1} for $l \leq 40^{o}$ and -5 to 85 kms^{-1} for $l > 40^{o}$ with a spectral resolution of 0.21 kms^{-1} and hence covers a large portion of the proposed 5 kpc ring [4] The use of spectral line observation over continuum observations allows the identification of multiple objects within the line of sight.

The results of this survey were analysed in Roman-Duval *et al.* (2010, 2009); Rathborne *et al.* (2009). Rathborne *et al.* (2009) outlines the procedures undertaken to identify the molecular clouds. The GRS is smoothed spatially to a pixel size of 6 arcseconds and a velocity resolution of 0.6 kms^{-1} in order to improve sensitivity. CLUMPFIND [5] was then used to identify 848 molecular clouds. The parameters used with CLUMPFIND were then adjusted in order to identify smaller structures (clumps and cores). Sources smaller than 50 voxels [6] were rejected. A total of 6126 objects were identified in this manner. Roman-Duval *et al.* (2009) derived the kinematic distance from the velocity of the associated CO spectral line. However, as all the GRS objects are by definition located in the inner Galaxy they are subject to the kinematic distance ambiguity (KDA).

Roman-Duval *et al.* (2009) attempts to resolve the KDA by using monoatomic hydrogen (HI) self-absorption. The principle is relatively straightforward. HI is ubiquitous through the Galactic Plane and is a strong emitter at 21 cm. The HI located within the GRS molecular clouds is cold, is in an unexcited state and hence a 21cm absorber. Given its widespread nature we would expect the HI spectra to be broadly Gaussian reflecting the rotation curve of the Galaxy. If a clump is at the near solution, 21cm radiation emitted from the far solution is absorbed and we see a dip in the HI spectrum at the same velocity as the clump. For a clump at the far solution, it may locally absorb some velocity coherent HI emission but it has no influence on the 21cm emssion at the near solution. Hence for clumps located at the far distance we see no dip in the HI spectra coincident with the CO velocity (Roman-Duval *et al.* 2009). This technique allowed the solution of the KDA for 750 of the 848 molecular clouds identified in the GRS.

Roman-Duval *et al.* (2009) derived the physical properties of 580 GRS molecular clouds using the GRS data and ^{12}CO observations undertaken by the Massachusetts-Stony Brook

[4] There is some evidence for example, Jackson *et al.* (2008),that the ring is in fact the nearest part of a tightly wound arm.

[5] CLUMPFIND is an IDL program designed to find structure within images, particularly images of molecular clouds.

[6] A Voxel being a pixel in spatial and velocity space.

Galactic Plane Survey (Solomon *et al.* 1987). This analysis confirmed the molecular cloud mass-radius relationship and the observation that they are gravitationally bound. It is noted from these observations that the bulk molecular content within the Milky Way is located in a ring 4-5kpc from the centre. As yet there is no published work concerning the smaller scale molecular clumps and/or cores found within the GRS.

1.3.3 The BOLOCAM Galactic Plane Survey

The BOLOCAM Galactic Plane Survey (BGPS) is a 1.1 mm continuum survey of the Galactic Plane covering $-10.5^{o} < l < 70.0^{o}$ and $|b| < 0.5^{o}$ (Dunham *et al.* 2011; Aguirre *et al.* 2011) with a sensitivity of 0.4 Jy using the BOLOCAM [7] instrument mounted upon the Caltech Submillimeter Observatory (CSO) at the Mauna Kea Observatory. The observational window is centred on 271.1 GHz (1.1 mm) and has a width of 47 GHz (0.17 mm) and therefore excludes the very bright CO J = 2 $\longrightarrow$ 1 emission line (Aguirre *et al.* 2011). The beam width of BOLOCAM is 31 arcseconds although due to optical distortion the effective beam width of the BGPS is of the order of 33 arcseconds (Aguirre *et al.* 2011). Observationally, 1.1 mm emission is associated with cold dust (of the order of 10 K) emission and therefore, dense regions (Enoch *et al.* 2010) such as molecular clumps and starless cores, although contamination by free-free emission from HII regions is possible Figure 1.4.

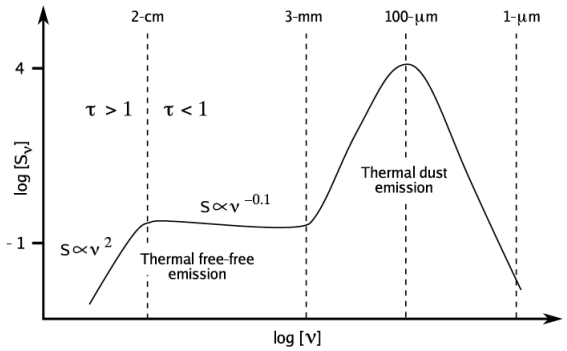


Figure 1.4: *Plot of frequency v intensity for a UCHII region.* Note that the tail end of the free-free emission extends into the 1.1 *mm* BGPS band, hence it is possible that the BGPS is contaminated by free-free emission from an UCHII region. Figure taken from (Purcell 2006)

The BGPS observations were made using a basket weave raster to remove the sky signal with each field being observed multiple times in order to improve signal to noise (Aguirre *et al.* 2011). A field rotator was employed to ensure that the scan was orthogonal to the array to avoid the detections gaps caused by the separations of the bolometers within the Bolocam array (Aguirre *et al.* 2011).

[7] BOLOCAM is a multi-element millilitre bolometer.

The BGPS team elected not to use existing extraction software such as CLUMPFIND or Source Extractor[8] as these require a prior. Instead a bespoke watershed algorithm was employed as the source extractor (Rosolowsky *et al.* 2010). The application of this method yielded 8358 sources (Rosolowsky *et al.* 2010). Dunham *et al.* (2010) used previous $NH_3(1,1)$ observations of objects within the BGPS field to obtain kinematic distances to a limited number of sources and therefore derive the dust mass and hence the total object mass. Dunham *et al.* (2010) used observations of star forming complexes with known distances to perform the same task. Dunham *et al.* (2010) concluded that the mean sizes, densities and kinetic temperatures were consistent with molecular clumps.

The BGPS sources are further categorised in Dunham *et al.* (2011) by their association with EGOs, the RMS [9], the Robitaille Red GLIMPSE [10] Sources (RGS) (Robitaille *et al.* 2008) and GLIMPSE sources with [4.5]-[8.0] > 0.4. By these associations they divide the BGPS sources into four categories (which should not be confused with the YSO classification discussed above). Group 0 members are considered starless as they show no sign of association in any of the cross matches, although this cannot be considered absolute without 24 μm observation which Dunham *et al.* (2011) lacks. Group 1 objects contain a GLIMPSE or RGS source but no associated RMS or EGO object. Group 2 objects are associated with RGS sources but not EGO or RMS objects, whilst Group 3 objects contain a RMS YSO-type object or an EGO. Hence, it appears that the BGPS contains a range of dusty objects from starless molecular clumps to early main sequence high mass stars.

1.3.4 The Galactic Legacy Infrared Mid-Plane Survey Extraordinaire

GLIMPSE is an infrared legacy survey conducted by the Spitzer Space Telescope using the IRAC camera and covers all four IRAC bands, 3.6, 4.5, 5.8 and 8.0μm, respectively. The initial survey (GLIMPSE I) covered $10^0 < l < 65^o$ $295^0 < l < 350^o, |b| < 1^o$. It was extended to $0^0 < l < 10^o$ $360^0 < l < 355^o$ in GLIMPSE II and then to $|b| < 3^o$ in GLIMPSE 3D (Archive 2011). The images from these observations are available from IPAC as either a native 1.2 arcsecond resolution or at a higher 0.6 arcsecond resolution. The 3σ magnitude detection limit is 15.5, 15.0,1 3.0 and 13.0 in bands 1 to 4, respectively. The Point Spread Function (PSF) photometry was undertaken using a modified version of DAOPHOT[11] and ALLSTARS[12] From the resultant photometry two point source catalogues were constructed. The Catalogue (GPSC) required a source be detected twice in one band and once in an adjacent band together with a minimum and maximum flux limits. The Archive (GPSA) uses a less stringent selection process and therefore is less reliable than the GPSC, but consequently more complete. There is, as yet, no extended source catalogue (Benjamin *et al.* 2003).

[8] Source Extractor or SExtractor is a program designed to detect objects with in images, in particular images within galactic surveys. It's purpose is similar to CLUMPFIND.

[9] The Red MSX Survey

[10] Galactic Legacy Infrared Mid-Plane Survey Extraordinaire

[11] DAOPHOT is a photometric package designed for crowded field photometry within astronomical images. It is part of the IRAF suite of astronomical software.

[12] ALLSTARS is the PSF photometry package held within DAOPHOT and as such is also part of the IRAF suite.

1.3.5 The UKIRT Infrared Deep Sky Survey Galactic Plane Survey

The UKIRT Infrared Deep Sky Survey Galactic Plane Survey (UKIDSS GPS) is a currently ongoing deep, near infrared survey of the northern Galactic Plane utilising the WFCAM[13] camera mounted on the UK Infrared Telescope (UKIRT). The survey covers $140^o < l < 230^o$ $b < |5|^o$ and $15^o < l < 107^o$ $b < |2|^o$ to a 5σ depth of 19.4 in J, 18.0 in H and 17.75 in K with the photometery being bootstrapped on the Two Micron All-Sky Survey (2MASS) (Lucas *et al.* 2008). The wide field of view of WFCAM means that there is always a bright star within the field in which to bootstrap the photometry, thereby avoiding issues with 2MASS uncertainties increasing with magnitude.

Photometric uncertainties are of the order of 0.2 mag in uncrowded fields. However, in crowded fields, such as those near the Galactic Centre, UKIDSS sensitivity drops off. Despite this, UKIDSS is significantly deeper than 2MASS (Skrutskie *et al.* 2006; Lucas *et al.* 2008). The data repository for UKIDSS is hosted by the Royal Observatory Edinburgh `http://surveys.roe.ac.uk/wsa/` and allows the query of the UKIDSS point source and image archives via a SQL database.

1.3.6 The Red MSX Survey

Attempts to identify massive star formation by colour selection using infrared surveys such as the Infrared Astronomical Satellite (IRAS) and MSX have been hampered by the lack of a tight definition of the colour space occupied by massive star formation but also contamination in the colour space by planetary nebula (PNe), evolved stars and other red objects (Lumsden *et al.* 2002). RMS (Urquhart *et al.* 2008) is a multi-wavelength examination of MSX objects within the colour space defined by $F_8 < F_{14} < F_{21}, F_{21}/F_8 > 2, F_8/F_K > 5$ and $F_K/F_J > 2$ (where K and J detections are available) with the additional criteria that the source is not extended (Mottram *et al.* 2007). This colour selection criteria was adapted in order to discover MYSOs, a term defined by the RMS team as an evolutionary stage of massive star formation, after the onset of hydrogen burning but before the ionisation of the star forming environment and the formation of an HII region (Mottram *et al.* 2007, 2010). Hence, MYSOs by this definition are infrared (IR) bright and radio quiet compact sources.

The RMS has to date been successful in identifying in excess of 500 MYSOs and a further 500-600 HII regions out of $\sim$ 2000 colour selected objects, the remainder being contamination from objects such as evolved stars and PNe (Urquhart *et al.* 2008). The survey includes a wide range of observations. Far infrared photometry of the objects was obtained using the IRAS infrared Galaxy Atlas and MiPSGAL[14] (Mottram *et al.* 2010).

Near infrared photometry of the targets was obtained from 2MASS (Mottram *et al.* 2011) and ^{13}CO observations at 480 μm from the Submillimetre Common-User Bolometer Array (SCUBA) legacy archive. These observations are supported by 1.2 mm observations taken using several instruments. By associating sources with the GRS clouds and the Southern Galactic Plane Survey the MSX team have been able to determine the distances

[13] WFCAM is a wide field near infrared camera.

[14] MiPSGAL is a 24 μm and 70 μm Galactic plane survey using the MiPS camera aboard Spitzer.

to the colour selected sources (see also Chapter 4) (Urquhart *et al.* 2011). This enables the determination of the spectral energy distribution of the sources and therefore their classification. Radio continuum observations (Urquhart *et al.* 2007) allowed the separation between sources with HII regions and those without. As of April 2011 the RMS had classified 646 objects as YSOs and 35 as HII/YSOs (Urquhart 2011).

1.3.7 The Peratto and Fuller - IRDC Catalogue

The Peretto & Fuller infrared dark cloud survey (Peretto and Fuller 2009) is a survey of the GLIMPSE and MiPSGAL survey data covering $10^o < |l| < 65^o$ and $|b| < 1^o$ with the aim of identifying IRDCs at a greater spatial resolution and sensitivity than the earlier MSX survey of Simon *et al.* (2006b) (Peretto and Fuller 2009). The process of identification involved the construction of opacity maps. Peretto and Fuller (2009) construct the opacity maps by the removal of point sources and replacing these with a representative background. The images were then smoothed at a half degree scale. The smoothing was then followed by the application of a bespoke "CLUMPFIND" like routine which identifies the cloud within the smoothed frames. Additionally, the H_2 density of the clouds were calculated by measuring the foreground and background emission which were also used in the construction of the opacity maps. This technique was checked against the 1.2 mm emission of a number of IRDCs from Rathborne *et al.* (2006) from which was also calculated the H_2 density, thereby confirming the reliability of the absorption only method for calculating opacity.

As a final check Peretto and Fuller (2009) constructed GLIMPSE 8 μm and MiPSGAL 24 μm mosaics which were visually inspected. They concluded that 90% of the 11,303 IRDCs within their catalogue were real detections. It was also noted that by applying different parameters to the cloud extraction process Peretto and Fuller (2009) were able to extract 20,000 to 50,000 cloud fragments. The authors also noted from the presence of 24 μm sources within the IRDC that between 20% and 68% of IRDCs are star forming.

It should be noted that the P&F survey only covers a limited section of the inner Galaxy and that the technique is unlikely to be successful in the outer Galaxy where there is less background emission for the IRDCs to be silhouetted against. We should also be aware that this technique does not eliminate holes in the background being misidentified as IRDCs which could be resolved with sub-millimetre follow-up. I also note that the cloud finding routine can be made to break the clouds into fragments and it is not clear what distinguishes a fragment from a cloud. It is therefore possible that some clouds within the catalogue are in fact fragments of larger clouds.

Despite these reservations the Peretto and Fuller (2009) IRDC catalogue is the most complete and reliable source of IRDC positions currently available and strongly suggests that at least some IRDCs host star formation sites.

Assuming that the mass of the objects is of the order of $1000M_\odot$ and that their lives are similar to that of a GMC. By taking the star formation efficiency rate to be 0.2, we see that the material identified in these surveys can account for all the star formation in the

Galaxy ($\sim$ 1M$_\odot$ yr^{-1}) which would seem to be unlikely as none of the surveys claim to contain the entire Galactic population of cores, clumps or IRDCs. Furthermore, as observed by Dunham *et al.* (2011) approximately half of the BGPS objects show no sign of star formation [15] suggesting that, for the BGPS sources at least, a significant population are quiescent and will not form stars. Hence, it is possible that not all GRS, BGPS or P&F IRDC objects will go on to form stars.

1.3.8 The Cyganowski Extended Green Object Catalogue

Extended Green Objects, also known as Green Fuzzies, are objects that were originally observed within GLIMPSE that show extended 4.5 μm emission with comparatively little emission in the 3.6m and 8.0 μm bands. As the 4.5m IRAC band is typically coded as green within GLIMPSE images these objects appear green, hence the name (Chambers *et al.* 2009; Cyganowski *et al.* 2008a).

Cyganowski *et al.* (2008a) undertook an extensive visual search of the GLIMPSE images in order to identify EGOs by their colour and extension. Point like sources with green colours were rejected as it was considered that these could be extincted stellar sources (Cyganowski *et al.* 2008a). The inspection of the GLIMPSE images resulted in the identification of more than 300 EGOs (Cyganowski *et al.* 2008a), which were divided into "possible" and "likely" candidates, with 70% of "likely" EGOs and 54% of "possible" EGOs being identified as associated with IRDCs. Cyganowski *et al.* (2008a) noted the strong association between EGOs and Class II methanol masers which was later confirmed in Cyganowski *et al.* (2009). The authors also concluded that EGOs were associated with Massive YSO outflows. Observations by Chen *et al.* (2009) and Cyganowski *et al.* (2009) confirmed that EGO are strongly associated with Class I methanol masers which themselves are known outflow tracers. A number of possible sources of the 4.5 μm emission have been suggested including shocked emission from H_2 or CO (Cyganowski *et al.* 2008a) and SiO absorption (De Buizer and Vacca 2010). To determine the exact nature of the emission spectrography is required. However, ground based spectrography in M band is extremely challenging, especially for extended objects with low surface brightness and in crowded fields. It may be possible to obtain spectra using long integrations with one of the small number of 8 m class telescopes with mid infrared capability such as VLT or Gemini. However, the problem is compounded in that EGOs have little or no associated near infrared emission which makes slit alignment somewhat problematic.

De Buizer and Vacca (2010) have obtained a number of M band spectra of two Cyganowski EGOs using Near InfraRed Imager and Spectrometer, (NIRI) on Gemini North. These observations resulted in H_2 being identified as the principle emitted for one source however H_2 was absent from the other. In De Buizer and Vacca (2010) it is suggested that extinction may cause th 3.6 μm and 8.0 μm band emission to be knocked back and thereby give the illusion of enhanced 5.8 μm and 4.5 μm band emission resulting in the false identification of an EGO.

[15] Although this number may reduce when observed at 24μm.

We see therefore that EGOs are tracing outflows given their association with Class I methanol masers and that It is also highly likely that these outflows are associated with massive objects based on the association with Class II Methanol Masers. However, the observation that there might be multiple processes producing EGOs suggests that further work is needed before these objects are truly understood in a star forming context.

1.4 Summary

We see that massive stars play an important role in not only the evolution of the Galaxy but also in the determination of its structure. Despite this, the processes that lead to the formation of massive stars is still poorly understood, which is partly due to the low number of formation sites currently identified. The large number of recent Galactic Plane surveys have resulted in a number of object classes apparently associated with Massive Star Formation, including RMS objects, IRDCs, UKIDSS Clusters, Roman-Duval Molecular Clouds, Class II Methanol Masers, EGO, BGPS objects and GRS objects. However, there are question as to how these objects relate to one another and how useful they are as tracers of further star formation.

In the following Chapters I attempt to determine the relationship between some of these star formation indicators and identify the most appropriate method for locating high mass star formation sites.

In doing so I will identify and overcome a principle limitation for the identification of massive star formation candidates imposed by the GLIMPSE catalogues and located the infrared counterparts of Class II Methanol Masers within IRAC colour magnitude space.

I will then illustrate how performing K=1 Nearest Neighbour analysis can identify the spatial and evolutionary relationship amongst these various the star formation indicators. Leading on from this I will show how a new and novel application of Minimum Spanning Trees within spatial and non-spatial parameter space can, with the implementation a new over sampling technique, identify both structure and clustering within the P&F IRDC catalogue, the GRS survey data and the BGPS catalogue. Thereby showing the relationship between these objects.

In Chapter 5 I report my recent James Clerk Maxwell Telescope (JCMT) HARP[16] observations of the young maser associated cluster [BDS2003] 107 and discuss the results of these observations in light of the recent paper of Motogi *et al.* (2010) and the observation from Chapter 3 that a small, yet statistically significant number of Class II Methanol Masers, appear to be cluster associated.

[16]HARP, formally known as HARP-B, is a 16 pixel Hetrodyne Spectral Line Receiver with a tuning range of $325 - 375\ GHz$.

Chapter 2

The Colours of MMB masers within GLIMPSE

"Innocence about science is the worst crime today." - Sir Charles Percy Snow

2.1 Introduction

Previous attempts at cross matching masers with GLIMPSE have been undertaken by Ellingsen (2006) and later Ellingsen (2007). In the first of these papers the author cross matched 56 masers detected in the ATCA survey (Ellingsen 2005) and the Mount Pleasant Survey (Ellingsen *et al.* 1996) with the GLIMPSE April 2005 release of the GPSC and GPSA (see Chapter 1 for descriptions of these). The cross match radius was set at 2 arcseconds, which is consistent with the GLIMPSE resolution (Benjamin *et al.* 2003).

Ellingsen discovered that 38 (68%) of the masers used had a GLIMPSE source within 2 arcseconds. For the sources with no detection Ellingsen concluded that the emission was either too faint or the source too extended to be in one of the GLIMPSE catalogues (Ellingsen 2006) although this conclusion was never investigated. It was also noted that there was a strong association between the masers and IRDCs which was taken as an indication that IRDCs are sites of massive star formation. The cross match led to the colour diagrams shown in Figure 2.1 which was noted to be similar to the colour space of the in-fall envelopes of a low mass class 0 protostar (see Chapter 1). The colour - magnitude space occupied by maser counterparts was determined to be [8.0] < 10 and [3.6]-[4.5] $>$ 1.3 using this method.

This colour-magnitude space was used in Ellingsen (2007) to identify possible sites of high mass star formation within the GPSC. Using this colour criteria Ellingsen identified 5675 candidates from the GPSC. From these, sources within 3.5 arcminutes of a maser listed in the catalogue of Pestalozzi *et al.* (2005) were excluded; this left 4878 GLIMPSE sources of which the 100 brightest at 8 μm and 100 reddest at [3.6]-[4.5] were selected for observation using the University of Tasmania Mt Pleasant 26 m and Ceduna 30 m radio telescopes

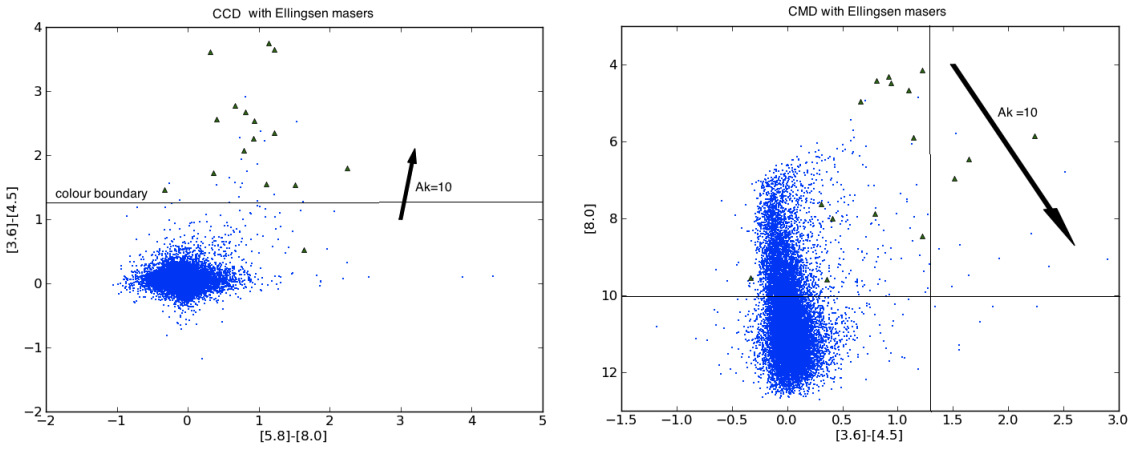


Figure 2.1: Ellingsen maser colour diagrams

The colour-colour and colour-magnitude diagrams from Ellingsen's cross match of GLIMPSE with the masers used in Ellingsen (2006). The small dots represent a random sample of sources drawn from the GLIMPSE point source catalogue. Only sources for which there is flux density information for all the IRAC bands have been included in the plots. The reddening vector is based on Gutermuth (2005) and the references therein. Note that Ellingsen has incomplete colours for some masers within the sample, hence the number of masers within the two plots differ due to this restriction.

with the aim of detecting 6.67 GHz Class II methanol maser.

These observations led to the detection of 38 masers, of which nine were new detections; which were associated with 27 GLIMPSE sources, although issues within the selection criteria led to there being effectively only 198 GLIMPSE sources as opposed to the original 200. Ellingsen concluded that the brighter and redder sources were less likely to be maser sites (Ellingsen 2007).

The large number of masers within the MMB catalogue[1] when comapsed to previous maser surveys, has enabled me to produce a maser GLIMPSE cross match an order of magnitude greater than that performed in (Ellingsen 2005). The MMB's high resolution allows for the accurate identification of maser counterparts within the GLIMPSE field and reduces the chances of false associations. Given the possibility that there is a maser luminosity infrared luminosity relationship, the high sensitivity of the MMB removes the possible selection of only high luminosity counterparts, which would artificially restrain the colour-magnitude space identified as being occupied by the maser counterparts. A number of previous maser surveys were targeted at either colour selected sources or at regions of known massive star formation. Hence, they may introduce a possible bias into the colour of the counterparts. As the MMB is un-targeted this colour bias will not be introduced. Furthermore, the use of a large survey such as the MMB allows me the opportunity to examine the conjecture that the non-detected masers sources within GLIMPSE are due to the source being too faint or extended. Hence we see the the MMB is an ideal survey to examine the relationship between Class II methanol maser and their infrared counterparts. By identifying the infrared counterparts to Class II methanol maser, we can determine the colour space occupied by masing bodies which will reveal, to some extent, their nature and the evolutionary stages covered by masing. This will also enable us to find sites of potential masing and therefore allow us to determine the number of objects that are at the correct evolutionary stage to mase, but do not do so. As stated in Chapter 1, there is a possible bias introduced by the followup of single dish observations by interferometric observations. In this work I use the maser more accurate positions from the interferometric observations.

2.1.1 Cross Matching GLIMPSE and the MMB

Given the relatively low number of masers cross matched in Ellingsen (2006) and Ellingsen (2007), it is clear that cross matching the much larger MMB survey with GLIMPSE is likely to obtain a much larger sample and consequently allowing the more accurate determination of the colour space occupied by maser infrared counterparts.

At the time this work was undertaken there were 812 masers within the MMB catalogue of which 776 lay in the GLIMPSE I-II-3D region. My initial cross matching of the 776 MMB masers and the GPSC used a 2 arcseconds matching radius for consistency with Ellingsen (2006, 2007). This resulted in 431 infrared counterparts (56%) being identified, of which 230 (29%) had fluxes in all four IRAC bands. Cross matching the MMB with the less reliable but more complete GLIMPSE Point Source Archive (GPSA) yields 520 (67%)

[1] 812 when this work was undertaken.

masers with GPSA counterparts within 2 arcseconds, of which 252 (32%) have fluxes in all four IRAC bands. The GPSA fails to list 249 (33%) maser counterparts in any band. The 67% figure for maser counterparts listed within the GPSA is in agreement with the 68% fraction found in Ellingsen (2006), which used both GPSC and GPSA.

The observation that some masers appeared to be missing GLIMPSE counterparts was made in Ellingsen (2006) but was not followed up. To address this I obtained GLIMPSE FITS [2] cut-out images with 12 arcminute diameters for all four IRAC bands (3.6, 4.5, 5.8 and 8.0 μm, respectively). These were obtained via the NASA/IPAC Gator database query tool [3] with each image centred on a MMB maser. A C-shell script utilising the SAO DS9 [4] command line generated RGB images with IRAC bands 1, 2 and 4 being coded as blue, green and red respectively, this being the traditional display for GLIMPSE images. This coding scheme was selected for consistency with previous work and to allow for the easy identification of EGOs. The images were overlaid with the MMB catalogue using the DS9 catalogue tool. The RGB images were used to visually classify the maser counterparts by environment and spatial extension in addition to the identification of masers associated with IRDC and Infrared Clusters.

As noted in Chapter 1, IRDCs, have high levels of extinction, often $A_v \geq 100$ and appear as dark "holes" within the diffuse 8.0 μm emission associated with Polycyclic Aromatic Hydrocarbons (PAHs) (Parsons *et al.* 2009; Simon *et al.* 2006b; Peretto and Fuller 2009). I therefore identified IRDCs by the absence of background 8.0 μm emission.

The existence of a cluster at the maser location is determined visually from the GLIMPSE and, where available, UKIDSS GPS images (Lucas *et al.* 2008). In order for a maser to be associated with a cluster the maser must lie within the bounds of a visually identified cluster. Visual identification of clusters was based on the presence of any candidate cluster containing $\geq$10 possible members, a process that might be expected to include some clusters that are merely random over densities rather than bona fide clusters. A further check was made by cross referencing with the Mercer (Mercer *et al.* 2005), Bica (Bica and Bonatto 2008) and Froebrich (Froebrich *et al.* 2007) cluster catalogues.

[2]FITS is a standard digital file format for astronomical images.

[3]`http://irsa.ipac.caltech.edu/`

[4]DS9 is a FITS multi-platform FITS view supplied by the Smithsonian Astrophysical Observatory.

2.1.2 Cross Match Results

Of the 776 masers within the GLIMPSE region, seven images could not be used for the visual inspection process due to either the maser being located near to the edge of the field or the image being of insufficient quality. On inspection we found that the remaining 769 maser counterparts are each located in one of four broad categories:-

- Type 1: Masers embedded within an IRDC with no IRAC counterpart. Hence, they are infrared-dark at IRAC wavelengths, possibly due to extinction (or the lack of infrared emission). This represents 5% (37) of the maser counterparts. Figure 2.2 shows a typical example.
- Type 2: Masers that are located within, or on the perimeter of, an IRDC but have a visible IRAC counterpart and therefore are infrared-bright. This class represents 21% (164) of the counterparts. Figure 2.3 shows a typical example.
- Type 3: Masers that are infrared-bright and not located within an IRDC. This class contains 62% (473) sources. Figure 2.4 shows a typical example.
- Type 4: Masers that are infrared-dark (they have no identifiable counterpart in any band) but are *not* located within an Infrared Dark Cloud. My examination indicates that 12% (95) of the objects are of this type.

The visual inspection also shows that 5% of the MMB population are embedded within IRDCs whilst 12% appear to have no infrared association (including IRDC). It is clear therefore that there is a significant population of MMB masers with IRAC counterparts that do not have GLIMPSE fluxes in all four IRAC bands because they are either extended in more than one band and hence not included in the GPSA/C or are infrared dark. As a consequence, previous IRAC colour-colour diagrams of maser counterparts have been limited to the restricted number of 6.67 GHz masers with GLIMPSE fluxes.

It was found from the visual inspection that 23% of maser counterparts showed excess emission in band 2 in the visual inspection when compared to bands 1 and 4, which I take as an indicator of an EGO. The number of clusters identified as being associated with masing sites was small, with only two clearly identifiable clusters being maser associated. Traditional aperture photometry such as that used in the construction of the GLIMPSE catalogues (Babler 2006) is only suitable for the objects that are compact, unresolved sources and preferable in regions that are neither crowded or have inconsistent backgrounds. As seen from the examination of the GLIMPSE images, many counterparts are extended and in regions that have a high source density and incongruent backgrounds; the very regions in which traditional aperture photometry can be unreliable. In order therefore to recover the fluxes of maser counterparts which are not detected by the GLIMPSE photometry pipeline, a photometric technique that overcomes these limitations is required. In the following section I discuss the development, testing, implementation and results of such a method.

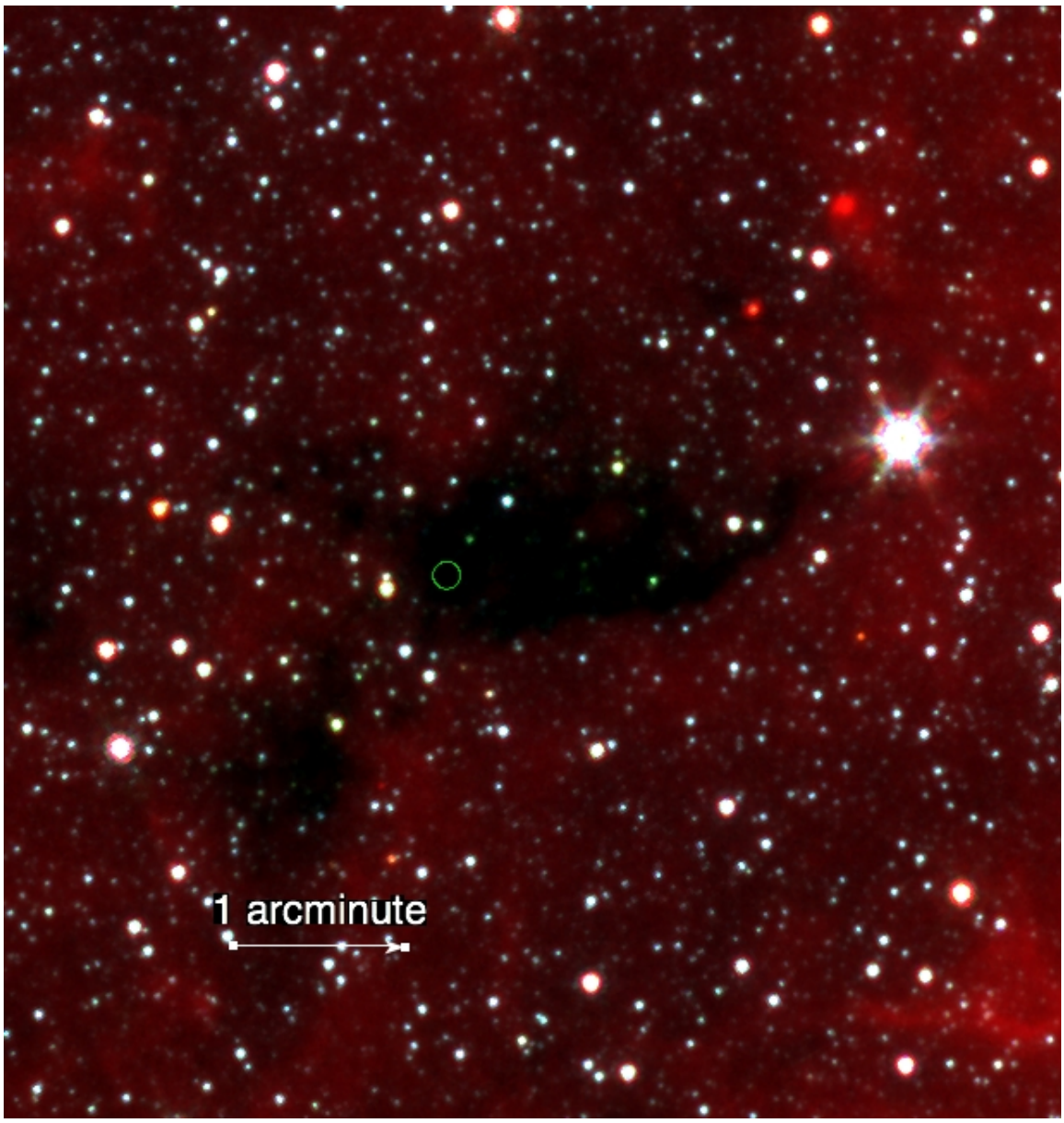


Figure 2.2: IRDC IR Dark Associated maser
IRAC RGB image of a maser counterpart, with red, green and blue being bands 4, 2 and 1 respectively. The maser position marked with the green circle is located within the IRDC, has no observable IRAC counterpart and is therefore considered a Type I source. The white arrow, representing 1 arcminute, is shown for scale.

2.2 Adaptive Non-Circular Aperture Photometry

From section 2.1.2 we see that the GLIMPSE catalogues only contain point source and that many MMB counterparts are not point like. Therefore, any attempt to define a colour

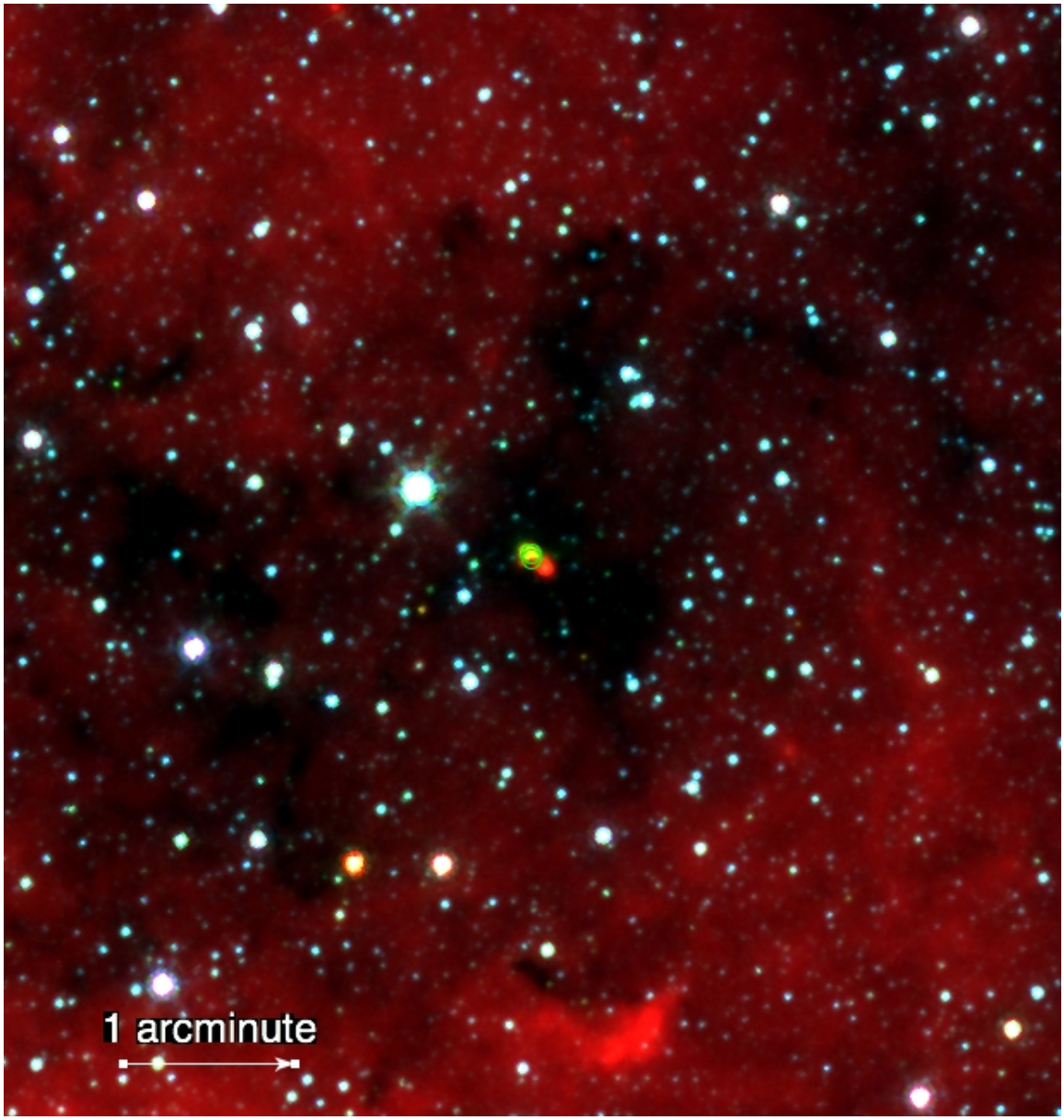


Figure 2.3: IRDC IR Bright maser

IRAC RGB image of a maser counterpart, with red, green and blue being bands 4, 2 and 1 respectively. The maser position is marked with the green circle. It is located within the IRDC but unlike Figure 2.2 it has a clear infrared counterpart and hence is considered a Type 2 maser. The white arrow, representing 1 arcminute, is shown for scale.

space for maser counterparts using solely the GPSA and GPSC such as those of Ellingsen (2006) and Ellingsen (2007) may be biased in that they only contain maser counterparts that are either distant, and hence unresolved, or are at an evolutionary stage were they are still relatively compact. Hence, the counterparts found within the GLIMPSE catalogues

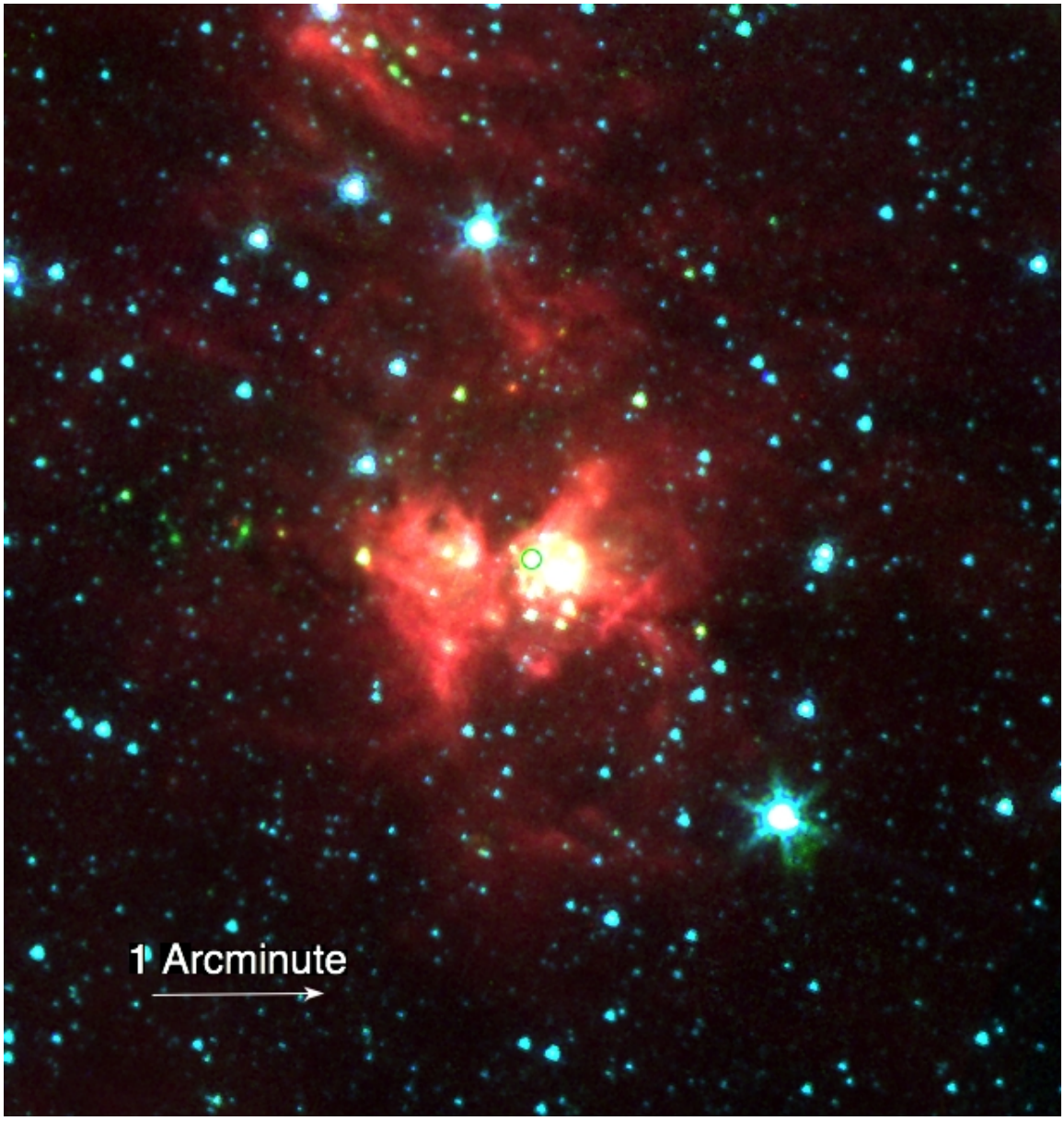


Figure 2.4: IR Bright maser
IRAC RGB image of a maser counterpart, with red, green and blue being bands 4, 2 and 1 respectively. The maser position is marked with the green circle. The maser appears to be associated with the bright source near the centre of the field and hence is classified as a Type 3 maser. The white arrow, representing 1 arcminute, is shown for scale.

may not be representative of the entire population.

Traditional photometry, such as that used with GLIMPSE attempts to reliably measure the flux of an object with no consideration given to any other information that might be

contained in alternative bands. However, as we see from section 2.1.2 many of the MMB infrared counterparts are extended, often but not exclusively, in bands 3 and 4. Hence, any attempt at finding the integrated flux risks contamination from background and foreground sources which will skew the colours towards the blue. Additionally, traditional photometry uses an annulus surrounding the aperture to define the background. However, in regions where the background is complicated this is unlikely to be representative of the local background. This is further confounded in regions of high stellar density by the high probability of contamination of the annulus by stellar sources. Furthermore, in the case of extended sources, large apertures may be required which may also be contaminated. For sources that are unresolved or show little extension it is possible to fit a reliable aperture using traditional photometry (although the background selection problem still persists).

2.2.1 GLIMPSE Photometry

To overcome these issues I have developed an adaptive non-circular aperture photometry (ANCAP) technique, based on that used by Cyganowski *et al.* (2008a) to measure the fluxes of EGOs within the GLIMPSE image archive. The process overcomes the issues with background selection by not restricting itself to an annulus. The issue with background contamination of the aperture is somewhat more problematic. I can accept that the aperture is contaminated and that the fluxes will be too large or I can remove the contamination, and by doing so some of the flux from the object, and accept that the fluxes are too low. The former method will result in colours skewed to the blue due to photosphere contamination. However, if we assume that the flux in the extended regions is relatively uniform [5] then the removal of a small amount of flux, although undesirable, should not overly affect the colour. For point sources, those found within the GPSA/C, the following method effectively becomes traditional photometry, however background subtraction implementation should result in more reliable results in crowded regions.

The region feature of the SAO DS9 FITS viewer was used to draw polygon apertures around the sources. The counterparts are identified by the presence of a co-located peak in all four IRAC bands within 2 arcseconds of the maser, a radius selected to be consistent with the previous GPSA cross matching of Ellingsen (2006) and IRAC's band 4 resolution (Benjamin *et al.* 2003). For reliability, if the peak cannot be identified in all four bands the photometry is not undertaken.

This process resolves an issue with very extended sources which may encompass multiple shorter waveband sources. By utilising all four bands to identify the counterpart the rate of misidentification is reduced. In the case where there is more than one possible counterpart I make no measurement. Figure 2.5 illustrates the process used in identifying the infrared counterpart. In this example there are two masers in the field, marked by the red and blue circles in each frame. The circles are 2 arcseconds in radius. If we used only the band 4 image to determine the association we would incorrectly associate both masers with the large "clover leaf" like object. However, looking at band 1 we see that in fact only the maser marked with the red circle is associated with a source with flux in all four

[5] It is likely that the extended sources are either infalling envelopes or UCHII regions in which case the assumption should hold. See Section 2.3

bands. In my example, Figure 2.5 bands 3 and 4 would be rejected for aperture size and shape determination, due to the likelihood of contamination. In this case band 2 would be used for this function.

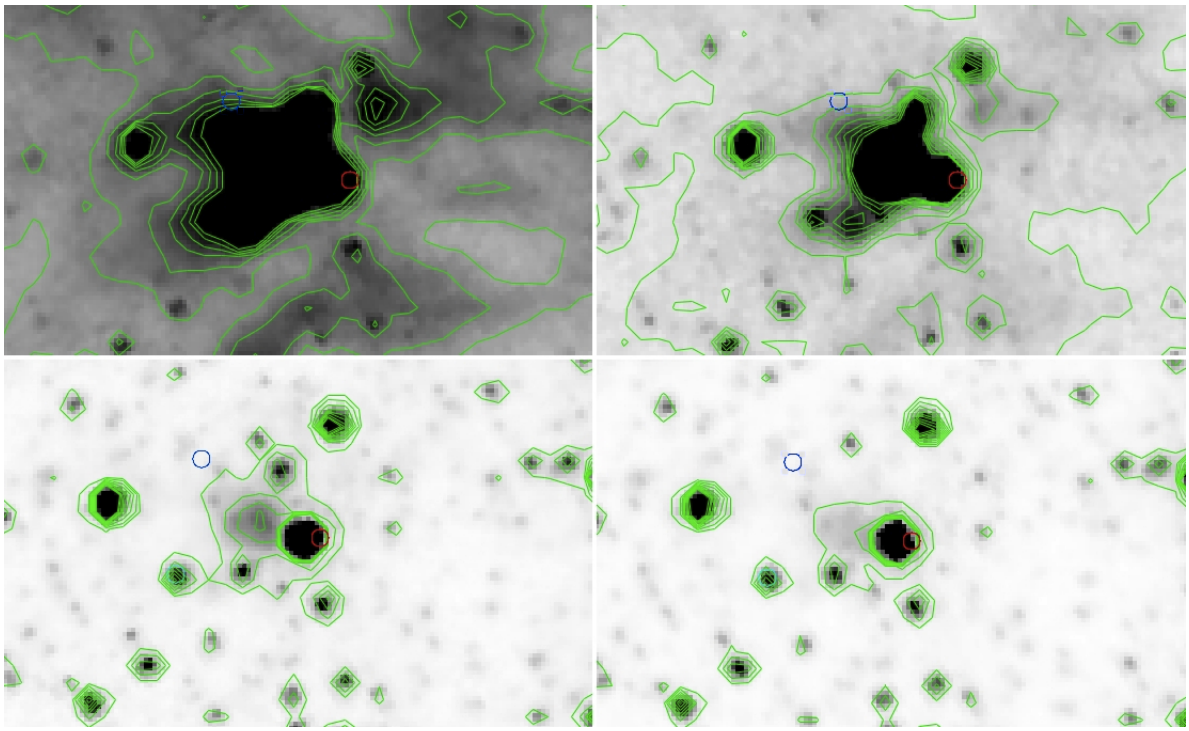


Figure 2.5: Select of infrared counterpart
The above images show an illustration of the process used to select the infrared counterpart. The images are taken in IRAC bands 4 to 1, left to right, top to bottom. The red and blue circles indicate the positions of two masers and are two arcsecond in radius. The green contouring is provided to help identify the sources. In this example bands 3 and 4 are unsuitable for the identification of the source and the determination of aperture size, shape and location. Hence in this example the aperture characteristics would be determined using band 2. Note that the maser within the blue circle will be considered to be infrared dark as it is not associated with a source, except with the extended source seen in band 4, which is likely to be coincidental.

The SAO DS9 program allows us to use an aperture of any size or shape. The aperture shapes are selected to avoid contamination from field sources and EGO whilst containing as much flux as possible, with preference given to band 4, when choosing the aperture shape. Hence, if using band 4 to select the aperture size would result in avoidable contamination band 3 would be used instead. An identical region was located in an area of the image representative of the background local to the maser source, for use as a background subtraction thereby addressing the aforementioned background subtraction problem as-

	Band 1	Band 2	Band 3	Band 4
Limit (mag)	7.0	6.5	4.0	4.0

Table 2.1: GLIMPSE Saturation Limits
GLIMPSE Point Source magnitudes Saturation limits taken from `www.astro.wisc.edu/sirtf/addendum4.pdf` and used to determine whether the maser counterparts are possibly saturated.

sociated with traditional aperture photometry. Both regions are constant in size, shape and position across all four bands. The regions were saved as d9.reg files for repeatability and consistency between bands. An alternative method was investigated whereby an all encompassing aperture would be used with any contamination masked out. However, during testing it was found that there were no discernible benefits for this alternative method which added complexity and reduced repeatability.

The DS9 funtools plugin [6] allows us to integrate the flux in both regions and thereby obtain a raw flux to which I applied unit, scale and aperture corrections as outlined in the IRAC cookbook [7] and described below. Figure 2.6 shows a set of IRAC images overlaid with the region which makes up the aperture and the background.

There is the possibility of saturation with the GLIMPSE fields. However, saturation, or near saturation, are not easily determined from the GLIMPSE images as the pixels values are not set to zero when saturated. It is likely that saturated sources will either be too extended for even this photometric technique or will have inconsistent colours, in which case they would be identified at either the pre or post measurement stage and removed from the sample if deemed from the individual pixel counts to be saturated. However, as an additional check, my results contain a "Saturation Warning" column which warns the user that a source may be saturated; although the user should be aware that many of the sources are extended and that this is only a warning based on point source saturation limits. The sources marked as saturated in the GPSC will be marked as possibly saturated within the ANCAP results and the extended sources, which are not in the GPSC, will be marked as possibly saturated if their magnitude exceeds the GLIMPSE limits for point source. The number of sources that carries this flag is approximately 8% in band 1, increasing to 15% in band 4. Table 2.1 shows the point source saturation limits.

Once the raw flux was obtained from Funtools I applied a correction factor of 1/108 which is partly derived from the IRAC cookbook. Within this correction factor there is an additional component which was obtained from the calibration exercise below. I have, as yet, been unable to clearly identify the reason behind this additional factor as it is not discussed in Cyganowski *et al.* (2008a) nor in any of the GLIMPSE documentation. It is very probable that this is caused by the GLIMPSE image processing pipeline which rescales the images from a 1.2 to a 0.6 arcsecond pixel size. However, this additional factor is needed

[6] (`http://www.cfa.harvard.edu/~john/funtools/ds9.html`)
[7] (`http://ssc.spitzer.caltech.edu/postbcd/irac_reduction.html`)

	Band 1	Band 2	Band 3	Band 4
Zero Point	280.9	179.7	115.0	64.12

Table 2.2: GLIMPSE zero-points
GLIMPSE zero-points as per the GLIMPSE cookbook (Meade *et al.* 2007) as used in the ANCAP photometry.

for consistent results with the GPSC and is constant between bands. A further check was made against a sample of EGO from Cyganowski *et al.* (2008a) from which this technique was derived. It was found that the results were consistent with Cyganowski *et al.* (2008a) if this factor was applied.

This process is then followed by aperture correction. Funtools reports the total area of the aperture and this is used to obtain an effective aperture radius, the radius of a circular aperture with the same area as reported by Funtools. Where needed I applied the required correction factor as per the GLIMPSE cookbook (Meade *et al.* 2007), although this correction has a modest effect on the flux (approximately 2%) and is much less than the measured photometric uncertainty.

Once the corrected fluxes are obtained the zero points are then applied in order to obtain a magnitude using Equation 2.1 and the zero points listed in Table 2.2 where f_o is the zero point and f_1 is the measured flux.

$$2.5\ log10(f_o/f_1) \tag{2.1}$$

If the source has a listed GPSC flux this was checked against the measured flux. If the measured and GPSC magnitudes were beyond the 0.2 mag uncertainty the images were checked to determine the nature of this magnitude offset. In the majority of such cases this was due to misidentification of the counterpart during the GPSC-MMB cross match. There were no such conflicts in magnitude that could not be resolved. An additional check was made once the colour diagrams were produced.

In order to confirm the reliability of this technique a GLIMPSE field representative of the general MMB population was selected, on which was performed the above aperture photometry technique on 100 sources with known GLIMPSE fluxes. It was found that 80% of the sources measured had band 1 GLIMPSE fluxes within 20% of those measured which is consistent with the 0.2 magnitude uncertainty within the GPSC (Benjamin *et al.* 2003). The measured fluxes in bands 2 to 4 depart from the GPSC fluxes more than in band 1 which can be attributed to the increased background variability and source extension. Note that the aperture used for the background subtraction is not the same as that used within the GPSC. However, from these test results we appear to obtain a photometric

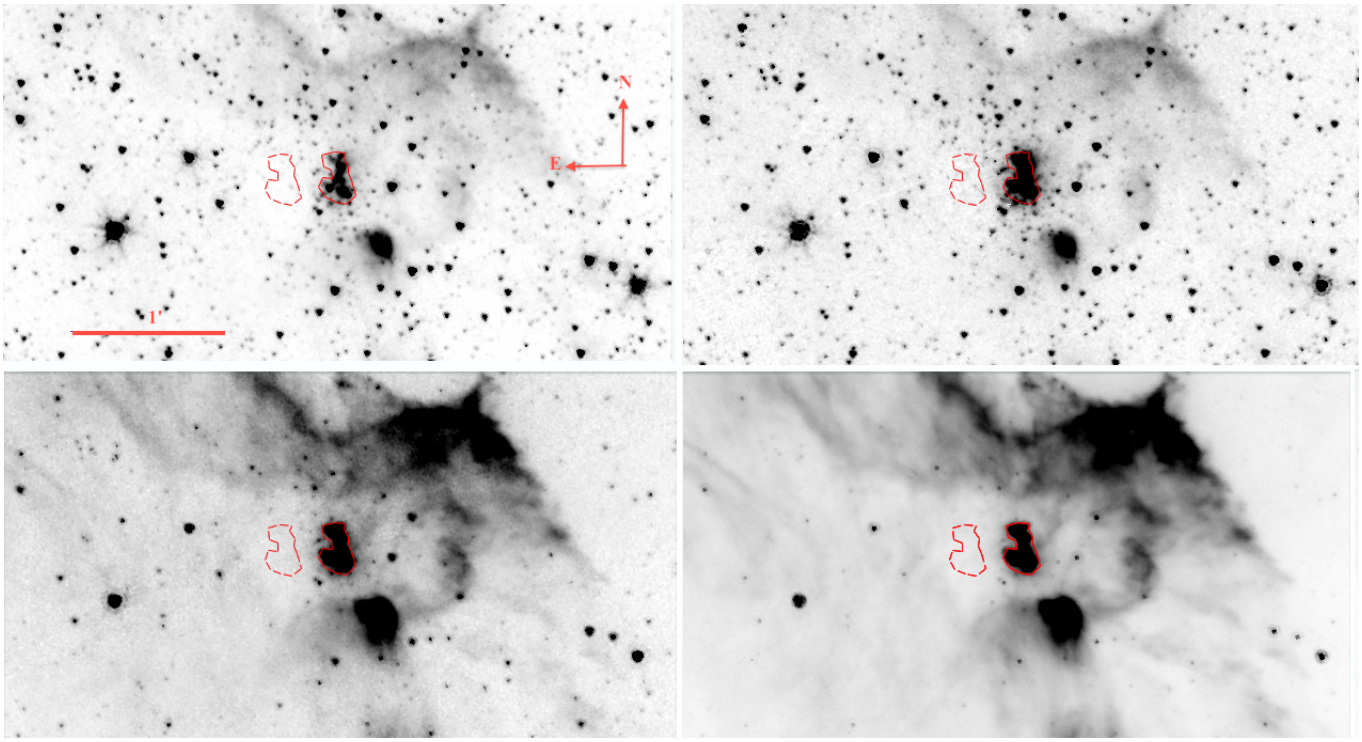


Figure 2.6: Non-circular aperture photometry
Example of the regions used in the non-circular aperture photometry. The above images are 3.6 μm *top left,* 4.5 μm *top right,* 5.8 μm *bottom left and* 8.0 μm *bottom right. The red vector has a length of 1 arcminute. The two red polygons are the apertures and background. The polygon in the centre of the image is the aperture and the broken polygon offset to the right is the region used as the background. The solid red line is the one arcminute scale line. The orientation of the images is shown by the red compass in the* 3.6 μm *image.*

accuracy of better than 0.2 mag at 2σ.

As an additional reliability check I compared the ANCAP final photometry against those sources that are also contained in the GPSC and the GLIMPSE Point Source Archive (GPSA). Within the sample of aperture photometry measurements there are ~ 200 sources in common with the GPSC and ~ 300 sources in common with the GPSA (the exact number in common varies from band to band). Plots of the aperture measured magnitudes (ANCAP) versus the GPSC and GPSA magnitudes at each waveband are shown in Figure 2.7. For the majority of sources in common with the GPSC and GPSA I measured a magnitude that is in either close agreement or brighter than that contained in the Catalogue or Archive (note that the typical magnitude errors in the GPSC and GPSA are ≤ 0.2 mag Benjamin *et al.* (2003)). The latter sources are consistent with those that the technique identifies as extended objects, hence the GPSC and GPSA have underestimated the total flux of these objects. I also identified a few objects ($\leq$10 in total) within the common sample where the aperture photometry measurements are *fainter* than those from the GPSC

or GPSA. In general these objects lie within regions of complex background and, as the background subtraction method differs substantially from that used by GLIMPSE, it is likely therefore that the discrepancy in measurement is caused by different background subtraction. In any case the fraction of sources affected in this way is small ($\leq$5%) and I am confident that the aperture photometry technique yields realistic and consistent results, compatible with the GPSC point source photometry, for the majority of sources.

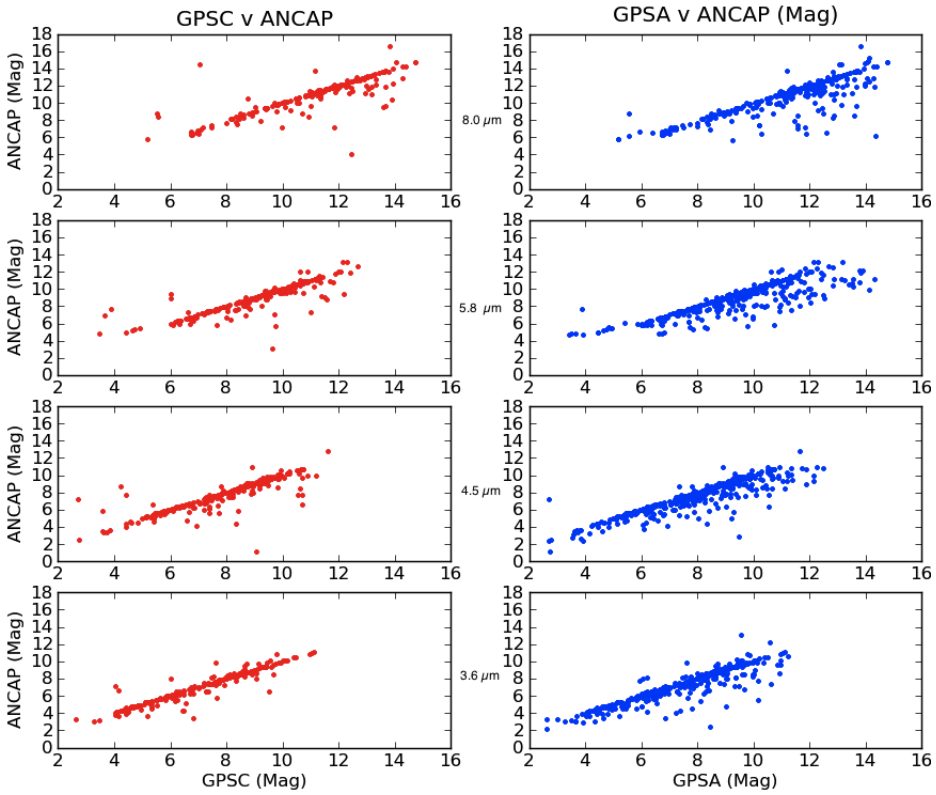


Figure 2.7: GPSC,GPSA v ANCAP magnitudes

Plot of ANCAP magnitudes for maser counterparts against magnitudes listed within the GPSC (left column) and the GPSA (right column). As can be seen the ANCAP magnitudes are in general in agreement with the catalogue. The population of objects that are brighter in ANCAP can be attributed to the small, fixed aperture size used by the GPSC/A being applied to resolved objects. The objects that are fainter in the ANCAP photometry are in general objects that have been mis-identified in the cross match.

Name	R (arcseconds)	Band 1	Band 2	Band 3	Band 4	Sat
G000.0916-00.6630	5.05	10.9	8.8	8.5	9.3	0000
G000.1669-00.4456	6.28	9.8	6.8	5.4	4.7	0000
G000.3154-00.2008	6.31	7.4	6.1	5.5	4.8	0100
G000.3157-00.2015	6.31	7.4	6.1	5.5	4.8	0100
G000.4962+00.1880	7.55	11.4	9.6	7.9	6.9	0000
G000.5463-00.8516	6.00	9.4	7.6	5.4	3.6	0000
G000.8358+00.1844	14.91	8.2	7.9	6.1	4.7	0000
G001.0081-00.2369	28.09	7.5	7.3	6.5	5.3	0000
G001.7191-00.0879	19.91	9.1	8.4	7.8	7.4	0000
G002.1434+00.0091	26.66	8.5	7.5	5.6	4.2	0000
G010.2048-00.3455	5.05	11.6	9.1	8.4	8.2	0000
G010.3205-00.2586	3.48	11.8	9.0	7.5	6.3	0000
G010.3423-00.1423	6.05	10.8	9.4	8.7	9.2	0000
G010.4723+00.0273	4.92	12.5	10.0	7.8	6.0	0000
G010.6265-00.3844	4.92	11.7	9.6	8.4	7.0	0000
G010.6287-00.3329	5.23	12.5	10.4	9.6	9.2	0000
G010.8856+00.1228	10.97	6.7	5.3	4.0	3.4	1111
G010.9583+00.0223	8.58	10.1	8.2	6.6	5.2	0000
G011.0341+00.0618	4.92	10.0	9.0	6.9	5.4	0000
G011.1093-00.1140	10.97	9.9	7.3	6.0	5.6	0000

Table 2.3: Sample GLIMPSE-MMB Photometry

using the ANCAP method outlined above. Col 1 - Maser Name from the MMB survey, Col 2 -Effective aperture radius arcseconds. For non-circular apertures this is the radius of the circle with the same area. Col 3- 3.5μm mag, Col 4- 4.5μm mag, Col 5 - 4.5μm mag, Col 6 - 8.0μm mag, Col 7 Saturation Bits, 0 is unsaturated, 1 is possibly saturation. Bits represent IRAC bands 1 to 4 left to right

2.3 Results

Using the method described above I successfully obtained aperture photometry in all four GLIMPSE bands toward a total of 512 masers. The remaining 257 masers either had no detectable counterpart due to being infrared dark, were located within an IRDC or were located in regions too confused to reliably identify an infrared counterpart. A total of 132 masers were found to have no detectable infrared counterpart and 125 masers were too confused to measure reliable fluxes. A simple positional association with the GPSC or GPSA results in a total of 219 and 253 masers with infrared counterparts measured at all four wavelength bands. Hence, the adaptive non-circular photometric method more than doubles the number of masers with known infrared counterparts measured at all 4 bands over catalogue-based searches. A sample of the source locations, the aperture used and the measured magnitudes are shown in Table 2.3.

Figure 2.8 shows the [3.6]−[4.5] versus [5.8]−[8.0] colour-colour plots of the maser coun-

terparts compared to a sample population of 15 000 sources drawn randomly from the GLIMPSE Point Source Catalogue. I showed the colours of maser counterparts drawn from the GPSC , GPSA) and the ANCAP photometry measured here. A reddening vector of $A_K = 10$ is displayed and was calculated based on that of Gutermuth (2005) and references therein. The maser counterparts are much redder than the general stellar population and show colours consistent with the smaller sample of 6.7 GHz masers investigated by Ellingsen (2006). All three methods show good agreement in the colours of the maser counterparts, although the ANCAP counterparts occupy a marginally wider range in colour space than the GPSC or GPSA sample. A small number of sources appear offset from the main body within this colour-colour space. I suspect that these are misidentified counterparts of infrared dark masers or are sources contaminated by either EGO emission or background/foreground stars. Figure 2.10 shows that EGO associated masers tend to be shifted to the blue in [5.8]-[8.0] which confirms the suspicion that at least some of the masers with unusual colours are contaminated by 5.8μm emission associated with EGO. Figure 2.9 illustrates the colour-space as defined by Ellingsen (2006) and Ellingsen (2007) suggesting that the colour space occupied by MMB masers is consistent with those from previous cross matches.

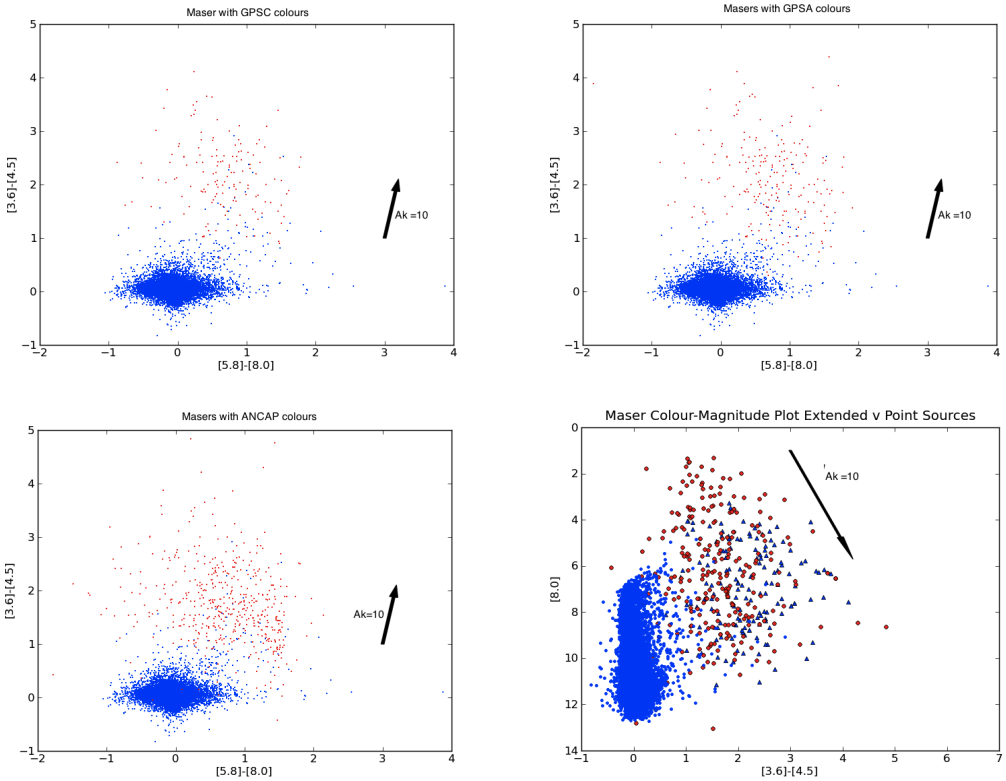


Figure 2.8: Colour plot of GPSC, GPSA, and ANCAP

[5.8]-[8.0] v [3.6]-[4.5] colour plot of GPSC, (top left) GPSA (top right), and ANCAP (bottom left) maser infrared counterparts respectively, with the counterparts shown as red dots and a random sample of GLIMPSE sources shown as blue dots for reference. This illustrates the broadening of colour space with increased source numbers and the general consistency of colour space between the three datasets. The plot on the bottom right compares the masers within with GPSC counterparts marked as blue triangles to these not found within the GPSC but with ANCAP fluxes, red triangles. Hence this plot compares the maser counterparts that are point sources, ie in the GPSC, to those that are extended. We see that the extended sources occupy a colour magnitude space that is about 0.5 mag bluer than the point sources. Reddening vector $A_k = 10$ *from Gutermuth (2005) and references therein.*

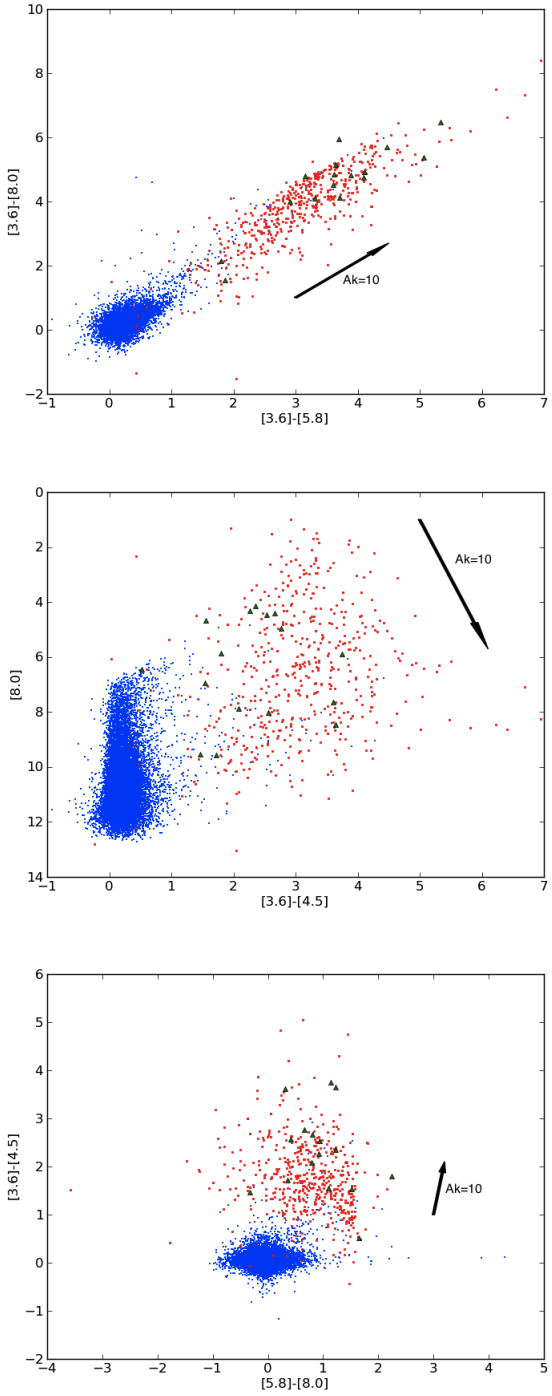


Figure 2.9: Colour plots of Ellingsen Masers

Colour-colour and Colour-magnitude plots showing the general GLIMPSE field population (blue), the maser counterparts as measured as part of this work (red) and the masers counterpart colours from Ellingsen (2006) (Green). Reddening vector $A_k = 10$ from Gutermuth (2005) and references therein. Note that the Ellingsen masers have incomplete colours and consequently not all Ellingsen masers are shown in every plot.

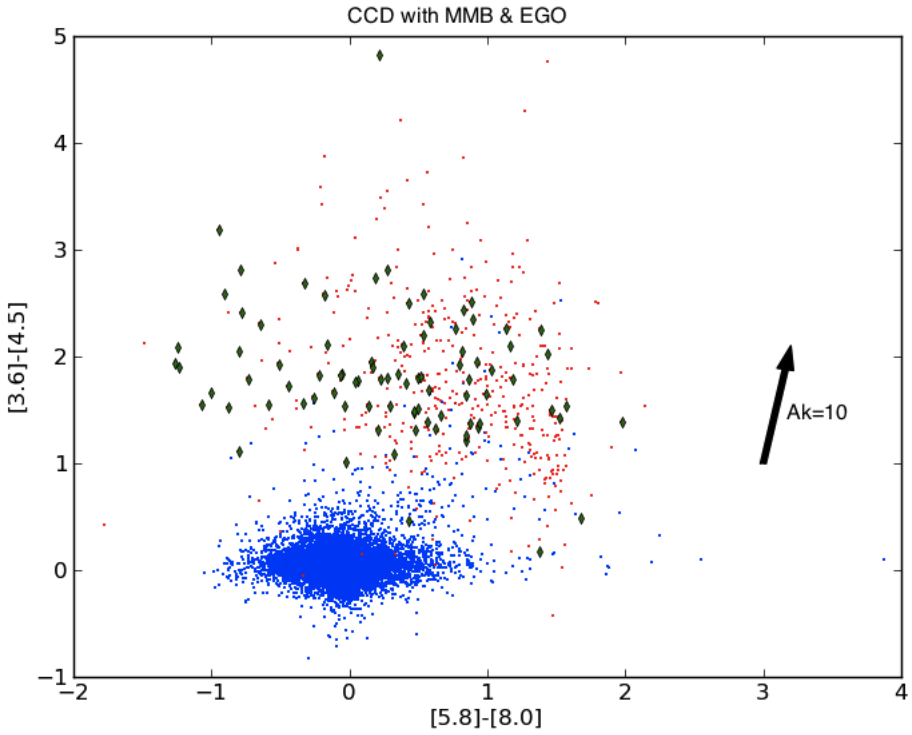


Figure 2.10: Colour-Colour Diagram of EGO associated masers

Colour-Colour Diagram showing the GLIMPSE field sources for reference (blue dots) the MMB masers (red dots) and the masers associated with an EGO (green diamonds) as identified from the visual inspection of the GLIMPSE images The offset towards the blue in [5.8]-[8.0] suggests that the EGOs maybe contaminating the photometry in some cases. The $A_k = 10$ reddening vector is from Gutermuth (2005) and references therein.

As discussed in Chapter 1, the current model for the production of Class II methanol maser involves infrared radiation from warm dust acting as a pump (Sobolev *et al.* 1997). This being the case we would expect there to be a relationship between infrared emission and maser emission. However, from Figure 2.11 we see that no such relationship is observed.

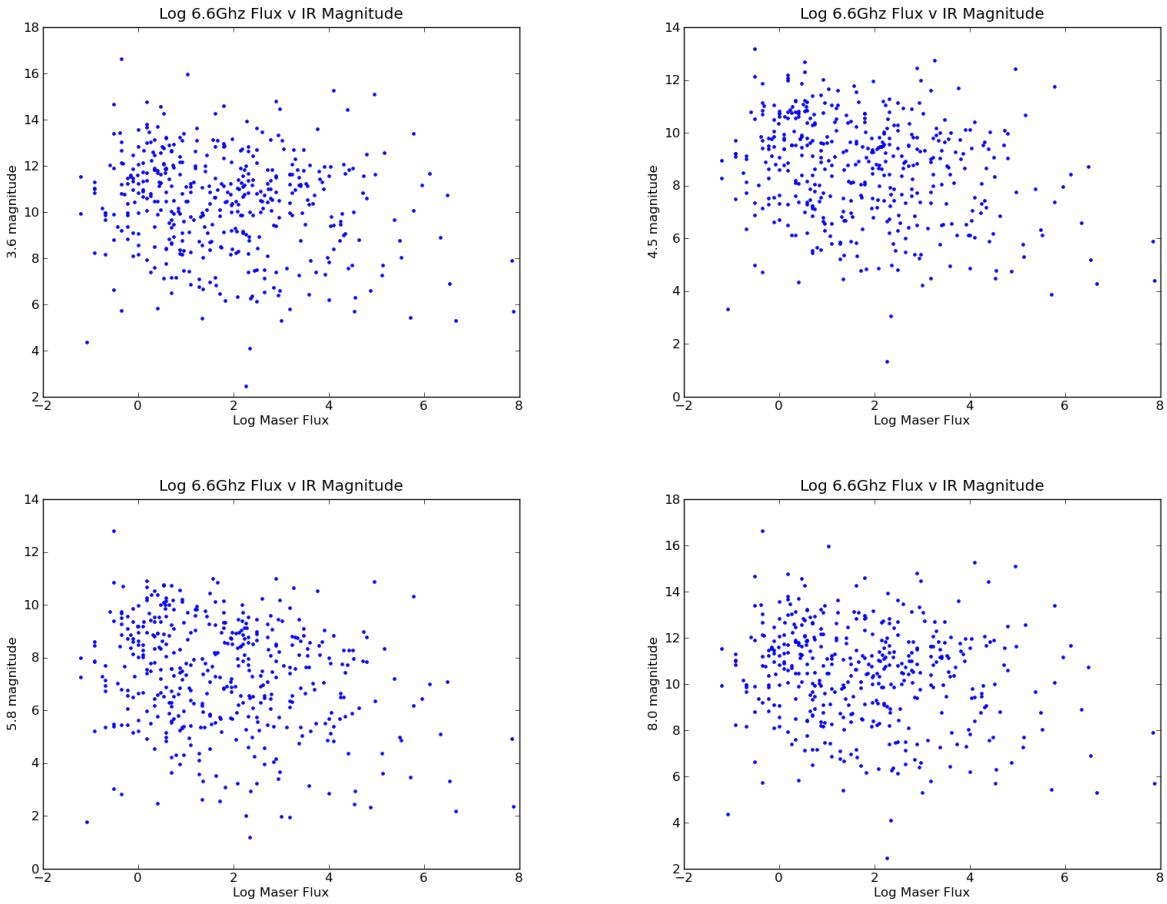


Figure 2.11: Radio v IR Flux

Plot of the log of the 6.67 GHz maser flux versus the infrared magnitude of the counterpart in each IRAC band. If the pumping rate is the controlling factor in maser brightness, then we might expect to see a correlation between maser luminosity and infrared luminosity. However, as can be seen there is no detectable correlation, suggesting that other factors are more dominant in determing maser brightness.

2.4 Discussion

The colour selection criterion derived in Ellingsen (2006) was based on the best available source of maser infrared counterpart fluxes, the GPSA, and as a consequence it contains a restricted sample size due to the GPSA's exclusion of extended objects. The procedure outlined in section 2.3 overcomes the restriction to point-like sources imposed by the GPSA and when combined with the large new MMB maser catalogue, provides an order of magnitude more counterparts than Ellingsen (2006). As such it should represent a more complete picture of the IRAC colours of maser counterparts.

We see from Figures 2.8 and 2.9 that we can expand the Ellingsen colour selection criteria to [3.6]-[4.5]>1 and [8.0]<12. In general the colour-magnitude space identified by Ellingsen (2006) is in agreement with my own observations. This indicates that the resolved and unresolved maser counterparts within GLIMPSE may form the same population. This is somewhat unexpected as I would have expected maser counterparts to evolve in colour-magnitude space as they become increasingly extended. This observation also suggests that the removal of point sources within extended objects is not overtly affecting the colour despite removing flux and therefore confirms that the ANCAP photometry is working as expected. This is in addition to the implication that the extended sources have broadly the same colour across their surfaces. These two observations indicate that in the wavelengths covered by GLIMPSE, we are observing the cocoon surrounding the hot core or MYSO and not the primary radiating source itself.

As noted in Ellingsen (2006) the masers appear to occupy a colour-space similar to that of Class 0 protostars as modelled in Whitney *et al.* (2003). However, it is also similar to the region occupied by HII regions as shown in Cohen *et al.* (2007) and we should not draw the conclusion that the maser counterparts represent a Class 0-like object with an in-falling envelope without first excluding the possibility that we are observing the surface of an HII region.

The observation that 83% of masers have an IRAC counterpart suggest that an IRAC colour selected survey would be successful in finding further massive star formation sites, although it would require the building of a GLIMPSE extended source catalogue in order to be complete. It is likely that the 5% of infrared dark masers associated with IRDCs have emissions at longer wavelengths and a tentative examination of the MiPSGAL 24 μm images suggest that this is the case. The remaining 12% of dark masers is somewhat more problematic, although it is likely that they are very distant and that the infrared counterpart is either too faint or too heavily extincted to be observed.

The observation that young infrared clusters do not seem to be associated with Class II methanol masers is suggestive that massive star formation has ceased once the cluster becomes identifiable as such. However, the visual identification of young clusters within GLIMPSE is complicated by the survey's relatively low saturation limit hence associations may be more common than the visual search suggests (see Chapters 3 and 5).

The observation that 23% of masers are EGO associated would seem to be in broad agree-

ment with the Class I and Class II methanol maser lifetime overlap shown in Ellingsen *et al.* (2007), if as suspected EGO are tracing outflows (Cyganowski *et al.* 2008a; Chen *et al.* 2009). However, this level of relationship is from a visual inspection only. This relationship is examined in greater detail in Chapter 3.

The lack of correlation between the maser brightness and infrared counterpart brightness would seem to indicate that infrared radiation is not the pumping source. However, this conclusion is incorrect. We would only see such a relationship if the masers are unsaturated (see Chapter 1), which we cannot assume. It is also likely that environment and geometry plays an important role in the determination of maser brightness. This would be especially true in saturated sources where coherence plays a role in limiting maser luminosity. Perhaps a more useful test be to see if there is a relationship between maser variability and variability in the infrared counterpart.

2.5 Chapter Summary

Previous attempts to determine the IRAC colour space occupied by Class II methanol maser indicated that 68% of masers had no IRAC counterpart. Cross referencing the more complete MMB with the GLIMPSE confirms this observation. However, inspection of the GLIMPSE images associated with the MMB masers indicates that only 17% of masers are infrared dark and that many maser counterparts are extended and therefore not contained within the GLIMPSE catalogues.

By implementing an adaptive non-circular aperture technique I have removed the limitations imposed by the GLIMPSE photometric pipeline and identified 512 maser counterparts with measurable fluxes in all four IRAC bands. This has enabled the production of a colour-colour plot with an order of magnitude more objects than previously obtained and a factor of two more than would be achieved by a GPSA MMB cross match alone. By doing so I have confirmed the previous work of Ellingsen (2006) that methanol masers infrared counterparts occupy the colour space [3.6]-[4.5]>1 and [8.0]<12 colour space that is occupied by both HII regions and Class 0 protostars.

From the visual inspection of the GLIMPSE images associated with MMB masers I have determined that 25% of masers are associated with EGO. I also report that Class II methanol maser are rarely associated with young infrared clusters suggesting that massive star formation has ceased once a cluster becomes infrared identifiable.

By matching observed maser flux with the flux of the infrared counterpart I show that there is no determinable relationship between the two. However, we cannot infer from this observation that the masers are not pumped by warm dust emission as other processes may be more dominant in determining maser brightness than the pump luminosity.

Chapter 3

Cross Matching the Massive Star Formation Indicators

"Research is what I'm doing when I don't know what I'm doing." - Wernher von Braun

In this Chapter I discuss the spatial cross matching of a number of catalogues discussed in Chapter 1 with the aim of improving our understanding of the relationship between these various indicators of star formation.

As part of this process I will be using a simple implementation of k Nearest Neighbour analysis (kNN). When applied to two populations, A & B such as in this Chapter, the algorithm simply finds the nearest member of population B for every member of population A, hence $k = 1$. In the case of this implementation I use the non-Euclidean sky distance although when the sky separation is very small it becomes in effect the Euclidean distance.

By undertaking this cross matching I can show how the various star formation indicators relate in terms of the Ellingsen Strawman evolution sequence. In addition the cross matching will allow the comparison of indicators that trace shared or overlapping massive star formation stages. The process will also give an indication of the reliability of each tracer in the identification of massive star formation sites.

Monte Carlos

In order to determine the relative strength of any association I performed a series of Monte Carlo simulations. Although each set of Monte Carlos is independent, it is only the parameters that differ. For each pair of datasets, which I shall term dataset A and dataset B, the positions of the A dataset objects are scattered by a random amount in l and b, with the angular distance between the original position and the scatter position never being greater than one degree (0.5 degrees in the case where the GRS is dataset A). This ensures that the object density is the same within the field and that objects stay approximately in the same star forming region. All scattered points are checked to ensure that they still lie within the original survey area. If they are not, they are re-scattered until they do so. The

angular separation between every B dataset source and the scattered A dataset source is calculated and the number of associations below a set matching radius is counted. This matching radius is 18 arcseconds for EGO and RMS sources, 33 arcseconds for BGPS and GRS sources, and 30 arcseconds and 60 arcseconds for clusters. The matching radius for IRDCs and Roman-Duval clouds is dependent on the individual source and is discussed below.

This process is iterated 100 times. With the maximum number of associations in any one Monte Carlo within the 100 iterations being reported as S_{Max} and the mean association rate over all iterations being reported as $\bar{\mathrm{S}}$. Hence if we have a random association rate in the live data we would expect an association rate equal to $\bar{\mathrm{S}}$, whilst an association rate greater than S_{Max} would represent a 3 sigma certainty that we are seeing real association between the two datasets.

Due to the extended nature of IRDCs, matching radius based on each IRDC's semi-major axis was used instead of using a fixed matching radius as above, an individual. Hence I am looking to see if the matching object lies within the IRDC boundaries.

Likewise, for Roman-Duval MMB cross match, I use an *lbv* box based on the velocity range of the cloud and semi-major and semi-minor axis. As I am using three dimensional parameter space in which to search, I also scatter the masers position by up to 5 kms^{-1} ensuring that the maser lies within the velocity range of the survey. Hence, in the case of the Roman-Duval MMB cross match, to be associated a maser must lie with the cloud in *lbv* space.

Note that for objects with a large number of sources, the P&F IRDCs, the BGPS and the GRS, I elect not to cross match but use a different method as discussed in Chapter 4.

Table 3.1 shows the specific setting used for each cross match and the resultant S_{Max} and $\bar{\mathrm{S}}$ resulting from the Monte Carlos, as well as the result from the cross match between the unscattered data.

3.1 The MMB and RMS

As both the MMB (Green *et al.* 2009a) and RMS (Urquhart *et al.* 2008) are potential tracers of high mass star formation, I spatially cross matched the RMS and MMB catalogues in order to determine the strength of the relationship, if any exists, between these objects. My search was limited to sources located within $10^o \leq l \leq 50^o$, $300^o \leq l \leq 350^o$,$-2^o \leq b \leq 2^o$; a region which contains 746 MMB masers and 1253 RMS objects and is the region of overlap between RMS and MMB. A matching radius equal to the beam width of MSX (18 arcseconds) was used (Hoare *et al.* 2005). The cross match resulted in 167 detections with the class distribution as shown in Table 3.2.

From the cross matching of the MMB and RMS surveys it can be seen that there is an apparent limited overlap between the results of the two survey techniques, with only 22%

A Source	B Source	Scatter (deg)	Association (arcsec)	S_{Max}	Actual	$\bar{S}$
MMB	RMS	1	18	2	167	0
MMB	EGO	1	18	1	118	0
RMS	EGO	1	18	1	60	0
MMB	GRS	0.5	33	31	88	5
RMS	GRS	0.5	33	30	95	4
EGO	GRS	0.5	33	10	27	2
EGO	BGPS	1	33	8	44	1
RMS	BGPS	1	33	12	159	3
MMB	BGPS	1	33	14	136	3
MMB	Cluster	1	30	2	26	0
MMB	Cluster	1	60	5	48	1
RMS	Cluster	1	30	3	52	0
RMS	Cluster	1	60	6	68	1
MMB	IRDC	1	IRDC SMA	2	195	0
BGPS	GRS		33		858	
EGO	IRDC	1	IRDC SMA	0	164	0
RMS	IRDC	1	IRDC SMA	2	242	0
MMB	Roman-Duval	1^o and 5 kms^{-1}	*lbv* box	88	195	71

Table 3.1: Table showing the live cross match results, the maximum number of cross matches and the mean number of cross matches achieved in 100 Monte Carlo simulations. Col:1 The primary scattered source, the A dataset. Col:2 The secondary source or B dataset. Col:3 The level of scatter in degrees but always restrained to lie within the survey area. Col:4 The association radius. Col:5 The maximum number of cross matches achieved in 100 Monte Carlo simulations. Col:6 Number of cross matches from the live data. Col 7: The mean number of associations found by the Monte Carlo simulations. No Monte Carlo simulation of the BGPS GRS cross match was undertaken due to high computational cost.

Class	Cross Matches	No in Region	%
Diffuse HII region	4	97	4
Diffuse HII region?	0	12	0
Evolved star	0	116	0
Evolved star?	0	61	0
HII region†	57	319	18
HII region?†	1	93	1
HII/YSO†	22	99	22
HII/YSO?†	1	20	5
Null	0	1	0
OH/IR Star	0	99	0
OH/IR Star?	0	5	0
Other	0	21	0
Other ?	0	2	0
PN	0	51	0
PN?	1	35	3
Proto-PN?	0	1	0
UCHII?†	0	1	0
Unclassified	1	28	4
Young/Old star	1	1	100
Young/Old star?	0	39	0
YSO†	69	181	38
YSO?†	10	41	24
Total	**167**	**1253**	

Table 3.2: RMS MMB cross match results.

Table showing the results of the 18 arcsecond cross match between the RMS and MMB surveys. † *denotes object classes I expect to be maser associated. Col:1 Object type, Col:2 Number found in cross match, Col:3 Number in cross match area and Col:4 Percentage of objects in cross match area with positive cross match.*

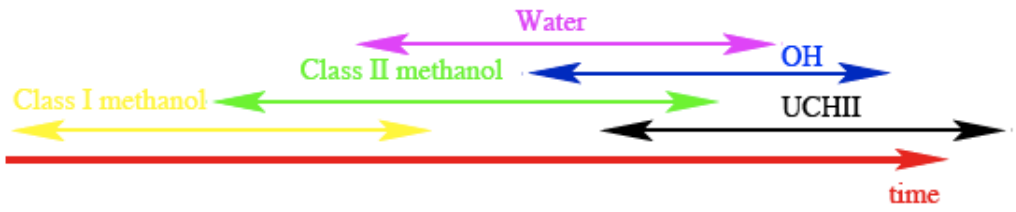


Figure 3.1: Strawman diagram
Strawman diagram showing the relationship of the statistical lifetimes of various star formation indicators (Ellingsen *et al.* 2007). Purple is Water Masers, green is Class II Methanol Masers, blue represents OH masers, black shows UCHII regions and yellow indicates class I methanol masers.

of MMB masers being associated with a RMS source and only 13% of RMS objects associated with a MMB maser. Reducing this to objects we can expect to be associated with star formation; objects classified within the RMS as HII regions, HII region?, HII/YSO, HII/YSO?, UCHII?, YSO and YSO?, we find that 21% of masers are associated with these objects whilst 21% of RMS objects within these groupings are associated with masers.

The Monte Carlo simulation between the MMB and the RMS results in $S_{max} = 2$ and $\bar{S}$=0 . This indicates that this association rate cannot be due to random alignments.

From the RMS catalogue [1] the number of YSO or YSO? classed RMS objects with luminosities compatible with massive stars (Lum $\geq 1000 L_{\odot}$) as identified by the RMS is of the order of 70%, hence it is unlikely that the lack of maser RMS-YSO association is due to misidentification of YSOs as MYSOs. Additionally, it is clear the objects classed as HII regions must contain massive young stars yet these objects are not strongly maser associated. Masing therefore appears to cease soon after the start of nuclear burning [2] and the subsequent formation of an UCHII region (Purcell 2006; Ellingsen *et al.* 2007). Hence the lack of association between RMS HII regions and masers is not entirely unexpected.

Figure 3.2 demonstrates that there are maser counterparts that are listed within the GPSC that are not within the RMS. A MMB/MSX catalogue cross match shows that the number of MMB masers with MSX association is 503, 72 more than the GPSC, although the GPSC is more sensitive to saturation and the search criteria is 18 arcseconds rather than 2 arcseconds. However, examination of the fluxes for the MSX sources associated with masers shows that the majority fail the MSX colour selection.

We see therefore that neither RMS surveys nor maser selected surveys are finding the entire high mass star formation population, rather they find two populations that have an overlap in the order of 20%. This is in agreement with the suggested lifetime overlap between Class II Methanol Masers and HII regions, see 3.1 (Ellingsen *et al.* 2007).

[1] `http://www.ast.leeds.ac.uk/cgi-bin/RMS/RMSDATABASE.cgi`
[2] Hence the formation of a MYSO as classified by the RMS.

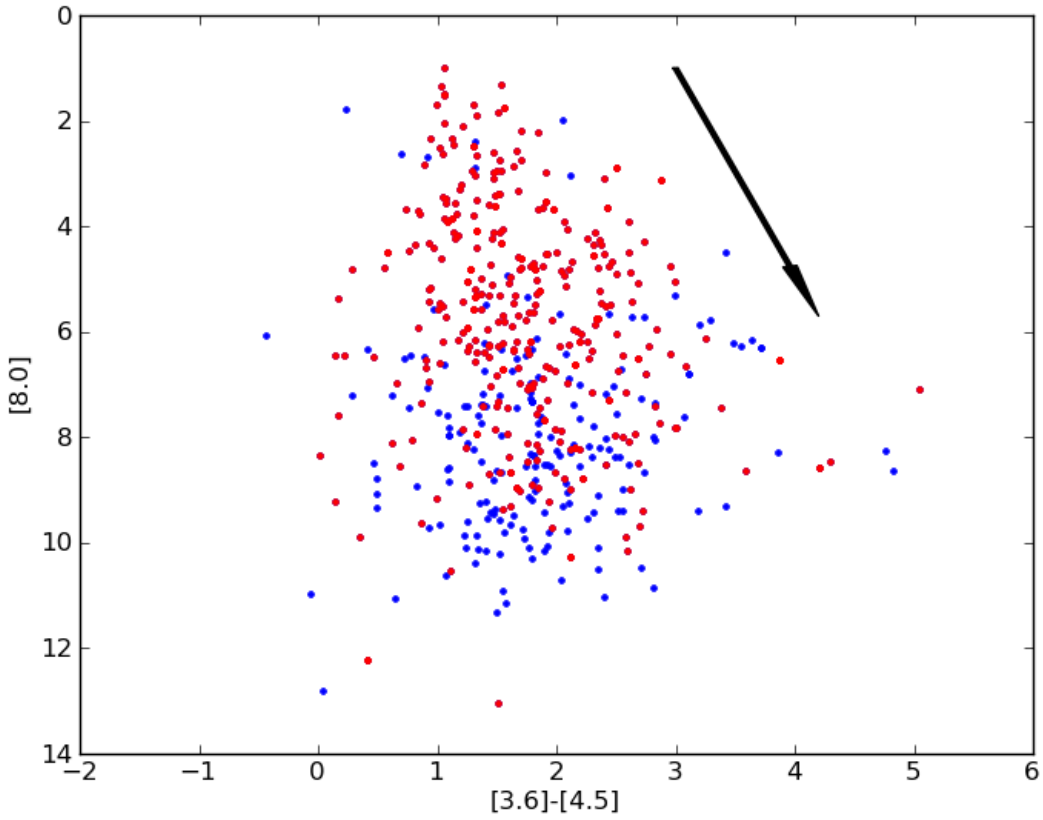


Figure 3.2: Colour - Magnitude diagram RMS v GPSC

Colour - Magnitude diagram comparing masers with counterparts within MSX (red) to the GLIMPSE masers (blue). As we can see there is no 8.0 μ *flux cut off within the MSX population that could suggest a sensitivity explanation behind the MMB objects missing from RMS. However, an examination of the MSX and GLIMPSE fluxes of the maser associated objects not with RMS shows that in fact most fail the MSX colour selection criteria. The* $A_k = 10$ *reddening vector is based on Gutermuth (2005) and the references therein.*

3.2 The MMB and Cyganowski et al EGO Catalogue

It has been suggested in Cyganowski *et al.* (2008a) that EGOs are strongly associated with Class II Methanol Masers and very strongly associated with Class I Methanol Masers (Chen *et al.* 2009). In this section I examine the strength of the relationship between EGOs and other massive star formation indicators.

When I undertook the visual inspection of the GLIMPSE field, the presence of 4.5 μm extended emission was noted for 23% of masers (see Chapter 2). However, both this search and the one produced by Cyganowski et al are arbitrary visual identification of emission in the GLIMPSE images. By performing a spatial cross match I have identified that 71 of the EGOs noted as associated with "likely MYSOs" were within 18 arcseconds of a MMB maser (53%) and 47 EGOs marked as "possible MYSOs" were within 18 arcseconds of a maser (28%). The 18 arcseconds matching radius was chosen for consistency with the RMS MMB cross match from section 3.1. The Monte Carlo simulation cross matching gives $S_{max} = 1$ and $\bar{S}$=0, showing that this association is non-random.

Cross matching the Cyganowski et al EGO catalogue and the RMS reveals that 17% of likely EGOs and 22% of possible EGOs are MYSO associated. Examination of the cross matched RMS objects suggests that the sample includes both MYSO and HII regions. This compares to the likely cross matches derived for the Monte Carlo simulations $S_{max} = 1$ and $\bar{S}$=0 which suggests this association is non random. The association with HII regions is somewhat difficult to explain given that we would expect outflows to have terminated by the time an HII region forms. This is due to the HII region destroying the disc and hence removing the launching mechanism.

3.3 MMB v Roman-Duval

As discussed in Chapter 1, the Roman-Duval Molecular Clouds catalogue (Roman-Duval *et al.* 2010) is a catalogue of molecular clouds whose physical characteristics are derived from the GRS (see Chapter 1) with distances from kinematics measurements and HI self absorption to resolve the near-far ambiguity.

As these objects are so extended any spatial cross match based on their central position would be meaningless. Instead I synthesised a box around each cloud using its listed line-width, semi-minor and semi-major axis as the dimension of the box l b v. I then searched for MMB masers that lay within the box. The region covered is $18^o < l < 50^o$ $|b| < 1^o$ which contains 752 Roman-Duval clouds and 296 MMB masers. I found that 195 masers within this region appear to be within a Roman-Duval cloud. We therefore can use this association to determine the distance to the 195 associated masers. Table 3.3 shows a sample of masers and their implied distance. This technique was previously used as a distance indicator for the RMS (Urquhart *et al.* 2011).
The number of associated masers is lower than I would expect as most, if not all masers, should be located within a molecular cloud. Although the association rate is significantly higher than random chance alignments as indicated by the Monte Carlo simulations, ($S_{max} = 88$ and $\bar{S}$=71), this discrepancy warrants further investigation.

In an attempt to understand why the association rate is so low I inspected the maser positions within the GRS images. The GRS datacubes were downloaded from the GRS website [3] as 2 by 2^o FITS datacubes with a velocity range of -5 to + 135 kms^{-1}. The datacubes were converted to NDF[4] format using the Starlink *CONVERT* package. They were then sliced into three cubes covering -5 to 40 kms^{-1} 40 to 85 kms^{-1} and 85 to 135 kms^{-1}. The cubes were then collapsed along the velocity axis to create two dimensional NDF files with the flux being integrated along the line of sight. The NDF files were then converted back to FITS format. These files were loaded into SAO DS9 and smoothed using a Gaussian with a kernel radius of three pixels selected to reduce clutter without overly reducing the resolution. The April 2010 version of the MMB catalogue was then overlaid and the images displayed as a linear staircase with blue being the lowest luminosity and red the highest.

A set of frames can be seen in Figure 3.3. Due to the low resolution of the GRS images (22 arcseconds (Jackson *et al.* 2006)), I consider maser sites rather than individual masers as we are unable to distinguish individual masers in all cases for the GRS images. I have identified 93 maser sites, of which 67 lie on an emission peak. I define a pixel as being in the peak emission if its pixel value is within 33% of the maximum pixel value within the image. An additional 24 masers lie on a cloud but not at the peak. Two masers appear not to be associated with any cloud.

We see therefore that Class II Methanol Masers are not as strongly associated with Roman-Duval clouds from a spatial cross match as they are from the inspection of the GRS images.

[3] `http://www.bu.edu/galacticring/`

[4] NDF is the standard file format for astronomical images used with the Starlink suite.

l	b	d (kpc)		l	b	d (kpc)
20.733	-0.06667	11.55		32.825	-0.3333	5.3
20.927	-0.05	2.45		33.4	0.0167	7.07
21.407	-0.25	5.9		33.633	-0.233	10.15
22.038	0.222	3.8		34.1	0.0167	3.78
22.335	-0.15	13.3		34.25	0.16	3.78
22.435	-0.1699	13.3		34.2667	0.15	3.78
23.365	-0.2833	5.22		34.4	0.216	3.78
23.437	-0.1833	6.65		35.39	0.02	6.93
23.9667	-0.1	4.78		37.0333	-0.0333	7.47
24.329	0.144	7.72		37.55	0.2	6.72
24.79	0.083	7.7		37.6	0.4333	6.72
24.9334	0.0667	3.33		40.2833	-0.2167	5.43
26.55	-0.3	8.02		41.122	-0.216	8.65
26.6	-0.0167	1.9		41.166	-0.183	8.65
27.2833	0.15	12.7		41.2333	-0.2	8.65
27.5667	0.1	5.25		41.87	-0.1	11.38
29.3158	-0.1744	3.22		42.4333	-0.2667	4.9
30.419	-0.2411	7.3		43.149	0.013	11.43
30.76	-0.058	6.22		43.165	0.013	11.43
32.0496	0.059	7.07		45.4667	0.1333	7.45
32.117	0.1	7.07		45.467	0.053	7.45
46.1167	0.3833	4.25		45.493	0.126	7.45
49.599	-0.25	5.53		45.8	-0.35	6.72

Table 3.3: MMB Roman-Duval cross match results.
Table showing the distances to 46 MMB masers derived from their association with Roman-Duval Clouds.

Inspection of the Roman-Duval catalogue shows that this is not due to the clouds being fragmented by the Roman-Duval *et al.* (2010) cloud identification method, rather it would appear that the detection threshold used to identify the cloud boundaries are set too high and consequently regions of star formation associated with the cloud are not being included. Consequently when spatial comparisons are made between the Roman-Duval cloud catalogue and sources that are likely to be associated with molecular clouds, such as Methanol Masers, there is a possibility of a false non-association, this being due to the Roman-Duval catalogue under−reporting the semi-major and semi-minor axis sizes.

3.4 Galactic Ring Survey

The Boston University and Five College Radio Astronomy Observatory Galactic Ring Survey (Jackson *et al.* 2006) contains 6187 identified "clumps" over $20^o < l < 50^o$ $|b| < 1^o$ (see Chapter 1 for the difference between GRS Clouds and Clumps), a region that contains

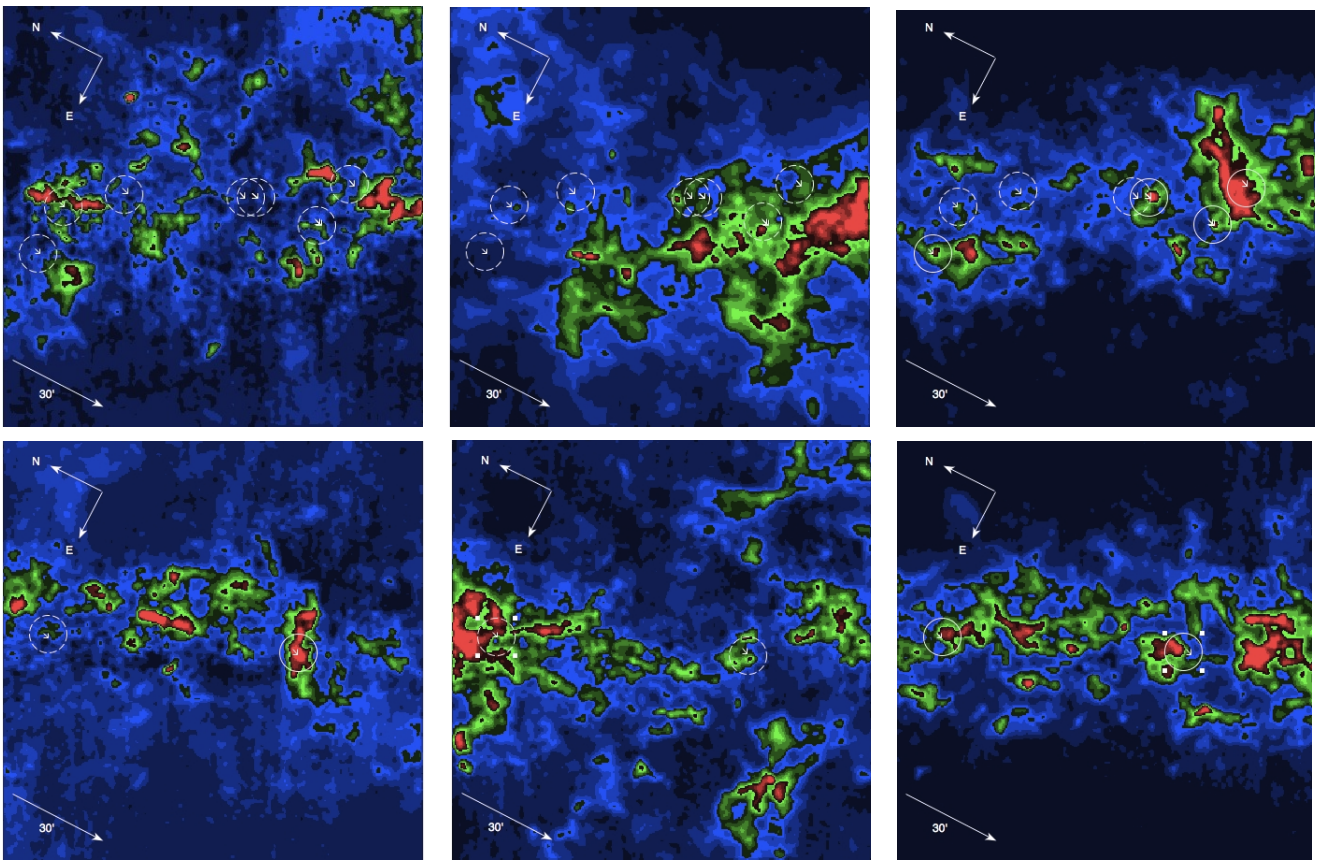


Figure 3.3: Sample set of GRS image, collapsed in the velocity dimension, covering 24^o to 28^o.
The left image covers -5 to 40 kms^{-1} the middle 40 to 85 kms^{-1} and the bottom 85 to 135 kms^{-1}. The MMB masers' positions are marked with white arrowheads and highlighted with white circles. Unbroken circles indicate masers associated with a cloud in that velocity whilst broken circles indicate unassociated masers. The colour scheme is red for peak emission moving to dark blue for no emission. The white compass is shown for orientation and the white arrow is the 30 arcminute scale.

255 MMB masers. I undertook a cross match between these two surveys using an associated distance of 33 arcseconds, for consistency with the BGPS. This resulted in 80 (31%) masers being associated with a GRS source this is significantly higher than the $S_{max} = 31$ and $\bar{S}$=5, as derived from the Monte Carlo simulations.

The same field contains 265 RMS objects of which 95 (36%) are within 33 arcseconds of a GRS clump, of which 12 are also maser associated. This compares to the $S_{max} = 30$ and $\bar{S}$=5, of the Monte Carlo simulation cross matches which indicates that this association is non-random.

The field also contains 54 EGOs of which 27 (50%) are within 33 arcseconds of a GRS

object, considerably higher than $S_{max} = 10$ and $\bar{S}$=2, suggested by the Monte Carlo simulations. We see therefore that GRS clumps are more strongly associated with EGOs than either RMS objects or masers. It is clear from the above results that the GRS is more strongly associated with early tracers, such as EGOs, than later tracers, such as RMS objects and Class II Methanol Masers.

3.5 BOLOCAM Cross Match

The BGPS survey catalogue (Aguirre *et al.* 2011) contains 3734 sources within $20^o < l < 50^o$ $|b| < 1^o$. Cross matching the MMB with the BGPS with a 33 arcsecond association radius (the approximate CSO beam width) we find 136 out of 255 (53%) masers are associated as compared to $S_{max} = 14$ and $\bar{S}$=4 from the Monte Carlo simulations. Fifty-nine of these sources are also GRS associated.

The RMS / BGPS cross match results in 159 (60%) RMS objects being located within 33 arcseconds of a BGPS object as compared to $S_{max} = 12$ and $\bar{S}$=3 from the Monte Carlo simulations.

The EGO BGPS cross match identifies 44 (77%) EGOs as being associated with BGPS objects; this compares to $S_{max} = 8$ and $\bar{S}$=1 from the Monte Carlo simulations. Hence, the BGPS is much more strongly associated with EGOs than the GRS despite having half the number density, although we should exercise caution as the GRS is much deeper, so the lower level of association may be due to EGO being unresolved or removed from the Cyganowski et al survey by extinction.

The BGPS and GRS cross match results are interesting with only 858 (23%) BGPS objects within 33 arcseconds of GRS objects. It therefore appears that the BGPS and the GRS are tracing different, but overlapping, populations. It could be argued that the difference between the surveys is due to depth, the GRS being a deeper survey. However, if this was the case we would expect the MMB to be more strongly associated with the GRS than the BGPS, which it is not. We are therefore left with several possible solutions. Firstly, the BGPS could be tracing later evolutionary objects than the GRS, or secondly, the BGPS being shallower and less sensitive, is preferentially selecting more massive objects which are more likely to be associated with EGOS, Class II Methanol Masers and RMS objects. Alternatively the proportion of GRS objects that are quiescent could be significantly higher than the BGPS.

3.6 Infrared Cluster Cross Match

As discussed in Chapter 1, my cluster catalogue consists of the merged Mercer (Mercer *et al.* 2005), Bica (Bica *et al.* 2003) , Froebirch (Froebrich *et al.* 2007) and UKIDSS (Lucas et al 2011 in prep) cluster catalogues. Over $10^o \leq l \leq 50^o, 300^o \leq l \leq 350^o - 2^o \leq b \leq 2^o$ used for the MMB RMS cross match (see section 3.2) there are 328 clusters listed, which reduces to 301 when possible duplications are removed. Cross matching the MMB and the cluster list with a radius of 60 arcseconds, the maximum size I expect a cluster to

present, yields 48 cross matches (6%). The RMS/Cluster cross match results in 85 (7%) associated RMS sources.

Distances determined by Lucas et al.(2011) for the majority of the UKIDSS GPS clusters have a mean of several kpc so it is unlikely that this very weak maser-cluster association can be explained by non-detection of the massive clusters in which we expect massive stars to form. A correction for incompleteness would be expected to raise the association rate somewhat but not by a factor of 10.

Hence, we see that infrared clusters are not strongly associated with either RMS objects or Class II Methanol Masers. This suggests that massive star formation ends before, or soon after the cluster becomes distinguishable as such. The nature of the small number of masers that are cluster associated is examined in more depth in Chapter 5.

3.7 IRDC Cross Match

A normal cross match with the P&F IRDC catalogue is meaningless as they are extended objects. Instead I use the method as outlined in the Chapter introduction whereby I use the IRDC's semi-major axis as the cross matching radius. For the MMB masers this yields 195 cross matches compared to $S_{max} = 25$ and $\bar{S}$=8 from the Monte Carlo simulations, confirming the observation in Chapter 2 that a significant population of masers are associated with IRDCs. The cross matching shows 195 MMB masers associated with IRDCs which is considerably higher than the $S_{max} = 25$ and $\bar{S}$=1, from the Monte Carlo simulations. This represents 25% of the masers within the field and is consistent with the visual inspection in Chapter 2.

The RMS cross match yields 242 associations or 28% of the RMS objects within the field. When compared to $S_{max} = 30$ and $\bar{S}$=15 from the Monte Carlo simulations we see that these are likely to be real associations.

3.8 Discussion

We can break the results of the cross match into two groups; the lifetimes of the indicators when compared to the Strawman lifetimes of Ellingsen *et al.* (2007) and the conflict between the GRS and BGPS results.

The MMB/RMS cross match result of 21% is consistent with the Ellingsen *et al.* (2007) lifetime of UCHII regions and Class II Methanol Masers. Hence the appearance of the MYSO region marks the beginning of the end of masing. The observation that approximately 50% of EGOs are Class II Methanol Maser associated is broadly in line with the lifetimes of Class I Methanol Masers, which are very strongly EGO associated (Chen *et al.* 2009). The observation that approximately 20% of EGOs are RMS associated, with an approximately evenly split between YSO and HII regions, is somewhat harder to explain

and is a departure from the Ellingsen Strawman sequence. There are a number of possible explanations. Accretion, and hence outflow, could continue after the nuclear burning starts if the inflow of material is enough to quench or at least hinder the formation of the HII region (Keto 2003). The RMS associated EGOs could be fossil outflows which are still excited enough to be dominant at 4.5 μm or there could be an alternative mechanism for producing EGOs that is linked to HII regions as opposed to outflows. The current spectographic observations of EGOs are currently too few and too inconclusive to determine the relationship between EGO and HII regions.

The spatial cross match between the MMB and the Roman-Duval molecular cloud catalogue (Roman-Duval *et al.* 2010) is surprising given that only 65% of MMB masers appear to be located within Roman-Duval molecular clouds. However, the visual inspection of the GRS images suggest that MMB masers *are* associated with molecular clouds. Inspecting the Roman-Duval catalogue shows that we are not seeing large clouds being fractured, rather it would appear that the processes that determine the edge of the cloud may be too aggressive and are missing regions that are star forming and hence dense.

The infrared cluster cross match shows that the link between clusters and masers is weak (as is the link with RMS objects) which suggests that massive star formation is usually over once the cluster becomes distinguishable in the near infrared. The link between masers and clusters is investigated in detail in Chapter 5.

The IRDC cross match indicates that the IRDC maser association is consistent with the visual inspection undertaken in Chapter 2. The high association rate between EGOs and IRDCs is a further indicator that EGOs are early evolution indicators. The lack of association between both the GRS and BGPS with the IRDC P&F catalogue may be due to IRDCs being blended out by foreground emission or the misidentification of holes in the 8.0 μm background as IRDCs. However, we should not reject the possibility that a sizeable fraction of GRS and BGPS objects are not embedded within IRDCs and that the "Russian Doll" model of star formation, as discussed in Chapter 1, is not correct.

Turning to the GRS and BGRS surveys. It is clear that the association rate between the two surveys is unexpectedly low if they are tracing the same population of objects. The BGPS is much more strongly associated with star formation indicators such as masers and EGOs, suggesting that the BGPS is tracing more star forming regions than the GRS, although there is clearly an overlap. It is possible that the lower sensitivity of the BGPS and lack of depth, when compared to the GRS, means that it is tracing more massive objects that are more likely to be forming massive stars and therefore more likely to be associated with massive star formation tracers. However, it is also possible that the BGPS is contaminated by free-free emission. Clearly this problem needs further investigation, ideally by cross matching the two surveys with the MiPSGAL catalogue once it becomes publicly available to see if the Clumps/Cores are warm.

3.9 Summary

In this chapter I have shown that the k=1 Nearest Neighbour function between the RMS, EGOs and the MMB is consistent with the lifetimes of Class II Methanol Masers and Class I Methanol Masers lifetimes as discussed in Ellingsen *et al.* (2007). We see that the association between Class II Methanol Maser and infrared cluster is low (as are RMS objects) indicating that high mass star formation has ceased once a cluster becomes distinguishable in the near infrared. I suggest that the association rate between Class II Methanol Masers and Roman-Duval molecular clouds is most likely due to the edge detection process used to determine the location of the clouds being over aggressive and not including some dense regions that are forming stars.

I have shown from the GRS BGPS cross match and association with other star formation tracers that there are unexpected differences between the GRS and BGPS given that we would expect them to tracing broadly the same objects. This indicates that the small number of objects in the BGPS do *not* represent a subset of the GRS objects. The observation that the BGPS is a stronger tracer of massive star formation than the GRS suggests that it is tracing more massive objects than the GRS. However, we should not rule out the possibility that the BGPS is contaminated by free-free emission. Further observations are required to resolve this problem.

In the following Chapter I expand on the cross matching of the BGPS the GRS and the P&F IRDC catalogue using more advanced methods drawn from graph theory and closely related to the kNN method used above.

Chapter 4

Star Formation at Large Scales

"Anyone who attempts to generate random numbers by deterministic means is, of course, living in a state of sin." - John von Neumann

4.1 Introduction

In Chapter 1 I introduced a large number of surveys that are believed to trace environments or astrophysical processes that are indicative of star formation and in some cases of specifically high mass star formation. In Chapter 3 I discussed these surveys further and looked at their associations with the MMB and to some extent with themselves. In general these associations are at fairly local levels, perhaps a single cluster or in some cases perhaps a segment of a Giant Molecular Cloud. I now ask the questions: "How do these various indicators relate to each other on a larger scale and what implications does this have in their application as high mass star formation tracers?".

In this chapter I will attempt to answer some of these questions. In order to do so I will introduce the concept of two and three dimensional minimum spanning trees (MST). Traditional dwo dimensional Minimum Spanning Trees have been used in astrophysics for a number of decades for the detection of clustering; (Cartwright and Whitworth 2004; Cartwright 2009; Allison *et al.* 2009; Bastian *et al.* 2009; Parker *et al.* 2011) and of extragalactic structure (Doroshkevich *et al.* 2004) with a brief study with three dimensional MST being undertaken in Cartwright (2009) and Barrow *et al.* (1985). However, the use of Minimum Spanning Trees is widely used to determine both clustering and structure within any parameter space, be that parameter space consisting of physical dimensions, units or non-physical values. For example, MST have been used in identifying disorders in biological membranes (Dussert *et al.* 1987), analysis of random resistor networks (Read 2005) and the identification of interconnections between financial markets (Coelho *et al.* 2007).

Minimum Spanning Trees are part of graph theory, the origins of which can be traced to the Seven Bridges of Königsberg problem. At the time of the inception of the problem

Königsberg (now Kaliningrad) had seven bridges crossing the Pregolya River. It was conjectured whether it was possible to undertake a walk where every bridge was crossed but each bridge was crossed only once (Bondy and Murty (2008) Section 3.3). The problem is illustrated in Figure 4.1.

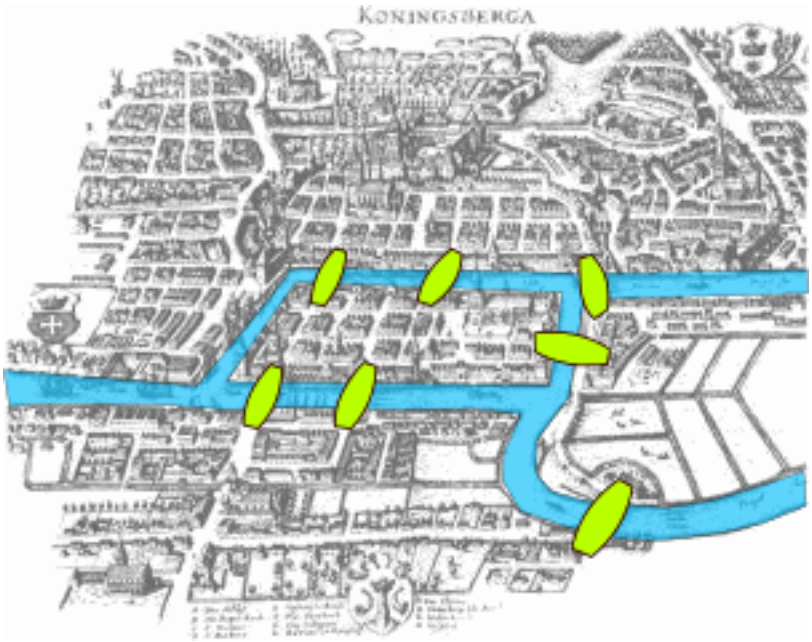


Figure 4.1: The Königsberg bridges
Layout of the Königsberg bridges showing the physical layout of the bridges around the time the Königsberg bridge problem was addressed by Euler. Source Wikipedia

The problem was eventually solved by Leonhard Euler (Bondy and Murty (2008) Section 3.3). Euler's approach was to consider the bridges as a series of points or *nodes*[1] connected by lines or *edges* and the number of edges attached to a node being the node's *degree*. This information was used to construct an image that was not constrained with being geographically consistent (Bondy and Murty (2008) Section 3.3). . This construct became what is now known as a *graph* (Bondy and Murty (2008) Section 3.3). Figure 4.2 shows the graph for the Königsberg bridge problem. In graph theory a graph is designated as G, the number of nodes within G is, n, with the number of edges being m; hence we can note the scale of a graph as $G(n, m)$. The degree of a node within G is denoted as $D_G(n)$ (Bondy and Murty (2008) Chapter 3).

Euler showed that for the walk to be successful the number of nodes in the graph with an odd degree must be either zero or two. The Königsberg graph contains three nodes

[1]The term *vertex* is also used in some texts.

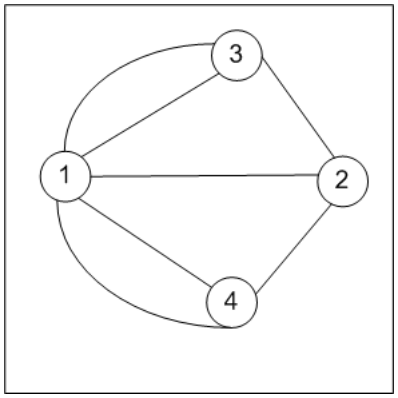


Figure 4.2: The Königsberg bridges graph
Graph representing the Königsberg bridge problem. It shows the non-geographical approach taken by Euler in solving the Königsberg bridge problem. Here each bridge is represented by an edge, with the intersections between bridges being nodes.

of three degrees and one node of five degrees and hence the Seven Bridges of Königsberg problem can be shown to have no solution (Bondy and Murty (2008) Section 3.3).

A related challenge is the "Travelling Salesman Problem" (TSP) (Bondy and Murty (2008) Chapter 2.) first put forward by Hamilton in the 19th century. The problem concerns a peripatetic worker travelling from client to client and asks the question: what is the most time efficient route in which each client is visited only once? This problem is closely associated with the Königsberg problem and can be tackled as an *undirected weighed graph* In a weighed graph each edge is assigned a value that reflects the *cost* of using that edge, in the case of the TSP the edge cost reflects the length of time that route would take. The graph is undirected when the edges have no direction and consequently both $A \longrightarrow B$ and $B \longrightarrow A$ are equally valid paths (Bondy and Murty (2008) Chapter 2).

Any path within a graph that connects every node but visits every node just once is known as a *Hamiltonian Cycle* and the solution to the Travelling Salesman Problem is the Hamiltonian Cycle with the lowest cost, where the cost is the total weighed path length for that edge. The efficient determination of the lowest cost Hamiltonian Cycle in any but the most simple graphs is extremely challenging and it is only recently using the application of low cost computing that this problem can now be easily solved (Bondy and Murty (2008) Chapter 11).

I now turn to the Minimum Spanning Tree Problem (MSTP), which is related to both the TSP and the Königsberg problem. The problem first arose when dealing with the efficient distribution of electrical power between cities (Graham and Hell 1985). For the power distribution company the solution needs to be the one of lowest cost. Treating each city as a node and each power line as an edge with each edge being weighted to reflect the cost of construction, power lost and maintenance for that route, we can then solve the problem of

finding the cheapest route by considering an undirected weighed graph $G(n, m)$ in which every city is connected once and once only and where there are no loops. Such a graph is known as a *Tree* and as such is denoted as T. The tree with the lowest cost is known as the "Minimum Spanning Tree" and is denoted as S. Although a graph may have many trees, if the edge costs are unique it will only contain one minimum spanning tree.

An important aspect of the MST is that if it is fragmented then the fragments are themselves minimum spanning trees. This is obvious given that the MST of a graph must be unique and hence any pathway between two nodes must also be unique and also be at the minimum cost.

The solution to the minimum spanning tree Problem can be attributed to a number of independent authors in different fields, working on the problem unaware of the activities of others (Graham and Hell 1985). Perhaps the best known of these solutions is Prim's algorithm (Prim 1957), which appears to have been independently discovered several times including by J. B Kruskal in 1956 (Graham and Hell 1985). For the work contained in this thesis I have chosen to implement Prim's algorithm, in the knowledge that faster, more efficient algorithms are available. However, the coding requirements associated with these more efficient algorithms are considerably higher than those associated with Prim's algorithm. Given that the number of nodes in each tree constructed as part of this thesis is small (less than 2000) and that fast, clustered, multi-processor computer time needed to solve large numbers of small minimum spanning trees is freely available the lower efficiency of Prim's algorithm does not increase the time required for this research.

The solution to an minimum spanning tree, as I will demonstrate in the following sections, allows identification of small scale (clustering) and large scale structure and has applications in financial market analysis (Gorski *et al.* 2008), wireless network design (Morgan and Grout 2006), disease outbreak control (Spada *et al.* 2004) and analysis of biological structure (Dussert *et al.* 1987). The process works in both spatial dimensions and in parameter space (Cartwright 2009).

It should be noted that the data within a minimum spanning tree is considered to be dimensionless and that the length (or cost) of any edge is the euclidean distance between the data points.

4.2 Minimum Spanning Trees in Astrophysics

The origins of the implication of minimum spanning trees to astronomy can be traced to the computer science paper, Zahn (1971) , in which it is shown that the MST process is an efficient cluster detection algorithm (Zahn 1971) and later work (Schmeja and Klessen 2006; Bastian *et al.* 2009; Cartwright and Whitworth 2004; Gutermuth *et al.* 2009). Of particular interest is section IX where Zahn discusses finding a particle track description. This is analogous to the use of MST for the identification of large scale structure, an implementation that is confirmed by the work of Barrow *et al.* (1985); Krzewina and Saslaw (1995); Doroshkevich *et al.* (2004); Schmeja (2011) and Plionis *et al.* (1992). The construction of an MST is related to two other commonly used structure detection methods,

Nearest Neighbour and Voronoi Tessellations.

The Nearest Neighbour (NN) method estimates the local source density ρ_k by measuring the distance to the k^{th} nearest object (Schmeja 2011; Casertano 1985). As in the construction of an MST, the NN method can be applied via graph theory. When $k = 1$ the graph of the NN function is a subset of the MST graph. Consider the graph $G(n, m)$ and a subtree of G $K(3, b < m)$. The NN graph and the MST of K will connect the nodes with the two shortest edges and hence, applying the NN function to K also generates the MST of K. As stated in section 4.1 any sub-tree of an MST must itself be an MST. Hence the three node NN tree K must be a subset of the MST for G. Note that because the NN function does not generate a spanning tree it cannot be used to build an MST.

An alternative graph theory cluster detection method is the construction of a Voronoi Tessellation. A Voronoi Tessellation is a graph comprising of a series of polygons each enclosing a nucleus, with the polygon representing the area *dominated* by the nucleus. Hence, if we draw a straight line x between two adjacent nuclei A and B then x will be intersected at its mid-point by the polygons enclosing both A and B (Schmeja 2011). Hence, over densities appear as a concentration of small cells (Schmeja 2011). Voronoi Tessellations are related to Delaunay triangulations, of which nearest neighbour graphs are a subset, hence we see that minimum spanning trees, nearest neighbour graphs and Voronoi Tessellations are all related processes.

In Zahn (1971) it is also noted that MSTs are deterministic and hence are independent of initial choices (although as we will see that does not hold for all variants of the MST as discussed later). It was also noted that MSTs are invariant under similarity transformations [2] from which we can conclude that Spatial Astrophysical MSTs are independent of the coordinate system used. Zahn (1971) also concludes that MSTs are relatively insensitive to small amounts of noise, although I would dispute this given the results in Schmeja and Klessen (2006) in which it was shown that the nearest neighbour method is more successful at detecting clusters projected onto non-uniform backgrounds.

The work outlined in Zahn (1971) is expanded in Barrow *et al.* (1985) with the introduction of pruning or k-branching was introduced, which is effectively a re-implementation of the hair removal used in (Zahn 1971). This mechanism involves removing all degree one nodes that connect to a node with a degree greater than two and results in an improved visibility of the central trunk of the MST. In addition to k-branching (Barrow *et al.* 1985) introduced the concept of fracturing.

In Barrow *et al.* (1985) the author compares the MST results of Zwicky and the CfA[3] Galaxy surveys with a model data generated as a Poisson distribution. Barrow *et al.* (1985) uses separating (Fracturing in graph theory) in an attempt to increase the structure signal over the noise. The level at which the fracturing occurs is critical to cluster detection.

[2] A similarity transformation is one that preserves angles and changes all distances in the same ratio.
[3] The Center for Astrophysics (CfA) Redshift Survey.

Typically the tree is fractured at the mean branch cost, multiples thereof or at the mean edge cost plus or minus a standard derivation (Barrow *et al.* 1985; Zahn 1971; Plionis *et al.* 1992; Schmeja and Klessen 2006). At this fracturing level the MST process is extremely adroit at discovering a cluster within a field containing a single cluster (Schmeja 2011). However, when there is more than one cluster within the field the mean branch length is driven down and consequently a fracturing level based on the mean edge cost may not be suitable for multi-cluster fields.

In Gutermuth *et al.* (2009); Bastian *et al.* (2007) and Billot *et al.* (2011) an alternative method is used to determine the fracturing level. In broad terms it is based on the assumption that there are two populations of objects within the field, clustered objects and field objects and that the mean edge costs between cluster members is lower than the mean edge costs between field objects. In Gutermuth *et al.* (2009) it is stated that this discongruity manifests as a kink in the edge cost cumulative distribution function (CDF) (see Figure 4.3). The fracture value can then be determined by fitting two straight lines to the CDF, one for small costs and one for the high costs. The fracture value being taken from the interception of the two fitted lines.

Figure 4.4 illustrates this procedure and compares it to fracturing at both the CDF determined level and at the mean for fields with one and two clusters. As can be seen for a field with a single cluster both systems perform approximately the same, however in the two cluster fields the CDF method underestimates the cluster membership, whilst fracturing at the mean overestimates the membership. Mean fracturing has the advantage that the mean edge cost can always be identified whilst, as discovered in Gutermuth *et al.* (2009), it is not always possible to identify the kink in the CDF.

An alternative method was utilised in Maschberger *et al.* (2010) whereby the fracturing level is adjusted so that known, visually identified clusters become distinguished. However, as noted in Barrow *et al.* (1985) visual identification of clusters is subject to psychological effects and it was for this very reason that the MST process was implemented for cluster detection. Hence ,any method that relies on the visual identification for the determination of the fracturing level should be treated with caution.

The process of fracturing creates a number of sub-trees, each a minimum spanning tree themselves. However, it should be noted that the number and structure of the sub-trees is dependent on the fracturing level, hence the sub-tree structure is not determinate unlike the MST. A process subsequent to fracturing was introduced in Plionis *et al.* (1992). Orphan Removal involves deleting all sub-trees with less then i nodes. This process removes clutter and highlights the main structure of the fragmented tree. The Orphan Removal process is expanded in section 4.4.

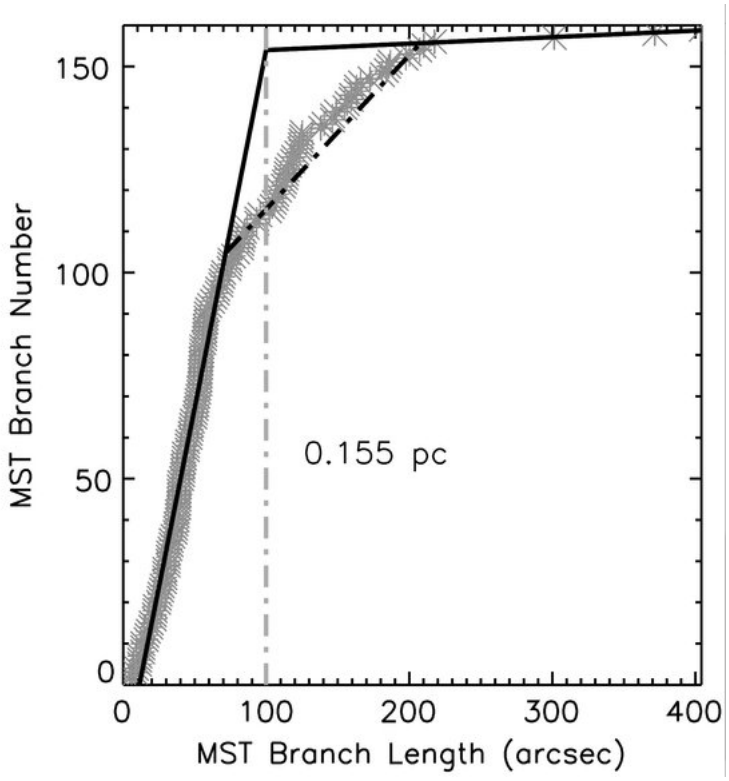


Figure 4.3: CDF Diagram

CDF diagram of the cumulative edge cost distribution for a cluster. This demonstrates the process involved in determining the fracturing level using the MST CDF. The process assumes that there is an inflection in the CDF caused by the difference in edge cost between the clustered and unclustered populations. The intersection of the two lines fitted to the region before and after the inflection point is used for the fracturing scale. However, as can be seen, it is difficult to determine where this inflection point may be and it is possible that a CDF has multiple inflection points. Hence, this process may not always be suitable. Image courtesy of Gutermuth *et al.* (2009)

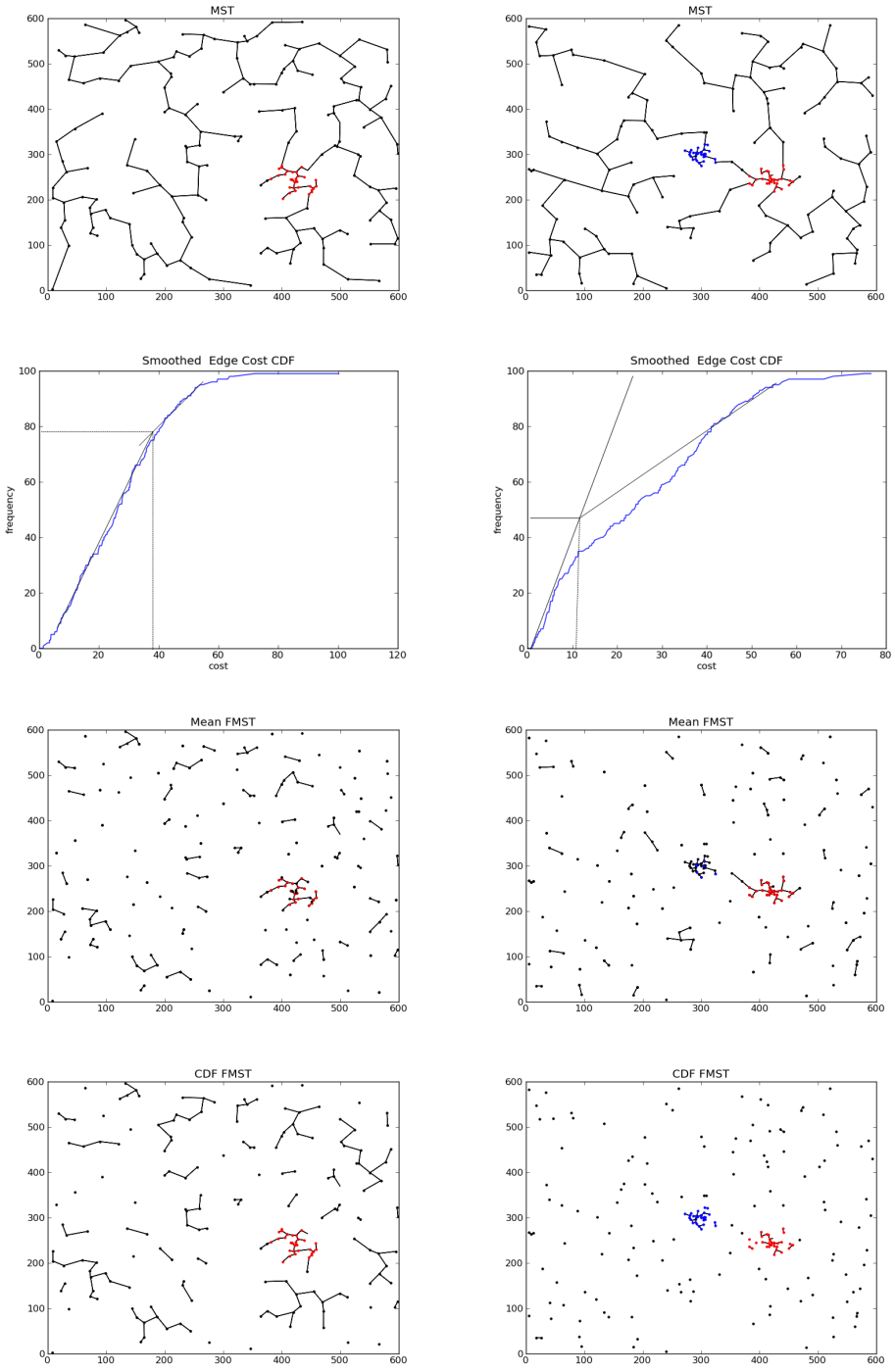


Figure 4.4: MST Fracturing

Plots showing the two fracturing processes, breaking at the mean and at the CDF inflection point. The top row shows two MSTs, one with one cluster, the other with two. The second row shows the CDF for each of the trees and the break point as determined from the CDF. The third row shows the two MSTs fractured at the CDF inflection point and shows that this fracturing technique tends to break the tree at too high a scale. The forth row shows the MSTs broken at the mean edge cost and shows that this technique tends to underestimate the fracturing scale in this case.

4.3 Quantifying MST

As mentioned in Barrow *et al.* (1985) humans are very unreliable detectors of patterns hence the need to extract meaningful statistical measurement from the trees produced from our datasets. As part of the post MST procedure a number of statistical measurements can be made which allows us to determine the level of clustering as well as the degree of large scale structure. Table 4.1 summarises this section.

The frequency distribution of edge cost over mean edge cost $l/\langle l \rangle$ [4] is introduced in Barrow *et al.* (1985). It is seen that in $l/\langle l \rangle$ for the live data there appears to be a relatively large number of small edges compared to large edges, whilst the Poisson data shows a Gaussian distribution centred on the mean edge cost. It is this discrepancy in distributions that allows Gutermuth *et al.* (2009) to determine a fracturing level from the Edge Cost CDF.

The mean edge cost, $\bar{m}$, is an indication of the level of clustering (Krzewina and Saslaw 1995) and is dependent on the density of nodes (Cartwright and Whitworth 2004). This being the case, to compare two fields a normalisation function needs to be applied. For a field of N nodes, with an area of A, the normalisation function of $(A/N)^{(1/2)}$ (Schmeja and Klessen 2006) is applied and for 3d fields with a volume V $(V/N)^{(1/3)}$ (Schmeja and Klessen 2006). Note that this is different to that applied in Cartwright and Whitworth (2004) wherein Cartwright et al appears to have mistakenly applied the normalisation function for the TSP problem (Schmeja and Klessen 2006); this error is corrected in Cartwright (2009).

In Cartwright and Whitworth (2004) the measurement Q was introduced. Q is the normalised mean edge cost of the MST $\langle l_{norm} \rangle$ over the mean internal-cluster edge cost $\langle L_{sub} \rangle$ which is not normalised. In Cartwright and Whitworth (2004) and Cartwright (2009) it was shown that Q is an excellent tool for identifying the degree of radial clustering, although in Cartwright (2009) it was suggested that Q may not be so indicative in two dimensional and velocity (2dv) spaces as in 2d space. However, I suggest this may be due to the random nature of the simulated velocity data used as this random element effectively introduces noise into the system, which as we see from section 4.2 reduces MST efficiency. In Bastian *et al.* (2007) it was shown that Q is affected by the elongation of the cluster. Bastian *et al.* (2007) suggests a correction factor of $Q_{intrinsic} = Q_{obs} - (-0.208 \times elongation)$, although this was observed to increase uncertainties for systems with high elongations. However, Bastian *et al.* (2007) also noted that Q was largely unaffected by extinction, which makes it suitable to be used in star formation regions where extinction levels can be high. Additionally, Q is also affected by contamination from foreground and background contamination (Bastian *et al.* 2007). This should come as no surprise, as was noted in section 4.2, minimum spanning trees perform less well in noisy fields. Notwithstanding, Q has been demonstrated as being a powerful discriminator of clustering in the distribution of gas in molecular clouds (Lomax *et al.* 2011), the analysis of cluster formation simulations (Maschberger *et al.* 2010), the detection of large scale Galaxy structure (Barrow *et al.* 1985; Doroshkevich *et al.* 2004; Plionis *et al.* 1992), the analysis of mass segregation in the Orion Nebula Cluster (Allison *et al.* 2009), investigations into the spatial evolution

[4] This would be m / $\bar{m}$ in my notation.

of clusters within the Large Magellanic Cloud (Bastian *et al.* 2009), detection of young clusters (Gutermuth *et al.* 2009) and the clustering of molecular clumps (Billot *et al.* 2011).

In Plionis *et al.* (1992) the statistical measure R_L was introduced. This being the ratio of the start-to-end cost (ie the length of the semi-major axis) to the total cost along an edge or sub-tree (the combined cost along the tree for the start and end points of the semi major axis). Plionis et al showed that if $R_L \simeq 1$ the sub-tree will be linear, if $R_L \ll 1$ the sub-tree will be highly curved. Hence, R_L is a measure of the linearity of a sub-tree. This process is also noted in Krzewina and Saslaw (1995) as a X/D test.

In Krzewina and Saslaw (1995) a number of new statistical tests are introduced. The edge angle test, $D\triangle$, measures the distribution of angles between intersecting edges within an MST. This is a measure of linearity within an MST structure or substructure as linear objects will have a high number of high value angles compared to non-linear objects. For a Poisson distribution the histogram of edge-edge angles should be flat (Krzewina and Saslaw 1995) except between 60^o and 90^o. It should be noted that if the structure is highly branched, and hence has a large number of nodes of degree greater than two, it will be detectable using this method.

As an additional test (Krzewina and Saslaw 1995) plots the number of sub-trees containing g nodes, $F(g)$, against $g \ / \ \bar{g}$ where $\bar{g}$ is the average number of nodes per sub-tree. Krzewina and Saslaw (1995) suggests that for a Poisson distribution the frequency curve should decrease with increasing G. Hence, this plot should indicate some measure of the field's tendency to cluster either linearly on large scales or non linearly at small scales. The Λ_i function allows the incorporation of velocity data into the MST data. However, as we have noted in section 4.1 the MST process can be expanded into parameter space to include measurements such as velocity and consequently Λ_i function is redundant.

An additional measure of elongation is made by Schmeja and Klessen (2006); ξ, which is the ratio of the radius of a convex hull fitted to a sub-tree, as derived from its area and the radius of the circle enclosing the convex hull, with $\xi = 1$ being spherical and $\xi = 3$ corresponding to an axis ratio of approximately 10 (Schmeja and Klessen 2006). A similar method is used in Billot *et al.* (2011) where the square of the radius of the convex hull R^2_{circ} is compared to the radius squared of the circle with the same area as the convex hull R^2_{hull}.

It is noted in Schmeja and Klessen (2006) that MST are computationally complex when working in large areas or areas with high numbers of object counts, when compared to alternatives such as the Nearest Neighbour or Voronoi Tessellations algorithms. It is also noted that MST require a priori assumptions (the fracturing level) and that ideally an adaptive method for determining the fracturing level should be used. However, as I will demonstrate, these problems are not insurmountable and show that the MST process produces a host of data concerning both the clustering and large scale structure of the field being examined.

Parameter	Notes
Mean edge cost $\bar{\text{m}}$	The mean edge cost within the entire MST.
Mean subtree edge cost $\bar{\text{s}}$	The mean edge cost within a subtree. This is a measure of clustering.
Edge cost over mean edge cost $l/\langle l \rangle$	The frequency distribution of the edge cost over mean edge cost. A measure of clustering.
Q	The normalised mean edge cost of the MST $\langle l_{norm} \rangle$ and over the mean internal-cluster edge cost. A clustering measurement.
R_L	The ratio of the start-to-end to the total cost along an edge or sub-tree. A measure of linearity.
X/D	See R_L above.
Edge to edge angle test, $D\triangle$	The distribution of angles between intersecting edges within an MST. This is a measure of linearity.
Node count distribution g $F(g)$	The sub-tree node count distribution. Tests linearity.
ξ - Schmeja elongation factor	Radius of a convex hull fitted to a sub-tree and the radius of the circle enclosing the convex hull. This is a measure of linearity.
$Q_{intrinsic}$ or Q^*	$Q_{obs} - (-0.208 \times elongation)$, adjusted Q value allowing for elongation.
Billot elongation factor - BE	Alternative to ξ which allows for the slight increase in hull size due to the requirement that the boundary nodes be contained within the hull.

Table 4.1: Overview of MST statistical tests. Table showing the Statistical MST tests, column 1 list the name of the parameter and column 2 the derivation.

4.4 Applying the MST process

I applied the MST process to three datasets; the BGPS, the GRS and the P&F IRDC catalogue (Chapter 1). The remaining datasets; the MMB, the RMS, the Cyganowski EGO catalogues, the Churchwell Bubble Catalogue, the Green Supernova Catalogue, the Roman-Duval cloud catalogue and the UKIDSS Cluster Catalogue have too small a number density for an MST to have any purposeful meaning. Sources in the range $20 < l < 50^o$ were used as this is the area of maximum overlap between the datasets.

In addition to spatial data (I use l and b coordinates listed with the datasets) I chose to use an additional factor for the GRS and P&F datasets. The GRS carries velocity for the sources therein which may be used as a third dimension within the MST, if we assume that objects that are spatially close and that are also close in velocity are physically close in three dimensions. For the P&F dataset I assume that IRDCs have a characteristic size. Figure 4.5 shows the size distribution of the Simon *et al.* (2006a) IRDC with known distances and hence diameters. The initial rise in number density will be due to incom-

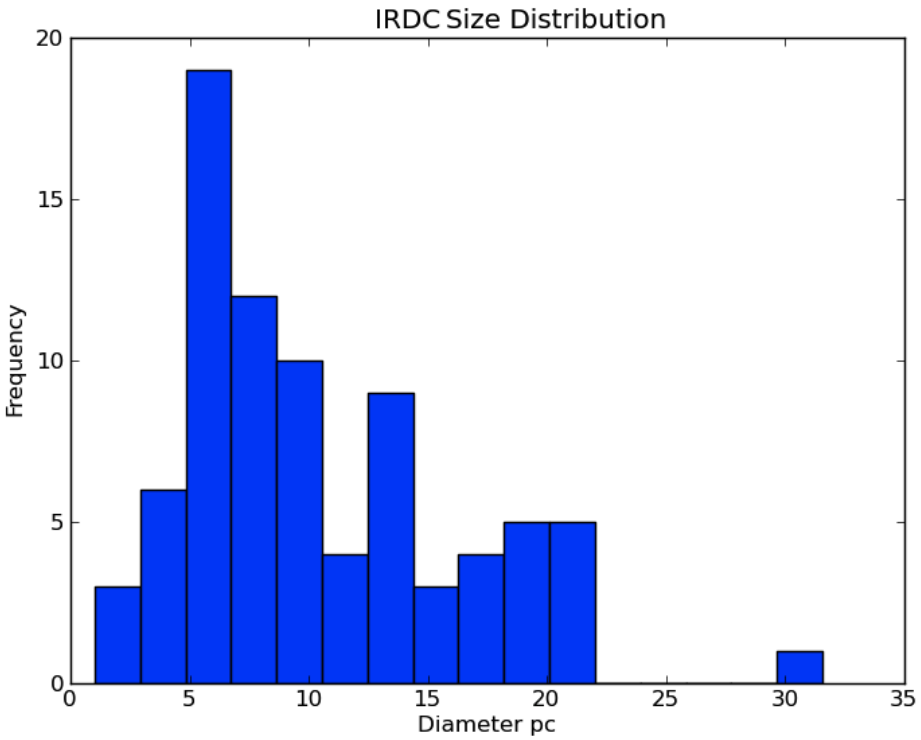


Figure 4.5: IRDC Size Distribution
Plot of the IRDC size distribution using the IRDCs from Simon et al. *(2006a) but using an independent method. This indicates the possibility that IRDCs have a characteristic size of the order of* 6 *pc and is in agreement with the observations with Simon* et al. *(2006a).*

pleteness for very small clouds that are unresolved. However, the plot strongly implies that IRDCs do indeed have a characteristic size of the order of $\sim 4\,pc$. This being the case, clouds with small angular diameters will tend to be more distant than those with large angular diameters.

For the production of the MST used within this work Python 2.6 was deployed in two environments. The first of these was a dual core Macbook running the Endthought 2.6 Python distribution which was used as a development environment for the running of the construction of the live data trees and for the tree plotting. The second environment utilised the Redhat 2.6.2 deployment of Python on a 600+ core Intel Xeon Cluster that was used for the production of Monte Carlo simulations and the analysis thereof.

As was discussed in section 4.1 the trees were produced using Prim's algorithm (Prim 1957) which is interpreted below in Algorithm 1. The code used to produce the MST was checked against a graph with a known MST as well as graphs produced by dry running the algorithm. Sample results from the two environments were cross checked to ensure

consistency.

Algorithm 1 *Prim's algorithm*

```
Load Tree
Count number of nodes N
Initiate EdgeCounter and set to zero
Select a random start node N_start
Create the list,Nodes, containing all nodes -
Remove N_start from Nodes
Create an empty list Edges
while N>EdgeCounter do
  Create a list (A) of all edges attached to N_start
  Create a list (B) of all members of A that end in a member of Nodes
  Add List B to List Edges
  Remove members of Edges not ending with a node in Nodes
  Find the member M of Edges with the lowest cost
  Set the end node of M to end
  Plot the Edge M
  Remove M from Edges
  remove end from Nodes
  N_start = end
  EdgeCounter = EdgeCounter + 1
end while
END
```

Although minimum spanning trees identify the scale of cluster they fail to show the location within the parameter space. In order to identify specific regions of clustering within a minimum spanning tree we need to $fracture$ the tree, thereby creating a Fractured Minimum Spanning Tree (FMST). The fracturing process (also known as *cutting or separating*) as discussed in section 4.2, involves removing all edges with lengths longer than a given value, typically the mean edge cost. During my code testing process, test trees were fractured at the mean, median, $\pm 1\sigma$ levels. The mean edge cost, $\bar{m}$, was identified as the most suitable level at which to fracture (as has previously been used Krzewina and Saslaw (1995)). I chose to use the mean edge cost over an edge cost determined from the edge cost CDF (see section 4.2) for a number of reasons. Firstly, the computation of the CDF and the fitting of two lines introduces additional complexity for relatively little gain in determining cluster size. Secondly, it is not possible to fit two lines to all my CDF while the Monte Carlos CDF generally do not show an inflection and the live data requires smoothing before a fitting can be made. However, I was able to calculate $\bar{m}$ for all fields. Additionally, the use of line fitting to the CDF is dependent on the line fitting algorithm and hence reduces repeatability.

As stated in section 4.2, fracturing produces graphs consisting of a series of unconnected sub-trees or bushes, which in themselves are minimum spanning trees (section 4.1). To these graphs I applied a further process known as Orphan Removal (or stripping) (see also section 4.1). For each sub-tree the number of connected nodes was counted, if the number of nodes is less than a given figure, [5], the sub-tree, its associated nodes and edges were removed from the tree. The tree that remains was a Stripped Fractured minimum spanning tree.

The Orphan Removal removes orphaned nodes (hence the name) as well as reducing the number of sub-trees that are chance connections. Furthermore, it also generates clusters and allows the statistical measure of the size distribution of the sub-trees. It further allows the mean edge cost to be measured within a cluster, rather than just within the tree as a whole.

Figure 4.6 illustrates the construction of an MST using Prim's algorithm.

There are a number of issues with this technique that required addressing before it was applied to my datasets. The computational complexity of Prim's algorithm is approximately $O(m(log(n))$ in big Oh notation. For my dataset $m = (n-1)^2$ hence the computational complexity becomes $O((n-1)^2(log(n))$ which reduces to effectively an $O(n^2)$ problem. For very large datasets this is clearly an obstacle. The obvious way of dealing with this is to break the data into a series of equal cells. This would reduce the computation complexity to $O(C(n/C)^2)$ where C is the number of cells. Unfortunately breaking the data up into cells exacerbates the second problem - edge effects.

In graph theory it is assumed that there are no nodes outside the boundaries of the graph. This is clearly not the case for these datasets, especially if the data is fragmented into cells to reduce complexity. Edge effects are briefly discussed in Cartwright and Whitworth (2004) but their implications are not explained. The consequence of edge effects is that false edges may be being made by nodes near the edge of the graph to nodes within the graph. For relatively densely populated fields the effect is small but for low density fields it may have a significant impact on the mean edge cost and hence the FMST. The solution is linked to the third problem - mixed clustering scales.

[5] I use the typically used four node cut off for this work.

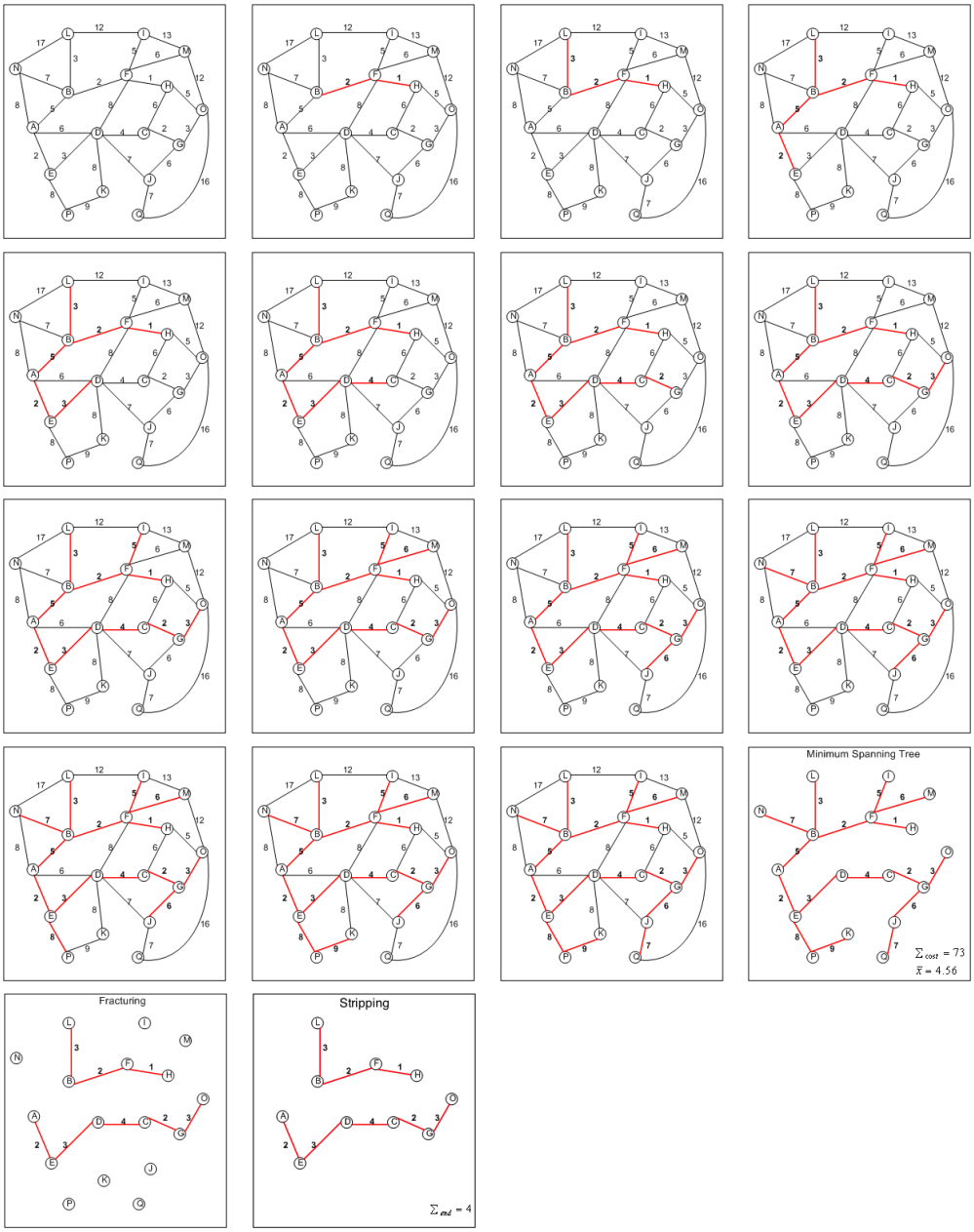


Figure 4.6: Tree Construction

Series of images that show the process construction of a minimum spanning tree using Prim's algorithm. The nodes are the number dots and the permitted edges are the black lines. As the tree is built selected edges turn from black to red. The image in column 4 row 4 is the final tree. The graph in column one row five is the mean fractured tree, with the final imaging being the stripped fractured tree.

Consider a graph which contains a single over density. The MST of the graph contains both members and non-members of the cluster. The edge costs between members is low compared to the non-members. Fracturing the tree at the mean edge costs separates the members from the non-members using this discrepancy. Now consider a graph that has two over densities (d_1 and d_2) with different mean edge costs within each over density. Such a situation can be considered to be two clusters, one distant one close, within the same field. In this case we have three populations; members of d_1 which have small edge costs, members of d_2 which have longer edge costs than members of d_1 and non-members. The effect of having two over densities is to reduce the fracturing efficiency because the mean edge cost is now too small. The method discussed in Gutermuth *et al.* (2009) to determine the fracturing cost by using the edge cost CDF would seem a suitable alternative. However, this would involve producing in the order of 8500 CDF for which there is no certainty that breaking between clustering and non-cluster objects could be determined. Furthermore, this CDF fitting does not fundamentally address the key issue; MSTs are most efficient when detecting single clusters (Schmeja 2011), hence if there is more than one cluster in the field then the field is inescapably too large. Reducing the cell size would help reduce the chances of multiple clusters occurring, however it would also increase the edge effects. Instead I chose to use a series of overlapping cells, such that every part of the field, except the first and last square degrees, are covered twice. We can therefore compare the overlaps, which contain different but similar trees. Hence, I effectively apply a factor of two over sampling interval to the MST. This allows me to address the issues of edge effects and multiple clusters, as they will be highlighted as differences between overlapping fields, whilst reducing the overall computation complexity. These cells are four degrees in size in l and two degrees in b. There is a two degree overlap in l between adjacent cells.

The correct weighing of any of the parameters that locate a node within the parameter space that will form the tree is an important aspect within the initial set-up prior to the generation of the MST. The weighing processes allows differing parameters to be used within the tree formation process and for uncertainties within those parameters to be accounted for. For example, consider a single two dimensional dataset where both axis are spatial. In normal situations both axis would have the same uncertainties but let us consider the situation where this is not so, for example if there was a x -axis encoder error. In this case two objects located closely in the y axis would be more likely to be related than those x axis. In order to account for this we would multiply the x axis values by a weight greater than one thereby effectively smearing the x axis. Hence when the separation distance between any two nodes is calculated we find that nodes located close in the y axis are more likely to be linked by an edge than those closely located on the x axis therefore reflecting the increased uncertainties within the x axis. This technique is discussed further in Hoffmann *et al.* (2008). However, the reader should be aware that other algorithms exist for the construction of trees where there is uncertainty in the edge costs such as those discussed in Erlebach *et al.* (2008).

As the positional uncertainties within my datasets are common for each spatial axis, within any one dataset the effective weight for these axes is one. The determination of the weight for the GRS velocity and P&F angular size is challenging for a number of reasons. Firstly, in velocity space for objects in the inner Galaxy, two objects may have the same velocity

but one will be on the near side of the Galaxy and one at the far. Secondly, the exact nature of the distribution of the velocities with the GRS is not fully understood. Thirdly, although I have demonstrated that IRDCs appear to have a characteristic size it is based on a relatively small sample of IRDCs with known distances. Hence, we need to weight velocity and angular diameter such that it only becomes important if two objects are spatially close. In order to determine this weighing factor I generated a number of MSTs for an IRDC and a GRS field with a range of weights. The trees showed that at low weights non-physical connections are made; for example objects separated by a degree are associated. This gave a minimum weighing level from which to work. Running further trial trees and examining the statistical output rather than the trees themselves, suggested that
$1km\ s^{-1} = 30$ in l and b and 1"=15" l and b in were appropriate weights for the GRS velocity axis and the IRDC diameter axis. This is discussed further in section 4.5.

It should be noted that applying the uncertainties post construction of the MST, such as used in Gutermuth *et al.* (2009) for the determination of the fracturing length, may produce erroneous results as it assumes that the MST is insensitive to the edge length at the level of uncertainty encountered. However, it is highly unlikely that this is the case. I therefore suggest that the use of edge length uncertainty, post tree construction, be avoided.

For a statistical comparison 200 Monte Carlo datasets were produced for each of the three datasets and for each cell resulting in 8400 Monte Carlo runs in total. The number density of objects within each cell was taken from the live data. Sources were distributed randomly in l and as a Gaussian in b, centred on $b = 0$. The points were generated by Python's built-in Gaussian and Random [6] functions. For the P&F and the GRS, the non-spatial dimensional data (angular size P&F, velocity for GRS) from the live data was randomly assigned to a Monte Carlo point. For the generation of the BGPS Monte Carlo data, each point was checked to ensure that it was more than a CSO beam width away from any other point before it was allowed into the Monte Carlo. If it did not match with the criteria the point was rejected and a replacement generated. For the GRS the same principle was applied but with the added criterion of a separation of $\pm 5\ kms^{-1}$ in velocity. Hence, two GRS objects could be in the same beam if they could be resolved in velocity. For the P&F the Monte Carlo points were generated, using the angular dimensions of the IRDCs, in such a way that no two IRDCs would overlap.

Prim's is a linear operation algorithm, it does not lend itself to parallel programming. Furthermore, it was discovered during the testing process that the built-in Python thread management was inefficient when compared to the native cluster process management. To account for this the minimum spanning tree code used on the live data was adjusted so that multiple incidences could be run in co-ordination with a single job stack. This allowed several hundred trees to be generated in parallel. Due to a fault within the cluster kernel, which has now been rectified, the job stack would be occasionally corrupted. This would result in an inability of a thread to move onto the next job and as a result would

[6] The Python random number generator is, of course, only pseudo random but in the context of this work it is sufficient.

crash. In the most extreme cases, for example when the problem occurred overnight, all the threads would crash. However, these crashes never corrupted the output and once the Job Stack File was re-built, all the threads could be restarted.

4.5 Statistical Testing and Results

In order to understand the correlation between the star formation indicators discussed in Chapter 1 I have extracted a number of statistical measurements from the MST. As discussed in sections 4.2 and 4.4 I have fractured the trees at the mean edge cost. I choose not to introduce K-branching (section 4.2) as it tends to enhance linear features. For comparison between the GRS and BGPS I have produced a series of flat (i.e. trees built using l and b only) MSTs for the GRS. Due to the GRS being a line based survey some of the GRS objects have the same spatial position but differing velocities. However, when I produced the flat GRS trees the region between $l = 30$ and 34^o the analysis code failed because effectively it was trying to fit a convex hull to a point caused by there being multiple sources in the same location within the datasets. To overcome this it was necessary to remove all duplicates from the source data. Hence, the flat GRS results has fewer data points between $l = 30$ and 34^o than the 2dv GRS.

For all four tree sets (P&F, BGPS, GRS and Flat GRS) and the three Monte Carlo datasets I extracted the mean edge cost for the entire tree, $\bar{\mathrm{m}}$, and the mean internal edge cost, $\bar{\mathrm{s}}$, from which we can derive Q (Cartwright and Whitworth 2004; Cartwright 2009) (section 4.3). A convex hull was fitted to each sub-tree in the two spatial dimensions using the Python code of Kutterer (2011) which implements the Graham Scan algorithm to perform this function. Once the convex hull and the data points that define it were determined [7], a simple algorithm was implemented that breaks the hull up into a series of triangles and therefore allows the determination of the convex hull's area. Additionally I determined the centre of each convex hull and hence the centre of each sub-tree by summing the boundary points and dividing by the number of such points. The length of the semi-major axis of the convex hull was determined by finding the boundary points with the greatest separation. The area of the convex hull was used to determine the circular radius of the sub-tree, which together with the semi-major axis was used to determine the elongation measurement ξ (section 4.3 and (Schmeja and Klessen 2006)). I also calculated the Billot Elongation factor (see section 4.3) which allows for the fact that in reality the convex hull should extend slightly beyond the boundary points. Finally, I calculated the adjusted Q value Q^* or $Q_{intrinsic}$ which is the Q value adjusted for elongation see (Bastian *et al.* 2007) and section 4.3. Table 4.3, Table 4.4 and Table 4.5 show a sample of the results from the MST. The full details of which can be found in the Appendix. In addition, I determined the edge angle distribution (section 4.3) for each of the MST datasets as illustrated in Figure 4.7, for comparison between datasets. I chose not to perform this function on the Monte Carlo data as the effect of random fields on the edge angle distribution has previously been investigated (section 4.3) and is computationally intensive.

[7] I shall use the term *boundary points for these objects*

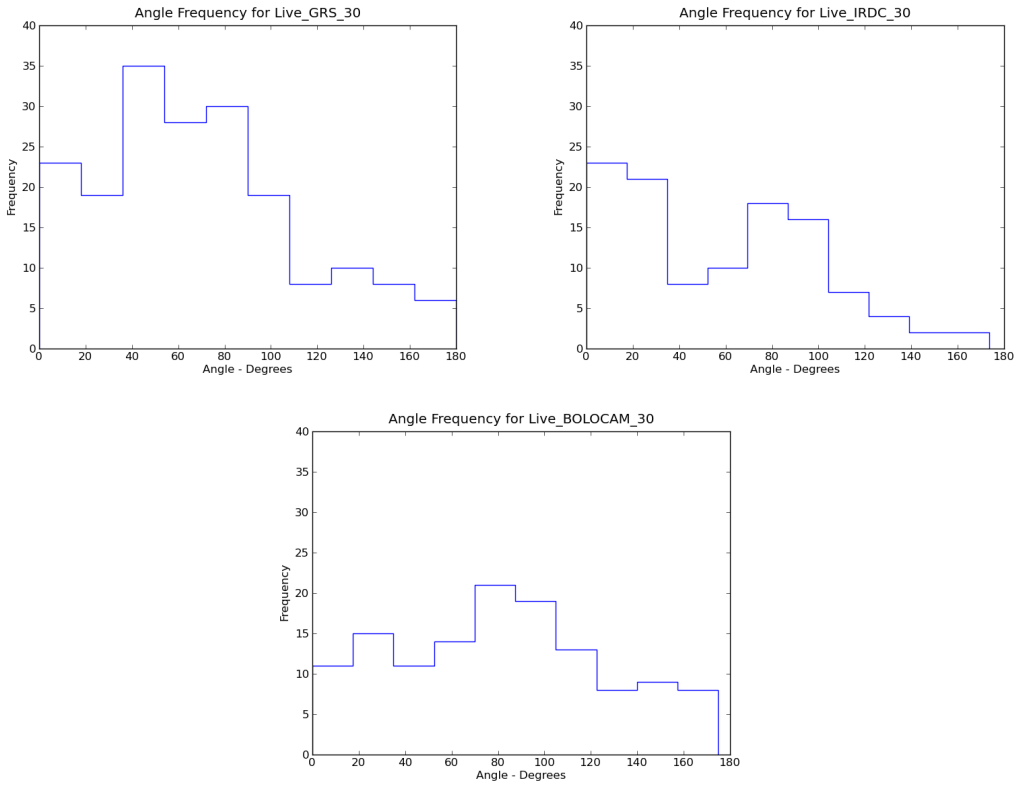


Figure 4.7: Angle Frequency Distribution

Plot of the node edge to edge angle distribution for the GRS, P&F and BGPS Stripped FMST The Angle Frequency Distribution measures the degree at which edges are preferentially aligned. In a random distribution of nodes the MST Angle Frequency Distribution will be flat. In this case we see that the distribution is not flat indicating a tendency for the edges to line parallel to the Galactic Plane. However, this may be due to the limiting scale height of the data and may not be indicative of real structure.

Due to the use of over sampling it was highly likely that hulls will be duplicated. To overcome this the sub-tree centres were cross checked. If any centres were within 30 arcseconds of another hull centre the sub-tree with the lowest number of nodes of the pair was removed from the list. This process was continued until no hull centres were located within 30 arcseconds of another hull. This process results in the identification of 268 GRS clusters, 198 IRDC clusters and 132 BGPS clusters consisting of 2382, 1577 and 969 objects, respectively. This represents approximately 39% of the GRS population, 57% of the IRDC population and 26% of the BGPS population.

It is interesting to note that the KS-testing of the GRS and Flat GRS trees, for parameters that are not strongly dependent on the mean edge cost, which will not be the same between the 2d and 2dv trees, show that they are tracing the same populations.

Figures 4.8, 4.9 and 4.10 show the convex hulls for the region $30 < l < 34^o$ prior to the duplicate hulls being stripped out; whilst Figure 4.11 shows the combined Stripped Fractured minimum spanning tree for the same region. The full set of plots can be found in the Appendix. It should be noted that the inclusion with the plots of the Churchwell GLIMPSE bubbles and the Green Supernova remnants is due to early results that suggested that cavities may exist within the live data which was initially thought to be related to these objects. However, examination of the Monte Carlo data initiates that these cavities are artefacts and not real structure.

In order to confirm that these results are detecting real structure as opposed to a random distribution each parameter is checked against the same parameter in the Monte Carlos using the two sample Kolmogorov-Smirnof (KS) statistic built into Python. The KS test compares two samples and assumes that they are drawn from the same population. The Scipy KS test returns two indicators KS and p. If p is small (less than 0.1) we can reject the hypothesis that our two samples are drawn from the same population. Table 4.2 shows the KS test results.

KS Test	$\bar{s}$	$\bar{m}$	Q	A	SMA	BE	ξ	Q^*
IRDC v MC	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.01	< 0.01
BGPS v MC	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.0056	0.0028	< 0.0001
GRS v MC	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.7164	< 0.0001	< 0.0001
BGPS v GRS	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.5922	< 0.0001	< 0.0001

Table 4.2: MST KS Test Results

Table showing the results of the MST KS tests. $\bar{s}$ is the inter-cluster mean edge cost, whilst $\bar{m}$ is the tree mean edge cost. Q is the Cartwright cluster parameter and Q^* is the Bastion adjusted Q parameter. A is the area of the convex hull enclosing the sub-tree in square degrees. SMA is the size of the semi-major axis of the convex hull in degrees. BE is the Billot Elongation Factor and ξ the Schmeja Elongation Factor.

BGPS MST Results										
N	$\bar{s}$ deg	$\bar{m}$ deg	Q	Area deg^2	SMA deg	l deg	b deg	BE	ξ	Q^*
8	0.0255	0.0383	0.6649	0.0010	0.0444	32.4361	0.2273	2.5390	0.3939	0.7469
32	0.0257	0.0383	0.6718	0.0050	0.1317	30.8116	-0.0327	3.3112	0.3020	0.7346
14	0.0205	0.0383	0.5351	0.0012	0.0410	32.8191	0.0427	2.1430	0.4666	0.6322
22	0.0263	0.0383	0.6867	0.0038	0.1024	33.6098	-0.0321	2.9519	0.3388	0.7572
14	0.0298	0.0383	0.7789	0.0033	0.0577	30.0522	-0.2707	1.7802	0.5617	0.8958
20	0.0252	0.0383	0.6594	0.0034	0.0800	30.4850	-0.4125	2.4214	0.4130	0.7453
36	0.0303	0.0383	0.7903	0.0046	0.0997	31.0989	0.0580	2.5977	0.3850	0.8704
10	0.0280	0.0383	0.7311	0.0013	0.0531	31.1024	0.2751	2.6549	0.3767	0.8095
10	0.0261	0.0383	0.6806	0.0010	0.0462	30.7476	0.1224	2.5803	0.3876	0.7612
12	0.0265	0.0383	0.6915	0.0015	0.0537	30.3341	-0.1340	2.4894	0.4017	0.7750
18	0.0290	0.0383	0.7566	0.0025	0.0698	30.6610	0.2305	2.4677	0.4052	0.8409
8	0.0214	0.0383	0.5596	0.0005	0.0375	33.6456	0.1713	2.9176	0.3427	0.6308
10	0.0278	0.0383	0.7259	0.0017	0.0396	30.8653	-0.1713	1.6993	0.5885	0.8483
26	0.0274	0.0383	0.7151	0.0033	0.0734	30.8110	0.1702	2.2722	0.4401	0.8066
20	0.0274	0.0383	0.7159	0.0069	0.0758	30.3952	0.0036	1.6224	0.6164	0.8441
8	0.0193	0.0383	0.5046	0.0011	0.0298	30.9562	-0.8529	1.5828	0.6318	0.6360
8	0.0291	0.0383	0.7594	0.0014	0.0330	30.6024	-0.0361	1.5584	0.6417	0.8928
14	0.0264	0.0383	0.6889	0.0055	0.0574	30.8296	-0.1134	1.3770	0.7262	0.8400
10	0.0241	0.0383	0.6288	0.0008	0.0434	30.3063	0.1800	2.6711	0.3744	0.7067

Table 4.3: BGPS MST Result - Part

Table showing the results of the BGPS MST. N is the number of nodes within the sub-tree, $\bar{s}$ is the inter-cluster mean edge cost, whilst $\bar{m}$ is the tree mean edge cost. Q is the Cartwright cluster parameter and Q^* is the Bastion adjusted Q parameter. The Area column is the area of the convex hull enclosing the sub-tree in square degrees. SMA is the size of the semi-major axis of the convex hull in degrees. l and b locates the centre of the convex hull. BE is the Billot Elongation Factor and ξ the Schmeja Elongation Factor.

GRS MST Results										
N	$\bar{s}$ deg	$\bar{m}$ deg	Q	Area deg^2	SMA deg	l deg	b deg	BE	ξ	Q^*
52	0.0275	0.0545	0.5051	0.0129	0.123	47.4642	-0.8442	1.923	0.52	0.6132
34	0.0138	0.0545	0.2525	0.0051	0.0626	49.4933	-0.3867	1.5549	0.6431	0.3863
10	0.0249	0.0545	0.4579	0.002	0.05	46.3675	-0.22	1.9817	0.5046	0.5629
24	0.0206	0.0545	0.3777	0.0073	0.0863	48.6575	0.2225	1.7906	0.5585	0.4938
8	0.0276	0.0545	0.5068	0.0012	0.0364	46.2175	-0.67	1.8625	0.5369	0.6185
14	0.0229	0.0545	0.4197	0.002	0.0354	46.2925	-0.2275	1.4191	0.7047	0.5662
18	0.0271	0.0545	0.4983	0.005	0.0901	48.65	0.0913	2.2708	0.4404	0.5899
26	0.0307	0.0545	0.5641	0.0201	0.0943	48.782	0.072	1.1794	0.8479	0.7405
10	0.0224	0.0545	0.4115	0.0015	0.0427	48.8525	0.2525	1.9885	0.5029	0.5161
10	0.0258	0.0545	0.4744	0.0027	0.0472	49.415	-0.3375	1.609	0.6215	0.6037
14	0.0236	0.0545	0.4325	0.0019	0.0412	47.3567	-0.885	1.6766	0.5965	0.5566
10	0.0237	0.0545	0.4347	0.0015	0.0515	48.7829	0.2929	2.3559	0.4245	0.523
12	0.0325	0.0545	0.5973	0.0038	0.0757	49.385	-0.2425	2.1756	0.4597	0.6929
18	0.0216	0.0545	0.3961	0.0022	0.0738	49.006	-0.306	2.7897	0.3585	0.4706
52	0.0275	0.0592	0.4646	0.0129	0.123	47.4642	-0.8442	1.923	0.52	0.5727
12	0.0259	0.0592	0.4376	0.0041	0.064	45.406	-0.74	1.7834	0.5607	0.5542
22	0.0154	0.0592	0.2608	0.0043	0.0541	45.708	-0.238	1.4619	0.6841	0.4031
14	0.0282	0.0592	0.4764	0.0026	0.0757	45.4225	0.055	2.6558	0.3765	0.5547
24	0.0237	0.0592	0.4007	0.0032	0.0814	46.3122	-0.2256	2.5503	0.3921	0.4823

Table 4.4: GRS MST Result - Part

Table showing the results of the GRS MST. N is the number of nodes within the sub-tree, $\bar{s}$ is the inter-cluster mean edge cost, whilst $\bar{m}$ is the tree mean edge cost. Q is the Cartwright cluster parameter and Q^* is the Bastion adjusted Q parameter. The Area column is the area of the convex hull enclosing the sub-tree in square degrees. SMA is the size of the semi-major axis of the convex hull in degrees. l and b locates the centre of the convex hull. BE is the Billot Elongation Factor and ξ the Schmeja Elongation Factor.

P&F MST Results										
N	$\bar{s}$ deg	$\bar{m}$ deg	Q	Area deg^2	SMA deg	l deg	b deg	BE	ξ	Q^*
20	0.0273	0.047	0.5811	0.004	0.0856	23.0121	-0.4176	2.4111	0.4148	0.6673
12	0.0183	0.047	0.3902	0.0008	0.0324	22.0512	0.22	2.0817	0.4804	0.4901
10	0.0268	0.047	0.5707	0.0024	0.053	24.1738	-0.0697	1.9069	0.5244	0.6798
28	0.0317	0.047	0.6733	0.004	0.126	23.2934	-0.1808	3.5338	0.283	0.7321
12	0.0317	0.047	0.6738	0.0032	0.0713	23.4846	0.0625	2.2159	0.4513	0.7676
10	0.0337	0.047	0.7175	0.0023	0.0391	23.2435	-0.2779	1.4425	0.6932	0.8617
10	0.0295	0.047	0.6284	0.0022	0.046	23.2072	-0.3924	1.7332	0.577	0.7484
12	0.0241	0.047	0.5125	0.0012	0.0601	24.1618	0.0439	3.0976	0.3228	0.5796
12	0.0341	0.047	0.7247	0.0041	0.0459	23.4516	-0.0791	1.2663	0.7897	0.889
30	0.0271	0.047	0.5762	0.0037	0.124	24.6411	0.1507	3.5951	0.2782	0.634
8	0.0325	0.047	0.6903	0.0026	0.039	24.8454	-0.2074	1.3508	0.7403	0.8443
24	0.0236	0.047	0.5011	0.0036	0.0703	24.8163	-0.1011	2.0619	0.485	0.602
12	0.0242	0.047	0.515	0.0017	0.0404	23.4746	-0.262	1.7458	0.5728	0.6341
18	0.0273	0.047	0.5808	0.0029	0.0627	24.594	-0.3141	2.0459	0.4888	0.6825
20	0.0325	0.047	0.691	0.0075	0.0773	24.5371	-0.2082	1.5808	0.6326	0.8226
12	0.0194	0.047	0.4131	0.002	0.0398	25.3918	-0.3154	1.586	0.6305	0.5443
10	0.0264	0.047	0.562	0.0014	0.0344	23.5818	0.1532	1.6351	0.6116	0.6893
12	0.0228	0.047	0.4849	0.0011	0.0319	24.6792	-0.1427	1.6955	0.5898	0.6075
10	0.0182	0.047	0.3874	0.0008	0.0299	22.1644	-0.6583	1.8465	0.5416	0.5

Table 4.5: P& F MST Result - Part

Table showing the results of the P& F MST. N is the number of nodes within the sub-tree, $\bar{s}$ is the inter-cluster mean edge cost, whilst $\bar{m}$ is the tree mean edge cost. Q is the Cartwright cluster parameter and Q^* is the Bastion adjusted Q parameter. The Area column is the area of the convex hull enclosing the sub-tree in square degrees. SMA is the size of the semi-major axis of the convex hull in degrees. l and b locates the centre of the convex hull. BE is the Billot Elongation Factor and ξ the Schmeja Elongation Factor.

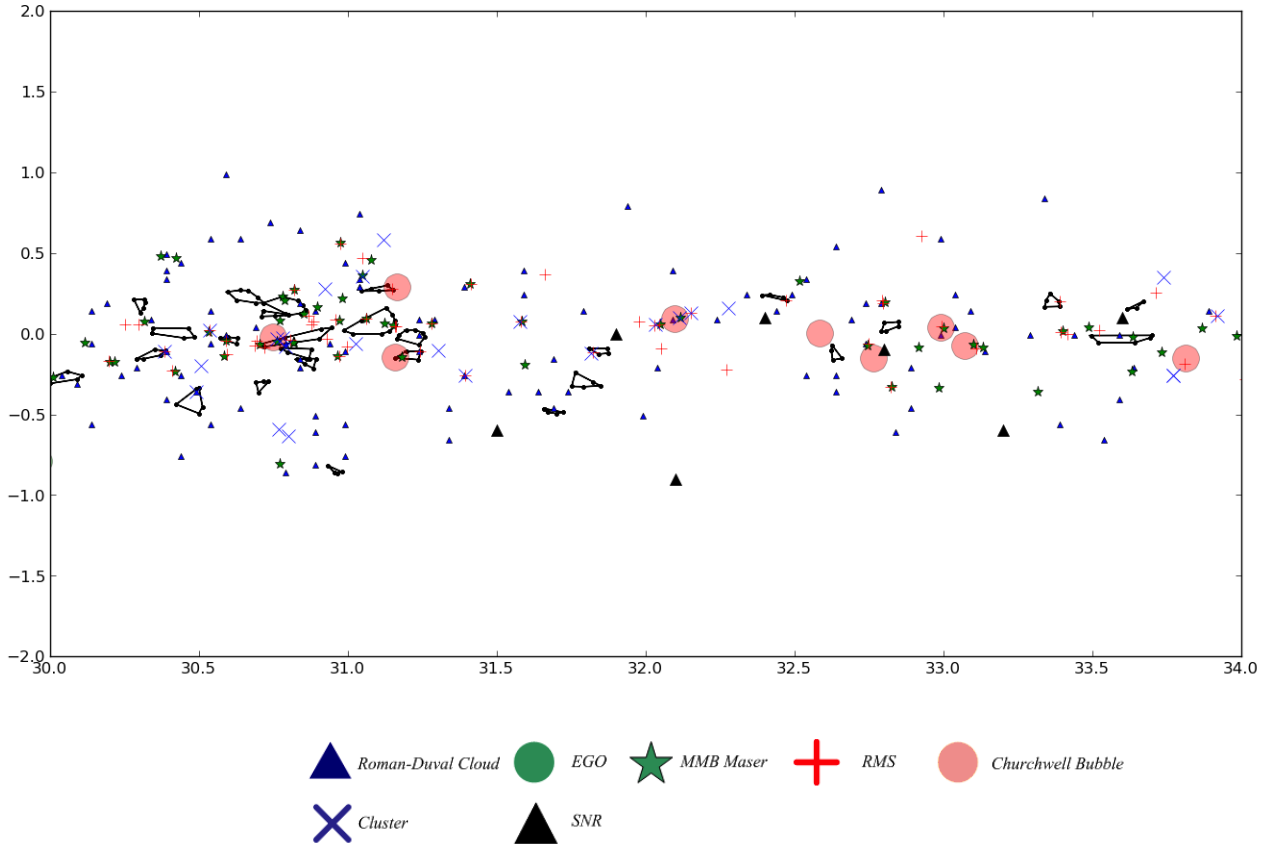


Figure 4.8: BGPS Convex Hull $30^o < l < 34^o$

Plot of the BGPS convex hulls generated from the MST contained within $30^o < l < 34^o$. The coloured marker show the location of star formation indicators that I expected to be associated with BGPS sources as shown in the key.

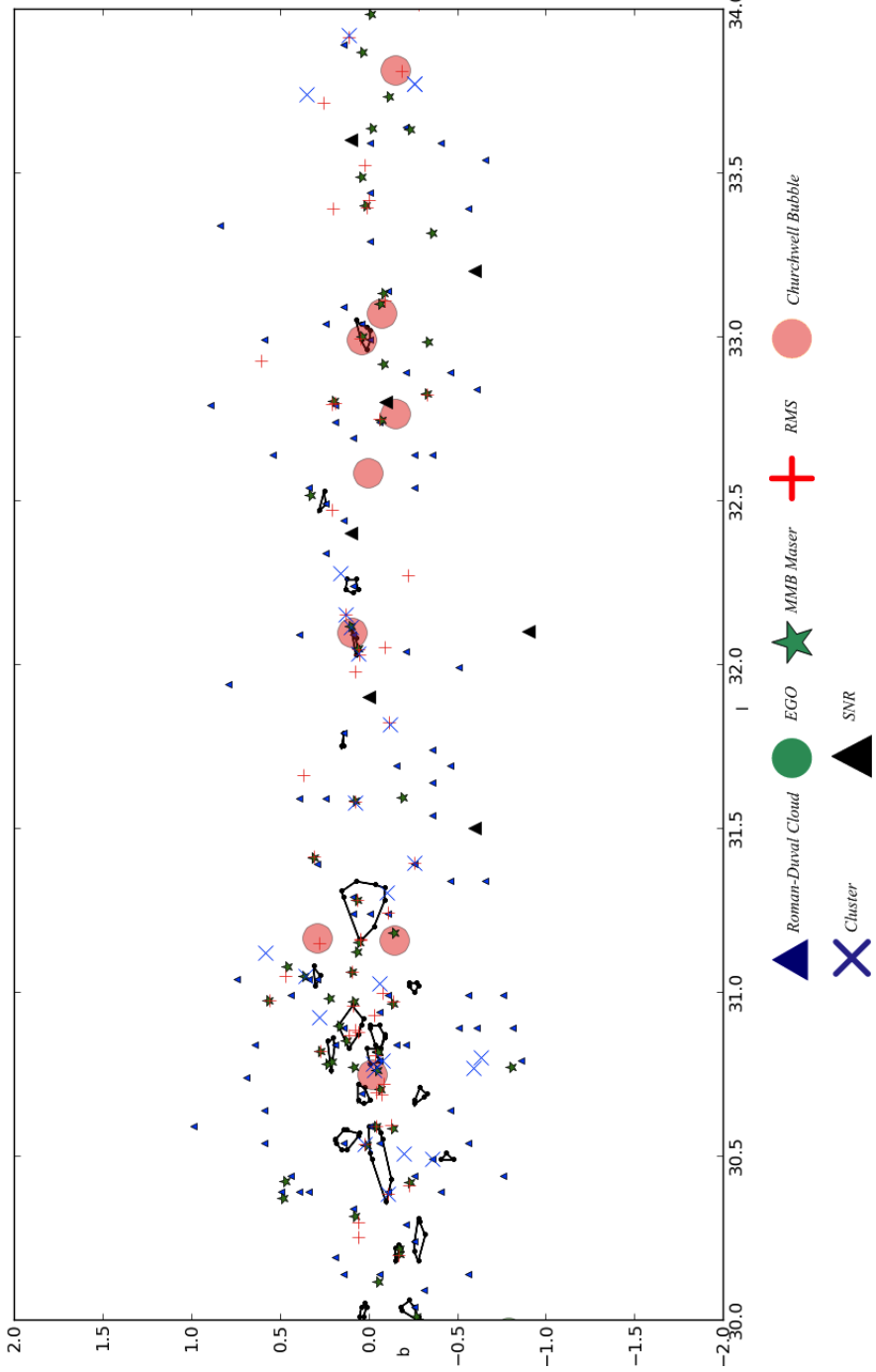


Figure 4.9: GRS Convex Hull $30^o < l < 34^o$

Plot of the GRS convex hulls generated from the MST contained within $30^o < l < 34^o$. The coloured marker show the location of star formation indicators that I expected to be associated with GRS sources as shown in the key.

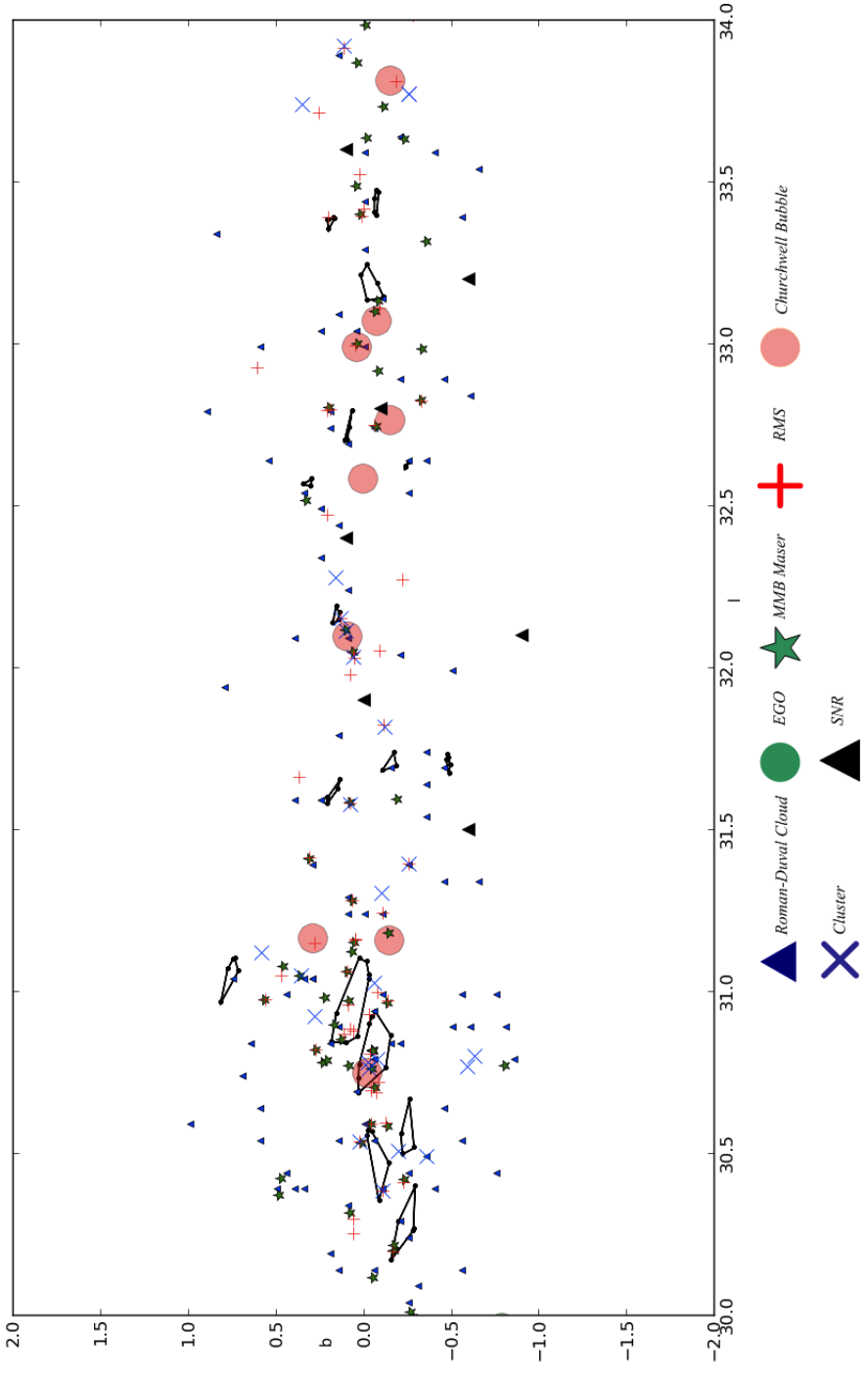


Figure 4.10: P&F Convex Hull $30^o < l < 34^o$

Plot of the P&F IRDC convex hulls generated from the MST contained within $30^o < l < 34^o$. The coloured marker show the location of star formation indicators that I expected to be associated with P&F IRDC sources as shown in the key.

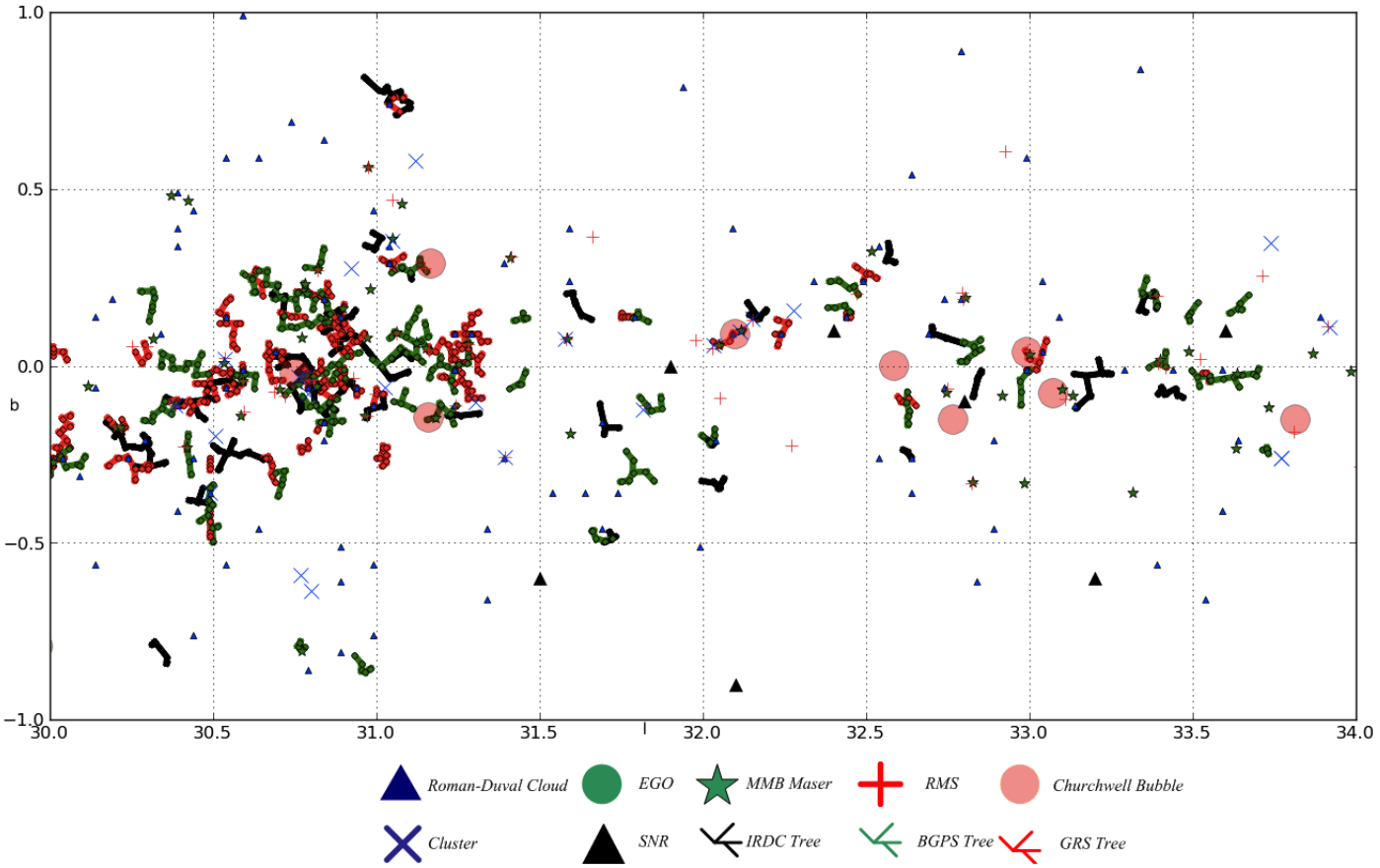


Figure 4.11: Stripped Fractured MST $30^o < l < 34^o$

Plot of the Stripped Fracture MST contained within $30^o < l < 34^o$ as generated from the GRS (red lines) BGPS (green lines) and P&F IRDC (black lines) catalogues. The coloured marker show the location of star formation indicators that I excepted to be associated with P&F IRDC sources as shown in the key. Note that we see inter group clustering, for example at $l = 30.75$ as well as isolated trees.

Mean Statistic	GRS 2dv	GRS 2d	% Change	P&F 2dd	P&F 2d	% Change	BGPS
Node Count	8.8881	8.6932	2	7.9646	7.8959	1	7.3409
$\bar{s}$	0.0214	0.0209	2	0.0302	0.0284	6	0.0272
$\bar{m}$	0.0426	0.0419	2	0.0612	0.0594	3	0.0458
Q	0.5038	0.4987	-1	0.6178	0.5992	-3	0.6324
Area	0.0030	0.0029	4	0.0045	0.0042	8	0.0035
SMA	0.1035	0.1001	3	0.1324	0.1275	4	0.1231
BE	1.9593	1.9593	0	2.0517	2.1337	4	1.9898
Xi	0.5482	0.5482	0	0.5848	0.5964	2	0.5848
Q*	0.6178	0.6487	5	0.6152	0.6447	3	0.7541

Table 4.6: MST Statistical Comparison

Comparison of the mean tree statistics between sources and three parameter and two parameter fracturing. This shows that the effect of the performing the P&F and GRS trees in three dimensional parameter space as an effect but it is small.

As an additional test on the P&F and GRS trees I recalculated the edge cost of each edge in the MST to the 2d edge cost. This was then used to calculate revised $\bar{m}$ and $\bar{s}$ and new Stripped Fractured MST for the P&F and GRS trees. The removal of the third dimension naturally reduces $\bar{m}$ and $\bar{s}$ and consequently has an impact on the convex hulls with several hulls that appear in the 2dv and 2dd trees being stripped out, 16 for the GRS and 11 for P&F. However, the remaining trees have centres consistent with the 2dv/2dd sample. Table 4.6 shows the comparison between the mean statistics for the live trees for both two parameter and three parameter fracturing. We see from Table 4.6 that using a two dimensional edge cost for fracturing and statistical purposes has relatively little impact on the statistical results.

To compare the convex hulls to our other dataset I performed an adaptive cross match with the cross match radius being tailored for each hull by using the hull's semi-major axis. Table 4.7 shows the association rate. At first glance it would appear that there is a marked difference between the three datasets. However, once these are hull-area normalised we find that the association levels are constant between the sets apart from a reduction of EGO associated with BGPS objects.

	GRS	*P&F*	BGPS
Maser	82	101	53
RMS	90	110	53
Cluster	40	64	33
Roman Duval	199	169	88
Green SNR	13	16	7
Churchwell Bubbles	32	47	23
EGO	15	27	8

Table 4.7: Hull associations
The above table shows the number of objects located within a convex hulls of each of our three datasets.

	GRS	*P&F*	BGPS
Maser	7	7.2	6.6
RMS	8	7.5	6.6
Cluster	3.6	4.4	4.0
Roman Duval	18	11.6	11
Green SNR	1.2	1.1	0.9
Churchwell Bubbles	2.9	3.0	2.9
EGO	1.4	1.8	0.9

Table 4.8: Area normalised hull associations
Table showing the number of objects located within a convex hulls of each of our three datasets, but adjusted for the total hull area of each object type.

Figure 4.12 shows the Q BE parameter space occupied by the P&F, GRS and BGPS hulls. They are characteristically flat in BE when compared to the Monte Carlo indicating that the hulls are generally oblate. The observation that Q for the P&F sources are slightly more distributed than Q for the GRS and significantly more than the BGPS indicates, that IRDC are more tightly clustered than the other two datasets, with the GRS being more clustered than the BGPS. This is shown quantitatively in Table 4.6.

In order to understand the relationship between the three star formation indicators and actual star formation I cross matched the convexed hulls with the WMAP[8] Galactic Free-Free emission survey of Murray (2010). For the cross matching radius I used the angular radius of the star formation region (SFR) taken from the distance and R_{SFR} listed in Murray (2010). This resulted in 54 GRS, 36 P&F and 31 BGPS convex hulls identified as being associated with Murray (2010) SFR. Hence, we are able to determine distances to these hulls.

[8]The Wilkinson Microwave Anisotropy Probe, millimetre space telescope.

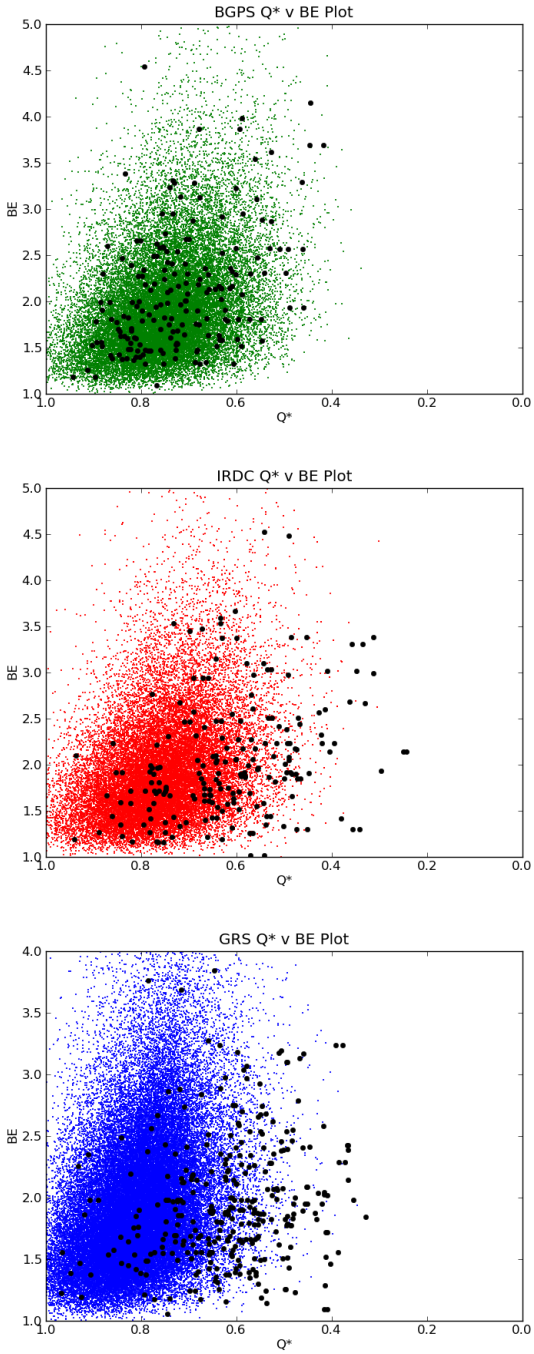


Figure 4.12: Q* v BE

Q* v BE plot for the P&F, BGPS and GRS convex hulls. The coloured background in each plot represents the Monte Carlo results for that datase with the black dots representing the live data. This is an effective plot of clustering against elongation. We see that the IRDC and GRS sample is strongly differentiated from the Monte Carlo data, whilst the BGPS data is less so. This most likely do to the contamination of the BGRS by dusty objects and HII regions.

The cross matching of the Roman-Duval cloud catalogue (Roman-Duval *et al.* 2010) with the convex hulls has allowed the identification of the approximate distance to 456 Hulls (see Appendix). From this I produced histograms of the hull semi major axis for the GRS, P&F and BGPS datasets, Figure 4.13.

This illustrates that the hulls have a typical scale for the IRDC ($\sim 6\ pc$), GRS objects ($\sim 10\ pc$) and BGPS objects ($\sim 8\ pc$).

4.6 Interpretation

It is clear from the KS test results (Table 4.2) that the GRS and P&F MST are tracing real structure within the star formation environment and that the clustering measurements Q, Q^* and $\bar{s}$ are all indicative of high levels of clustering within the GRS.

It is interesting to note that the GRS elongation parameter BE shows a very strong association with the associated Monte Carlo dataset. At first glance this seems to refute the assumption that the GRS MST is tracing structure, however when I perform a KS test between the 2d GRS and the 2dv Monte Carlo it returns values consistent with there being no association. This suggests that the issues lies with the GRS velocity data. The GRS is the deepest of the three surveys and is the only one to have multiple objects located in the same spatial position (as discussed in the section above). I suggest that the random appearance of the elongation indicators may be due to the large number of objects at a wide range of distances blurring the adjusted effective hull radius that is used to determine BE.

The distribution of the BGPS semi-major axis, Q and BE suggest that the BGPS objects are more like the Monte Carlo results than either the P&F or GRS trees. This is highlighted by the observation that KS tests between the BGPS and the Flat GRS trees indicates that they are unlikely to be drawn from the same population. The nature of the BGPS means that it is both shallower and less sensitive than the GRS. Consequently, it will preferentially detect the more massive and/or closer objects. It will also detect colder objects than the GRS because of CO depletion due to freeze out.

An additional issue with the BGPS is that the 1.1 mm band isn't exclusively occupied by molecular clumps and therefore contamination by free-free emission *might* be an issue. Given this, it is entirely possible that the BGPS is tracing a number of populations one of which is a subset of the GRS. Hence, there is some association with the GRS, as seen from Table 4.6 but any clustering or elongation signature is being lost due to blending with other sources with differing clustering and elongation parameters.

The results from the P&F trees indicate that although there is a high degree of clustering, as measured by Q and Q^*, there is also an indication that the P&F hulls are more linear than the GRS or BGPS (Figure 4.12). We should be wary when visually inspecting the trees and hull plots for linear structures as the MST tends to produce linear structures,

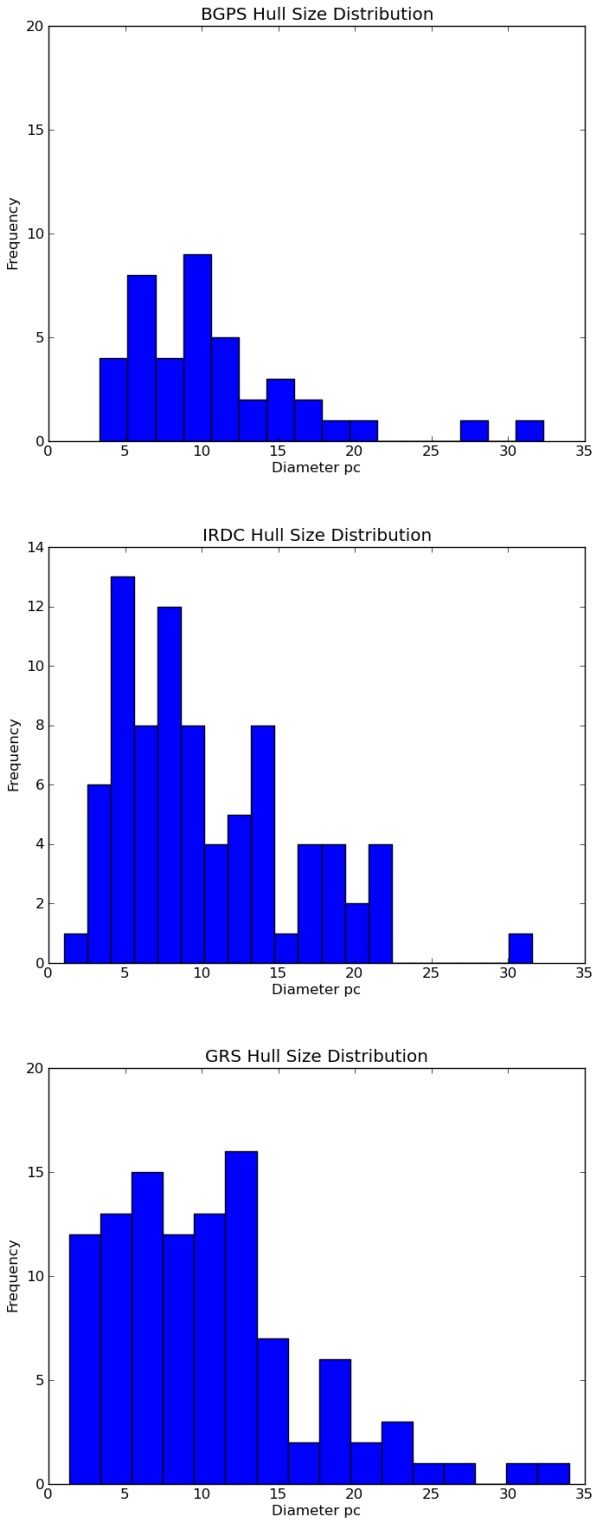


Figure 4.13: Semi Major Axis diameters for convex hulls as derived from their distance from the cross matching with the Roman-Duval Molecular Cloud Catalogue and the semi-major axis of the convex hull.

also star formation is concentrated on the plane and this will also tend to introduce elongated structures.

The edge to edge angle distribution plots for the MST for each of the three fields is confused (see Appendix and Figure 4.7) although they do not show the flatness we would expect for a random distribution Krzewina and Saslaw (1995). In general the GRS, BGPS and P&F edge angle distributions looks similar within each given field (with a few exceptions for example, IRDC 24 plot) suggesting that the three datasets have to some extent consubstantial structure. The peaking of the edge to edge angle distribution plots at $\sim 90^o$ could indicate that we are seeing structures aligned with the plane of the Galaxy. Given that the characteristic size of the hulls is considerably larger that the star formation scale height this might be interpreted as a sign of the propagation of star formation down the spiral arms. However, given that star formation tends to lie very close to the the Galactic plane and that MST have a tendency to produce linear structures it is likely that this is not the case.

Table 4.6 suggests that having the third dimension for the GRS and P&F trees is having only a minor effect. This could be for a number of reasons. Firstly, the weighting for the velocity and angular dimension parameters could be too high. This is unlikely as high weights were tested and rejected during the testing process. Alternatively, the *resolution* of the third parameters could be so low that effectively the third dimensions are binned into a small number of bins. This can also be rejected as both the GRS and P&F data have a good resolution in the non-spatial dimension. The last explanation and the most likely, is that the P&F and GRS objects cluster tightly around the third dimension and spatially. Examination of the live data suggests that this, to some extent is, true. This being the case it is likely that two dimensional MST of the P&F and GRS dataset would have performed competitively to the three dimensional trees produced but for a small amount of contamination from background objects. Notwithstanding, inspection of the convex hull plots shows that some GRS hulls overlap indicating that the 2dv is picking up some structure that would be lost in 2d plots.

On a larger scale, by examining the convex hulls and combined Stripped FMST plots we see that there are regions that are densely populated, that contain all three populations, whilst there are small pockets of objects, often consisting of a single species, located away from the bulk of the star formation and often away from the mid-plane. For example, the large GRS tree located at $l = 34.5$ $b-0.75$, the P&F tree located at $l = 20$ $b = -0.75$ or the BGPS pair of trees at $l = 38.3$ $b = -0.2$. One could attribute the phenomena of isolated objects as being due to the stripping and fracturing routine. However, examination of the raw data indicates that these are indeed isolated groups of objects. It is possible that additional species are co-located with these isolated groups and fall below the detection limit of the surveys. However, additional observations will be required to resolve this issue.

The association rate between the convex hulls and other objects within this thesis is shown in Tables 4.7 and 4.8. This suggests that IRDC and GRS and BGPS objects are similarly associated with these objects and are indeed associated with star formation.
The hull size frequency distributions, Figure 4.13, shows that the IRDC hulls have a

typical size of IRDC ($\sim 8.5\ pc$), GRS objects ($\sim 9.3\ pc$) and BGPS objects ($\sim 9\ pc$) [9]. The observation that the convex hulls of the three datasets tend not to overlap but rather reside in the same star forming region, strongly suggests that they do not always form part of a "Russian Doll" evolution sequence. Rather they are density enhancements that *may* go on to form stars. This is consistent with the observed number of IRDC and the number of objects in the GRS and BGPS which are too high to be consistent with the currently accepted Galactic star formation rate of $\sim 1\ M_{\odot}$ (Robitaille and Whitney 2010).

4.7 Summary

In this chapter I have discussed minimum spanning tree theory and application and have shown its link to the Nearest Neighbour and Voronoi Tessellation processes. I have shown and discussed the application of the MST process to three datasets; the GRS, the BGPS and the Peretto and Fuller IRDC catalogue and have shown the advantages of using an over sampling method in overcoming some of the problems associated with MST production. I show that for these datasets fracturing at the mean edge cost is preferential to fracturing at the edge cost cumulative distribution frequency inflection. I introduce the concept of fitting convex hulls to the resultant Stripped Fractured minimum spanning tree. The production of the hulls and trees allows for the extraction of a number of clustering and elongation parameters. This is in addition to the construction of large scale structure plots.

Due to the tendency of the human eye to see structure where no structure is present I undertook a large number of Monte Carlo simulations. The results of these I compared to the live data using a series of KS tests. I was therefore able to distinguish between real clustering and elongation signatures and those that are random.

Inspection of the hull plots show that giant molecular clouds contain substructure and that the IRDC, GRS objects and BGPS objects are associated with star formation. I have identified 268 GRS structures, 198 IRDC structures and 132 BGPS structures..

Statistical analysis of a number of cluster and elongation parameters when compared to the Monte Carlo data indicates that IRDC, GRS objects show clustering whilst the IRDC appear elongated. Association of the convex hulls with Roman-Duval clouds and Murray SFR gives the distance to a number of hulls, analysis of which suggests that there is a characteristic size for all three hull types. The statistical analysis of the BGPS suggests it *may* be contaminated by warm, dusty, objects.

The observation that the convex hulls of the tree datasets tend not to have similar centres local or morphologies, suggests that they do not form part of a "Russian Doll" like sequential evolution sequence. Rather they are density enhancements that *may* go on to form stars.

[9] As determined from the median value.

Chapter 5

Massive Cluster Bica 107

" Some people think that stars rule their lives. Some other people think otherwise" - I am Kloot

5.1 Introduction

As observed in Chapter 3, 26 MMB masers (3% of the population) were within 30 arcseconds of an infrared cluster. This is much higher than I would have expected from random chance as indicated by the cross matching Monte Carlo (see Chapter 3). It is currently not clear if high mass stars form first, last or concurrently with low mass stars (Beuther 2011). The presence of a Class II methanol maser within a visually identifiable cluster may provide a clue to this problem. Additionally, their existence in such an environment may be indicative of triggered star formation. Hence, the existence of Class II methanol maser associated clusters is worthy of further investigation.

In this chapter I discuss my investigation of the Class II 6.67 GHz methanol maser Caswell CH3OH 005.90-00.43 and the nature of its apparent association with the massive Galactic Cluster [BDS2003] 107 (henceforth Bica 107). I will present the results of new JCMT HARP CO 3-2 observations of Bica 107 and its surroundings in relation to previous targeted infrared observations, the UKIDSS and GLIMPSE infrared surveys, the BOLOCAM Galactic Plane Survey, MiPSGAL and the 90 cm, 20 cm and 6 cm Radio Galactic Plane surveys.

Bica 107 was selected for examination as it has been observed in detail by both Borissova *et al.* (2005) and Hanson and Bubnick (2008) and has been covered by UKIDSS, GLIMPSE and MiPSGAL in the infrared and a number of radio surveys. It is located close to the well known UCHII region G005.89-0.39 (Kim and Koo 2001; Feldt *et al.* 2003; Acord *et al.* 1998; Hunter *et al.* 2008) which has a recently revised distance obtained by parallax (Motogi *et al.* 2010).

A key challenge in the investigation of infrared clusters is the determination of their distance. This maybe done indirectly by association with objects with known distances such as HII regions or molecular clouds, which requires the existence of such an object, or alternatively, it may also be undertaken directly by the separation of clusters from foreground and background contamination and the performance of follow up spectrography in order to identify the spectral type of bright members. Unfortunately, the decontamination of clusters is non-trivial in very young infrared clusters, as I will illustrate.

5.2 Local Environment

Figure 5.1 shows the UKIDSS GPS colour composite image of Bica 107 and its surroundings. The well known Ultra Compact HII region G005.89-0.39 and the ionizing source, Feldt's Star, lies three arcminutes to the east for Bica 107 (Kim and Koo 2001; Feldt *et al.* 2003; Acord *et al.* 1998). Parallax observations of water masers associated with G005.89-0.39 with the VERA interferometer [1] (Motogi *et al.* 2010) indicates a revised distance of 1.28 kpc and spectral class of O8-O.8.5 for its ionizing source. The observation of a crescent shaped chain of water masers suggests the possible existence of a disc which is in agreement with the assumption that G005.89-0.39 contains a young massive star (Motogi *et al.* 2010).

One arcminute south east of Bica 107 lies the young cluster [BDS2003] 108 (Bica *et al.* 2003), which is yet to be independently investigated. The associated nebulosity suggests that this is a young object with a possible outflow. The 2MASS photometry of Bica 108 (Skrutskie *et al.* 2006) reports a K magnitude of 10.1 suggesting that if this cluster is at a similar distance to Bica 107, it too contains at least one massive star.

The Class II methanol maser Caswell CH3OH 005.90-00.43 Caswell *et al.* (1995) is located 43 arcseconds to the west of Bica 107. The maser had a flux of 8 Jy at the time of discovery, although the maser is variable and had faded slightly when observed later by the MMB. The radial velocity was measured at 10 kms^{-1} in the MMB.

5.3 Targeted Observations

Bica 107 was first observed in the Two Micron All Sky Survey (2MASS) (Skrutskie *et al.* 2006) search conducted by Bica *et al.* (2003). This search resulted in the detection of 167 new infrared clusters and cluster candidates. A distance of 3 kpc was determined for Bica 107 using the kinematic velocity of the associated nebula. This initial detection was followed up with observations using the 6.5 m Baade telescope by Borissova *et al.* (2005). Photometry was performed on the image using ALLSTAR and DAOPHOT II which in turn were calibrated using 2MASS photometry. 2MASS was also used for objects that were saturated in the Baade observations. From this photometry Borissova *et al.* (2005) produced a H-Ks v Ks Colour-Magnitude diagram which was statistically decontaminated

[1] VERA is a large Japanese radio interferometer.

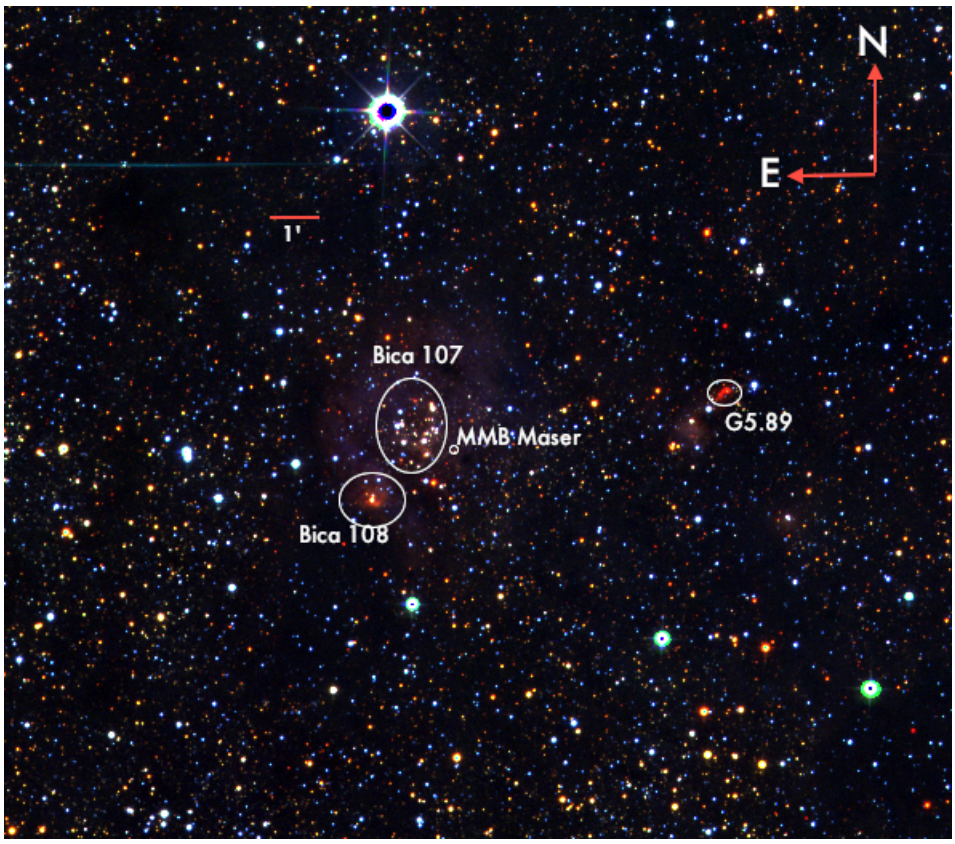


Figure 5.1: Bica 107
UKIDSS GPS image of Bica 107. Red=K, Green=H and Blue=J. To the right of the image is the UCHII region G5.89-0.39. Near the centre lies the infrared clusters Bica 107 and 108 and the Class II methanol maser Caswell CH3OH 005.90-00.43, the red compass in the top left corner show orientation while the short red line is the one arcminute scale.

(see the following section for more details) and used for extinction determination. Applying the 10th brightest star technique (Dutra *et al.* 2003), whereby the 10th brightest star is assumed to be a B0V, a distance to the cluster of 5.8 kpc was estimated. An upper mass of 4500$M_\odot$ was also calculated along with an extinction of A_v=18.2.

Spectroscopic observations of the cluster were made in 2005 by Hanson and Bubnick (2008) using the SpeX spectrograph mounted on the Infrared Telescope Facility (IRTF). This allowed the identification of the brightest cluster member as a mid-O dwarf and a revised extinction for the cluster of A_v ~25 from the colour excess. Due to problems mapping the Spex imager to 2MASS colours, no independent photometry of the cluster was made. However, from the spectral identification and extinction measurement a distance of $\leq 2kpc$ was estimated and from the observation of CO the band-head of spectroscopically identified B stars, a cluster age of approximately 10^6 years was suggested (Hanson and Bubnick 2008).

5.4 UKIDSS GPS

Given the uncertainty inherent in automated photometry within crowded fields with variable backgrounds such as that found in the vicinity of Bica 107, I undertook independent point spread function (PSF) photometry of the UKIDSS Bica 107 fields using IRAF/DAOPHOT, for each UKIDSS band using 15 sources within PSTselect [2] for the determination of the PSF. The process was repeated, with the detection thresholds adjusted in order to gain the maximum number of detections whilst avoiding false detections. The photometry was then plotted as a histogram and compared to the UKIDSS photometry of the same region. A correction factor was then applied to each band so that the peak in the magnitude distribution matched that of UKIDSS. Hence, I have *bootstrapped* my photometry onto the UKIDSS photometry. The strap values were 1.85, 1.725 and 0 magnitudes for J, H and K, respectively. Figure 5.2 shows the distribution of the bootstrapped photometry for each band.

Figure 5.3 shows the plot of the bootstrapped photometry against the UKIDSS archive (apermag1) photometry for J, H and K. As can be seen in the J band the two sets of photometry are in good agreement, however, for faint sources in H and K, the two systems scatter. This highlights the difficulty of performing precision photometry in such environments, a similar plot between 2MASS and UKIDSS shows that the 2MASS photometry deviates more from UKIDSS than the bootstrapped photometry, although this is to be expected given 2MASS's lower sensitivity, different filter profiles and resolution (Lucas *et al.* 2008; Skrutskie *et al.* 2006).

[2]PSTselect is a routine within the IRAF/DAOPHOT package design to determine the point spread function with astronomical images.

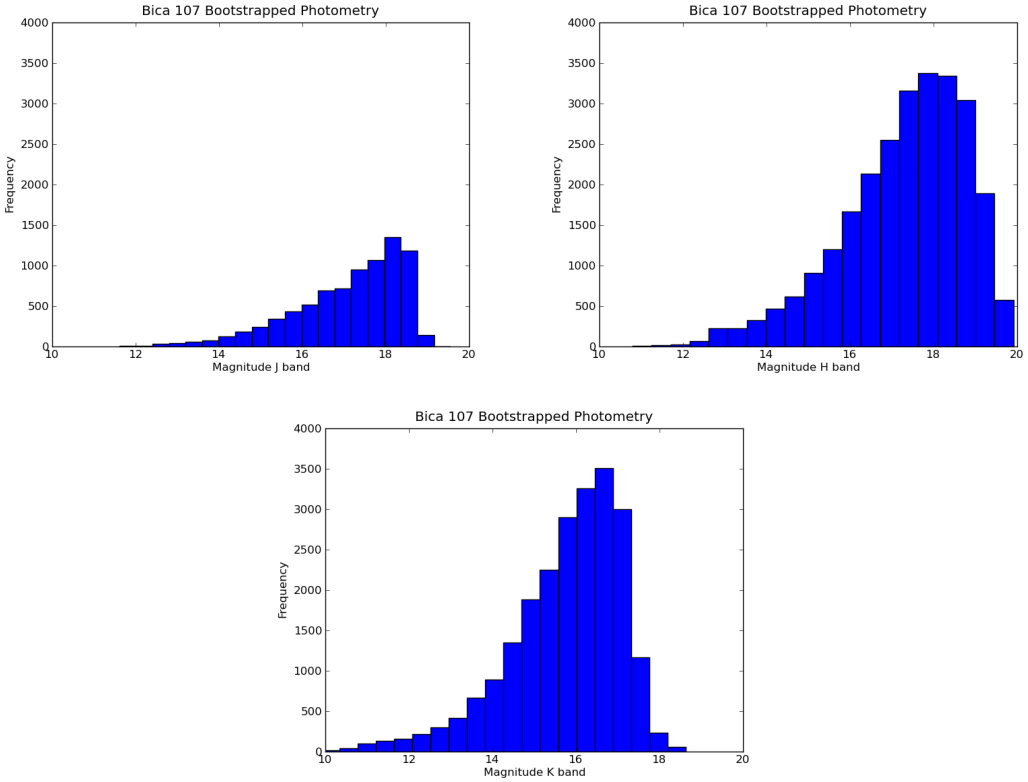


Figure 5.2: Plot showing the magnitude distribution of Bica 107 and the surrounding 6 arcminutes, for the UKIRT J,H and K filters. The magnitude are taken from my own PSF photometry from the UKIDSS images which is bootstrapped onto the UKIDSS archive photometry of the same region.

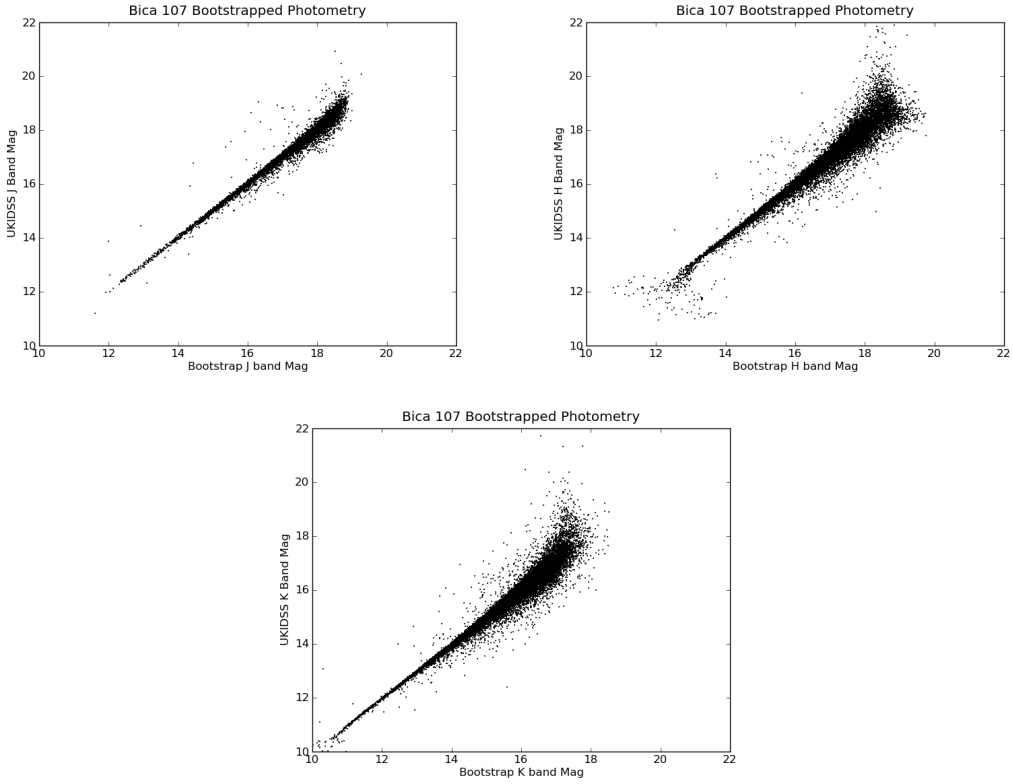


Figure 5.3: Bootstrapp v UKIDSS photometry
Plot of the my own J, H and K PSF photometry for 6 arcseconds around Bica 107 using the UKIDSS images. This images compared this photometry to the UKIDSS automated photometry for the same region. The deviation at lower magnitudes is most likely due to the limits of the standard UKIDSS photometry system at the time this work was undertaken.

From my bootstrapped photometry I produced the colour-magnitude and colour-colour diagrams shown in Figure 5.4. The colour-magnitude diagram shows a possible reddened group between H-K = 1 and H-K = 1.8, whilst the colour-colour digram of the same region, 5.4, shows a possible break at H-K = 1, J-H = 1.1. However, in both diagrams objects determined as cluster members in Hanson and Bubnick (2008) appear outside this range.

If we assume that the ionizing source within Bica 107 is indeed an O5, as derived spectroscopically (Hanson and Bubnick 2008), then we only need to determine the extinction to derive the distance. However, the derivation of extinction is non-trivial in such a crowded and young star formation region.

The traditional approach to identifying cluster members would be statistical decontamination. In this method the photometry is divided into two regions. Region A encompasses

the cluster while region B covers an annulus enclosing but not including the cluster. The areas of regions A and B are the same. A colour-colour diagram is then plotted for A and B. The colour-colour diagram for both regions is then divided into cells. We then count the number of objects in each cell in each field and then randomly remove the number of objects in each in B from the corresponding cell in A (Borissova *et al.* 2006). The implementation of the procedure is not possible in Bica 107 as the number density inside the cluster is lower than that outside the cluster due to the high levels of extinction associated with the cluster region and the high surface brightness of the nebulosity.

The alternative NICE and NICER algorithms outlined in Lada *et al.* (2007) could be implemented. However, both methods require that the level of extinction is relatively constant (Lombardi 2009) which does not appear to be the case for Bica 107.

The procedure used in Froebrich *et al.* (2010) to identify the red clump branch in old stellar systems uses a probability of membership function (PCCM) which may be able to identify cluster members. The procedure requires the selection of two regions, the cluster region and a control region located away from the cluster that is representative of the local background.

The colour space separation, r_{ccm}, between every object in the cluster region is then determined using Equation 5.1 where i is the first object and $j \neq i$. We find the 10th lowest r_{ccm} R^{10}_{ccm} within the field, for each member.

$$r_{ccm} = \left(\frac{1}{2}[J^i - J^j]^2 + [JK^i - JK^j]^2 + [JH^i + JH^j]^2\right)^{1/2} \tag{5.1}$$

The probability for each object being a cluster member, P_{ccm} is then determined from Equation 5.2, where $Nccm$ is the number of objects within the control area under R^{10}_{ccm} for the object in question, A_{cl} is the area of the cluster region and A_{com} the area of the control region. Where the P_{ccm} is less than zero, it is treated as zero.

$$P_{ccm} = 1.0 - \frac{N_{ccm}A_{cl}}{10A_{con}} \tag{5.2}$$

Applying this function to Bica 107 results in the colour-magnitude diagram Figure 5.5. The upper group at K$<$ 14 contains the objects identified in Hanson and Bubnick (2008) as spectral type B2 (noted as objects 17 and 79 in Hanson and Bubnick (2008)). The lower group at K=15 and J-K are possible low mass members of this group. However, the scatter in extinction values using this method is still large. Only objects with PCCM$>$ 0.5 are plotted. Froebrich *et al.* (2010) rejects the introduction of a spatial component into the test stating that it makes the test unreliable for clusters in dense regions. Alternative PCCM cutoffs do not significantly improve the isolation of clusters members within the CMD.

An alternative method is required that uses both position and extinction to assess if a star is a cluster member. The Minimum Spanning Tree (MST) discussed in Chapter 4 is one such method. For the implementation of MST on Bica 107 I used the l, b positions

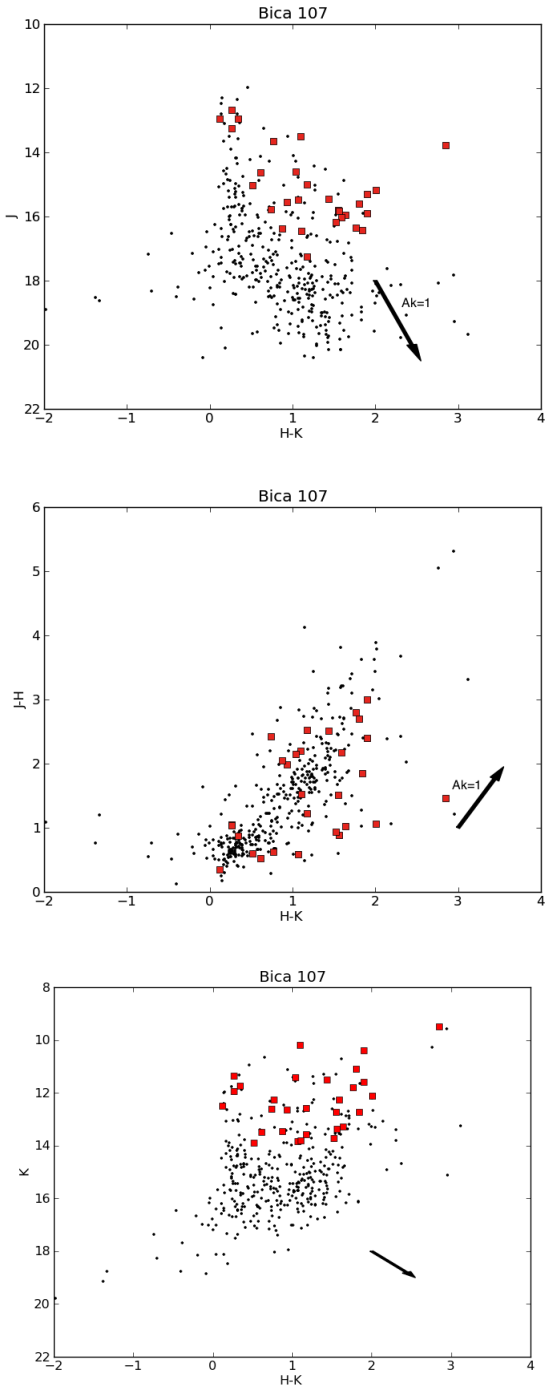


Figure 5.4: Bica107 CCD & CMD

Colour-magnitude and colour-colour diagrams of Bica 107 using bootstrapped photometry. The reddening vector represents $A_k = 1$ and is based on Gutermuth (2005). In the CMD (left) we see a reddened population which we also see as a break in the CCD (right) at $J - H = 1$. We suspect that this change in reddening in due to a molecular cloud associated with the wider star forming environment. The red squares represent the 2MASS photometry used in Hanson and Bubnick (2008)

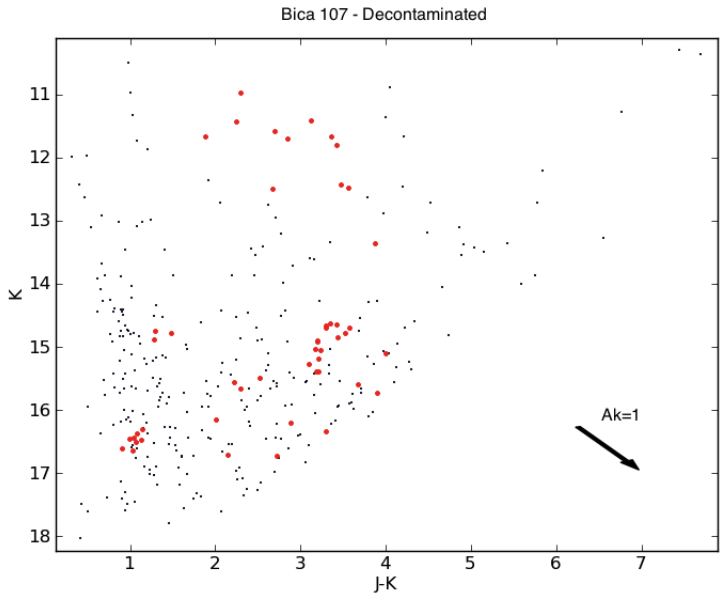


Figure 5.5: Bica 107 decontaminated
J-K v K colour-magnitude diagram of Bica 107. Black dots show the position of all cluster members while the red circles show the objects with a cluster membership probability greater than 0.5. The B2 stars identified in (17 and 79) are identified as cluster members and lie in the group at the upper part of the plot. The B3, also in the cluster (89) is not identified. The central O5 (25) is saturated in the UKIDSS K band and hence is not plotted. The reddening vector is based on Gutermuth (2005). This image was constructed using the decontamination code developed by Nithin Shajan and Philip Lucas with my own Python wrapper.

and the A_k, extinction as measured from my photometry to form 2dE parameter space. I assume that a main sequence, unextincted stars have an $H - K$ colour of 0.2, hence A_k can be measured as $H - K - 0.2$. The MST produced was fractured at the mean and stripped as discussed in Chapter 4. Figure 5.6 shows the FMST for Bica 107. The process identifies the central core of the cluster, marked as sub-tree A, but fails to include known cluster members such as the B3 marked in red. Extinction levels in sub tree A appear to be Ak $\approx$ 1.2 to 1.75, however it is possible that items with much higher or lower levels of extinction, due to being deeply embedded, would not be included as part of this sub tree even if they are spatially close. This is shown by Bica 107 (17) which appears to be a cluster member but its relatively low extinction means that it is excluded from the main sub-tree. The tree is weighted such that one arcsecond of separation corresponds to $0.2A_k$ of extinction.

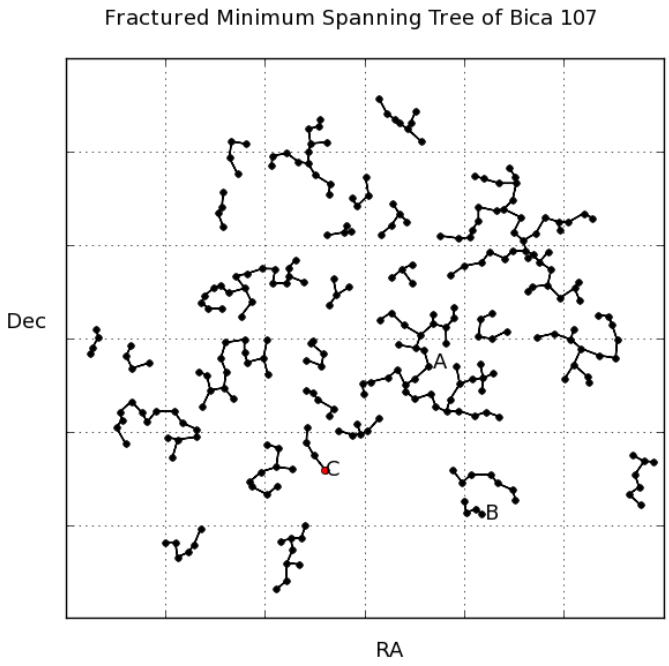


Figure 5.6: Bica 107 FMST
The Fractured Minimum Spanning Tree for Bica 107 and the surrounding one arcminute, based on UKIDSS photometry. I use RA,DEC and extinction to form the MST and collapse the extinction axis to form this plot. The sub tree marked A maps to the central part of the cluster. The two sub trees marked B are associated with Bica 108, whilst the red dot marks the position of the B2 star Bica 107 (17).

5.5 GLIMPSE & MiPSGAL

The two arcminute query of the GPSC yielded 482 detections of which only 19 have detections in all four IRAC bands. Of these 19 the majority are found outside of Bica 107. The Spitzer IRAC image taken as part of GLIMPSE (Figure 5.7) for the region surrounding Bica 107 is dominated locally at 8.0 and 5.8 μm by Polycyclic Aromatic Hydrocarbon (PAH) emission from the bubble being generated by Bica 107 and at larger scales by the bubble, CN72 (Churchwell *et al.* 2007). In IRAC bands 1 and 2 where PAH emission is less dominant, the individual stellar components become distinguishable.

The GLIMPSE field for Bica 107, Figure 5.7, contains numerous saturated sources with limited photometric detections. Morphologically we see a large extended 8.0 μm source approximately 15 arcminutes in diameter which encompasses Bica 107 and 108, G5.89 and Caswell CH3OH 005.90-00.43. At 4.5 and 3.6 μm we see nebulosity associated Bica 108 which may be indicative of a blown bubble. Bica 108 lies on the edge of this possible bubble and is seen as an extended source in all four IRAC bands. There appears to be a NE-SW bipolar outflow associated with Bica 108 although we should not rule out that

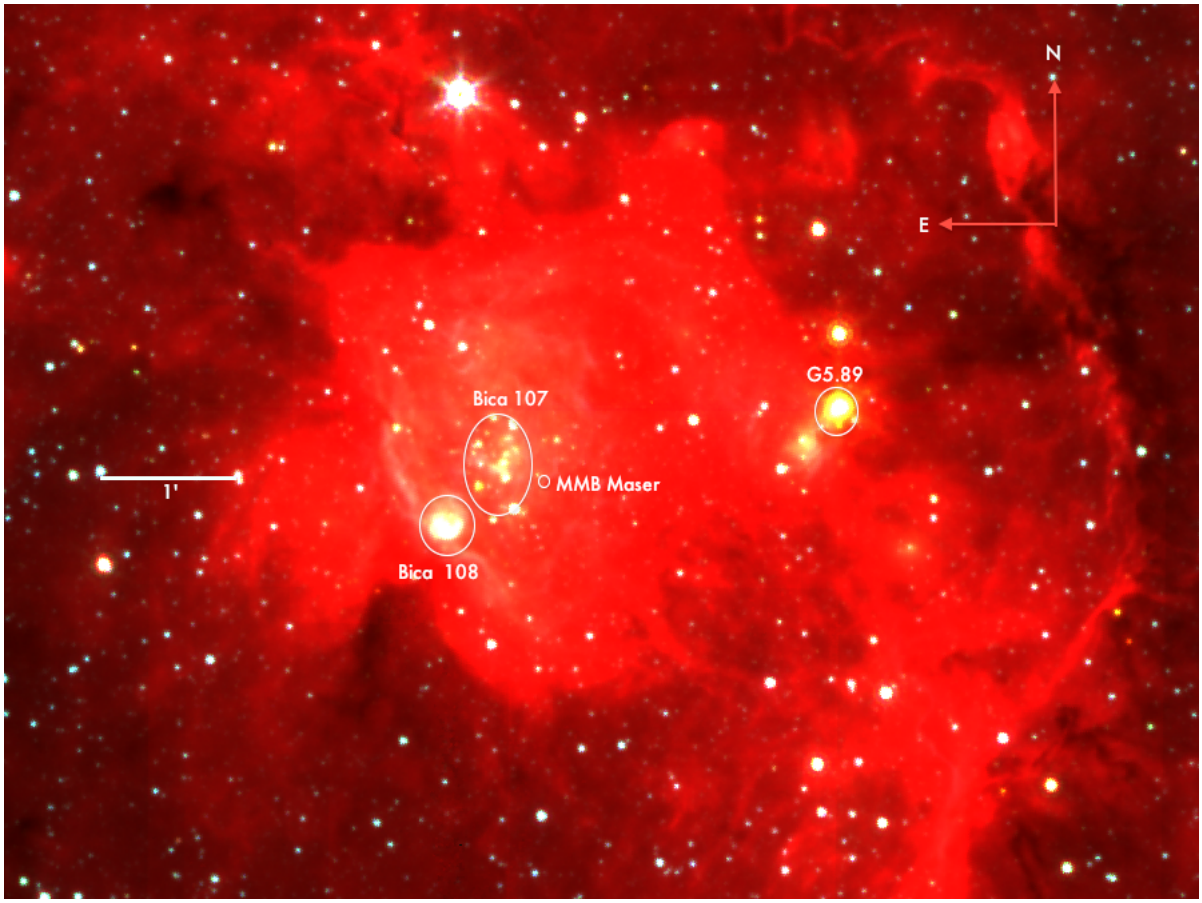


Figure 5.7: Bica 107 GLIMPSE
Colour composite image of the young infrared cluster Bica 107 (centre) with the cluster Bica 108, the UCHII region G5.89 and the Class II methanol maser, Caswell CH3OH 005.90-00.43 also shown. The images from the GLIMPSE survey. Red is 8.0 μm, green is 4.5μm and blue 3.6μm. The white bar is the one arcminute scale.

this may be instead a filament illuminated by Bica 107, which would appear to be more likely given the lack of a velocity gradient associated with the extended source. Caswell CH3OH 005.90-00.43 appears to lie on the edge of the Bica 107 bubble.

I obtained the 24 μm MiPSGAL image covering the Bica 107 region from IRSA. The 24 μm source catalogue is yet to be published. However, from Figure 5.8 we see that all the targets under investigation in this paper are within regions of saturation.

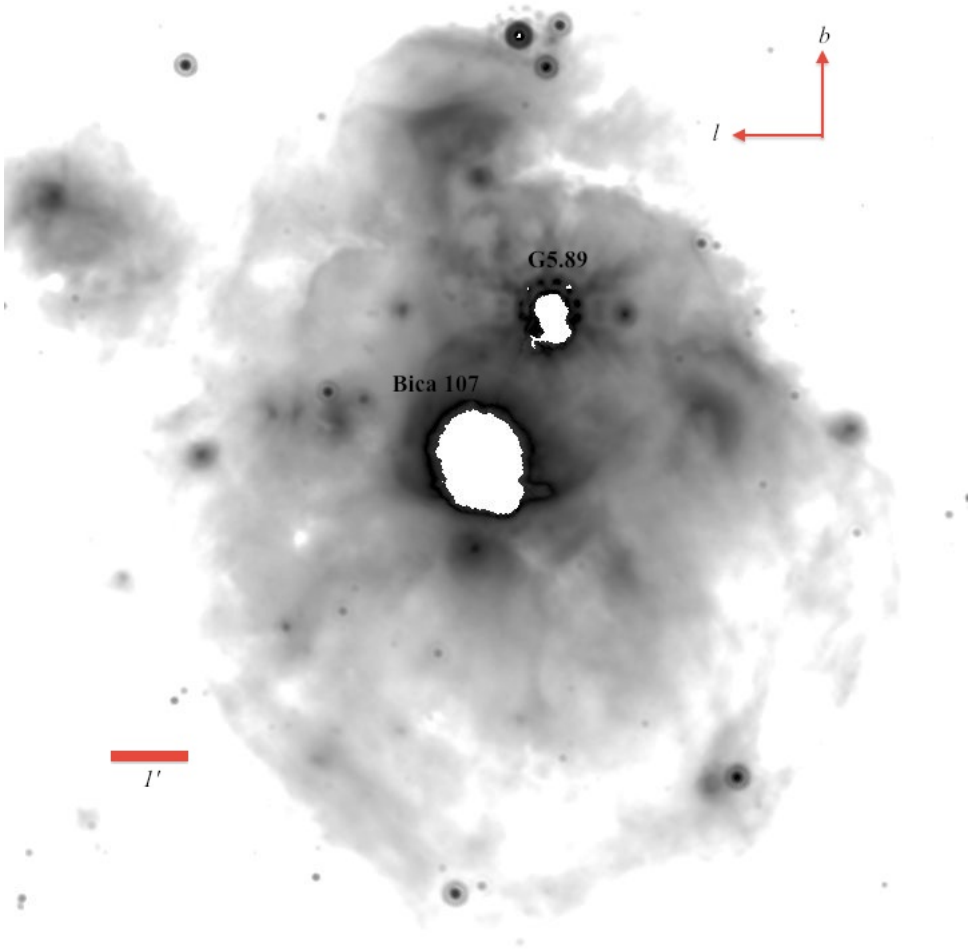


Figure 5.8: Bica 107 MipSGAL
24 μm MIPSGAL image of Bica 107 (bottom) and G5.89, (top); both sources appear saturated. The inrared cluster, Bica 108 and the methanol maser Caswell CH3OH 005.90-00.43, are located with by the Bica 107 saturation. The red line is the one arcminute scale.

5.6 Millimetre and Radio Observations

As discussed in Chapter 1, the BOLOCAM Galactic Plane Survey (BGPS) is a 1.1 mm continuum survey of the Galactic plane using the BOLOCAM instrument mounted on the

CSO. The survey resolution is 33 arcseconds with a 1σ sensitivity of the order of 11 to 53 mJy (Aguirre *et al.* 2011). The BGPS is sensitive to cold dense dust, such as that found in areas of potential star formation, and hence acts as a sign post to these regions. Figure 5.9 shows the BGPS synthesised image for the region surrounding Bica 107.

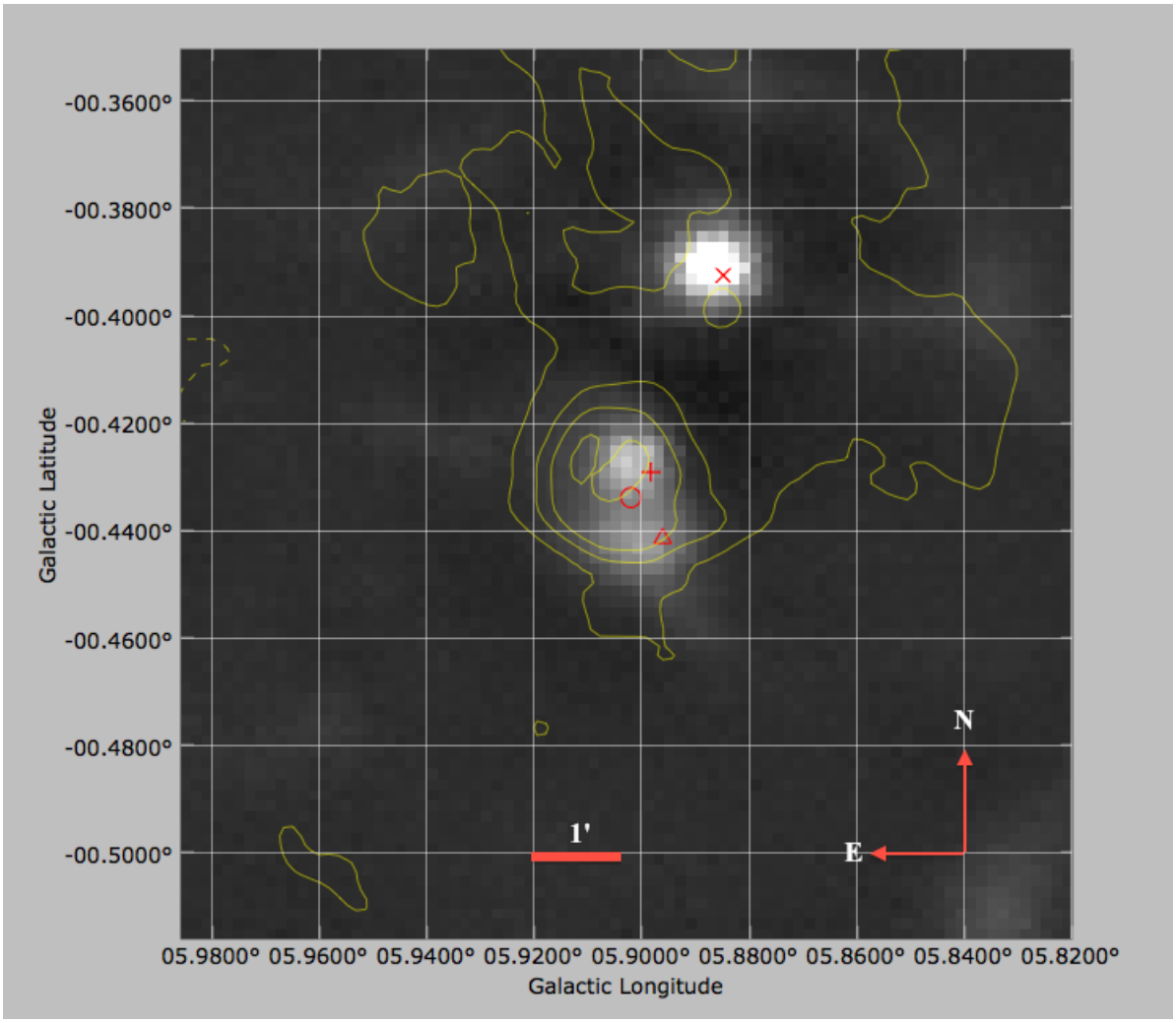


Figure 5.9: Bica 107 at 1.1 mm.
BOLOCAM 1.1 mm continuum image of Bica 107 and its surroundings overlaid with the 90cm VLA contour. The circle marks the position of infrared cluster Bica 107, the triangle nfrared cluster Bica 108, the plus sign shows the location of the Class II methanol maser CH3OH G005.90-00.43 and the X G5.89. The contour lines are drawn to identify the sites of peak radio emission. The red line is the one arcminute scale line.

The 1.1mm BGPS image shows three objects coincident with Bica 108, Bica 107 and G5.89 respectively. There is no clear indication of a BGPS source associated with Caswell CH3OH 005.90-00.43; however, the maser lays at the edge of the source associated with Bica 107.

Bica 107 has been observed by the VLA at 6 cm in configuration C (White and Helfand 2005), 20 cm (White and Helfand 2005; Helfand *et al.* 2006) in configurations B, C & D and 90 cm (Helfand *et al.* 2006) in configurations B & C. At these wavelengths the detected flux is mostly free-free emission from the HII regions associated with Bica 107, Bica 108 and G5.89 as the alternative extended free-free sources such as SNR and PNe are to be found in star forming regions. Figure 5.10 shows the radio composite at 6, 20 and 90 cm as red, green and blue respectively. The 6 cm image is truncated close to the edge of Bica 107 and appears to be resolved out. For these reasons I consider the 6 cm flux measurement of the HII region associated with Bica 107 to be unreliable. Resolution is 25 arcseconds at 90 cm, and approximately six arcseconds at 20 cm, with a sensitivity of $1-2\ mJ$ per beam.

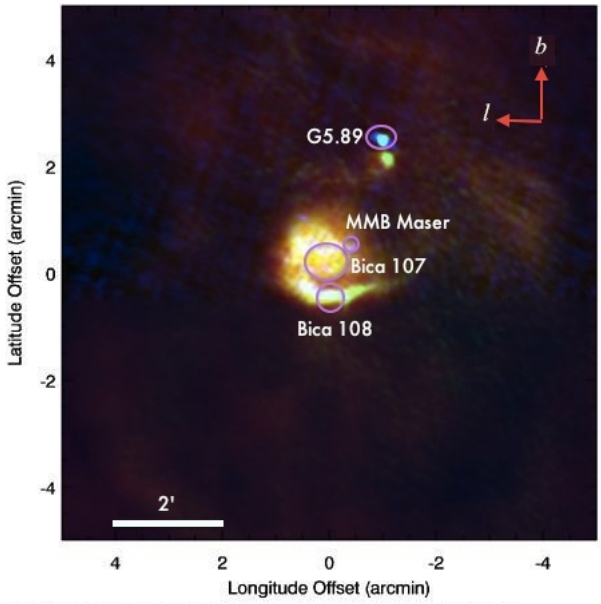


Figure 5.10: Radio colour composite image of Bica 107.
Also in this image is Bica 108, the Class II methanol maser Caswell CH3OH 005.90-00.43 and the UCHII region G5.89. Red is 90 cm, green 20 cm and blue 6 cm (White and Helfand 2005; Helfand *et al.* 2006). The 6 cm image is truncated near to the edge of Bica 107. The white line is the two arcminute scale line.

Within the VLA 90 cm we observe two sources of free-free emission, the UCHII region associated with G5.89 and the emission associated with Bica 107 that appears to have

overrun Bica 108. Both HII regions are centred on their proposed ionizing sources.

If we assume that the region being ionised is dense enough that the ionizing photons cannot escape then the radio flux from HII region is dependent on the number of Lyman α photons being produced by the ionising source, which is in turn highly dependent on spectral type (Panagia (1973), Stahler and Palla (2004) Section 15.1.1). Consequently, the determination of spectral type using free-free emission from an HII region is relatively insensitive to distance and the presence of multiple ionising sources.

I can therefore use the 90 cm flux ratio between G5.89 and Bica 107 to determine the Lyman α ratio and from this, attempt to confirm the previously published spectral types. Measuring the fluxes from the VLA 90 cm image using SAO DS9 and the Funtools plugin I determined that the 90 cm flux ratio between G5.89 and Bica 107 is of the order of 16.5:1 $\pm$4.

Equation 5.3 (Carpenter *et al.* 1990) shows the relationship between free-free emission and Lyman α flux, where v is the frequency being observed, S_v is the measured flux and d is the distance to the source. For the measurements taken for G5.89 and Bica 107, d and v are the same so the ratio in Lyman α between the two objects should be the same as the 90 cm flux ratio, if we assume that both objects are optically thin at 90 cm.

$$N_i = 9.0 \times 10^{43} \left(\frac{S_v}{mJy}\right) \left(\frac{d}{kpc}\right)^2 \left(\frac{v}{5\ Ghz}\right)^{0.1} photon\ s^{-1} \tag{5.3}$$

Table 5.1 shows the expected Lyman α flux over a range of spectral classes (Panagia 1973). We can assume that given both objects have HII regions they are Zero Age Main Sequence (ZAMS) stars. Assuming that Hanson and Bubnick (2008) is correct in their identification of the most massive star as a O5, which from Table 5.1 has $N_i = 10^{49.62}$, by applying the measured 90 cm flux ratio we find that G5.89 should contain an object with $N_i \sim 10^{48.4}$ which equates to a O8 which is in agreement with Motogi *et al.* (2010).

Log N_i				
SP	ZAMS	V	III	I
O4	49.93	49.93	49.93	49.93
O5	49.62	49.71	49.71	49.77
O5.5	49.35	49.5	49.53	49.66
O6	49.08	49.02	49.15	49.43
O6.5	48.82	49.02	49.15	49.43
O7	48.62	48.86	49.05	49.37
O7.5	48.51	48.70	49.98	49.34
O8	48.35	48.59	48.90	49.30
O8.5	48.21	48.45	48.83	49.22
O9	48.08	48.32	48.78	49.12
09.5	47.84	48.08	48.53	48.97
B0	47.36	47.63	47.94	48.53
B0.5	46.23	46.50	46.80	47.60
B1	45.29	45.52	45.87	46.78
B2	44.65	44.89	45.25	46.18
B3	43.69	43.91	44.30	45.57

Table 5.1: Lyman α photons by spectral type. Table taken from Panagia (1973) Indicating the number of Lymanα photons emitted per second for a range of spectral classes.

5.7 Sub millimetre observations

In May 2009 I undertook sub millimetre observations of Class II methanol maser sites which included Bica 107 and the surrounding region. These observations were undertaken with the HARP-B instrument mounted upon the James Clark Maxwell Telescope (JCMT and were undertaken as part of Project M09AU13 which called for ^{12}CO, ^{13}CO and C^{18}O J= 3 $\longrightarrow$ 2 observations of eight star forming complexes. Weather limitations resulted in only part of the Minimum Schedulable Block (MSB) being observed during the May run. The balance of the MSBs were placed in the queue. At the time of writing a number of M09AU13 targets have yet to be fully observed, which includes Bica 107. Hence, for this reason, this section deals only with the ^{12}CO observation during May 2009.

5.7.1 Observation

The observation consisted of a pair of orthogonal basket weaved raster-scan maps 17x17 arcminutes in size, performed in position switching mode. In the aforementioned mode the telescope scans the survey area in a series of parallel tracks, moving every $\sim 30\ s$ to an offset position, free of CO emission, to be used as a sky reference. Once the initial scan is complete a second scan is made perpendicular to the first. The ACSIS pipeline converts the two scans in NDF[3] datacubes with axis one and two being spatial and axis three velocity, for visual inspection during the run. During the data reduction process (see below) the two datacubes are "basket-weaved" to form a single cube with reduced noise when compared to the individual cubes. The beam size of JCMT is 14.6 arcseconds for the ^{12}CO observations. Tau at the start of the run, was determined from the CSO to be 0.115 which rose to 0.116 by the end of the observation. This represents Grade 4 weather. During this observation it was noted that there was some movement in the K mirror, which was attributed to relaxation of the telescope due to the recent removal of SCUBA2[4]. IRAS 16293-2422 was used as a calibrator and total integration time for the single observation of Bica 107, was 1475 s.

5.7.2 Data Reduction

The initial cubes were reduced by the ORACDR_ACSIS[5] pipeline by Canadian Astronomy Data Centre (CADC) prior to download. The pipeline removes bad pixels and creates the basket-weaved datacube using the STARLINK programs MAKECUBE and SETBB and ORACD. and outputs the data as a FITS datacube.

Once downloaded the FITS cube was converted back to a NDF format for use within STARLINK. Pixels located around Bica 107, Bica 108, Caswell CH3OH 005.90-00.43 and G5.89 were selected and the spectra associated with these pixels extracted using GAIA.

[3]NDF is a image file format used by the STARLINK package.

[4]SCUBA2 is a next generation BOLOMETER.

[5]An automated script written by the Joint Astronomy Centre Hawaii in for the production of semi-reduced data for initial viewing.

The velocity axis was then re-binned to 5 kms^{-1} per voxel using COMPAVE[6] in order to reduce noise. This was followed by the removal of the baseline using MFITTREND[7] and SUB[8] , again to reduce noise. The image was then chopped in the velocity dimension so that the velocity range covered was reduced to ± 15 kms^{-1}, the region of emission as determined by the examination of the spectra. This datacube was then examined in order to determine the large scale CO morphology of the region. The chopped image was then used to create the channel map (Figure 5.12) using CHANMAP[9] and a 3D rendering using GAIA[10]. It was then collapsed and converted back to FITS to create the white light image 5.11 which was produced using the aplpy Python module.

5.7.3 Interpretation

The HARP observations show a complex interconnected structure, particularly between 0 and 10 kms^{-1} (See Figure 5.12).

More locally, the ^{12}CO 3-2 spectrum at the peak of the 90 cm emission (Helfand *et al.* 2006) located over Bica 107 and hence centred on the massive star Bica 107 (25), shows a single line peaking at 6 kms^{-1} with a FWHM of 4 kms^{-1} which can be seen as a bright spot in the middle right frame of Figure 5.12. The width of this line is too high to be thermal line broadening so must be bulk motion. This suggests that we are observing a CO cocoon (Churchwell 2002; Thompson *et al.* 2006; Crowther and Conti 2003) being driven out by the expanding HII region visible in the radio and coincident, although not entirely so, with the CO emission.

The spectra located at Bica 108 contain two peaks; a strong peak at 6 kms^{-1} which appears to be the overrunning Bica 107 HII region, and a second smaller narrower peak at 14 kms^{-1} which may be the HII region surrounding Bica 108, the Bica 108 outflow or, less likely, a background object. G5.89 has been discussed in great detail elsewhere ((Hunter *et al.* 2008; Acord *et al.* 1998; Feldt *et al.* 2003; Thompson 1999) and references therein). However, we note that the CO 3-2 spectrum at the peak of the 90 cm emission shows the well known double feature associated with G5.89's massive outflow. This double peak appears centred on 11 kms^{-1}. The class II methanol maser Caswell CH3OH G005.90-00.43 lies close to the edge of the CO emission associated with Bica 107. Its CO spectrum contains one line which is presumably the emission line from the CO cocoon associated with Bica 107.

[6]COMPAVE is a STARLINK package which averages the values in rectangular boxes which reduced noise and file size.

[7]A STARLINK routine which fits a line to a dataset.

[8]SUB is a STARLINK routine for subtracting, amongst other things, baselines.

[9]CHANMAP is a STARLINK routine that creates a channel map fF by compressing slices along an axis.

[10]GAIA is a FITS and NDF display package.

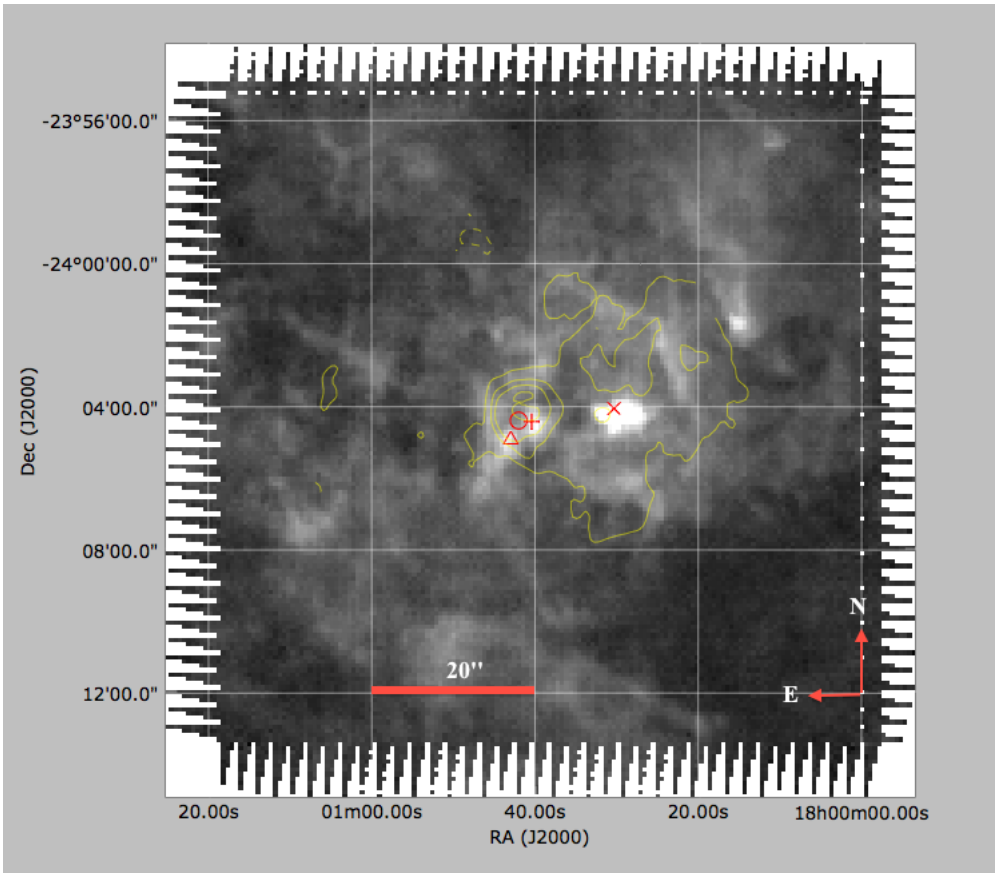


Figure 5.11: JCMT HARP CO 3-2 image of Bica 107 and its environment collapsed between -15 and 15 kms^{-1} (the central peak associated with Bica 107). The overlaid contours are generated from the VLA 90 cm map of the same region using aplpy and are shown in order to identify the peak of the 90 cm emission. The circle marks the position of Bica 107, the triangle Bica 108, the plus CH3OH G005.90-00.43 and the X G5.89. The red line is the 20 arcsecond scale line.

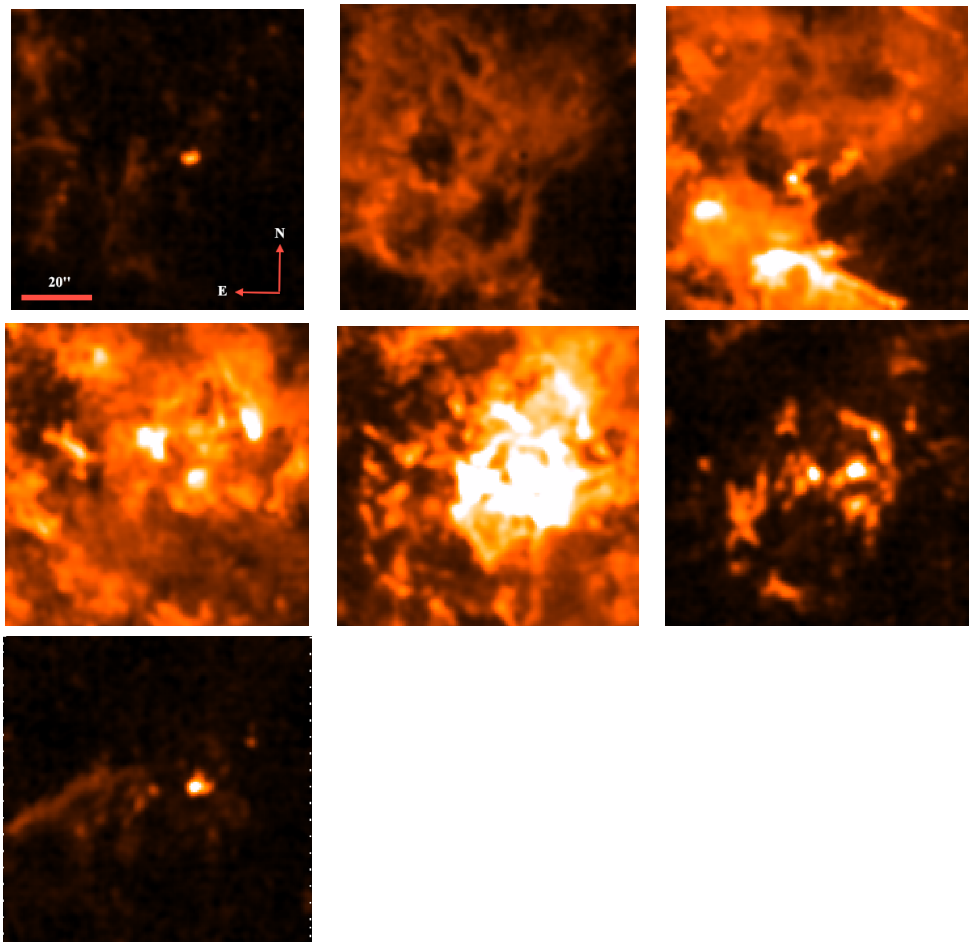


Figure 5.12: Bica 107 Channel Map
Mapping the 12CO integrated in 5kms^{-1} blocks covering $-$20kms^{-1} top left, to 15kms^{-1} bottom right. G5.89 is particularly prominent appearing as a bright source in four slices covering 20kms^{-1}. The scale and orientation is the same as Figure 4.11. The redline is the 20 arcsecond scale line.

5.8 Discussion

From my Minimum Spanning Tree, Figure 5.6 and Froebrich decontaminated colour-magnitude diagram Figure (5.5) it is clear that we cannot obtain a realistic extinction value towards Bica 107. Therefore, any attempt to obtain a distance using a method such as the 10th bright star technique is flawed. Standard statistical decontamination requires a nearby region with no local extinction that is representative of foreground and background density. In the case of Bica 107 there is no such region that we can use to decontaminate the cluster as it it located in a much larger star formation structure.

The NICE and NICER methods require that local extinction is uniform. For clusters more evolved than Bica 107, extinction uniformity is likely. However, in the case of very young

clusters, where the members are at differing stages of evolution, extinction uniformity cannot be guaranteed. The Froebrich PCCM function allows us to identify the most likely cluster members based on their colour but does not take into account any spatial information. The Minimum Spanning Tree method allows us to identify cluster members using their position in both colour space and positionally. However, it also fails to successfully isolate cluster members from the background. Significantly, the photometric uncertainties in regions of high, often variable extinction, that additionally are very crowded, are high. This high level of uncertainty in turn leads to high levels of uncertainties in individual stellar extinction values and therefore contributes to the difficultly in decontamination.

Kinematically derived distances are problematic for targets within the inner Galaxy where we have the distance ambiguity; this is further complicated for targets close to the Galactic center where the radial velocity is small and systematic errors are of the order of kiloparsecs. In Acord *et al.* (1998) a distance is derived to G5.89 by using the observed expansion rate and the observed line width. However, this method requires that the line profile be Gaussian and that the line broadening contributed by turbulence and bulk motion is less than twice the thermal contribution. It is assumed that the G5.89 Ultra Compact HII region is therefore similar to other more evolved HII regions. However, we see from Motogi *et al.* (2010) that G5.89 is extremely young and we may not be able to assume that its characteristics are consistent with more evolved objects. The VERA parallax measurements of associated masers towards G5.89 gives a measurement independent of kinematics and object evolution, although there is the possibility that the masers observed are not associated. However, the Motogi *et al.* (2010) distance to G5.89 of $1.28^{+0.09}_{-0.08}$ kpc appears to be the most reliable to date.

From the new CO observations taken as part of this thesis and outlined above, it is clear that Bica 107 and 108 and G5.89 are associated both in projection and in velocity. Hence, we can safely assume they are part of the same star forming complex and therefore at approximately the same distance. The re-categorisation of Feldt's star as approximately an O8.5 (Motogi *et al.* 2010) is consistent with our measurement of the flux ratio at 90 cm of the G5.89 and Bica 107 HII regions. This re-categorisation was made in part due to the revised distance of $1.28^{+0.09}_{-0.08}$ kpc to G5.89 determined by the VERA parallax observations of Motogi *et al.* (2010).

I can therefore revise the distance to Bica 107 and 108 to $\sim$ 1.28 kpc. At this distance the complex will be less that 1.5 pc in diameter. Caswell CH3OH 005.90-00.43's velocity, consistent with it being associated with Bica 107, indicates that it is a site of further star formation. Its position close to the edge of the Bica 107 HII region suggests that the massive star formation associated with Caswell CH3OH 005.90-00.43 is being triggered by Bica 107 (see Chapter 1) and that this is the first sign of a second generation of massive star formation within the cluster. The maser does not represent continuing massive star formation within the cluster, rather is an example of very local triggering. Due to the saturation of the MiPSGAL image we are unable to confirm the existence of a 24 μm source coincident with Caswell CH3OH 005.90-00.43; however, I suggest that future observations will confirm this. The observation that Caswell CH3OH 005.90-00.43 is not associated *directly* with Bica 107 needs following up for the remaining 26 masers associated with a

cluster. However, this will require extensive multi-facility observations and falls out of the scope of this thesis. It is however discussed further in Chapter 7.

From my JCMT CO observations it is clear that Bica 107 and 108, together with G5.89 and Caswell CH3OH 005.90-00.43, form an extended, very young, massive OB association, with multiple stages of massive star formation. The OB association consists of at least one mid-O and two late-O class stars, in addition to a number of B class stars and as such represents an "Orion like cluster" with a mass of the order of $10^3 M_{\odot}$. At a distance of 1.28 kpc this would be one of the nearest massive ultra young clusters. I note that in Hanson and Bubnick (2008) it is observed that Bica 107 (17), which has been identified as spectral class B2, appears to have a low extinction and apparent magnitude consistent with a distance of 1 kpc, which supports my revised distance to Bica 107. The ionized region surrounding Bica107 at 90 cm is 120 arcseconds in diameter, which at 1.28 kpc would represent a diameter of 0.7pc indicating that Bica 107 is indeed very young.

5.9 Chapter Summary

In this chapter I have investigated the nature of Bica 107 and its relationship to the Class II methanol maser Caswell CH3OH 005.90-00.43. I have shown that the methods for identification of cluster members using colours and magnitudes are unreliable given the high level of extinction associated with the Bica 107 and the highly variable nature of the extinction. I have been able to confirm that Bica 107 is spatially related to the UCHII region G5.89, the young infrared cluster Bica 108 and Caswell CH3OH 005.90-00.43, using radio and sub-millimetre observations. By doing so I have identified that Bica 107 is a very young infrared cluster at a distance of 1.28 kpc. Additionally, I can confirm that Caswell CH3OH 005.90-00.43 is not associated with the star formation in Bica 107, rather it is part of the larger star formation ongoing in this region of which Bica 107, Bica 108 and G5.89 are also a part. The location of Caswell CH3OH 005.90-00.43 seems to imply that it might be an example of triggered star formation. Hence, it appears that, at least for Caswell CH3OH 005.90-00.43, the association between maser and cluster appears to be one based on the large scale structure of star formation rather than that of a single cluster. However additional, multi-wavelength observations of other cluster related masers will be required to completely rule out the possibility that some masers are near-infrared cluster related.

Chapter 6

Discussion and Conclusions

In Chapter 2 I identified a problem caused by the lack of a GLIMPSE extended source archive. I addressed this problem by modifying the existing photometric technique from Cyganowski *et al.* (2008b). The ANCAP technique utilises all four IRAC bands in the determination of aperture size and positioning.

Indications from the initial testing and the work undertaken on MMB maser counterparts show that the technique has a photometric accuracy similar to the GPSC for unresolved objects. However, users should be aware that the process assumes that the object being measured has relatively constant colours across its surface and that the technique is designed to obtain colours rather than high photometric precision.

Not withstanding these caveats ANCAP has applications involving a number of extended sources including HII regions, supernova remnants and planetary nebulae.

The use of graph theory and especially the application of minimum spanning trees in astrophysics has increased in recent years, both as a method of detection of clustering, for example, and for the detection of structure. Within the cluster community a number of MST parameters, for example Q and Q*, have been shown to be excellent techniques for quantifying clustering in both clustering models and in live observational data.

When used for identification of large scale structure there are a number of inherent issues which are currently unaddressed. Uncertainties with the node parameters are either ignored or are addressed within the fracturing routine. Clearly ignoring the uncertainties is undesirable, however accounting for them post tree construction fails to take in to account the fact that the tree structure in completely dependent on the node parameters and small changes in those parameters may lead to large changes in the tree structure.

In this thesis I have weighed the parameter space to reflect the uncertainties in the velocity axis of the GRS dataset. This technique is discussed in Hoffmann *et al.* (2008). However, the reader should be made aware that other algorithms exist for the construction of trees

where there is uncertainty in the edge costs such as those discussed in Erlebach *et al.* (2008). Although these algorithms fall outside the scope of this thesis, clearly they should be considered when dealing with datasets where the distribution of edge costs is narrow compared to the uncertainties.

The integration of non-dimensional data within the tree has been largely ignored in astrophysics, except for the work of Cartwright (2009). Billot *et al.* (2011) slices and collapses the velocity axis in order to integrate the HiGAL velocity data within the MST. In Chapter 4 I showed that it is possible to integrate the velocity data into a minimum spanning tree by the current use of weighing. However, this does not solve the problem with the KDA. Although this is unlikely to be an issue in a field with such a low density as the MMB, it is a possibility within the GRS data where there are an order of magnitude more objects and the observing instrument has a large beam width.

One of the key problems when using spanning trees in general to identify structure is the determination of the fracturing point. In this work I fracture at the mean. This was chosen as it is the traditional method to determine the fracturing level and other methods, for example breaking at the median or a standard deviation above or below, were found to be unsatisfactory in testing. Breaking using the CDF such as used by Billot *et al.* (2011) and Gutermuth *et al.* (2009) assumes that an inflection point can be found. It is also assumed that the edge length at the inflection point is representative of the edge lengths within all clusters and not the edge length of a single cluster with a dominate characteristic; for example the cluster is very close or very large. The observation by Billot *et al.* (2011) that 90% of HiGAL objects are located within their generated convex hulls appears high and is suggestive of a problem with CDF breaking.

This brings us to a limitation of using minimum spanning trees within large scale surveys. Where there is a single cluster within the field, in general MST performs well at identifying clustering members. However, when there are multiple clusters within the field it is often impossible to have a "one size fits all" fracturing scale. Hence, performing MST analysis in large surveys by constructing a single tree can produce erroneous results. An iterative process, whereby clusters are provisionally identified so that they can be isolated and individual trees built for each cluster, may be the solution. The process, outlined in Chapter 4, whereby I use small overlapping fields may be successful in some cases but it is clear that there are still multiple clusters within the fields.

Despite these concerns, it is clear that graph theory will continue to be an important tool in determining structure within astrophysical data as further developments from mathematics and computer science are applied.

From Chapters 1 and 2 we see that Class II Methanol Masers can trace star formation whilst still deeply embedded within an IRDC or too distant to have an identifiable IRAC counterpart. They have a significant advantage over colour selection in that they are exclusively found in association with massive star formation and therefore do not need follow-up observations. They are also very bright and emit at optically thin wavelengths and hence can be observed throughout the Galaxy. However, the association between

massive stars is poorly understood and is it not clear what proportion of massive stars do not mase or exhibit multiple masers.

From the cross matching of the MMB and GLIMPSE it is clear that a large number of masers, and by implication a large number of massive star forming sites, do not have counterparts with the GPSA/C. However, inspection of the images shows that this is due to the criteria used when building the catalogue and not due to a lack of counterparts. Hence, any colour space derived from cross matching a maser catalogue with GLIMPSE is biased towards objects that are compact because the GPSC/A does not include extended objects and, hence, either young or distant ones. If the compact sources are indeed younger we would, therefore, expect there to be a colour difference between the sources within the GPSC/A and those sources that missing.

By using the ANCAP photometric technique I show that the colour magnitude space occupied by Class II methanol maser counterparts is [3.6]-[4.5]>1 and [8.0]<12. When compared to the colour magnitude space of [3.6]-[4.5] > 1.3 and [8.0] < 10, as determined by Ellingsen (2007), [3.6]-[4.5] for maser counterparts found within the GPSC/A, we see a remarkable level of agreement. Given that the maser sample provided by the MMB is an order of magnitude larger than that used in Ellingsen (2006) it is natural to get some scatter in the population but we clearly do not see any discernible colour difference between compact and extended sources.

We see that the ANCAP photometric process, which removes flux from extended objects in order to reduce contamination, is not significantly changing the colours of extended objects, which implies the maser infrared counterparts have effectively uniform colour across their surface. Secondly, we see that there is no significant colour evolution of the counterparts. The observation that the colour space occupied by maser counterparts is also occupied by Class 0 protostars and HII regions provides us with a clue. Looking at the Strawman evolution sequence (Figure 3.1) we see that maser counterparts should cover the region from the hot core to the formation of the UCHII region. Figure 1.5 indicates that at shorter wavelengths the UCHII region emission is dominated by dust emission. It is known that UCHII regions and hot cores are surrounded by a dust cocoon (Crowther and Conti 2003; Thompson *et al.* 2006). We can therefore explain the results of the photometry by assuming that we are observing the dust and PAH emission from this cocoon; the colours of that are independent from that of the hot core or MYSO, within the evolutionary boundaries of the Class II Methanol Maser.

Turning to EGOs. It is clear that at least some of these objects are tracing outflows, given their strong association with Class I Methanol Masers (Chen *et al.* 2009). However, the association with RMS objects and particularly RMS objects classified as HII regions is problematic as the evolutionary gap between Class I Masers and UCHII regions is too large for EGOs to be 90% associated with Class I Masers and to be associated with RMS objects (especially HII regions). Additionally, we would expect accretion, and hence outflow, to have ended by the time the UCHII had formed. This is further complicated by the observations of De Buizer and Vacca (2010) that suggests that EGOs may also be caused by SiO absorption. It could be argued that the RMS associated EGOs emission are the

fossil remnant of earlier outflows but without spectral observations comparing RMS and non RMS related EGO this has yet to be established. The strong association of EGOs with Class I Methanol Masers, BPGS sources and to a lesser extent GRS sources suggests that EGOs are an early evolutionary tracer. However, whether or not EGOs are exclusively massive star formation tracers is yet to be established and as such should not yet be considered a reliable massive star formation indicator.

In Chapter 3 I performed a spatial cross match between the RMS and MMB. Looking at only RMS objects that we expect to associated with massive star formation we see that 21% of masers are RMS associated as 21% of RMS objects are maser associated.

Assuming an evolution order of YSO $\rightarrow$ HII/YSO $\rightarrow$ HII within the RMS we see that maser association declines from 35% for YSOs to 14% for HII regions. This is in agreement with the maser evolution of Ellingsen *et al.* (2007) and suggests that the formation of a HII region result in the end of masing, most likely do the disruption of the masing environment. However, this observation fails to address the lack of association between RMS YSO and masers. A large population of RMS YSOs do not have associated masers, $\sim$65% . This can be explained by noting that, as stated in Chapter 1, it is possible that many MYSOs either do not exhibit masing, either because they do not have an environment conducive for masing to occur, or that masing does occur but the orientation of the coherent pathways means that the maser is not visible.

It is clear that there is a very large population of infrared bright masers that do not appear to be associated with RMS objects. Some 83% of masers are infrared bright and 62% of masers were classed as Type 3 in the visual inspection discussed in Chapter 2. Hence a significant fraction of masers show physical characteristics that would seem to infer that they should be within the RMS. Figure 3.2, which shows the GLIMPSE colour-magnitude space of sources associated with RMS objects and those not associated, show that, within GLIMPSE colour space, there is not a significant difference between the two populations. Cross matching the MMB with MSX results in 502 positive cross matches. Therefore a significant fraction of infrared bright masers are not within the MSX Catalogue. A small proportion of these are due to multiple masers falling within MSX's beam width. It is not clear why the balance are missing from the MSX Catalogue. Given the location of most of the masers close to the Galactic plane, it is possible that they fail the MSX Catalogue selection criteria due to confusion.

The number of MMB counterparts listed within the MSX is considerably higher than the number of MMB counterparts associated with RMS objects. Examining the colour cut criteria shows that 23% of objects fail the $F_8 < F_{14}$ colour cut with a small proportion, 3%, failing one of the other RMS colour cuts. The mean separation between sources that fail the colour cut and MMB masers is 10.5 arcseconds, compared to 6.5 arcseconds for those that do not. Hence, it is very likely that the colour cut is being failed due to photospheric contamination of the MSX beam by bright foreground stars. This will push the colour of the object to the blue and hence out of the RMS colour cut.

We see therefore that it is highly likely that the RMS is missing a significant fraction of

MYSOs due to the limitations of the MSX catalogue and additionally to colour contamination as a result of MSX's large beam width.

The generation of a GLIMPSE extended source catalogue, together with its narrower beam width, should ensure that a GLIMPSE based colour selected search for MYSOs does not suffer these restrictions.

In Chapter 3 I performed a cross match between the MMB and the Roman-Duval Cloud Catalogue both as a computational cross match and a visual inspection of the GRS images.

The initial observation that not all masers are in Roman-Duval Clouds is not confirmed by the visual inspection which shows that masers are located in molecular clouds. There are a number of potential explanations for this discrepancy. The beam width of the GRS is large, of the order of 0.5 *pc* at 5 *kpc* and it is possible that there are density enhancements below this scale that would not be seen by the BU-FCRO. A more likely explanation is that the edge detection algorithm used, in this case CLUMPFIND, may be set too aggressively so that regions that are lower density, but still star forming, are being excluded. This should be considered when performing cross matches with the Roman-Duval Cloud catalogue in order to derive distances. For large structures, such as the convex hulls in Chapter 4, this should not present a problem. However, for small objects such as those found in the RMS, it may be prudent to perform a visual cross match with the GRS images as well as a computational cross match with the Roman-Duval Cloud catalogue.

In Chapter 4 I introduced the concept of using minimum spanning trees in the identification of both clustering and large scale structure. This is the first time that minimum spanning trees have been used in astrophysics to determine both clustering and large scale structure simultaneously. It is also the first time oversampling has been used to address fracturing scale, and the first time in astrophysics that uncertainties have been addressed at the pre-tree formation stage.

We see from comparison between the live data and the Monte Carlos data that all three datasets show strong clustering as shown by $\bar{s}$, Q and Q^*, with the P&F IRDCs being more tightly clustered than either the BGPS or GRS objects. We are aware from Chapter 1 that when observed at great resolution (either physically or analytically by changing contouring parameters) there is a tendency for the clouds to fracture into small structures which would explain why the IRDCs are more strongly clustered despite their extended nature.

Looking at the convex hulls shown in Figures 4.8, 4.9 and 4.10 we are tempted to assume that the hulls are elongated and lie parallel with the Galactic plane. Both the P&F and BGPS convex hulls have Billot Elongation factor and ξ inconsistent with a random distribution, so we must assume that P&F IRDCs and BGPS form linear structures which, from Figure 4.3, would appear to tend to lie on the plane, however we should be cautionary as this may be an effect of scale height. The GRS KS-Test shows that the GRS is not elongated, however it is possible that any structure is being blended out by the higher distances of the GRS objects when compared to the P&F IRDCs or the BGPS.

The observation that the characteristic scale for the convex hulls of the P&F IRDCs, BGPS objects and GRS objects is similar is interesting, especially when we note that although the three dataset hulls tend to cluster together in star forming regions they tend *not* to overlap. This uniformity of hull sizes between the three tracers suggests that there exists a star formation scale below that of a GMC but above that of the IRDCs, with a scale length of the order of 5-8pc. This may be related to the large-scale shocks within GMC and the filamentary structure seen within GMC.

From the tree plots (as opposed to the hull plots) we see that although the three datasets tend to cluster together in regions, which are within GMC or at tangent points, we see that they do not always overlap. In some cases we see isolated trees with no corresponding tree from the other two datasets. We could explain the presence of such structures within the IRDC datasets as IRDCs that are yet to develop core. Alternatively these could be spurious IRDCs, regions of lower stellar density or reduced PAH emission that appears to be an IRDC, but in fact is not.

The existence of BGPS and GRS trees without IRDC trees suggests that the Russian Doll model of star formation does not always hold or that these objects are associated with IRDCs but the associated IRDC is not detected due to foreground emission. We also see that the BGPS is much more tightly correlated with MMB masers, RMS objects and EGO than with GRS objects, this despite my assertions that the BGPS is contaminated with free-free emission (although this would explain some but not all of the RMS association) and that it has a lower number density. It is likely therefore that the BGPS is a stronger tracer of massive star formation than the GRS, a property that is most likely linked to its lower sensitivity. This lower sensitivity (of an order of a factor of 20) manifests itself as tendency for BGPS source to be more massive and therefore more likely to be star forming than GRS objects.

Although the three datasets have structurally some physical similarities as determined from the minimum spanning trees we see that the existence of one does not imply the existence of the others even though all three should be tracing the very early stages of star formation. This could be due to the observational limits of the surveys. However, if IRDCs, GRS objects and BGPS objects were reliable indicators of ongoing star formation we would expect them to have a higher level of correlation than we see in Chapters 3 and 4.

As previously discussed we see that, for Bica 107 at least, the presence of a maser is indicative of a second, possible triggered, generation of star formation and shows that we are not seeing high mass star formation, after the low star formation. However, we also see that decontamination of a very young cluster is challenging given the high levels and variability of extinction. Both the colour only decontamination of Froebrich *et al.* (2010) and my own attempt, using using both spatial and colour data with minimum spanning trees, fail because of the extreme variation in extinction. Alternative applications of graph theory, such as robust minimum spanning trees, which allow edge costs to lie in an interval instead of having a fixed value, may be more successful in the decontamination of such clusters.

By using multiple wavelength observations, especially in the sub millimetre, over the larger star formation region it is possible not only to associate the cluster and maser in velocity space but also to place the cluster in the context of the wider star forming environment. In the case of Bica 107 this has allowed me to confirm not only the special type of the largest star, but also, by association with G5.89 to produce a reliable distance.

The lack of association between the three surveys and other star formation indications (Chapters 3 and 4) highlights that, as stated previously, these are not star formation indicators, rather indicators of the potential of star formation occurring. Class II Methanol Masers are an excellent massive star formation tracing being both bright and exclusively associated but it is still not clear if all massive stars mase or if the filling factor as illustrated by the RMS MMB cross match. From the same cross match we see that in all likelihood the RMS is failing to find all the MYSOs in the Milky Way because of the limitations of the MSX catalogue and because the large beam width of MSX results in contamination of MYSOs which pushes them out of the RMS colour cut.

EGOs at first sight appear to be a suitable massive star formation tracer, but the numbers known are low and their nature is uncertain. It is clear that, certainly while visible at IRAC wavebands, massive star formation sites have relatively stable colours. Therefore, a colour selected survey with follow up to determine luminosity would be the most productive approach to identifying massive star formation if the limitations of the GPSC/A can be overcome.

Graph theory would appear to have the potential to determine the large and small scale structure within the Milky Way. However, the application of graph theory in astrophysics is still in its infancy and without a fuller understanding of the tools available and how to apply them we will be unable to unlock its full potential. Not withstanding this I have shown how the creation of MST can be used to identify structure and potentially to decontaminate clusters. We see that we can identify structure with the BGPS, P&F IRDC and GRS surveys with a characteristic scale shared between all three datasets. We see that the apparent lack of correlation between the GRS and BGPS can be attributed to sensitivity, with the lower sensitivity of the BGPS preferentially identifying denser sites which are therefore more likely to be star forming. Likewise it appears that the lack of correlation between the BGPS and with GRS with IRDCs may be due to the possibility of faux IRDCs in the sample and/or obscured IRDCs.

In conclusion:-

- The IRAC colour magnitude space occupied by Class II Methanol Masers' infrared counterparts is [3.6]-[4.5]>1 and [8.0]<12, which is consistent with previous but more limited survey cross matches.
- There exists a large population of maser counterparts that are not listed within the GLIMPSE catalogues due to their extended nature.
- The colour space occupied by maser counterparts does not significantly differ between

compact and extended sources, indicating that colour does not significantly change between the counterpart becoming distinguishable and the onset of the UCHII region.

- Approximately 17% of masers have no infrared counterpart at IRAC wavelengths. Of these, 5% are embedded in IRDC.
- In the order of 20% of masers are associated with a massive star formation associated RMS object. This is consistent with the Ellingsen Strawman evolutionary diagram and indicates that masing can continue after the onset of nuclear burning.
- The observation that $\sim$ 20% of EGOs are associated with an RMS object is in conflict with the previous observations that 90% of EGOs are associated with Class I Methanol Masers.
- It is likely that the Class II Methanol Masers found associated with infrared clusters are either chance alignment or are examples of trigger star formation.
- IRDC, GPS objects and BGPS objects are not suitable star formation tracers given their large numbers and low association rate with other star formation indicators.
- There is a significant population of likely MYSOs that do not exhibit masing either due to environmental or orientation reasons.
- The RMS is failing to identify all MYSOs in the Milky Way due to the limitations of the MSX catalogue and the large beam width of MSX. This large beam width is resulting in photospheric contamination of MYSOs and thereby moving them outside of the RMS colour cut.
- Massive star formation is best traced by performing high resolution, colour selected surveys with follow up observations to determine luminosity.
- IRDCs, BGPS objects and GPS objects, appear strongly clustered and, in the case of IRDCs and BGPS objects appear to have a linear structure which may tend to align with the Galactic plane.
- The convex hull size of IRDCs, BGPS objects and GPS objects appear similar and suggests that there may be structure above the scale of IRDCs but below the scale of GMC.
- We see that Bica 107 is an example of an Orion like cluster with a Compact HII region powered by a O5V. From association with the UCHII region G5.89 I have established a distance to the cluster of 1.28 kpc.
- We see that, when applied correctly, taking into account uncertainties, edge effects and mixed clustering scales, graph theory is a powerful tool in determining structure and clustering.

Returning to my original objectives. We see that the relationship between Class II methanol masers and RMS objects is broadly in line with the Ellingsen Strawman diagram with masers and RMS objects having a slight evolutionary overlap with the processes associated with the formation of a RMS object marking the end of masing. On the other

hand, the relationship I reported between EGO and RMS objects and Class II methanol masers seems to be inconsistent with previously the reported association rate between EGOs and Class I methanol masers. Given our currently poor understanding of EGOs this inconsistency cannot currently be resolved.

The relationship as determined by K=1 NN analysis in Chapter 3 and the MST work in Chapter 4, between IRDCs, GRS objects and BGPS objects, suggests that they do not consistently form an evolutionary sequence. The association rate of these objects between other star formation indicators suggests that the BGPS are more strongly associated with massive star formation than P&F IRDCs or GRS objects. However, it is clear that all three object types have a considerable quiescent population and therefore are poor massive star formation indicators.

We see therefore, that I have identified the relationship, or in some cases a lack of a relationship, between the star formation indicators discussed in the Introduction.

In Chapter 2 I identified that maser infrared counterparts appear consistent in colour during their evolution. In the light of the association rate and evolutionary coverage of the various star formation indicators discussed in Chapter 1, I have identified colour selection as the most appropriate method for locating high mass star formation sites.

Chapter 7

Future Work

The key conclusion from this thesis is that the most promising method for the identification of mass star formation sites is the route taken by the RMS survey team for the identification of MYSO, i.e. colour selection with follow-up observations to remove contamination and to determine luminosities.

GLIMPSE should be an ideal survey for the colour selection of mass star formation sites as it has comparatively high resolution, sensitivity and coverage, however it is fettered by the lack of an extended source catalogue, for which there are currently no plans to produce. I have demonstrated that it is possible to determine the flux of extended sources with GLIMPSE with some ease. However, the definition of source boundaries is extremely challenging especially in regions with background emission. This is not a problem confined to astrophysics with similar problems occurring in medical imaging and computer science. A number of potential solutions exist, such as, for example, the algorithms of Leung *et al.* (2003). An alternative approach may be to train a neural network to identify extended objects using the ANCAP photometry discussed in Chapter 2. A similar approach is being used by the Galaxy Zoo team to characterise galaxies (Miller and Lintott 2010).

The GLIMPSE region has been (or is being) surveyed by a number of additional surveys, including MiPSGAL, Hi-GAL, JCMT GPS, UKIDSS GPS and CORNISH and hence many of the follow-up observations required for a colour selected survey potentially have already been undertaken. There is the additional benefit that the extended source catalogue could also be used to identify objects extended sources with excesses 4.5 μm emission and thereby the identification of EGO.

We see therefore that a GLIMPSE extended source catalogue, although challenging to produce, would be a powerful tool for the identification of massive star formation sites and EGO.

Over the last few years the rate of study of EGOs or "Green Fuzzies" has increased substantially. However, apart for the Gemini observations of Jim De Buizer we still have no

spectra. In June 2010, a small number of EGOs were observed in service mode in M band using the ISAAC spectrometer mounted upon the VLT. These observation was undertaken as myself as the PI. Unfortunately, these observations were made at the incorrect slit size and by the time this was recognised the targets were no longer in an optimal position for observation. However, it is possible that these targets may be re-observed by the VLT in which case they may provide a clue as to the nature of EGO.

As discussed in Chapter 5, at least one association between maser and cluster is due to triggered star formation. However, it would be prudent to examine all the maser associated clusters. Some of these clusters have already been observed by JCMT as part of larger surveys. However others require further observations in order to understand more fully the relationship between Class II Methanol Masers and young infrared clusters.

As discussed in Chapter 4, Prim's algorithm was chosen as it is easy to implement although alternative, faster but somewhat more complex to code, algorithms do exist. For very large datasets, for example the Hi-GAL catalogue, it maybe necessary to re-write the MST tree algorithm in order to improve performance. The most similar work currently being undertaken is that of Billot *et al.* (2011) on the Hi-GAL survey data.. It would be very interesting therefore to pass Billot's Hi-GAL dataset through my own code to see how the two methods compare and to see if the structure seen in Billot *et al.* (2011) is retained. An initial "quick look" at an early Hi-GAL field, without creating convex hulls, suggests that the structure will be retained. Hence, this appears a promising avenue of work.

REFERENCES

Acord, J. M., Churchwell, E., and Wood, D. O. S. (1998). *ApJ*, **495**, L107.

Aguirre, j. E., Ginsburg, A. G., Dunham, M. K., Drosback, M., Bally, J., Battersby, C., Bradley, E. T., Cyganowski, C., Dowell, D., Evans, N. J., I., Glenn, J., Harvey, P., Rosolowsky, E., Stringfellow, G. S., Walawender, J., and Williams, J. P. (2011). *ApJS*, **1**, 4–.

Allison, R. J., Goodwin, S. P., Parker, R. J., Portegies Zwart, S. F., de Grijs, R., and Kouwenhoven, M. B. N. (2009). *MNRAS*, **395**, 1449–1454.

Archive, N. I. I. S. (2011). Glimpse spitzer data. Website. `http://irsa.ipac.caltech.edu/data/SPITZER/GLIMPSE/`.

Babler, B. L. (2006). Description of point source photometry steps used by glimpse. Website. `www.astro.wisc.edu/sirtf/glimpse_photometry_v1.0.pdf`.

Ballesteros-Paredes, J., Klessen, R. S., Mac Low, M.-M., and Vazquez-Semadeni, E. (2007). Molecular Cloud Turbulence and Star Formation. *Protostars and Planets V*, pages 63–80.

Barrow, J. D., Bhavsar, S. P., and Sonoda, D. H. (1985). *MNRAS*, **216**, 17–35.

Bastian, N., Ercolano, B., Gieles, M., Rosolowsky, E., Scheepmaker, R. A., Gutermuth, R., and Efremov, Y. (2007). *MNRAS*, **397**, 1302–1312.

Bastian, N., Gieles, M., Ercolano, B., and Gutermuth, R. (2009). *MNRAS*, **392**, 868–878.

Baumgardt, H. and Klessen, R. S. (2011). *MNRAS*, **413**, 1810–1818.

Benjamin, R., Churchwell, E., Babler, B., Bania, T., Clemens, D., Cohen, M. Dickey, J., Indebetouw, R., and et al. (2003). *PASP*, **810**, 953–964.

Bertoldi, F. and McKee, C. F. (1992). *ApJ*, **395**, 140–157.

Beuther, H. (2011). *Bulletin de la Societe Royale des Sciences de Liege*, **80**, 200–210.

Bica, E. and Bonatto, C.and Camargo, D. (2008). *MNRAS*, **385**, 349–360.

Bica, E., Dutra, C. M., Soares, J., and Barbuy, B. (2003). *A&A*, **440**, 223–232.

Billot, N., Schisano, E., Molinari, S.and Pestalozzi, M., Elia, D., Polychroni, D., Noriega-Crespo, A., Mottram, J., Anderson, L., Stringfellow, G., Thompson, M., and Testi, L. (2011). Clustering properties of far-infrared sources in hi-gal sdp fields. in-prep.

Blitz, L. and Williams, J. P. (1999). In C. J. Lada & N. D. Kylafis, editor, *NATO ASIC Proc. 540: The Origin of Stars and Planetary Systems*, pages 3–.

Bondy, J, A. and Murty, U. (2008). *Graph Theory: An Advanced Course (Graduate Texts in Mathematics)*. Springer.

Bonnell, I. A. (2000). Competitive Accretion in Clusters and the IMF. In *Star Formation from the Small to the Large Scale*, volume 445, page 273.

Bonnell, I. A., Bate, M. R., and Zinnecker, H. (1998). *MNRAS*, **298**, 93–102.

Borissova, J., Ivanov, V. D., Minniti, D., Geisler, D., and Stephens, A. W. (2005). *A&A*, **435**, 95–105.

Borissova, J., Ivanov, V. D., Minniti, D., and Geisler, D. (2006). *A&A*, **455**, 923–930.

Broadley, M. R., White, P. J., Hammond, J. P., Zelko, I., and Lux, A. (2007). *New Phytologist*, **173**, 677702.

Burke, B. F. and Graham-Smith, F. (2009). *An Introduction to Radio Astronomy*. Cambridge University Press.

Carey, S. J., Clark, F. O., Egan, M. P., Price, S. D., Shipman, R. F., and Kuchar, T. A. (1998). *ApJ*, **508**, 721–728.

Carpenter, J., Snell, R., and Schleorb, F. (1990). *ApJ*, **1990**, 147–164.

Cartwright, A. (2009). *MNRAS*, **400**, 1427–1430.

Cartwright, A. and Whitworth, A. P. (2004). *MNRAS*, **348**, 589–598.

Casertano, S. (1985). *ApJ*, **298**, 80–94.

Caswell, J. L., Vaile, R. A., Ellingsen, S. P., Whiteoak, J. B., and Norris, R. P. (1995). *MNRAS*, **272**, 96–138.

Chabrier, G. and Hennebelle, P. (2010). *ApJ*, **725**, L79–L83.

Chambers, E. T., Jackson, J. M., Rathborne, J. M., and Simon, R. (2009). *ApJ*, **181**, 360–390.

Chen, X., Ellingsen, S., and Shen, Z. (2009).

Chrysostomou, A. and Lucas, P. W. (2005). *Contemporary Physics*, **46**, 29–40.

Chu, Y.-H. and Gruendl, R. A. (2011). *Bulletin de la Societe Royale des Sciences de Liege*, **80**, 297–309.

Churchwell, E. (2002). *Annual Review of Astronomy and Astrophysics,*, **40**, 27–62.

Churchwell, E., Watson, D. F., Povich, M. S., Taylor, M. G., Babler, B. L., Meade, M. R., Benjamin, R. A., Indebetouw, R., and Whitney, B. A. (2007). *ApJ*, **670**, 428–441.

Clemens, D. P., Sanders, D. B., and Scoville, N. Z. (1988). *ApJ*, **327**, 139–155.

Coelho, R., Gilmore, C. G., Lucey, B., Richmond, P., and Hutzler, S. (2007). *Physica A*, **376**, 455–466.

Cohen, M., Green, A. J., Meade, M. R., Babler, B., Churchwell, E. B., and et al (2007). *MNRAS*, **374**, 979–998.

Crowther, P. A. and Conti, P. S. (2003). *MNRAS*, **343**, 143–163.

Cyganowski, C. J., Whitney, B. A., Holden, E., Braden, E., Brogan, C. L., Churchwell, E., Indebetouw, R., Watson, D. F., Babler, B. L., Benjamin, R., Gomez, M., Meade, M. R., Povich, M. S., Robitaille, T. P., and Watson, C. (2008a). *ApJ*, **136**, 2391–2412.

Cyganowski, C. J., Whitney, B. A., Holden, E., Braden, E., Brogan, C. L., Churchwell, E., Indebetouw, R., and Watson, C. e. a. (2008b). *ApJ*, **136**, 2391–2412.

Cyganowski, C. J., Brogan, C. L., Hunter, T. R., and Churchwell, E. (2009). *ApJ*, **702**, 1615–1647.

De Buizer, J. M. and Vacca, W. D. (2010). *AJ*, **140**, 192–202.

Deharveng, L. and Zavagno, A. (2008). In *Massive Star Formation: Observations Confront Theory*, volume 387 of *Astronomical Society of the Pacific Conference Series*, page 338.

Deharveng, L., Zavagno, A., and Caplan, J. (2005). *A&A*, **433**, 565–577.

Doroshkevich, A., Tucker, D. L., Allam, S., and Way, M. J. (2004). *A&A*, **418**, 7–23.

Dunham, M. K., Rosolowsky, E., Evans, N. J., I., Cyganowski, C. J., Aguirre, J., Bally, J., Battersby, C., Bradley, E. T., Dowell, D., Drosback, M., Ginsburg, A., Glenn, J., Harvey, P., Merello, M., Schlingman, W., Shirley, Y. L., Stringfellow, G. S.and Walawender, J., and Williams, J. P. (2010). *ApJ*, **717**, 1157–1180.

Dunham, M. K., Robitaille, T. P., Evans, N., J., Schlingman, W. M., Cyganowski, C. J., and Urquhart, J. (2011). A mid-infrared census of star formation activity in bolocam galactic plane survey sources. Website. eprint arXiv:1102.1032.

Dussert, C., Rasigni, M., Palmari, J., Rasigni, G., Llebaria, A., and Marty, F. (1987). *J theor. Biol*, **125**, 317–323.

Dutra, C. M., Bica, E., Soares, J., and Badbuy, B. (2003). *A&A*, **400**, 533–539.

Elitzur, M. (1992). *Astronomical Masers*. Kluwer Academic Publishers.

Ellingsen, S. (2005). *MNRAS*, **359**, 1498–1516.

Ellingsen, S. (2006). *ApJ*, **638**, 241–261.

Ellingsen, S. P. (2007). *MNRAS*, **377**, 571–583.

Ellingsen, S. P., von Bibra, M. L., McCulloch, P. M., Norris, R. P., Deshpande, A. A., and Phillips, C. J. (1996). *MNRAS*, **280**, 378–396.

Ellingsen, S. P., Voronkov, M. A., Cragg, D. M., Sobolev, A. M., Breen, S. L., and Godfrey, P. D. (2007). *Astrophysical Masers and their Environments, Proceedings of the International Astronomical Union, IAU Symposium*, **242**, 213–217.

Enoch, M. L., Young, K. E.and Glenn, J., Evans, N. J., Golwala, S., Sargent, A. I.and Harvey, P., Aguirre, J., Goldin, A., Haig, D., Huard, T. L., Lange, A., Laurent, G., Maloney, P., Mauskopf, P., Rossinot, P., and Sayers, J. (2010). *ApJ*, **638**, 293–313.

Erlebach, T., Hoffmann, M., Krizanc, D., Mihal'ák, M., and Raman, R. (2008). Computing Minimum Spanning Trees with Uncertainty. *ArXiv e-prints*.

Feldt, M., Puga, E., Lenzen, R., Henning, T., Brandner, W., Stecklum, B., Lagrange, A.-M., Gendron, E., and Rousset, G. (2003). *ApJ*, **599**, L91–L94.

Froebrich, D., Scholz, A., and Raftery, C. L. (2007). *MNRAS*, **374**, 399–408.

Froebrich, D., Schmeja, S., Samuel, D., and Lucas, P. W. (2010). *MNRAS*, **409**, 1281–1288.

Goedhart, S., Langa, M. C., Gaylard, M. J., and van der Walt, D. J. (2009). **398**, 995–1010.

Gorski, A. Z., Drozdz, S., and Kwapien, J. (2008). Minimal spanning tree graphs and power like scaling in forex networks. arXiv.

Graham, R. and Hell, P. (1985). *Sarnoff Symposium, 2006 IEEE*, pages 1–4.

Green, J. A., Caswell, J. L., Fuller, G. A., Avison, A., Breen, S. L., Brooks, K., Burton, M. G., Chrysostomou, A., Cox, J., Diamond, P. J., Ellingsen, S. P., Gray, M. D., Hoare, M. G., Masheder, M. R. W., McClure-Griffiths, N. M., Pestalozzi, M., Phillips, C., Quinn, L., Thompson, M. A., and et al. (2009a). *MNRAS*, **392**, 783–794.

Green, J. A., Caswell, J. L., Fuller, G. A., Avison, A., Breen, S. L., Brooks, K., Burton, M. G., Chrysostomou, A., and et al. (2009b). *MNRAS*, **394**, 783–794.

Gutermuth, R. (2005). *The Initial Configuration of Young Stellar Clusters.* Ph.D. thesis, University of Rochester.

Gutermuth, R. A., Megeath, S. T., Myers, P. C., Allen, L. E., Pipher, J. L., and Fazio, G. G. (2009). *ApJ*, **184**, 18–83.

Hanson, M. M. and Bubnick, B. F. (2008). *PASP*, **120**, 150–159.

Helfand, D. J., Becker, R. H., White, R. L., Fallon, A., and Tuttle, S. (2006). *ApJ*, **131**, 2525–2537.

Heyer, M., Krawczyk, C., Duval, J., and M., J. J. (2011). Re-examining larson's scaling relationships in galactic molecular clouds. Website. arXiv:0809.1397v2.

Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., Urquhart, J. S., Busfield, A. L., Sheret, T. L., Clarke, A. J., Moore, T. J. T., Allsopp, J., Burton, M. G., Purcell, C. R., Jiang, Z., and Wang, M. (2005). *Massive star birth: A crossroads of Astrophysics, IAU Symposium Proceedings of the international Astronomical Union*, **227**, 370–375.

Hoffmann, M., Erlebach, T., Krizanc, D., Mihal'ák, M., and Raman, R. (2008). Computing minimum spanning trees with uncertainty. In *25th International Symposium on Theoretical Aspects of Computer Science (STACS 2008)*, pages 277–288.

Hunter, T. R., Brogan, C. L., Indebetouw, R., and Cyganowski, C. J. (2008). *ApJ*, **680**, 1271–1288.

Jackson, J. M., Rathborne, J. M., Shah, R. Y., Simon, R., Bania, T. M., Clemens, D. P., Chambers, E. T., Johnson, A. M., Dormody, M., Lavoie, R., and Heyer, M. H. (2006). *ApJ*, **163**, 145–159.

Jackson, J. M., Finn, S., Rathborne, J., Chambers, E., and Simon, R. (2008). *ApJ*, **680**, 349–361.

Jiménez-Serra, I., Caselli, P., Tan, J. C., Hernandez, A. K., Fontani, F., Butler, M. J., and van Loo, S. (2010). *MNRAS*, **406**, 187–196.

Kauffman, J. and Pillani, T. (2010). *ApJ*, **723**, L7–12.

Keto, E. (2003). *ApJ*, **599**, 1196–1206.

Keto, E. (2007). *ApJ*, **666**, 976–981.

Kim, K. and Koo, B. (2001). *ApJ*, **549**, 979–996.

Klessen, R. S., Spaans, M., and Jappsen, A.-K. (2005). Gravoturbulent star formation: Effects of the equation of state on stellar masses. In R. Cesaroni, M. Felli, E. Churchwell, & M. Walmsley, editor, *Massive Star Birth: A Crossroads of Astrophysics*, volume 227, pages 337–345.

Kraus, S., Hofmann, K.-H., Menten, K. M., Schertl, D., Weigelt, G., Wyrowski, F., Meilland, A., Perraut, K., Petrov, R., Robbe-Dubois, S., Schilke, P., and Testi, L. (2010). **466**, 339–342.

Krumholz, M. R. (2011). Star formation in molecular clouds. Website.

Krumholz, M. R., Klein, R. I., and McKee, C. F. (2005). In *Protostars and Planets V*, pages 8271–.

Krzewina, L. G. and Saslaw, W. C. (1995). *MNRAS*, **278**, 869–876.

Kutterer, A. (2011). Convex hull with graham scan. Website. `http://blog.kutterer.at/?p=17`.

Lada, C., Alves, J. F., and Lombardi, M. (2007). In *Protostars and Planets V*, pages 3–15.

Lada, C. J. and Lada, E. A. (2003). *Ann. Rev. A&A*, **41**, 57–115.

Larson, R. B. (1981). *MNRAS*, **194**, 809–826.

Leung, C. C., Chen, W. F., Kwok, P. C. K., and Chan, F. H. Y. (2003). In *International Conference on Image Processing*, pages 1057–1060.

Lomax, O., Whitworth, A. P., and Cartwright, A. (2011). Statistical comparison of clouds and star clusters. Astro-Ph.

Lombardi, M. (2009). *A&A*, **493**, 735–745.

Lucas, P. W., Hoare, M. G., Longmore, A., Schröder, A. C., Davis, C. J., Adamson, A., and Bandyopadhyay, R. M. e. a. (2008). *MNRAS*, **391**, 136–163.

Lumsden, S. L., Hoare, M. G., Oudmaijer, R. D., and Richards, D. (2002). *MNRAS*, **336**, 621–636.

Maschberger, T., Clarke, C. J., Bonnell, I. A., and Kroupa, P. (2010). *MNRAS*, **404**, 1061–1080.

McKee, C. F. and Ostriker, E. C. (2007). *Annual Review of Astronomy & Astrophysics*, **45**, 565–687.

McKee, C. F. and Tan, J. C. (2003). The Formation of Massive Stars from Turbulent Cores. *ApJ*, **585**, 850–871.

Meade, M. R., Whitney, B. A., Babler, B. L., Indebetouw, R., Bracker, S., Cohen, M., Robitaille, T., Benjamin, B., Mark, W., and Churchwell, E. (2007). Glimpsei - v2.0 data release. Website. `http://irsa.ipac.caltech.edu/data/SPITZER/GLIMPSE/doc/glimpse1_dataprod_v2.0.pdf`.

Menten, K. M. (1991). *ApJ*, **380**, L75–78.

Mercer, E. P., Clemens, D. P., Meade, M. R., Babler, B. L., Indebetouw, R., and Churchwell, E. B. e. a. (2005). *ApJ*, **635**, 560–569.

Miller, S. and Lintott, C. (2010). Galaxy Zoo, Training Artificial Neural Networks with Fuzzy Logic. In *American Astronomical Society Meeting Abstracts #215*, volume 42, page 438.

Minier, V., Ellingsen, S. P.and Norris, R. P., and Booth, R. S. (2003). *AA*, **403**, 1095–1100.

Morgan, M. and Grout, V. (2006). Sarnoff symposium, 2006 ieee. pages 1–4.

Motogi, K., Sorai, K., Habe, A., Honma, M., Kobayashi, H., and Sato, K. (2010). New distance and revised natures of high mass star formation in g5.89-0.39. astro-ph.

Mottram, J. C., Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., Urquhart, J. S., Sheret, T. L., Clarke, A. J., and Allsopp, J. (2007). *A&A*, **476**, 1019–1111.

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

Mottram, J. C., Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., Urquhart, J. S., Meade, M. R., Moore, T. J. T., and Stead, J. J. (2011). *A&A*, **525**, A149.

Murray, N. (2010). Star formation efficiencies and the lifetimes of giant molecular clouds. Website. arXiv:1007.3270v1.

Panagia, N. (1973). *AJ*, **73**, 929–934.

Parker, R. J., Bouvier, J., Goodwin, S. P., Moraux, E., Allison, R. J., Guieu, S., and Gdel, M. (2011). On the mass segregation of stars and brown dwarfs in taurus. astro-ph.

Parsons, H., Thompson, M. A., and Chrysostomou, A. (2009). *MNRAS*, **399**, 1506–1522.

Perault, M., Omont, A., Simon, G., Seguin, P., Ojha, D., Blommaert, J., Felli, M., Gilmore, G., Guglielmo, F., Habing, H., Price, S., Robin, A., de Batz, B., Cesarsky, C., Elbaz, D., Epchtein, N., Fouque, P., Guest, S., Levine, D., Pollock, A., Prusti, T., Siebenmorgen, R., Testi, L., and Tiphene, D. (1996). *A&A*, **315**, L165–L16.

Peretto, N. and Fuller, G. A. (2009). *A&A*, **505**, 405–415.

Peretto, N., Fuller, G. A., André, P., and Hennebelle, P. (2008). In *Massive Star Formation: Observations Confront Theory*, volume 387 of *Astronomical Society of the Pacific Conference Series*, pages 50–.

Pestalozzi, M. R., Elitzur, M., Conway, J. E., and Booth, R. S. (2004). *ApJ*, **603**, L113–L116.

Pestalozzi, M. R., Minier, V., and Booth, R. S. (2005). *A&A*, **432**, 737–742.

Pineda, J., Rosolowsky, E., and Goodman, A. (2009). *Apj*, **699**, L134–L138.

Plionis, M., Valdarnini, R., and Jing, Y. (1992). *ApJ*, **398**, 12–32.

Prim, R. (1957). *Bell Syst. Tech. J.*, **36**, 1389.

Pudritz, R. E. and Banerjee, R. (2005). The disc-jet connection. In *Massive Star Birth: A Crossroads of Astrophysics*, pages 163–173.

Purcell, C. (2006). *What's in the brew? A study of the molecular environment of methanol masers and UCHII regions.* Ph.D. thesis, University of New South Wales.

Rathborne, J., Jackson, J., and Simon, R. (2006). *ApJ*, **641**, 389–405.

Rathborne, J. M., Johnson, A. M., Jackson, J. M., Shah, R. Y., and Simon, R. (2009). *ApJ*, **182**, 131–142.

Read, N. (2005). *Physical Review E*, **72**, Not supplied.

Robitaille, T. P. and Whitney, B. A. (2010). *ApJ*, **710**, L11–L15.

Robitaille, T. P., Meade, M. R., Babler, B. L., Whitney, B. A., Johnston, K. G., Indebetouw, R., Cohen, M., Povich, M. S., Sewilo, M., Benjamin, R. A., and Churchwell, E. (2008). *ApJ*, **136**, 2413–2440.

Roman-Duval, J., Jackson, J. M., Heyer, M., Johnson, A., Rathborne, J., Shah, R., and Simon, R. (2009). *ApJ*, **699**, 1153–1170.

Roman-Duval, J., Jackson, J. M., Heyer, M., Rathborne, J., and Simon, R. (2010). *ApJ*, **723**, 492–507.

Rosolowsky, E., Dunham, M. K., Ginsburg, A., Bradley, E. T. Aguirre, J., Bally, J., Battersby, C., Cyganowski, C., Dowell, D., Drosback, M., Evans, N. J.and Glenn, J., Harvey, P., Stringfellow, G. S., Walawender, J., and Williams, J. P. (2010). *ApJ*, **188**, 123–138.

Sandford, II, M. T., Whitaker, R. W., and Klein, R. I. (1982). *ApJ*, **260**, 183–201.

Schmeja, S. (2011). Identifying star clusters in a field: A comparison of different algorithms. arXiv.

Schmeja, S. and Klessen, R. S. (2006). *A&A*, **449**, 151–159.

Schneider, N. and Brooks, K. (2004). *PASA*, **21**, 290–301.

Shu, F. H. (2001). The X-Wind Theory for the Origin of Chondritic Meteorites. In *Eleventh Annual V. M. Goldschmidt Conference*, pages 370–.

Shu, F. H., Adams, F. C., and Lizano, S. (1987). *Ann Rev A&A*, **25**, 23–81.

Simon, R., Rathborne, R., Jill, M., Jackson, J., and Chambers, E. (2006a). *ApJ*, **635**, 227–236.

Simon, R., Rathborne, J. M., Shah, R. Y., Jackson, J. M., and Chambers, E. T. (2006b). *ApJ*, **652**, 1325–1335.

Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., Carpenter, J. M., and Wheelock, S. e. a. (2006). *AJ*, **131**, 1163–1183.

Sobolev, A. M., Cragg, D. M., and Godfrey, P. D. (1997). *A&A*, **324**, 211–220.

Sobolev, A. M., Ostrovskii, A. B., Kirsanova, M. S., Shelemei, O. V., Voronkov, M. A., and Malyshev, A. V. (2005). In *Massive star birth: A crossroads of Astrophysics, IAU Symposium Proceedings of the international Astronomical Union 227*, pages 174–179.

Solomon, P. and Rivolo, A, R. (1989). *ApJ*, **339**, 919–925.

Solomon, P. M., Rivolo, A. R., Barrett, J., and Yahil, A. (1987). *ApJ*, **319**, 730–741.

Spada, E., L., S., Sourdis, J., Garbuglia, A., Poggi, V., De Fusco, C., and Mele, A. (2004). *Journal of Clinical Microbiology*, **42**, 4230–4236.

Stahler, S. and Palla, F. (2004). *The Formation of Stars*. Wiley-VCH.

Stutz, A. M., Rubin, M., Werner, M. W., Rieke, G. H., Bieging, J. H., Keene, J., Kang, M., Shirley, Y. L., Su, K. Y. L., Velusamy, T., and Wilner, D. J. (2008). *ApJ*, **687**, 389–405.

Thompson, M. A.and MacDonald, G. H. (1999). *A&A*, **135**, 531–546.

Thompson, M. A., Hatchell, J., Walsh, A. J., MacDonald, G. H., and Millar, T. J. (2006). *A&A*, **453**, 1003–1026.

Urquhart, J. S. (2011). The rms database server. Website. `http://www.ast.leeds.ac.uk/cgi-bin/RMS/RMS_DATABASE.cgi`.

Urquhart, J. S., Busfield, A. L., Hoare, M. G., Lumsden, S. L., Clarke, A. J., Moore, T. J. T., and Mottram, J. C.and Oudmaijer, R. D. (2007). *A&A*, **461**, 11–23.

Urquhart, J. S., Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., and Moore, T. J. T. (2008). In *Massive Star Formation: Observations Confront Theory ASP Conference Series*, volume 387, page 381.

Urquhart, J. S., Moore, T. J. T., Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., Rathborne, J. M., Mottram, J. C.and Davies, B., and Stead, J. J. (2011). *MNRAS*, **410**, 1237–1250.

Vazquez-Semadeni, E. (2010). Molecular cloud evolution. Website. arXiv:1009.3962.

Ward-Thompson, D. and Whitmorth, A. (2011). *An Instorduction to Star Formation*. Cambridge University Press.

White, R.L. Becker, R. and Helfand, D. (2005). *ApJ*, **130**, 586–596.

Whitney, B. A., Wood, K., Bjorkman, J. E., and Cohen, M. (2003). *ApJ*, **598**, 1079–1099.

Yorke, H. W. (2004). Theory of Formation of Massive Stars via Accretion. In *Star Formation at High Angular Resolution*, volume 221, page 141.

Yorke, H. W. and Bodenheimer, P. (1999). *ApJ*, **525**, 330–342.

Zahn, C. T. (1971). *IEEE*, **C-20**, 68–86.

Zhang, Q. (2005). Massive star disks. In *Massive Star Birth: A Crossroads of Astrophysics Proceedings IAU Symposium*, volume 227.

Zinnecker, H. and Yorke, H. (2007). *Ann.Rev.Astron.Astrophys.*, **45**, 481–563.

APPENDIX A

GLIMPSE photometry results

MMB ANCAP Photometry Results

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G284.3516-00.4191	156.0454	-57.8774	12.6	11.4	10.7	9.9	0 0 0 0
G298.1774-00.7949	182.2398	-63.2738	9.8	8.4	6.4	5.5	0 0 0 0
G298.2133-00.3429	182.4799	-62.8336	13.1	11.2	10.5	10.2	0 0 0 0
G298.2621+00.7393	182.9485	-61.7725	7.5	6.3	5.8	5.4	0 1 0 0
G298.6324-00.3621	183.3818	-62.9169	8.8	7.5	6.1	5.2	0 0 0 0
G298.7229-00.0864	183.6646	-62.6572	9.4	7.6	6.2	5.4	0 0 0 0
G299.0131+00.1282	184.3525	-62.4844	10.0	9.4	8.7	8.1	0 0 0 0
G299.7716-00.0049	185.9540	-62.7070	10.9	9.1	7.3	6.2	0 0 0 0
G300.5039-00.1759	187.5149	-62.9469	9.5	8.1	7.7	7.9	0 0 0 0
G301.1365-00.2259	188.8964	-63.0424	7.9	6.5	4.9	3.6	0 1 0 1
G302.0336+00.6254	190.9310	-62.2329	13.6	10.9	10.5	10.5	0 0 0 0
G302.4548-00.7406	191.7860	-63.6084	8.1	6.6	5.2	4.1	0 0 0 0
G303.8462-00.3634	194.8890	-63.2208	12.3	10.6	10.2	10.1	0 0 0 0
G304.3672-00.3360	196.0409	-63.1723	9.4	7.3	5.5	4.0	0 0 0 0
G304.8867+00.6354	197.0498	-62.1728	9.4	8.0	7.0	6.4	0 0 0 0
G305.2083+00.2061	197.8071	-62.5782	13.4	11.7	10.3	9.5	0 0 0 0
G305.4748-00.0964	198.4406	-62.8579	8.7	7.1	6.3	5.8	0 0 0 0
G305.6153-00.3437	198.7971	-63.0916	11.1	10.0	9.7	10.5	0 0 0 0
G305.7992-00.2449	199.1801	-62.9758	12.7	10.8	9.8	8.4	0 0 0 0
G305.8225-00.1147	199.2030	-62.8439	10.7	8.8	8.1	8.0	0 0 0 0
G305.8870+00.0171	199.3147	-62.7064	9.8	8.0	7.3	7.1	0 0 0 0
G306.3221-00.3337	200.3459	-63.0082	11.4	8.9	7.4	6.7	0 0 0 0
G307.1330-00.4766	202.1595	-63.0446	11.3	9.9	9.4	9.2	0 0 0 0
G308.0561-00.3960	204.1340	-62.8182	10.8	9.5	7.8	6.3	0 0 0 0
G308.7154-00.2161	205.4733	-62.5199	13.6	10.7	9.0	8.1	0 0 0 0
G309.9205+00.4789	207.6741	-61.5862	5.3	4.3	2.2	1.4	1 1 1 1
G309.9014+00.2307	207.7544	-61.8322	14.8	12.4	11.0	10.1	0 0 0 0
G310.1804-00.1220	208.5074	-62.1096	10.8	9.1	8.6	8.1	0 0 0 0
G311.2297-00.0318	210.6145	-61.7540	9.7	7.2	5.6	4.8	0 0 0 0
G311.6284+00.2655	211.2467	-61.3582	10.2	8.3	6.8	6.1	0 0 0 0
G311.5508-00.0552	211.2794	-61.6878	11.9	9.9	8.7	8.3	0 0 0 0
G311.6427-00.3800	211.6615	-61.9731	13.6	11.2	9.4	8.0	0 0 0 0
G312.1082+00.2618	212.2055	-61.2236	11.2	9.5	9.3	9.7	0 0 0 0
G312.0712+00.0818	212.2425	-61.4066	9.9	9.5	7.8	6.3	0 0 0 0
G312.3073+00.6610	212.3540	-60.7835	9.9	8.6	7.4	5.9	0 0 0 0
G312.5013-00.0837	213.2040	-61.4342	10.7	9.8	8.0	6.7	0 0 0 0
G312.5965+00.0447	213.3098	-61.2827	11.5	9.8	9.0	8.4	0 0 0 0
G312.5982+00.0453	213.3126	-61.2816	11.6	10.0	8.5	7.3	0 0 0 0
G312.6983+00.1263	213.4577	-61.1734	12.5	10.6	9.5	8.5	0 0 0 0
G312.7021-00.0870	213.6047	-61.3747	12.7	10.6	9.3	8.8	0 0 0 0
G313.4693+00.1904	214.9206	-60.8631	10.6	7.6	6.1	5.3	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G313.5769+00.3250	215.0358	-60.7002	9.6	8.1	6.5	5.3	0 0 0 0
G313.7053-00.1899	215.6448	-61.1408	10.5	8.1	6.4	5.7	0 0 0 0
G313.9942-00.0843	216.1283	-60.9412	13.5	10.9	10.2	9.4	0 0 0 0
G313.7671-00.8625	216.2572	-61.7495	8.1	6.4	5.5	4.8	0 1 0 0
G314.2205+00.2726	216.3037	-60.5273	8.9	7.1	5.7	3.7	0 0 0 1
G314.3198+00.1120	216.6092	-60.6420	5.8	4.5	2.0	1.9	1 1 1 1
G315.8025-00.5754	219.9436	-60.7110	10.2	8.8	7.1	6.3	0 0 0 0
G316.4118-00.3084	220.8473	-60.2169	12.8	11.3	9.5	8.0	0 0 0 0
G316.3814-00.3789	220.8509	-60.2937	11.0	10.0	8.1	6.6	0 0 0 0
G316.4839-00.3104	220.9807	-60.1886	12.7	9.7	7.6	6.4	0 0 0 0
G317.0287+00.3613	221.3981	-59.3496	10.8	9.4	6.8	5.3	0 0 0 0
G317.0609+00.2565	221.5430	-59.4307	9.9	9.1	7.7	6.5	0 0 0 0
G317.7010+00.1098	222.7987	-59.2839	9.1	7.4	6.3	5.3	0 0 0 0
G317.4657-00.4025	222.8320	-59.8474	11.3	8.7	7.7	7.2	0 0 0 0
G318.0497+00.0870	223.4278	-59.1479	8.2	6.7	5.8	5.1	0 0 0 0
G318.4715-00.2135	224.4286	-59.2211	10.9	8.7	7.5	7.0	0 0 0 0
G318.9480-00.1960	225.2308	-58.9811	6.9	5.2	3.3	2.7	1 1 1 1
G319.1630-00.4209	225.8073	-59.0751	7.8	6.6	6.2	4.2	0 0 0 0
G319.8358-00.1967	226.7277	-58.5500	11.0	9.1	7.9	6.7	0 0 0 0
G320.4271+00.1027	227.4163	-57.9949	11.6	9.4	7.5	6.0	0 0 0 0
G320.4236+00.0894	227.4232	-58.0081	12.2	10.0	8.5	7.6	0 0 0 0
G320.1233-00.5037	227.5007	-58.6717	9.1	6.8	5.3	4.6	0 0 0 0
G320.2854-00.3086	227.5792	-58.4213	10.9	9.7	8.7	7.8	0 0 0 0
G320.6252+00.0984	227.7422	-57.8982	10.8	9.2	7.9	6.7	0 0 0 0
G320.2437-00.5617	227.7567	-58.6605	6.5	5.4	4.2	3.6	1 1 0 1
G320.2634-00.5345	227.7626	-58.6271	13.2	13.1	13.4	12.8	0 0 0 0
G320.7796+00.2485	227.8478	-57.6903	7.6	8.0	7.5	6.1	0 0 0 0
G321.0297-00.4847	228.9658	-58.1883	9.1	8.1	6.8	5.6	0 0 0 0
G321.0328-00.4832	228.9693	-58.1855	7.3	6.5	5.5	4.5	0 0 0 0
G321.1483-00.5294	229.2016	-58.1639	10.0	8.7	8.0	7.4	0 0 0 0
G323.4588-00.0789	232.3305	-56.5230	5.3	4.2	2.0	1.5	1 1 1 1
G323.7993+00.0173	232.7380	-56.2504	6.8	5.3	3.5	2.9	1 1 1 1
G323.7401-00.2626	232.9394	-56.5139	7.9	5.9	4.9	4.5	0 1 0 0
G323.7934-00.3972	233.1587	-56.5935	10.7	9.5	7.5	6.2	0 0 0 0
G324.7163+00.3418	233.7395	-55.4566	10.9	8.9	8.5	9.8	0 0 0 0
G324.9147+00.1578	234.2132	-55.4897	12.6	9.8	8.6	7.4	0 0 0 0
G324.7892-00.3782	234.5952	-55.9964	10.1	8.6	7.6	6.8	0 0 0 0
G325.6591-00.0217	235.4529	-55.1908	11.2	9.7	8.9	7.4	0 0 0 0
G326.4748+00.7027	235.8193	-54.1207	6.2	4.9	2.9	2.4	1 1 1 1
G326.4756+00.6947	235.8288	-54.1265	8.8	7.7	7.2	6.6	0 0 0 0
G326.6081+00.7991	235.9000	-53.9631	11.5	9.5	9.1	9.7	0 0 0 0
G326.6622+00.5207	236.2623	-54.1509	6.4	4.9	2.6	1.7	1 1 1 1
G327.1196+00.5108	236.8864	-53.8773	7.6	5.9	4.4	3.3	0 1 0 1

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G326.4484-00.7482	237.3276	-55.2810	6.8	5.2	3.6	3.0	1 1 1 1
G327.4022+00.4445	237.3313	-53.7539	7.8	5.8	4.5	3.7	0 1 0 1
G327.3917+00.1994	237.5770	-53.9518	9.2	7.5	6.7	6.2	0 0 0 0
G327.3945+00.1967	237.5836	-53.9521	7.0	5.8	4.4	3.7	1 1 0 1
G327.5899-00.0942	238.1534	-54.0552	11.3	10.4	7.9	6.5	0 0 0 0
G328.1643+00.5872	238.1765	-53.1640	10.0	8.1	6.7	5.6	0 0 0 0
G327.6184-00.1111	238.2093	-54.0502	8.9	7.1	5.7	4.7	0 0 0 0
G327.7105-00.3940	238.6374	-54.2101	9.7	9.1	6.3	4.8	0 0 0 0
G327.9447-00.1147	238.6413	-53.8456	8.2	7.1	5.7	4.6	0 0 0 0
G328.3851+00.1309	238.9472	-53.3752	9.1	7.6	6.7	6.0	0 0 0 0
G328.8083+00.6330	238.9519	-52.7185	8.0	6.1	4.9	3.5	0 1 0 1
G328.8090+00.6328	238.9529	-52.7182	8.0	6.1	4.9	3.5	0 1 0 1
G327.8083-00.6337	239.0288	-54.3318	6.9	5.8	4.7	3.9	1 1 0 1
G328.9417+00.5653	239.1924	-52.6847	10.9	9.5	8.5	7.6	0 0 0 0
G328.1401-00.4316	239.2406	-53.9635	10.5	9.4	8.9	8.6	0 0 0 0
G328.2541-00.5322	239.4990	-53.9668	8.9	6.6	5.1	4.3	0 0 0 0
G329.7188+01.1642	239.5295	-51.7257	8.9	7.9	7.5	7.5	0 0 0 0
G329.4690+00.5026	239.9196	-52.3909	11.0	9.5	9.0	8.7	0 0 0 0
G330.0697+01.0638	240.0643	-51.5738	9.2	7.5	5.9	4.6	0 0 0 0
G329.2722+00.1147	240.0907	-52.8134	7.6	6.1	4.5	3.1	0 1 0 1
G329.0663-00.3076	240.2914	-53.2674	9.4	6.8	5.4	4.5	0 0 0 0
G329.5556+00.1808	240.3727	-52.5778	11.7	9.9	9.2	8.9	0 0 0 0
G329.1833-00.3139	240.4459	-53.1954	10.8	8.4	7.7	8.5	0 0 0 0
G329.6101+00.1136	240.5131	-52.5926	8.8	6.5	5.4	4.2	0 1 0 0
G329.4051-00.4594	240.8840	-53.1585	9.5	7.3	6.6	6.6	0 0 0 0
G329.4067-00.4594	240.8860	-53.1575	9.5	7.3	6.6	6.6	0 0 0 0
G330.2830+00.4931	240.9294	-51.8634	9.6	8.6	6.0	4.4	0 0 0 0
G330.2259+00.2897	241.0789	-52.0536	12.5	11.1	9.8	8.7	0 0 0 0
G332.2948+02.2797	241.4238	-49.1918	11.6	10.1	9.0	7.9	0 0 0 0
G331.0586+00.3754	241.9868	-51.4333	12.8	12.3	10.7	9.1	0 0 0 0
G331.1339+00.1565	242.3132	-51.5441	11.8	9.5	8.3	7.8	0 0 0 0
G330.9527-00.1822	242.4682	-51.9160	10.1	7.6	5.7	4.7	0 0 0 0
G331.7100+00.6026	242.5074	-50.8256	10.3	8.8	7.7	6.7	0 0 0 0
G331.4246+00.2642	242.5390	-51.2679	11.6	9.8	7.9	6.4	0 0 0 0
G331.1195-00.1181	242.5960	-51.7557	12.2	9.6	8.3	8.0	0 0 0 0
G330.8747-00.3828	242.5962	-52.1163	8.2	7.5	5.2	3.7	0 0 0 1
G331.1319-00.2439	242.7490	-51.8396	10.7	9.3	6.8	5.3	0 0 0 0
G331.5423-00.0665	243.0376	-51.4299	8.4	6.9	5.6	4.3	0 0 0 0
G331.5430-00.0662	243.0381	-51.4293	8.4	6.9	5.6	4.3	0 0 0 0
G331.4420-00.1867	243.0520	-51.5861	14.4	9.4	7.7	7.1	0 0 0 0
G331.3418-00.3464	243.1102	-51.7712	9.4	6.7	5.1	4.3	0 0 0 0
G331.5557-00.1212	243.1134	-51.4606	11.6	10.1	8.0	6.5	0 0 0 0
G331.4367-00.3039	243.1750	-51.6753	7.1	5.8	4.4	3.8	0 1 0 1

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G332.3637+00.6067	243.2561	-50.3769	9.9	8.3	7.2	6.5	0 0 0 0
G332.2953-00.0937	243.9392	-50.9316	9.1	7.6	6.9	6.7	0 0 0 0
G332.5829+00.1469	244.0036	-50.5588	11.9	10.1	9.1	8.5	0 0 0 0
G332.3523-00.1167	244.0296	-50.9087	8.4	7.0	5.9	5.0	0 0 0 0
G332.0935-00.4207	244.0685	-51.3071	4.1	3.1	1.2	1.0	1 1 1 1
G332.5599-00.1479	244.3005	-50.7868	10.5	9.0	8.7	9.6	0 0 0 0
G332.9600+00.1353	244.4418	-50.3050	8.5	8.2	7.8	7.2	0 0 0 0
G333.3150+00.1053	244.8709	-50.0781	9.0	6.6	5.3	4.3	0 0 0 0
G333.1632-00.1007	244.9278	-50.3314	8.2	7.2	4.9	3.4	0 0 0 1
G332.6532-00.6212	244.9312	-51.0601	10.4	9.1	9.3	7.4	0 0 0 0
G333.1841-00.0908	244.9401	-50.3097	8.6	7.0	6.3	6.3	0 0 0 0
G332.7009-00.5875	244.9476	-51.0026	6.6	5.0	3.0	2.6	1 1 1 1
G333.2339-00.0602	244.9620	-50.2530	10.6	7.3	6.0	5.8	0 0 0 0
G333.8510+00.5265	245.0038	-49.4019	14.7	13.2	12.8	11.3	0 0 0 0
G332.7255-00.6205	245.0123	-51.0089	11.1	9.2	8.7	8.5	0 0 0 0
G333.3873+00.0317	245.0316	-50.0798	10.5	9.6	7.3	5.9	0 0 0 0
G332.9865-00.4871	245.1575	-50.7305	6.2	4.8	3.1	2.6	1 1 1 1
G332.8129-00.7006	245.2005	-51.0043	9.3	8.8	6.9	6.5	0 0 0 0
G333.0680-00.4469	245.2040	-50.6446	8.7	5.9	4.1	3.1	0 1 0 1
G333.1260-00.4397	245.2609	-50.5986	6.7	5.4	3.3	2.9	1 1 1 1
G333.1088-00.5004	245.3091	-50.6536	8.3	6.6	5.4	4.8	0 0 0 0
G332.9419-00.6858	245.3292	-50.9028	10.0	7.7	6.5	5.9	0 0 0 0
G333.4660-00.1643	245.3341	-50.1635	11.4	9.0	8.1	7.3	0 0 0 0
G333.7605-00.2258	245.7261	-49.9985	8.6	6.9	6.2	5.6	0 0 0 0
G333.9306-00.1347	245.8118	-49.8136	11.2	9.1	7.3	6.3	0 0 0 0
G333.6825-00.4367	245.8741	-50.2024	11.8	10.9	8.5	6.9	0 0 0 0
G334.1382-00.0231	245.9148	-49.5876	11.6	9.1	7.4	6.5	0 0 0 0
G334.3068-00.0792	246.1588	-49.5068	11.7	10.8	10.4	9.6	0 0 0 0
G334.6351-00.0148	246.4405	-49.2271	12.0	10.0	8.6	7.9	0 0 0 0
G334.9349-00.0981	246.8510	-49.0698	13.2	12.0	9.0	7.4	0 0 0 0
G334.9325-00.3066	247.0784	-49.2160	11.6	10.9	10.4	11.1	0 0 0 0
G335.0595-00.4274	247.3464	-49.2075	9.7	7.6	6.6	6.2	0 0 0 0
G335.4260-00.2395	247.5233	-48.8124	9.8	7.0	5.9	5.7	0 0 0 0
G335.5562-00.3067	247.7332	-48.7639	10.5	9.3	8.2	7.9	0 0 0 0
G335.5851-00.2846	247.7387	-48.7277	10.8	9.1	8.5	8.5	0 0 0 0
G335.5852-00.2901	247.7450	-48.7314	8.8	6.8	6.1	5.8	0 0 0 0
G335.8239-00.1775	247.8680	-48.4803	11.5	9.8	8.9	8.3	0 0 0 0
G336.4636-00.1570	248.5000	-47.9981	11.5	8.9	8.0	8.1	0 0 0 0
G336.8253+00.1394	248.5409	-47.5314	11.5	9.7	8.1	7.3	0 0 0 0
G336.4094-00.2570	248.5550	-48.1058	10.6	9.6	7.3	5.7	0 0 0 0
G336.5258-00.1565	248.5625	-47.9521	12.0	10.8	9.8	8.9	0 0 0 0
G336.4326-00.2622	248.5843	-48.0923	11.2	9.3	8.6	7.6	0 0 0 0
G336.4965-00.2707	248.6584	-48.0511	11.7	10.8	8.8	7.3	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G336.8224+00.0280	248.6595	-47.6089	14.5	12.0	10.2	8.4	0 0 0 0
G336.7030-00.0988	248.6782	-47.7827	12.6	9.8	8.3	7.2	0 0 0 0
G336.8643+00.0049	248.7268	-47.5937	11.8	9.6	8.8	8.2	0 0 0 0
G336.9939-00.0273	248.8916	-47.5199	11.8	9.7	8.8	8.2	0 0 0 0
G337.2013+00.1142	248.9440	-47.2713	10.8	8.9	8.1	7.9	0 0 0 0
G336.9410-00.1563	248.9800	-47.6459	11.2	8.7	7.6	7.0	0 0 0 0
G336.9832-00.1832	249.0517	-47.6328	6.4	5.3	3.4	2.5	1 1 1 1
G337.1323-00.0675	249.0736	-47.4447	9.7	9.4	7.7	6.4	0 0 0 0
G336.8298-00.3749	249.1091	-47.8753	10.3	8.4	7.6	6.7	0 0 0 0
G337.0520-00.2258	249.1672	-47.6106	8.5	6.3	5.2	4.7	0 1 0 0
G337.2018-00.0941	249.1718	-47.4112	11.9	10.8	9.4	8.6	0 0 0 0
G337.2628-00.0697	249.2057	-47.3496	11.6	10.4	9.2	8.2	0 0 0 0
G337.2579-00.1007	249.2347	-47.3740	11.8	9.5	8.8	9.4	0 0 0 0
G337.6864+00.1373	249.3976	-46.8966	12.9	10.3	9.1	8.5	0 0 0 0
G337.1530-00.3948	249.4536	-47.6490	11.8	9.8	9.1	8.7	0 0 0 0
G337.7097+00.0894	249.4725	-46.9113	14.3	11.5	10.1	9.4	0 0 0 0
G338.2803+00.5423	249.5378	-46.1842	10.6	8.8	8.2	7.6	0 0 0 0
G337.6125-00.0597	249.5398	-47.0833	13.1	10.7	9.4	8.2	0 0 0 0
G337.6320-00.0788	249.5797	-47.0815	10.2	8.9	7.0	5.6	0 0 0 0
G337.7053-00.0531	249.6235	-47.0099	12.6	10.7	8.3	6.9	0 0 0 0
G337.9973+00.1364	249.7021	-46.6660	16.0	11.7	9.7	8.4	0 0 0 0
G337.4044-00.4020	249.7105	-47.4668	9.1	7.2	5.8	4.5	0 0 0 0
G338.1400+00.1778	249.7955	-46.5320	11.1	9.7	9.3	8.8	0 0 0 0
G336.9578-00.9770	249.9065	-48.1829	10.0	8.6	8.0	7.7	0 0 0 0
G338.0691+00.0110	249.9081	-46.6959	9.3	7.2	5.6	4.8	0 0 0 0
G338.0748+00.0118	249.9127	-46.6911	10.7	8.9	6.5	5.0	0 0 0 0
G337.0967-00.9287	249.9907	-48.0470	8.0	6.7	6.0	5.6	0 0 0 0
G338.4723+00.2891	249.9955	-46.2098	9.9	6.9	5.4	5.1	0 0 0 0
G338.9262+00.6337	250.0581	-45.6416	14.6	10.7	9.1	8.3	0 0 0 0
G338.1605-00.0636	250.0779	-46.6772	11.9	9.4	8.0	7.5	0 0 0 0
G338.4973+00.2073	250.1079	-46.2454	11.9	10.4	9.8	10.9	0 0 0 0
G337.8438-00.3746	250.1111	-47.1204	5.4	4.5	2.6	2.3	1 1 1 1
G337.2996-00.8740	250.1306	-47.8589	9.9	8.6	8.2	8.2	0 0 0 0
G338.9197+00.5500	250.1417	-45.7020	7.8	6.1	5.7	4.6	0 1 0 0
G338.5615+00.2176	250.1582	-46.1905	13.7	13.2	10.0	8.8	0 0 0 0
G338.4319+00.0576	250.2075	-46.3936	12.0	10.1	9.0	8.2	0 0 0 0
G338.8495+00.4092	250.2262	-45.8478	8.2	7.3	5.2	3.8	0 0 0 1
G338.3960-00.0071	250.2434	-46.4633	12.6	10.9	9.6	9.0	0 0 0 0
G337.9201-00.4562	250.2752	-47.1173	10.4	9.0	7.4	6.2	0 0 0 0
G338.5665+00.1102	250.2793	-46.2579	12.9	10.0	8.6	7.3	0 0 0 0
G338.9020+00.3945	250.2919	-45.8182	12.1	10.6	8.5	7.0	0 0 0 0
G338.4605-00.2447	250.5646	-46.5718	12.1	9.7	8.3	8.0	0 0 0 0
G338.3252-00.4089	250.6147	-46.7818	10.9	8.8	8.1	8.3	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G338.3918-00.4027	250.6720	-46.7275	9.5	7.8	7.3	7.4	0 0 0 0
G339.0636+00.1520	250.7065	-45.8566	10.8	9.8	7.1	5.6	0 0 0 0
G338.8753-00.0839	250.7844	-46.1536	6.6	5.4	3.7	3.2	1 1 1 1
G338.9346-00.0620	250.8167	-46.0945	7.6	6.4	5.4	4.2	0 1 0 0
G339.2822+00.1359	250.9296	-45.7022	12.7	11.2	10.0	8.7	0 0 0 0
G339.2044-00.0183	251.0237	-45.8620	11.9	10.4	9.6	9.4	0 0 0 0
G339.0530-00.3146	251.2041	-46.1703	15.1	12.4	10.9	9.7	0 0 0 0
G339.9086+00.2396	251.4007	-45.1601	16.6	11.9	9.7	8.2	0 0 0 0
G339.7616+00.0543	251.4648	-45.3924	10.0	8.4	6.9	5.9	0 0 0 0
G339.5824-00.1270	251.4951	-45.6464	11.3	9.6	9.0	9.3	0 0 0 0
G339.6221-00.1205	251.5250	-45.6120	6.3	4.8	2.9	2.7	1 1 1 1
G340.1821-00.0471	251.9616	-45.1378	10.8	9.6	7.5	6.0	0 0 0 0
G340.2492-00.0458	252.0216	-45.0857	7.4	6.2	4.6	3.3	0 1 0 1
G340.0545-00.2435	252.0579	-45.3620	9.6	6.9	5.8	5.4	0 0 0 0
G339.9863-00.4248	252.1930	-45.5309	11.5	10.0	9.5	13.0	0 0 0 0
G339.9489-00.5387	252.2833	-45.6329	10.4	8.1	7.0	6.2	0 0 0 0
G339.9799-00.5384	252.3114	-45.6088	10.9	8.8	6.8	5.1	0 0 0 0
G340.2488-00.3717	252.3753	-45.2957	9.2	7.1	5.5	3.9	0 0 0 1
G340.5175-00.1519	252.3807	-44.9485	10.5	9.4	6.9	5.5	0 0 0 0
G340.5429-00.1620	252.4147	-44.9355	11.2	10.7	10.0	9.3	0 0 0 0
G340.6547-00.2354	252.5954	-44.8967	12.1	9.4	8.5	8.6	0 0 0 0
G341.2757+00.0618	252.8309	-44.2290	10.0	7.5	6.2	5.5	0 0 0 0
G340.0344-01.1103	252.9920	-45.9324	11.4	9.8	9.5	9.8	0 0 0 0
G341.2177-00.2118	253.0743	-44.4478	7.3	5.8	4.4	3.4	0 1 0 1
G341.2384-00.2701	253.1559	-44.4688	11.2	9.5	9.0	8.5	0 0 0 0
G342.2507+00.3078	253.4244	-43.3194	8.4	7.3	5.2	3.9	0 0 0 1
G341.9897-00.1029	253.6370	-43.7809	11.2	9.4	8.7	8.8	0 0 0 0
G342.3677+00.1401	253.7048	-43.3343	11.2	11.1	9.0	7.6	0 0 0 0
G340.9703-01.0219	253.7388	-45.1514	12.7	9.5	8.4	9.4	0 0 0 0
G342.4463-00.0720	253.9998	-43.4063	11.1	9.8	8.4	7.4	0 0 0 0
G345.0117+01.7972	254.1951	-40.2358	10.0	6.5	5.0	4.5	0 0 0 0
G345.0095+01.7924	254.1983	-40.2405	8.0	5.5	3.8	2.9	0 1 1 1
G342.9542-00.0186	254.3778	-42.9763	13.1	11.8	11.0	9.9	0 0 0 0
G343.3539-00.0667	254.7677	-42.6929	7.4	6.5	4.2	2.7	0 1 0 1
G345.4977+01.4669	254.9285	-40.0600	8.2	5.6	4.0	3.9	0 1 1 1
G343.9295+00.1249	255.0455	-42.1220	12.1	10.3	8.9	7.7	0 0 0 0
G343.7564-00.1629	255.2079	-42.4356	12.0	10.5	9.8	9.4	0 0 0 0
G343.5015-00.4723	255.3267	-42.8269	9.6	7.4	6.5	6.2	0 0 0 0
G344.4185+00.0435	255.5359	-41.7862	7.2	5.9	4.4	2.9	0 1 0 1
G344.4214+00.0452	255.5365	-41.7829	6.5	5.5	4.0	2.6	1 1 1 1
G344.5814-00.0235	255.7405	-41.6983	9.4	7.6	6.7	5.8	0 0 0 0
G344.2275-00.5687	256.0324	-42.3110	14.6	12.3	10.9	9.5	0 0 0 0
G345.5047+00.3483	256.0955	-40.7394	6.6	4.8	2.3	2.2	1 1 1 1

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G345.1981-00.0302	256.2479	-41.2127	7.2	5.6	4.3	3.4	0 1 0 1
G345.0030-00.2235	256.2954	-41.4851	8.8	6.3	5.0	3.6	0 1 0 1
G345.0034-00.2244	256.2968	-41.4853	8.8	6.3	5.0	3.6	0 1 0 1
G345.1314-00.1743	256.3468	-41.3530	11.1	9.3	8.0	7.0	0 0 0 0
G345.8235+00.0439	256.6696	-40.6694	7.4	6.1	5.0	4.4	0 1 0 0
G346.0362+00.0481	256.8334	-40.4969	12.0	10.2	9.3	9.0	0 0 0 0
G345.9851-00.0202	256.8649	-40.5788	13.5	10.9	9.7	9.9	0 0 0 0
G346.2315+00.1190	256.9129	-40.2981	11.9	10.5	9.4	8.4	0 0 0 0
G346.4810+00.1321	257.0947	-40.0904	9.5	7.5	5.9	4.9	0 0 0 0
G345.9493-00.2682	257.0985	-40.7560	12.2	11.0	10.4	10.1	0 0 0 0
G345.4244-00.9510	257.4107	-41.5846	11.9	9.8	8.0	6.9	0 0 0 0
G347.2300+00.0157	257.7966	-39.5576	11.3	9.5	8.1	7.4	0 0 0 0
G347.5833+00.2127	257.8613	-39.1563	8.4	7.5	5.2	3.7	0 0 0 1
G347.6279+00.1485	257.9622	-39.1581	10.5	8.7	7.5	6.3	0 0 0 0
G347.8171+00.0178	258.2419	-39.0823	13.6	11.1	9.8	9.0	0 0 0 0
G347.9025+00.0516	258.2713	-38.9932	8.6	7.1	5.8	4.7	0 0 0 0
G347.8632+00.0186	258.2760	-39.0444	8.4	6.8	5.7	4.9	0 0 0 0
G348.6540+00.2443	258.6349	-38.2713	10.2	8.5	7.3	6.7	0 0 0 0
G348.8846+00.0964	258.9589	-38.1701	10.4	9.3	8.6	8.0	0 0 0 0
G348.7233-00.0775	259.0199	-38.4024	11.9	10.9	9.9	9.1	0 0 0 0
G349.0919+00.1055	259.1025	-37.9961	11.1	8.0	6.8	6.8	0 0 0 0
G349.0919+00.1049	259.1031	-37.9964	11.1	8.0	6.8	6.8	0 0 0 0
G349.0673-00.0173	259.2114	-38.0873	8.8	7.9	6.3	5.2	0 0 0 0
G349.1509+00.0214	259.2328	-37.9966	13.2	11.4	10.6	10.3	0 0 0 0
G348.8917-00.1801	259.2510	-38.3247	10.0	8.4	6.8	5.4	0 0 0 0
G349.8836+00.2305	259.5515	-37.2778	9.7	7.4	6.0	5.2	0 0 0 0
G349.7985+00.1084	259.6156	-37.4176	14.0	11.2	10.1	10.8	0 0 0 0
G350.1156+00.2203	259.7296	-37.0939	12.0	10.7	10.2	10.1	0 0 0 0
G348.5788-00.9198	259.7942	-39.0067	9.7	7.7	7.1	7.3	0 0 0 0
G348.5498-00.9788	259.8350	-39.0643	12.7	10.5	9.1	8.6	0 0 0 0
G350.3405+00.1409	259.9726	-36.9552	11.7	11.0	10.0	8.5	0 0 0 0
G351.1605+00.6967	259.9896	-35.9647	6.6	5.1	4.3	3.4	1 1 0 1
G348.7270-01.0372	260.0273	-38.9525	11.9	9.1	8.3	8.0	0 0 0 0
G351.2515+00.6516	260.0995	-35.9158	7.5	5.4	3.5	3.0	0 1 1 1
G350.4700+00.0289	260.1802	-36.9129	13.8	12.2	10.9	9.6	0 0 0 0
G350.3560-00.0677	260.1981	-37.0617	11.3	9.8	9.5	10.2	0 0 0 0
G351.4169+00.6452	260.2224	-35.7837	5.7	4.4	2.3	1.7	1 1 1 1
G349.5790-00.6789	260.2727	-38.0485	10.1	8.8	6.7	5.3	0 0 0 0
G350.7756+00.1381	260.2858	-36.5997	12.1	11.0	10.7	10.6	0 0 0 0
G350.6860-00.4909	260.8693	-37.0302	9.8	7.4	5.7	4.5	0 0 0 0
G351.6881+00.1708	260.8938	-35.8295	11.1	9.8	8.3	7.2	0 0 0 0
G351.3824-00.1810	261.0399	-36.2804	10.5	8.8	7.5	6.4	0 0 0 0
G352.0833+00.1670	261.1718	-35.5052	12.1	10.8	8.9	7.4	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G352.1108+00.1761	261.1815	-35.4773	11.0	9.4	8.4	7.5	0 0 0 0
G350.0111-01.3417	261.2773	-38.0669	4.4	3.3	1.8	1.5	1 1 1 1
G351.5809-00.3529	261.3547	-36.2128	11.8	9.7	8.5	8.1	0 0 0 0
G353.2731+00.6411	261.5066	-34.2543	6.5	4.8	2.6	2.2	1 1 1 1
G353.4636+00.5618	261.7147	-34.1405	9.4	7.7	7.2	7.0	0 0 0 0
G353.3780+00.4377	261.7816	-34.2807	11.0	10.5	9.4	8.5	0 0 0 0
G352.5252-00.1583	261.8059	-35.3210	10.8	8.9	7.7	6.8	0 0 0 0
G352.6040-00.2254	261.9280	-35.2928	13.2	11.2	10.3	10.7	0 0 0 0
G352.8550-00.2006	262.0733	-35.0703	12.8	10.8	9.2	7.8	0 0 0 0
G353.3701-00.0909	262.3095	-34.5806	11.7	10.8	10.3	9.7	0 0 0 0
G352.1324-00.9438	262.3426	-36.0834	8.6	6.8	6.0	5.5	0 0 0 0
G353.4292-00.0896	262.3478	-34.5306	12.4	10.3	8.5	8.1	0 0 0 0
G353.3632-00.1659	262.3808	-34.6279	7.5	6.6	4.3	2.8	0 0 0 1
G353.5375-00.0911	262.4219	-34.4412	10.5	8.2	6.9	6.3	0 0 0 0
G352.6302-01.0671	262.8080	-35.7358	7.7	5.3	3.6	3.1	0 1 1 1
G354.2055-00.0381	262.8125	-33.8542	11.0	9.5	7.7	6.2	0 0 0 0
G354.7243+00.3001	262.8148	-33.2349	7.5	5.5	4.2	3.0	0 1 0 1
G354.3077-00.1100	262.9523	-33.8081	13.3	11.1	9.5	8.2	0 0 0 0
G355.6420+00.3979	263.3142	-32.4130	12.0	10.7	9.4	8.1	0 0 0 0
G355.3430+00.1476	263.3702	-32.8001	9.9	8.6	6.6	5.1	0 0 0 0
G355.3437+00.1472	263.3711	-32.7997	9.9	8.6	6.6	5.1	0 0 0 0
G355.5453-00.1030	263.7512	-32.7661	9.1	7.6	6.5	5.7	0 0 0 0
G356.6619-00.2631	264.6215	-31.9108	10.4	8.1	7.1	6.5	0 0 0 0
G357.5576-00.3211	265.2383	-31.1831	14.6	11.9	10.0	8.5	0 0 0 0
G357.5585-00.3212	265.2389	-31.1825	12.8	11.1	8.6	7.1	0 0 0 0
G357.9651-00.1636	265.3339	-30.7540	9.9	7.4	5.8	4.9	0 0 0 0
G357.9239-00.3369	265.4799	-30.8806	10.2	8.0	6.6	6.0	0 0 0 0
G358.9057+00.1056	265.6440	-29.8130	7.1	7.0	6.5	6.4	0 0 0 0
G358.9309-00.0295	265.7918	-29.8627	7.0	5.9	4.5	3.9	1 1 0 1
G359.1375+00.0314	265.8570	-29.6548	8.5	7.6	7.0	6.5	0 0 0 0
G358.3707-00.4680	265.8831	-30.5697	11.2	8.9	7.6	7.2	0 0 0 0
G358.3863-00.4831	265.9076	-30.5643	8.2	6.6	5.4	4.1	0 0 0 0
G359.6148-00.2432	266.4129	-29.3917	7.7	6.2	5.9	6.3	0 1 0 0
G358.8412-00.7372	266.4345	-30.3093	10.7	9.6	8.9	8.0	0 0 0 0
G000.4962+00.1880	266.5165	-28.4147	11.4	9.6	8.0	7.0	0 0 0 0
G000.8358+00.1844	266.7203	-28.1263	8.3	8.0	6.2	4.8	0 0 0 0
G000.3154-00.2008	266.7880	-28.7710	7.4	6.1	5.5	4.8	0 1 0 0
G000.3157-00.2015	266.7889	-28.7711	7.4	6.1	5.5	4.8	0 1 0 0
G000.1669-00.4456	266.9394	-29.0248	9.8	6.8	5.4	4.7	0 0 0 0
G000.0916-00.6630	267.1079	-29.2016	10.9	8.8	8.5	9.3	0 0 0 0
G001.1469-00.1245	267.2022	-28.0198	8.7	6.8	5.5	4.5	0 0 0 0
G001.0081-00.2369	267.2304	-28.1966	7.5	7.3	6.5	5.4	0 0 0 0
G358.2630-02.0608	267.4068	-31.4883	8.8	7.2	6.6	6.2	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G001.7191-00.0879	267.4993	-27.5103	9.2	8.4	7.9	7.4	0 0 0 0
G000.5463-00.8516	267.5605	-28.9087	9.4	7.6	5.4	3.6	0 0 0 1
G002.1434+00.0091	267.6506	-27.0963	8.5	7.6	5.7	4.3	0 0 0 0
G002.7034+00.0401	267.9416	-26.5991	9.9	9.5	9.4	9.9	0 0 0 0
G002.5919-00.0302	267.9455	-26.7309	8.4	5.7	7.9	7.9	0 1 0 0
G002.5206-00.2204	268.0882	-26.8892	11.6	9.8	9.1	9.1	0 0 0 0
G003.2532+00.0185	268.2748	-26.1369	13.4	11.1	9.3	8.2	0 0 0 0
G003.5022-00.2002	268.6253	-26.0332	12.2	10.8	10.1	10.2	0 0 0 0
G003.3119-00.3986	268.7088	-26.2976	10.4	8.3	6.8	6.4	0 0 0 0
G003.4421-00.3484	268.7338	-26.1599	13.4	9.9	8.4	8.6	0 0 0 0
G004.6762+00.2763	268.8264	-24.7793	7.8	6.2	4.8	4.1	0 1 0 0
G004.4343+00.1289	268.8323	-25.0624	13.1	11.2	10.9	10.1	0 0 0 0
G004.3933+00.0786	268.8574	-25.1232	10.9	10.5	10.4	12.2	0 0 0 0
G004.5688-00.0791	269.1054	-25.0510	11.3	9.7	8.6	7.9	0 0 0 0
G004.8662-00.1715	269.3582	-24.8401	13.4	12.1	10.9	9.6	0 0 0 0
G005.6771-00.0274	269.6666	-24.0659	11.5	10.1	8.6	7.4	0 0 0 0
G005.8852-00.3934	270.1277	-24.0676	5.8	4.3	2.5	1.8	1 1 1 1
G006.8808+00.0931	270.2058	-22.9618	13.1	11.6	9.9	8.6	0 0 0 0
G006.5393-00.1085	270.2119	-23.3583	10.4	8.8	7.3	6.3	0 0 0 0
G006.6097-00.0823	270.2251	-23.2842	6.4	5.8	5.0	4.5	1 1 0 0
G006.1889-00.3581	270.2590	-23.7863	9.7	7.9	7.2	7.1	0 0 0 0
G006.5877-00.1919	270.3170	-23.3576	10.4	9.3	8.5	7.8	0 0 0 0
G006.7952-00.2573	270.4906	-23.2097	7.9	6.6	5.2	4.1	0 0 0 0
G007.6010-00.1389	270.8101	-22.4503	6.3	4.9	3.2	3.0	1 1 1 1
G008.8316-00.0280	271.3570	-21.3236	10.6	9.0	8.8	9.4	0 0 0 0
G009.2146-00.2017	271.7202	-21.0743	12.7	10.1	9.0	9.0	0 0 0 0
G008.8723-00.4926	271.8139	-21.5149	6.4	5.0	3.1	2.6	1 1 1 1
G009.9860-00.0276	271.9588	-20.3157	12.7	12.8	10.6	11.0	0 0 0 0
G010.4799+00.0330	272.1578	-19.8545	13.9	10.5	9.1	9.3	0 0 0 0
G010.4723+00.0273	272.1592	-19.8639	12.5	10.0	7.8	6.0	0 0 0 0
G010.3423-00.1423	272.2500	-20.0598	10.8	9.4	8.7	9.2	0 0 0 0
G010.8856+00.1228	272.2833	-19.4561	6.7	5.4	4.1	3.5	1 1 0 1
G010.3205-00.2586	272.3471	-20.1353	11.8	9.0	7.5	6.3	0 0 0 0
G010.2048-00.3455	272.3685	-20.2785	11.6	9.1	8.4	8.2	0 0 0 0
G010.9583+00.0223	272.4138	-19.4411	10.1	8.2	6.6	5.3	0 0 0 0
G011.0341+00.0618	272.4160	-19.3556	10.0	9.0	6.9	5.4	0 0 0 0
G010.6287-00.3329	272.5749	-19.9013	12.5	10.4	9.6	9.2	0 0 0 0
G011.1093-00.1140	272.6177	-19.3748	10.0	7.3	6.1	5.7	0 0 0 0
G010.6265-00.3844	272.6218	-19.9281	11.7	9.6	8.4	7.0	0 0 0 0
G012.8887+00.4892	272.9642	-17.5249	12.1	8.4	6.6	6.2	0 0 0 0
G012.0254-00.0314	273.0078	-18.5321	8.9	6.6	5.4	4.8	0 0 0 0
G011.9031-00.1023	273.0113	-18.6735	13.2	10.8	10.2	11.0	0 0 0 0
G011.9041-00.1411	273.0477	-18.6913	12.0	9.4	7.8	6.7	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G011.9361-00.1499	273.0720	-18.6674	12.9	10.6	10.0	10.5	0 0 0 0
G012.1992-00.0335	273.0977	-18.3808	10.5	7.7	6.3	6.0	0 0 0 0
G012.1116-00.1262	273.1391	-18.5021	13.1	12.0	10.2	8.8	0 0 0 0
G012.2654-00.0512	273.1475	-18.3312	12.8	10.7	9.7	9.8	0 0 0 0
G012.2092-00.1023	273.1663	-18.4050	11.8	11.1	9.1	8.0	0 0 0 0
G012.2026-00.1073	273.1677	-18.4132	9.9	8.9	6.9	5.5	0 0 0 0
G012.1811-00.1225	273.1708	-18.4394	12.6	10.1	8.9	8.4	0 0 0 0
G012.2018-00.1199	273.1789	-18.4199	12.1	11.2	10.1	8.9	0 0 0 0
G011.9916-00.2720	273.2133	-18.6773	8.4	6.0	4.8	4.3	0 1 0 0
G012.5257+00.0162	273.2168	-18.0704	8.3	6.3	5.4	4.8	0 1 0 0
G012.7762+00.1281	273.2399	-17.7970	12.1	10.7	9.5	9.4	0 0 0 0
G012.6247-00.0167	273.2971	-17.9993	11.4	7.9	6.5	6.3	0 0 0 0
G012.9044-00.0311	273.4511	-17.7608	11.5	9.4	8.5	9.0	0 0 0 0
G012.6806-00.1823	273.4781	-18.0296	11.7	8.4	7.0	6.1	0 0 0 0
G012.9086-00.2603	273.6647	-17.8667	5.4	3.9	3.5	1.3	1 1 1 1
G013.7129-00.0836	273.9041	-17.0755	10.8	8.0	6.8	6.8	0 0 0 0
G013.6961-00.1563	273.9627	-17.1249	11.1	9.1	7.6	6.7	0 0 0 0
G014.5206+00.1552	274.0864	-16.2515	13.0	12.0	10.7	9.7	0 0 0 0
G014.3895-00.0203	274.1824	-16.4503	11.1	9.9	7.8	6.4	0 0 0 0
G014.6041+00.0167	274.2548	-16.2439	12.6	10.3	9.2	9.1	0 0 0 0
G014.4422-00.1509	274.3283	-16.4660	10.8	8.0	6.8	6.8	0 0 0 0
G015.0942+00.1915	274.3368	-15.7296	7.1	6.0	4.6	4.1	0 1 0 0
G013.6567-00.5992	274.3511	-17.3701	7.3	5.8	4.7	3.6	0 1 0 1
G018.3414+01.7681	274.4922	-12.1236	5.7	4.5	2.4	2.1	1 1 1 1
G014.9908-00.1214	274.5722	-15.9690	10.3	8.7	7.2	6.4	0 0 0 0
G016.8549+00.6412	274.7899	-13.9660	11.8	11.2	8.8	7.2	0 0 0 0
G015.6066-00.2553	274.9973	-15.4897	11.0	9.2	8.5	9.2	0 0 0 0
G015.6653-00.4990	275.2490	-15.5528	11.3	9.2	8.6	8.9	0 0 0 0
G016.3019-00.1960	275.2826	-14.8485	9.9	7.4	6.1	5.5	0 0 0 0
G016.5853-00.0507	275.2880	-14.5301	12.7	9.3	7.7	7.4	0 0 0 0
G016.1118-00.3034	275.2881	-15.0668	10.5	8.1	6.5	5.7	0 0 0 0
G016.9763-00.0048	275.4362	-14.1635	11.3	9.8	8.3	7.2	0 0 0 0
G016.6621-00.3312	275.5811	-14.5942	11.1	9.7	7.7	6.4	0 0 0 0
G017.6381+00.1568	275.6096	-13.5034	2.5	1.3	2.0	2.3	1 1 1 1
G017.8616+00.0745	275.7921	-13.3447	10.8	9.5	7.9	6.6	0 0 0 0
G018.0727+00.0768	275.8916	-13.1571	9.4	6.7	5.4	5.1	0 0 0 0
G018.4397+00.0447	276.0972	-12.8478	9.1	8.2	7.5	7.1	0 0 0 0
G018.4604-00.0042	276.1514	-12.8524	9.9	8.0	6.3	4.8	0 0 0 0
G018.6675+00.0251	276.2241	-12.6557	11.9	9.5	8.4	8.5	0 0 0 0
G018.2621-00.2439	276.2738	-13.1398	9.0	7.5	6.5	5.8	0 0 0 0
G019.0086-00.0293	276.4366	-12.3795	11.4	7.9	6.4	6.2	0 0 0 0
G019.4958+00.1154	276.5382	-11.8810	11.0	8.9	8.2	8.4	0 0 0 0
G018.8337-00.3003	276.5986	-12.6606	8.7	7.4	5.8	4.8	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G019.3651-00.0301	276.6075	-12.0644	10.5	8.7	8.1	8.1	0 0 0 0
G019.6136+00.0112	276.6885	-11.8254	12.0	10.7	9.8	10.4	0 0 0 0
G018.8884-00.4745	276.7827	-12.6933	13.0	10.4	9.3	9.4	0 0 0 0
G019.7549-00.1284	276.8819	-11.7653	8.8	7.3	5.5	4.2	0 0 0 0
G020.2370+00.0654	276.9357	-11.2484	11.7	8.7	7.4	7.8	0 0 0 0
G020.237014+0.065404	276.9357	-11.2484	11.7	8.7	7.4	7.8	0 0 0 0
G020.239295+0.064806	276.9373	-11.2466	9.2	7.0	6.2	6.0	0 0 0 0
G019.7005-00.2669	276.9813	-11.8779	11.0	10.4	8.3	7.0	0 0 0 0
G020.0810-00.1354	277.0430	-11.4799	11.3	8.4	7.2	7.7	0 0 0 0
G020.9267-00.0497	277.3658	-10.6911	10.0	7.4	5.8	4.9	0 0 0 0
G020.9633-00.0746	277.4056	-10.6702	10.1	9.2	6.7	5.2	0 0 0 0
G021.0227-00.0629	277.4231	-10.6122	12.4	9.3	7.6	6.6	0 0 0 0
G022.0383+00.2221	277.6446	-9.5801	12.8	10.2	9.3	10.2	0 0 0 0
G021.562665-0.032743	277.6503	-10.1197	6.3	5.0	3.0	2.5	1 1 1 1
G021.5624-00.0329	277.6503	-10.1200	6.3	5.0	3.0	2.5	1 1 1 1
G021.4065-00.2544	277.7765	-10.3607	10.7	9.9	8.0	6.4	0 0 0 0
G023.1260+00.3949	277.9990	-8.5359	11.8	10.7	8.9	7.6	0 0 0 0
G022.3347-00.1549	278.1225	-9.4917	13.6	11.7	10.5	9.4	0 0 0 0
G022.4352-00.1694	278.1826	-9.4092	13.1	10.9	9.6	8.8	0 0 0 0
G022.435285-0.169352	278.1826	-9.4091	13.1	10.9	9.6	8.8	0 0 0 0
G023.389027+0.185324	278.3097	-8.3994	5.2	4.2	2.0	1.7	1 1 1 1
G023.4399-00.1823	278.6633	-8.5237	11.6	9.8	8.8	9.0	0 0 0 0
G023.4366-00.1844	278.6636	-8.5276	15.3	10.4	8.8	8.6	0 0 0 0
G023.0097-00.4106	278.6678	-9.0106	10.1	7.4	6.2	6.5	0 0 0 0
G023.009789-0.410657	278.6679	-9.0106	10.1	7.4	6.2	6.5	0 0 0 0
G023.3645-00.2908	278.7255	-8.6405	14.3	12.7	10.7	11.1	0 0 0 0
G024.3287+00.1444	278.7837	-7.5843	12.2	8.5	6.9	6.3	0 0 0 0
G024.329204+0.144411	278.7839	-7.5838	12.2	8.5	6.9	6.3	0 0 0 0
G023.7065-00.1981	278.8015	-8.2943	12.3	10.2	8.8	10.3	0 0 0 0
G024.789871+0.083705	279.0520	-7.2027	11.6	7.8	6.4	6.5	0 0 0 0
G024.789635+0.083290	279.0523	-7.2031	11.6	7.8	6.4	6.5	0 0 0 0
G024.7897+00.0832	279.0524	-7.2031	11.6	7.8	6.4	6.5	0 0 0 0
G024.789791+0.083231	279.0524	-7.2030	11.6	7.8	6.4	6.5	0 0 0 0
G024.850186+0.087388	279.0767	-7.1475	8.9	8.1	5.8	4.3	0 0 0 0
G024.8499+00.0872	279.0767	-7.1478	11.6	7.8	6.4	6.5	0 0 0 0
G025.7096+00.0439	279.5131	-6.4041	11.2	8.0	6.4	5.9	0 0 0 0
G027.784513+0.056883	280.4566	-4.5538	7.6	6.5	4.8	3.5	0 1 0 1
G028.1475-00.0040	280.6775	-4.2589	11.0	8.7	7.1	5.7	0 0 0 0
G028.2007-00.0492	280.7421	-4.2322	7.2	5.5	3.7	2.9	0 1 1 1
G028.3213-00.0106	280.7630	-4.1073	14.8	10.5	8.9	8.6	0 0 0 0
G028.321132-0.010706	280.7630	-4.1075	14.8	10.5	8.9	8.6	0 0 0 0
G028.607632+0.017793	280.8688	-3.8397	8.4	7.4	5.0	3.5	0 0 0 1
G028.2823-00.3584	281.0553	-4.3009	11.4	9.6	8.8	8.4	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G028.281911-0.358556	281.0553	-4.3014	11.4	9.6	8.8	8.4	0 0 0 0
G028.3050-00.3875	281.0916	-4.2940	8.8	7.0	5.7	4.7	0 0 0 0
G028.6872-00.2826	281.1731	-3.9061	5.7	4.7	2.8	2.5	1 1 1 1
G028.832271-0.252722	281.2129	-3.7634	10.0	8.3	7.9	8.9	0 0 0 0
G028.8323-00.2527	281.2129	-3.7634	10.0	8.3	7.9	8.9	0 0 0 0
G030.3709+00.4827	281.2613	-2.0588	11.8	10.0	8.5	7.3	0 0 0 0
G030.370407+0.482393	281.2613	-2.0594	11.8	10.0	8.5	7.3	0 0 0 0
G030.4228+00.4655	281.3003	-2.0204	13.7	12.1	10.2	8.7	0 0 0 0
G030.423532+0.465847	281.3003	-2.0196	13.7	12.1	10.2	8.7	0 0 0 0
G030.9726+00.5622	281.4654	-1.4871	8.2	6.4	5.4	4.8	0 1 0 0
G030.972735+0.562207	281.4654	-1.4870	8.2	6.4	5.4	4.8	0 1 0 0
G029.8646-00.0432	281.4983	-2.7493	6.1	4.8	3.2	3.0	1 1 1 1
G031.076397+0.457418	281.6060	-1.4426	14.5	12.5	10.4	9.0	0 0 0 0
G030.8185+00.2729	281.6525	-1.7563	7.2	5.8	5.6	5.3	0 1 0 0
G030.7810+00.2306	281.6730	-1.8089	11.0	9.9	7.6	6.2	0 0 0 0
G030.788127+0.203550	281.7004	-1.8150	11.2	9.0	8.1	7.4	0 0 0 0
G030.7901+00.2045	281.7004	-1.8128	11.9	10.5	9.8	9.5	0 0 0 0
G030.198626-0.168824	281.7628	-2.5094	5.7	4.6	2.6	2.1	1 1 1 1
G030.205880-0.165254	281.7629	-2.5014	6.2	6.0	3.3	1.8	1 1 1 1
G030.0106-00.2730	281.7696	-2.7243	11.8	10.0	8.5	7.3	0 0 0 0
G030.009672-0.273440	281.7696	-2.7253	13.7	10.6	8.9	7.6	0 0 0 0
G030.0095-00.2736	281.7697	-2.7255	13.7	12.0	10.9	9.9	0 0 0 0
G030.9799+00.2159	281.7770	-1.6387	8.6	6.7	5.6	5.2	0 0 0 0
G030.979752+0.215760	281.7770	-1.6389	8.6	6.7	5.6	5.2	0 0 0 0
G029.603488-0.625207	281.8975	-3.2472	10.0	9.8	9.6	9.2	0 0 0 0
G030.7603-00.0524	281.9155	-1.9565	10.5	8.1	6.6	5.4	0 0 0 0
G031.0605+00.0940	281.9223	-1.6226	11.6	11.6	9.9	8.3	0 0 0 0
G031.1219+00.0630	281.9778	-1.5821	11.0	9.4	6.7	5.1	0 0 0 0
G031.2823+00.0624	282.0516	-1.4396	11.8	10.0	8.3	7.0	0 0 0 0
G031.182483-0.148057	282.1934	-1.6245	13.0	10.7	9.4	8.4	0 0 0 0
G031.593744-0.192306	282.4204	-1.2786	13.0	11.1	9.3	7.7	0 0 0 0
G031.5943-00.1920	282.4204	-1.2780	13.0	11.1	9.3	7.7	0 0 0 0
G030.769882-0.804566	282.5898	-2.2910	11.2	9.3	8.7	8.5	0 0 0 0
G033.633757-0.020969	283.1982	0.6151	8.1	6.6	5.6	4.9	0 0 0 0
G034.257214+0.153055	283.3276	1.2493	6.8	4.8	2.8	2.0	1 1 1 1
G034.254793+0.138667	283.3393	1.2406	7.2	6.5	4.4	2.6	0 0 0 1
G033.317240-0.359814	283.3554	0.1789	11.5	10.1	8.3	6.7	0 0 0 0
G035.022567+0.348252	283.5030	2.0194	8.5	6.6	6.4	7.7	0 0 0 0
G038.258475-0.073436	285.3594	4.7055	10.0	8.6	7.4	6.2	0 0 0 0
G040.4254+00.7003	285.6651	6.9863	10.4	8.8	7.1	5.7	0 0 0 0
G040.6226-00.1383	286.5068	6.7768	8.9	6.6	5.3	4.1	0 0 0 0
G045.4727+00.1335	288.5307	11.2044	10.0	8.3	7.0	6.3	0 0 0 0
G045.4670+00.0530	288.6006	11.1619	11.8	9.3	7.8	7.2	0 0 0 0

Name	RA	Dec	Band 1	Band 3	Band 3	Band 4	Sat Warn
G049.4697-00.3707	290.9079	14.4998	11.2	9.2	8.7	9.2	0 0 0 0
G049.4898-00.3877	290.9331	14.5095	10.7	8.7	7.1	5.6	0 0 0 0

APPENDIX B

Minimum Spanning Trees

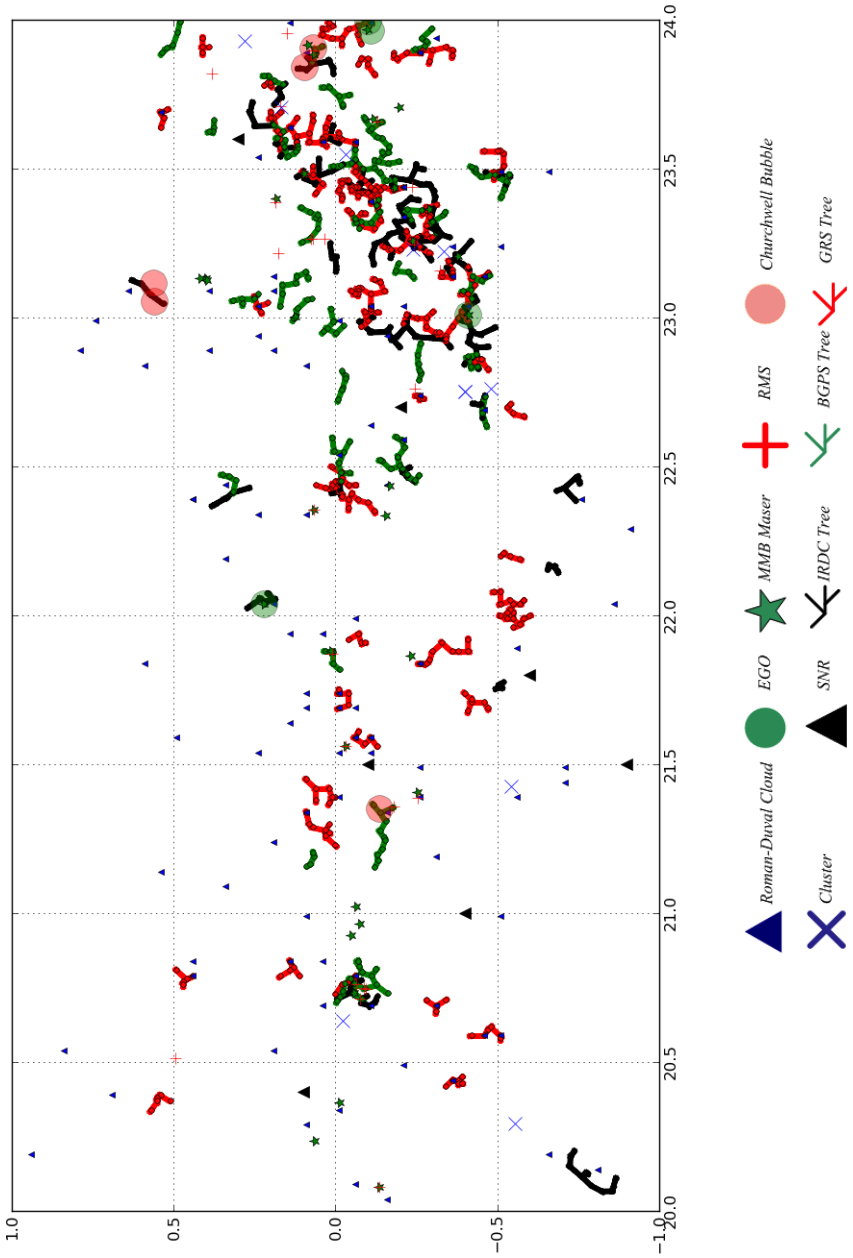


Figure B.1: Stripped Fractured MST $20^o < l < 24^o$
Plot of the Stripped Fracture MST contained with $20^o < l < 24^o$..

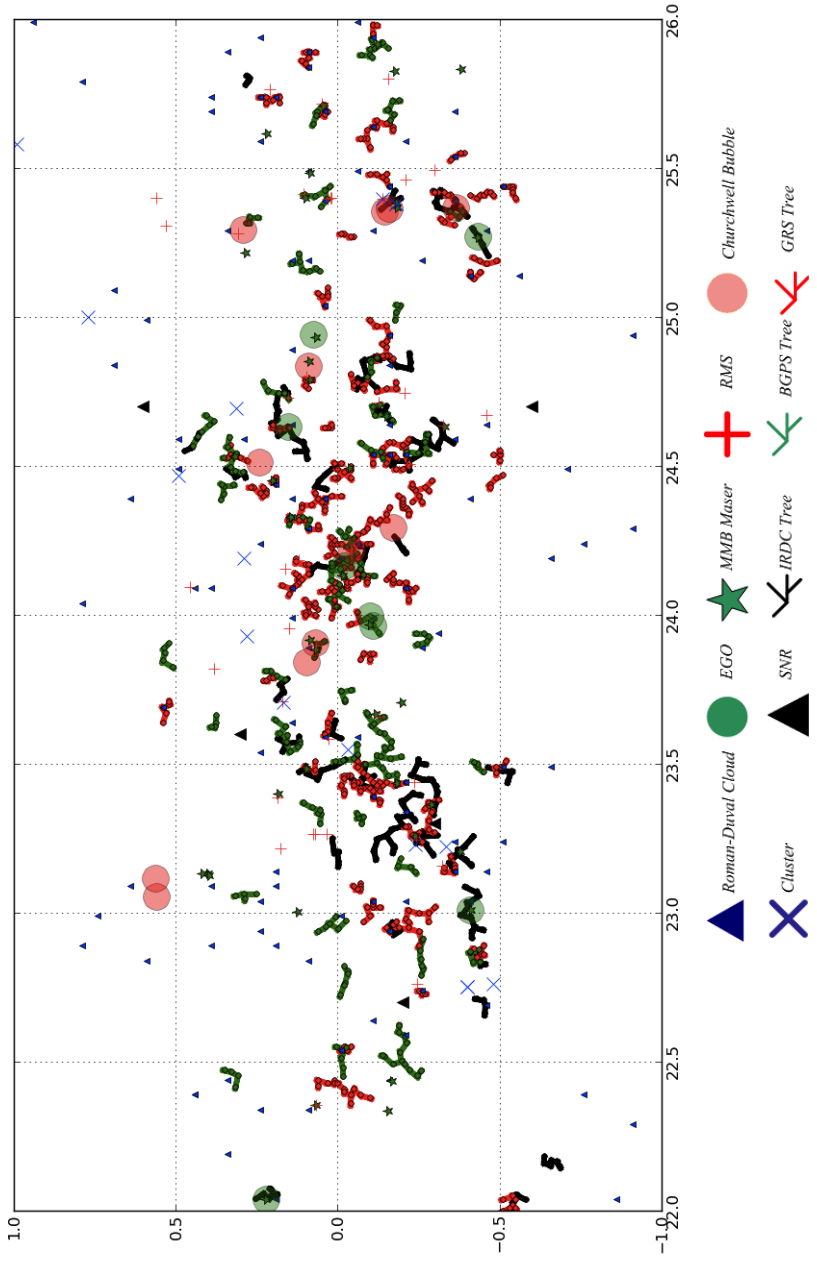


Figure B.2: Stripped Fractured MST $22^o < l < 26^o$
Plot of the Stripped Fracture MST contained with $22^o < l < 26^o$..

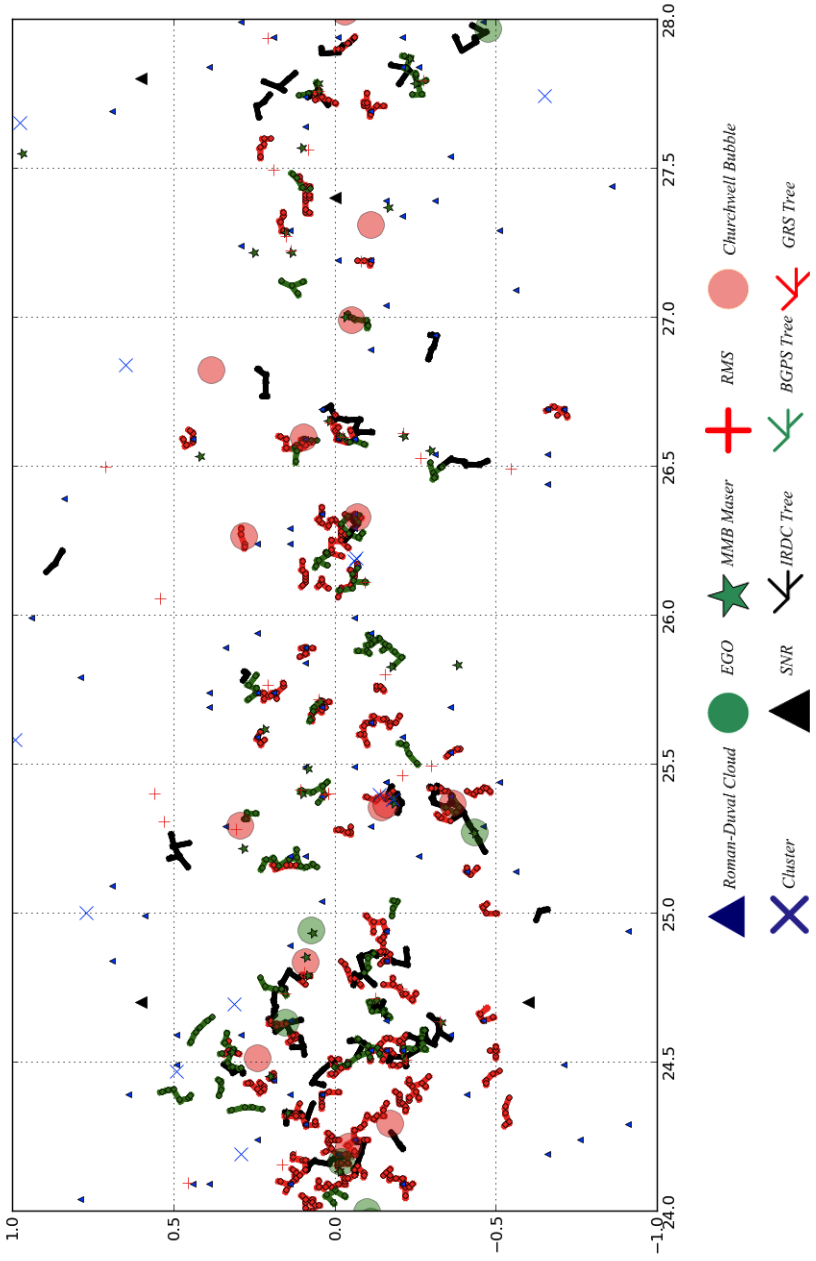


Figure B.3: Stripped Fractured MST $24^o < l < 28^o$
Plot of the Stripped Fracture MST contained with $24^o < l < 28^o$ ..

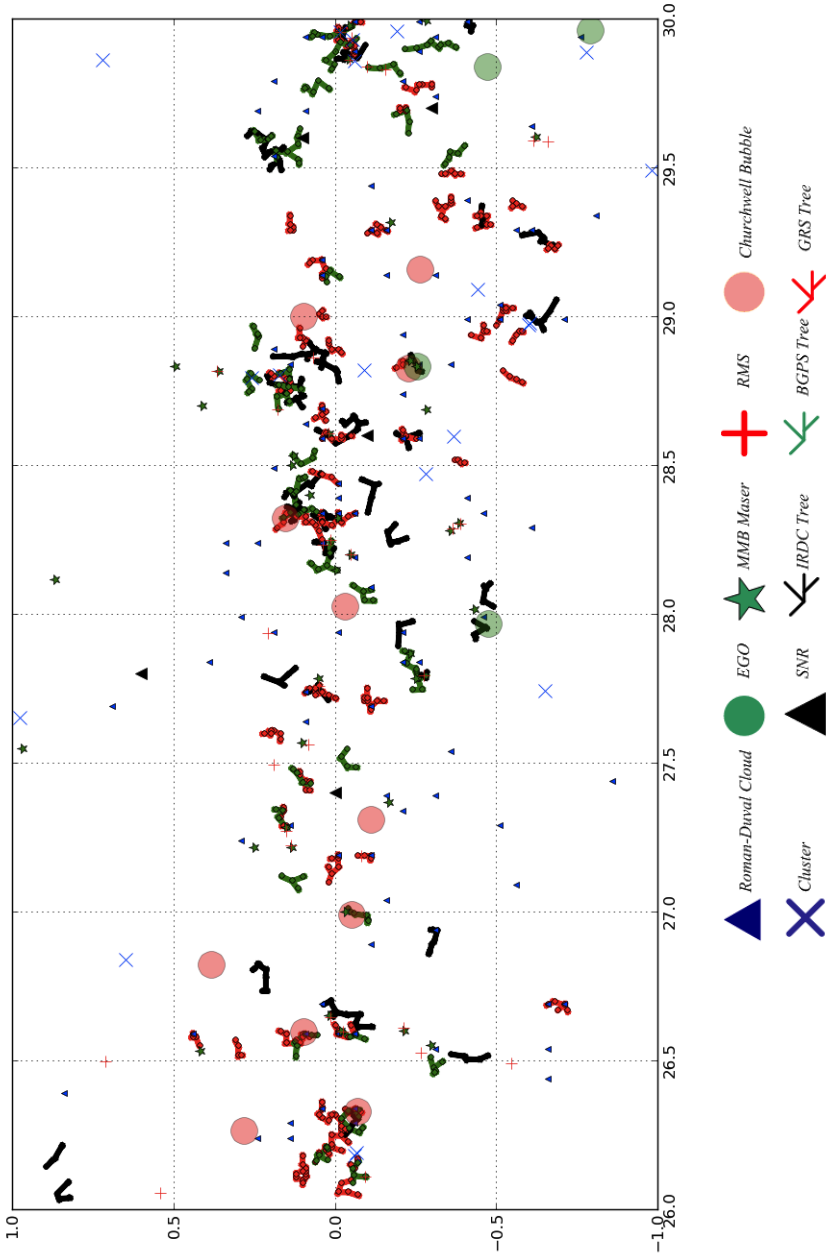


Figure B.4: Stripped Fractured MST $26^o < l < 39^o$
Plot of the Stripped Fracture MST contained with $26^o < l < 30^o$ ..

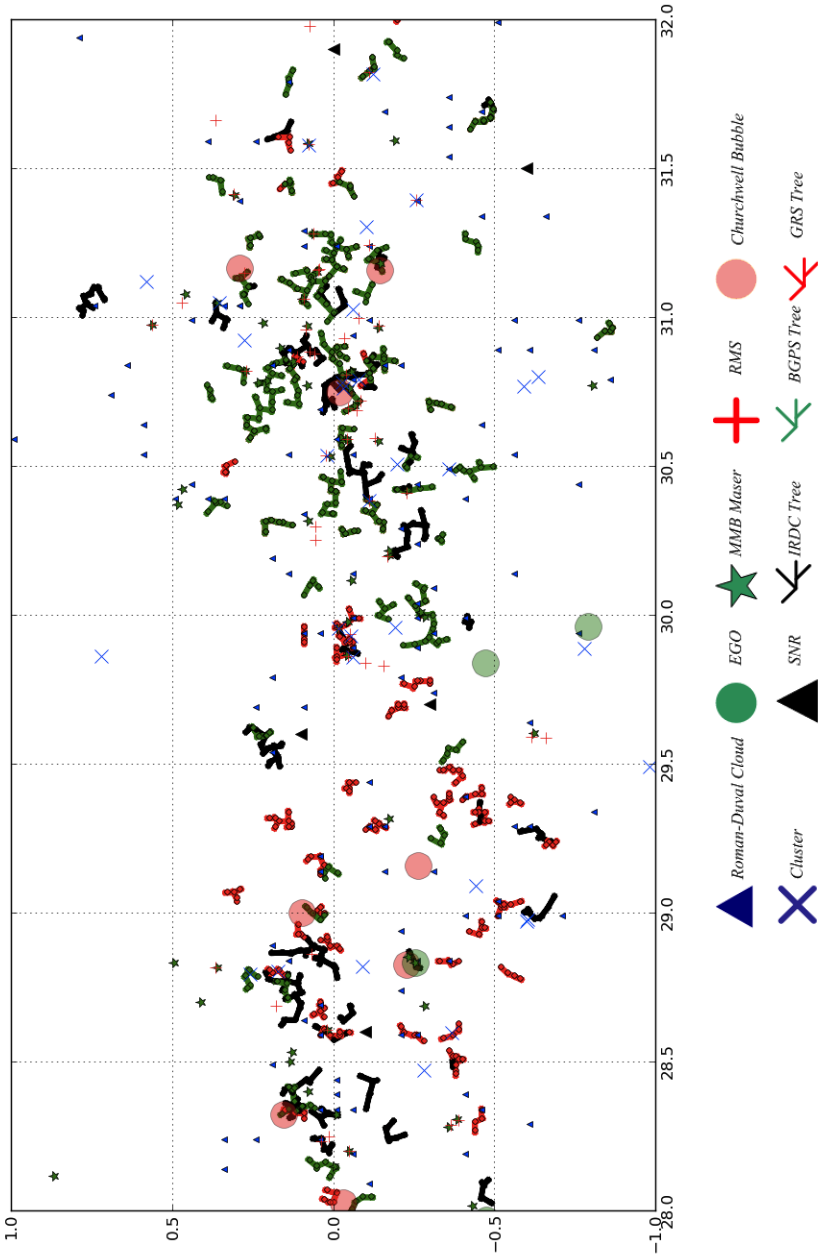


Figure B.5: Stripped Fractured MST $28^o < l < 32^o$
Plot of the Stripped Fracture MST contained with $28^o < l < 32^o$..

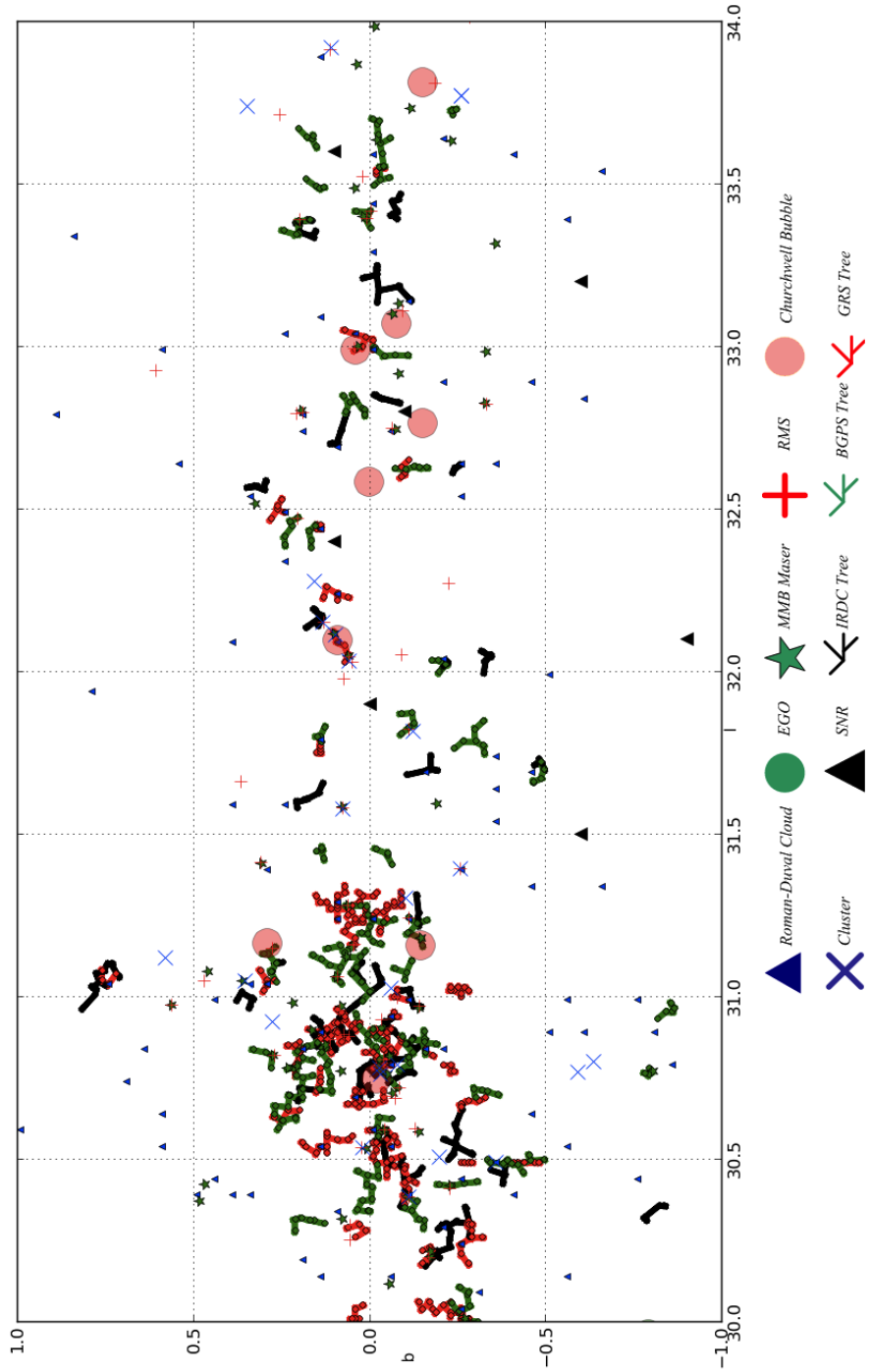


Figure B.6: Stripped Fractured MST $30^o < l < 34^o$
Plot of the Stripped Fracture MST contained with $30^o < l < 34^o$ ..

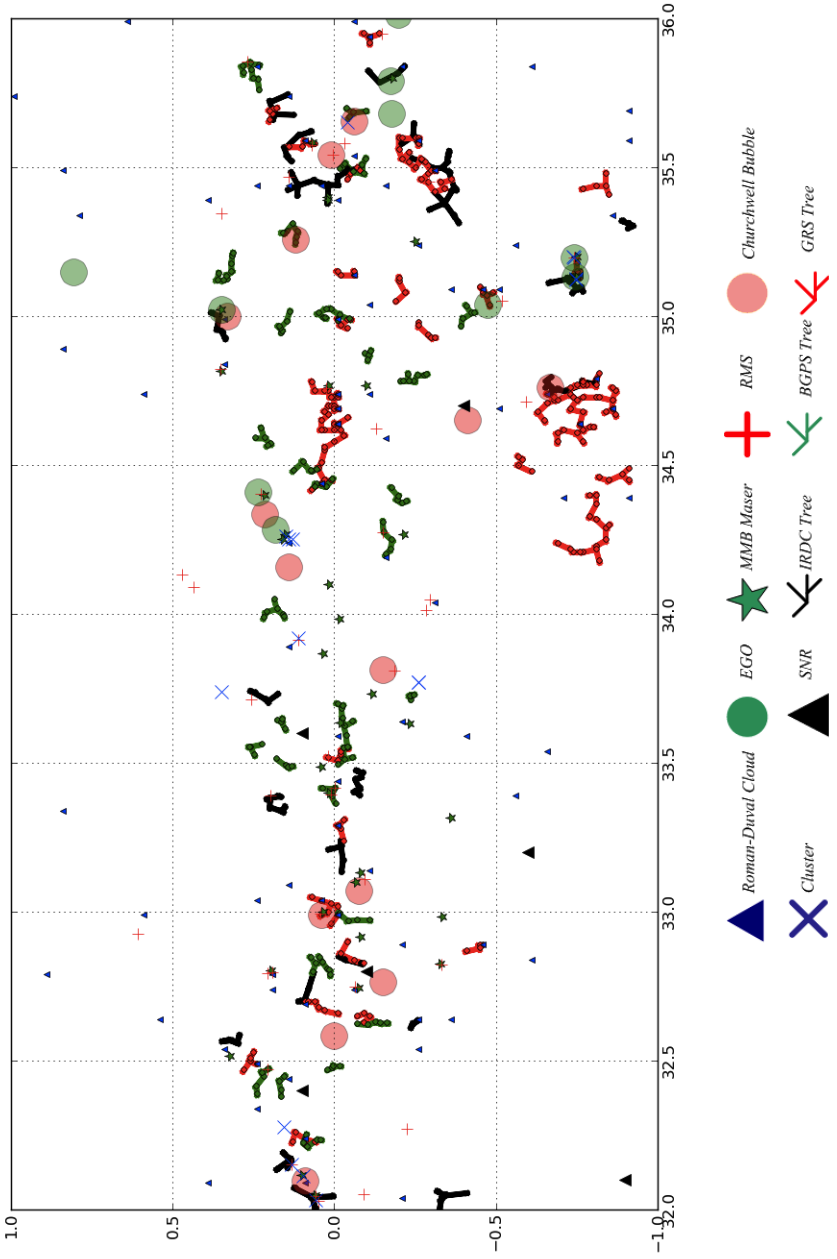


Figure B.7: Stripped Fractured MST $32^o < l < 36^o$
Plot of the Stripped Fracture MST contained with $32^o < l < 36^o$..

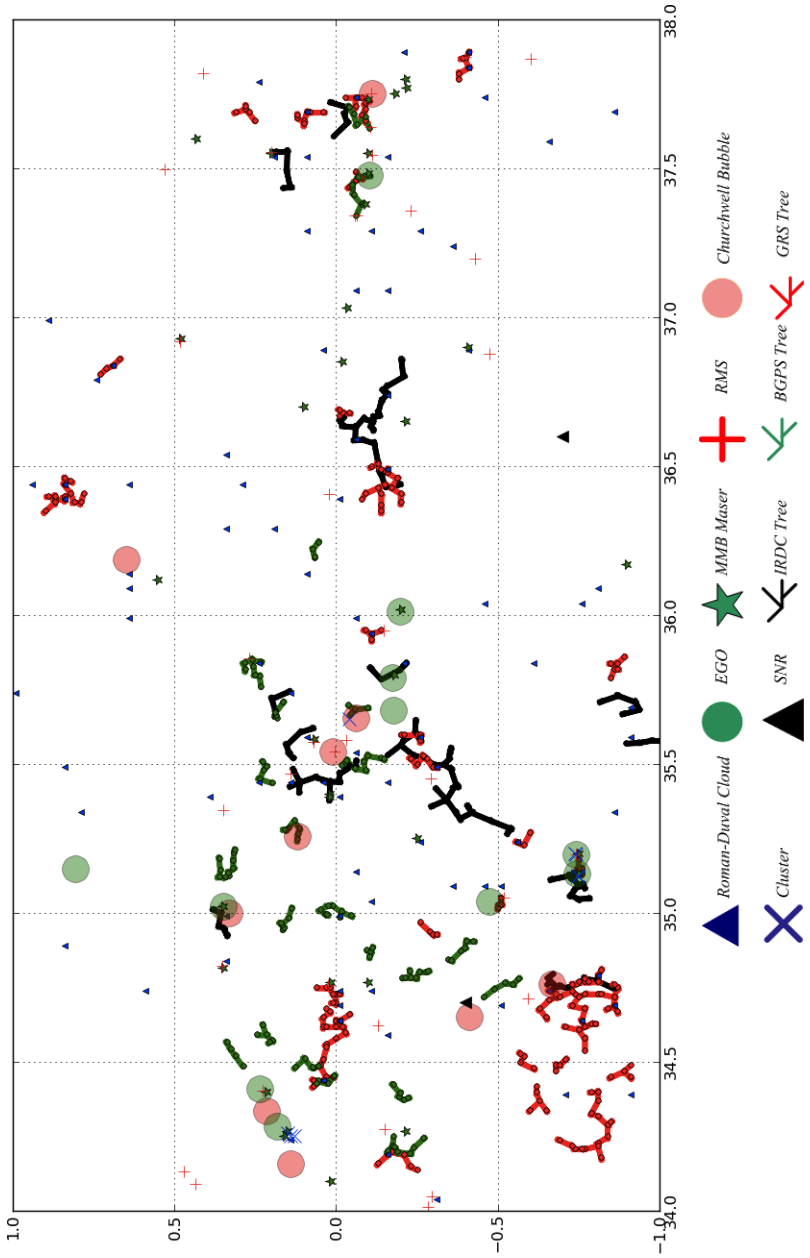


Figure B.8: Stripped Fractured MST $34^o < l < 38^o$
Plot of the Stripped Fracture MST contained with $30^o < l < 34^o$ ..

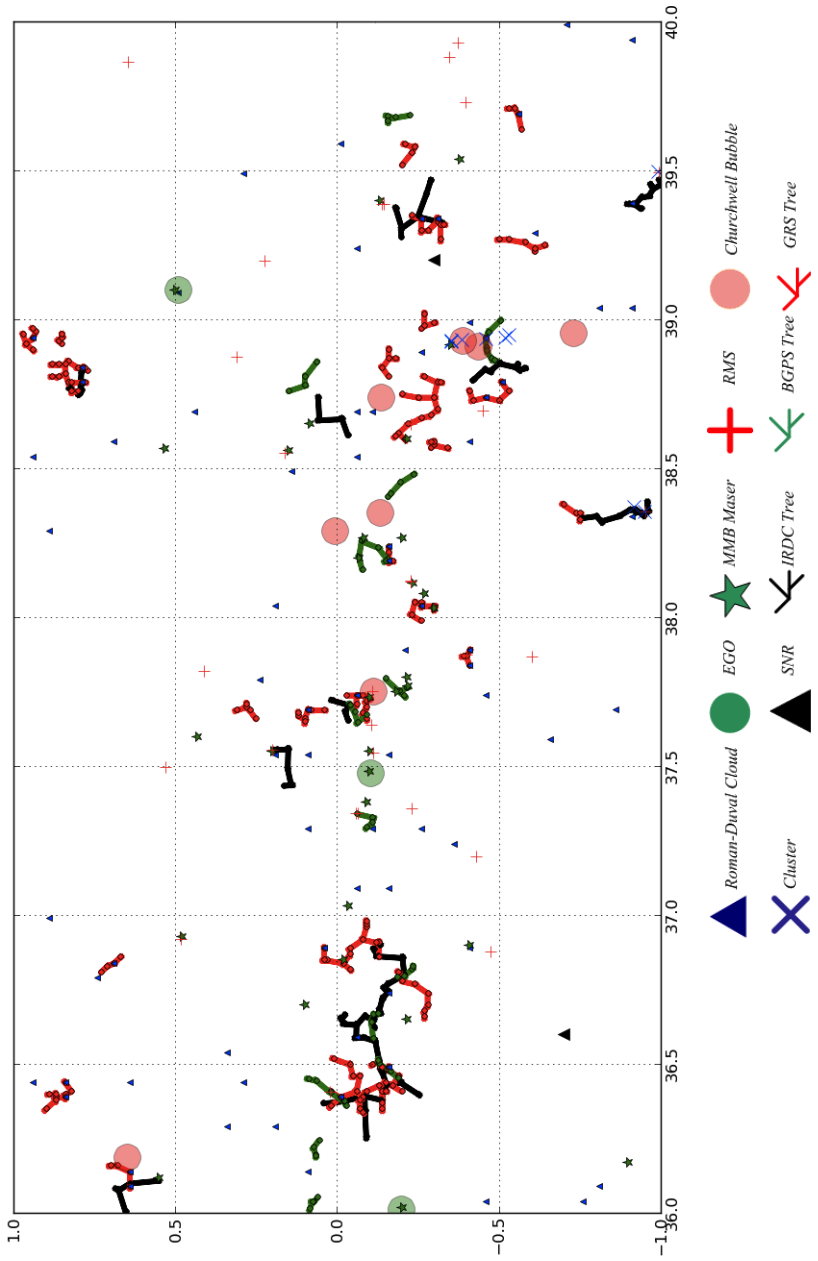


Figure B.9: Stripped Fractured MST $36^o < l < 40^o$
Plot of the Stripped Fracture MST contained with $30^o < l < 34^o$..

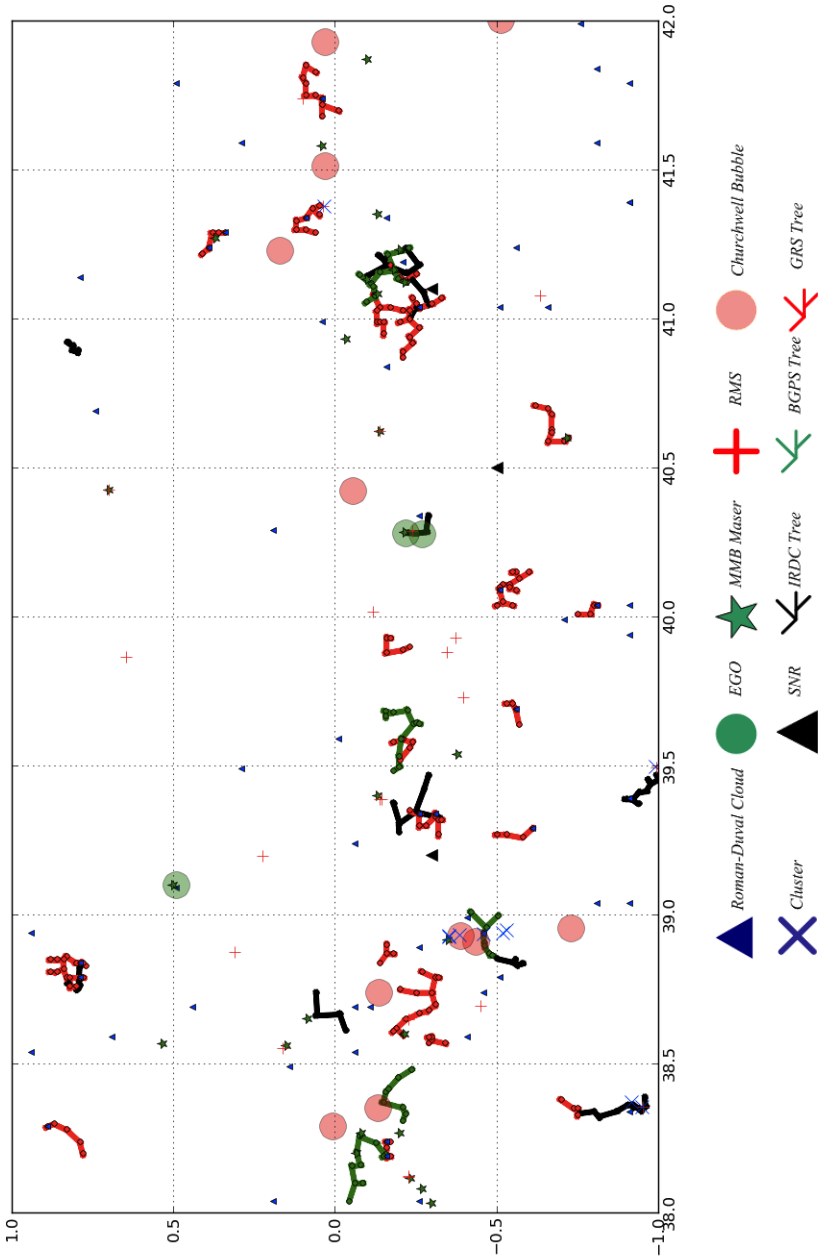


Figure B.10: Stripped Fractured MST $38^o < l < 42^o$
Plot of the Stripped Fracture MST contained with $30^o < l < 34^o$ ..

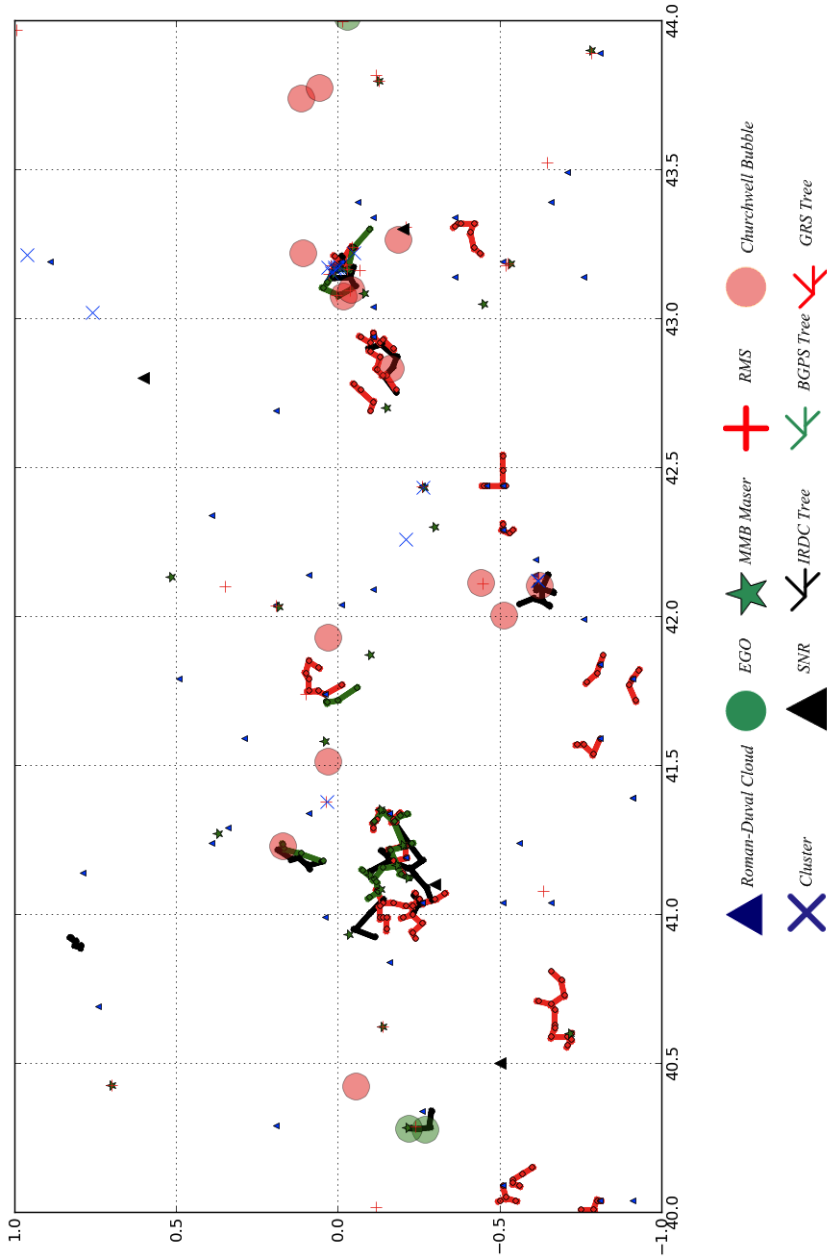


Figure B.11: Stripped Fractured MST $40^o < l < 44^o$
Plot of the Stripped Fracture MST contained with $30^o < l < 34^o$..

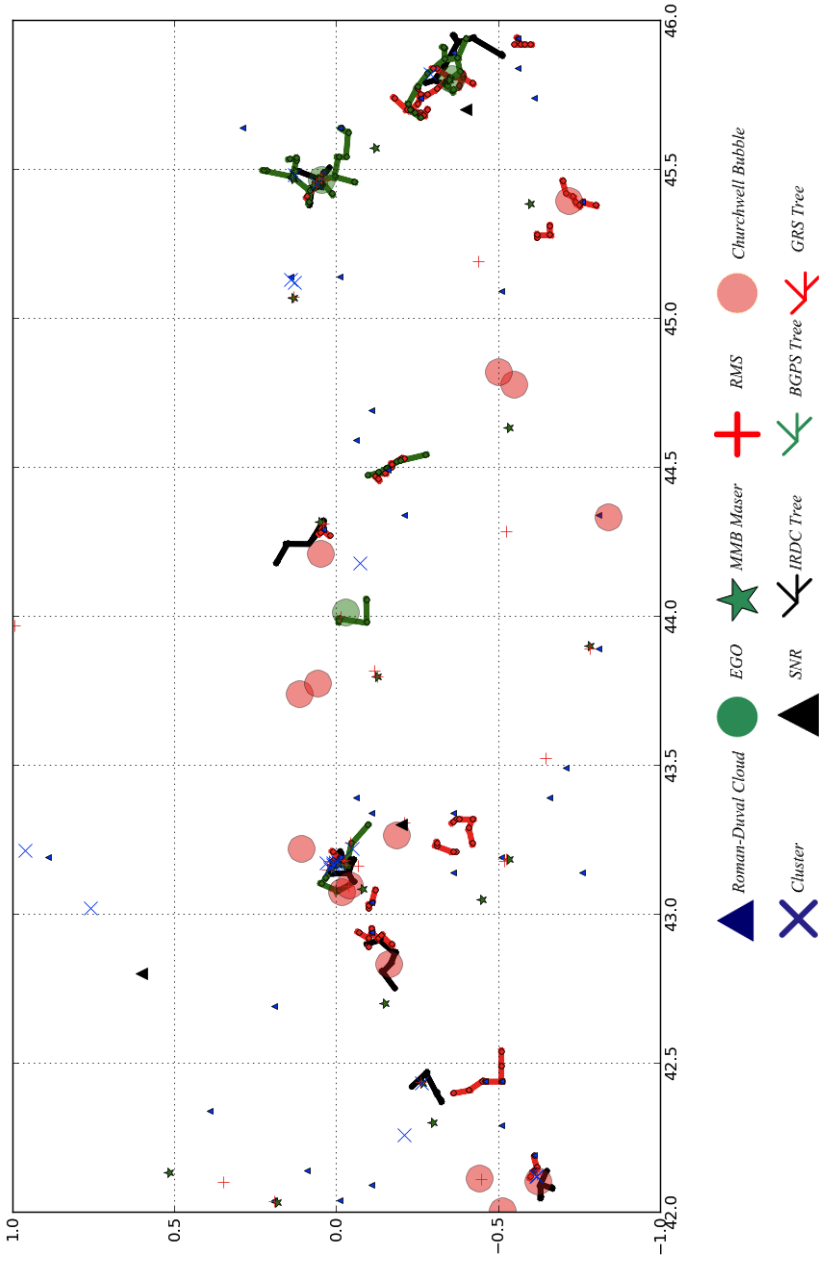


Figure B.12: Stripped Fractured MST $42^o < l < 46^o$
Plot of the Stripped Fracture MST contained with $30^o < l < 34^o$ ..

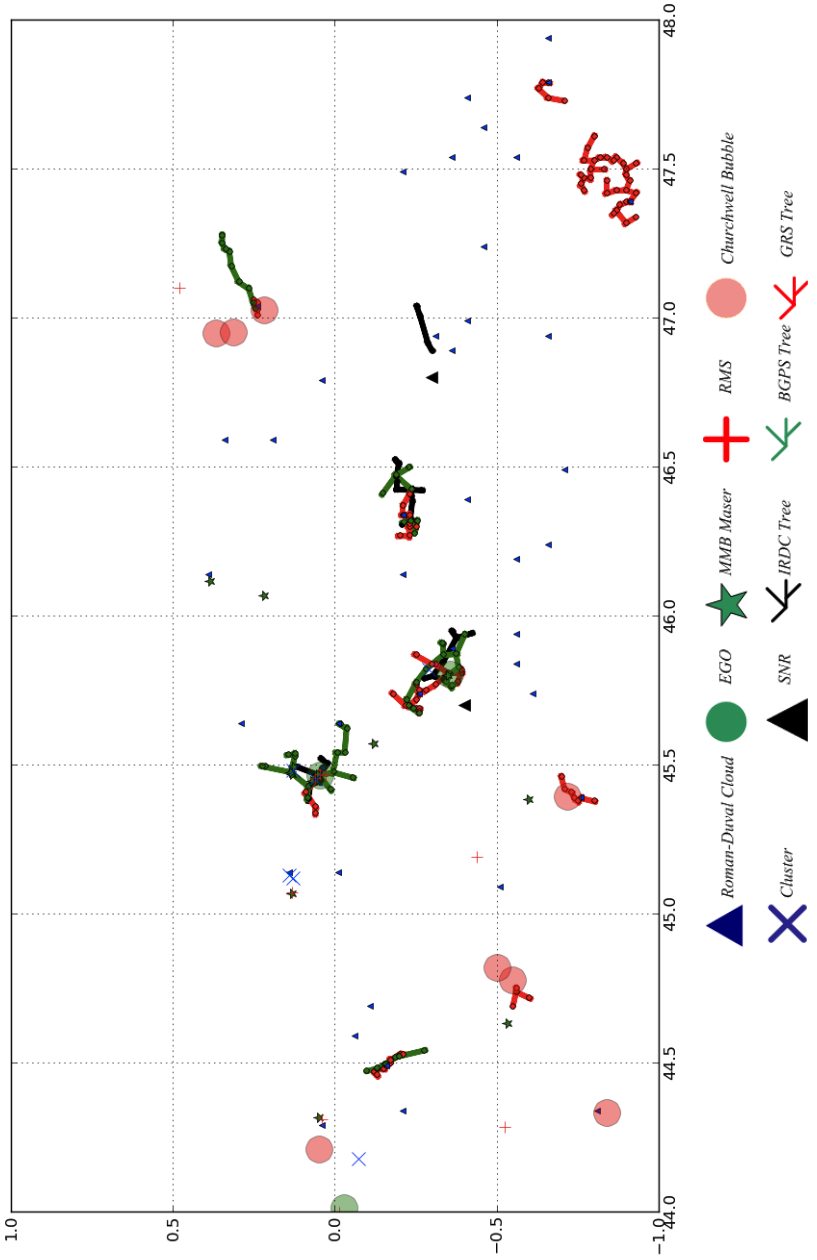


Figure B.13: Stripped Fractured MST $44^o < l < 48^o$
Plot of the Stripped Fracture MST contained with $30^o < l < 34^o$..

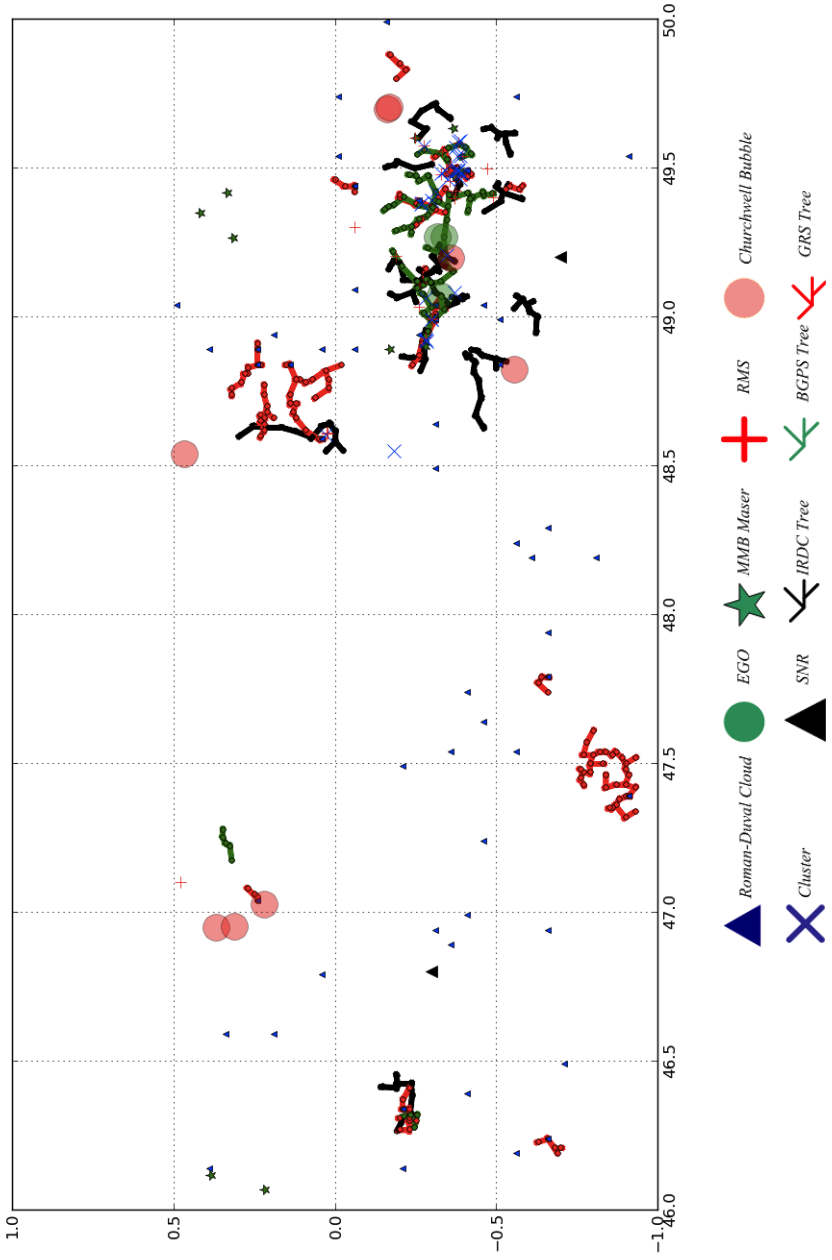


Figure B.14: Stripped Fractured MST $46^o < l < 50^o$
Plot of the Stripped Fracture MST contained with $30^o < l < 34^o$ ..

APPENDIX C

Convex Hulls

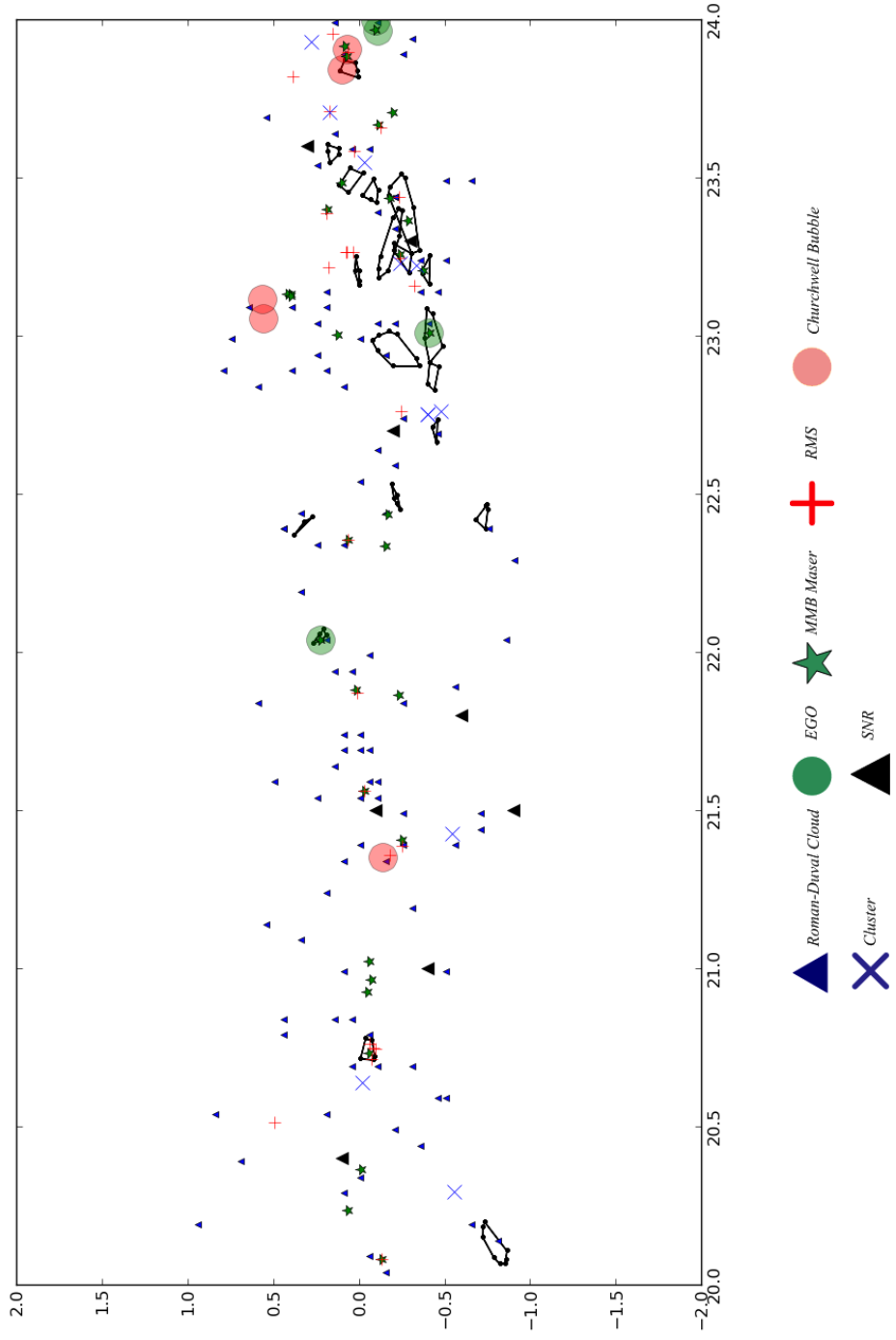


Figure C.1: Convex Hulls for P&F $20^o < l < 24^o$

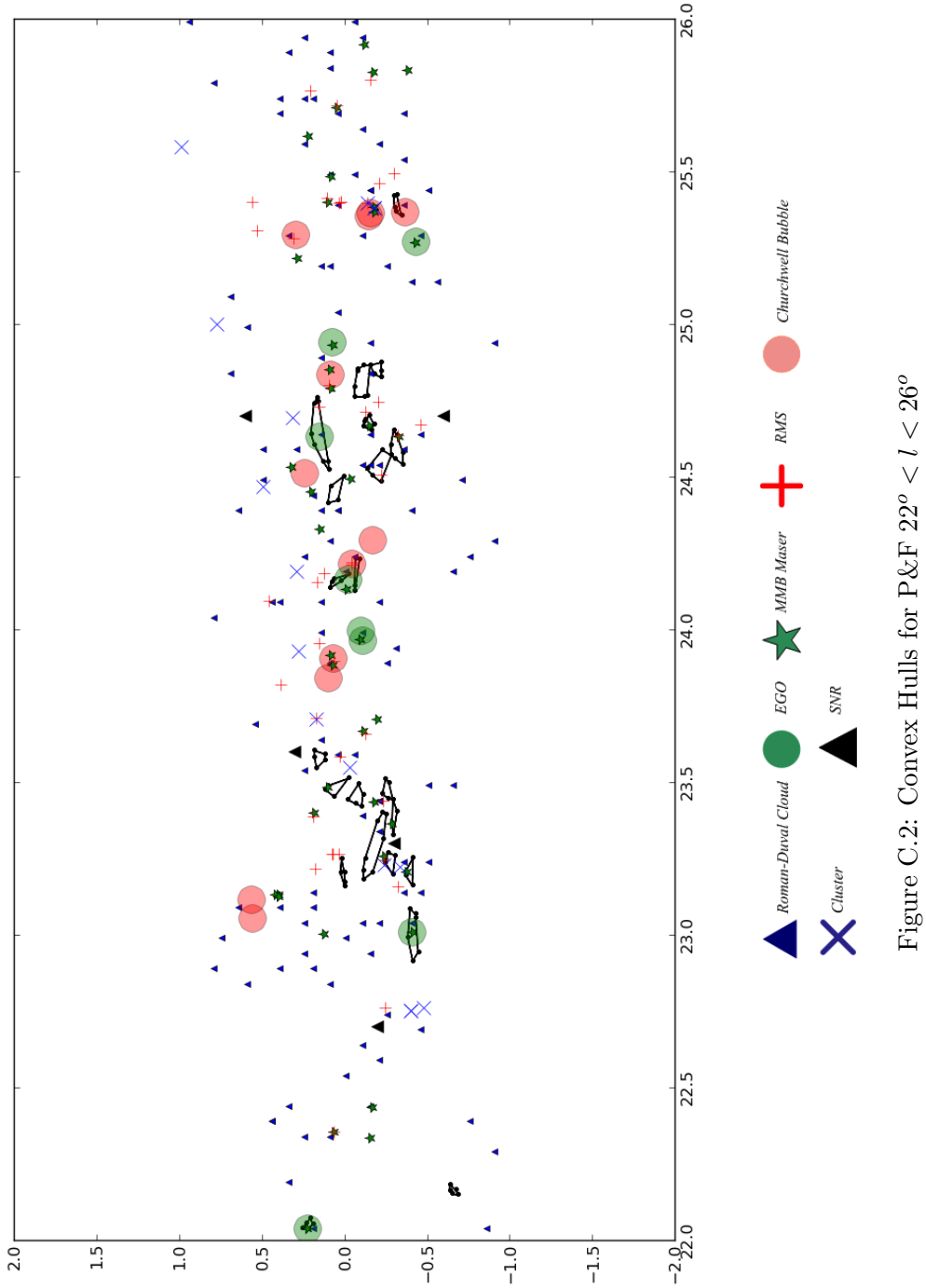


Figure C.2: Convex Hulls for P&F $22^o < l < 26^o$

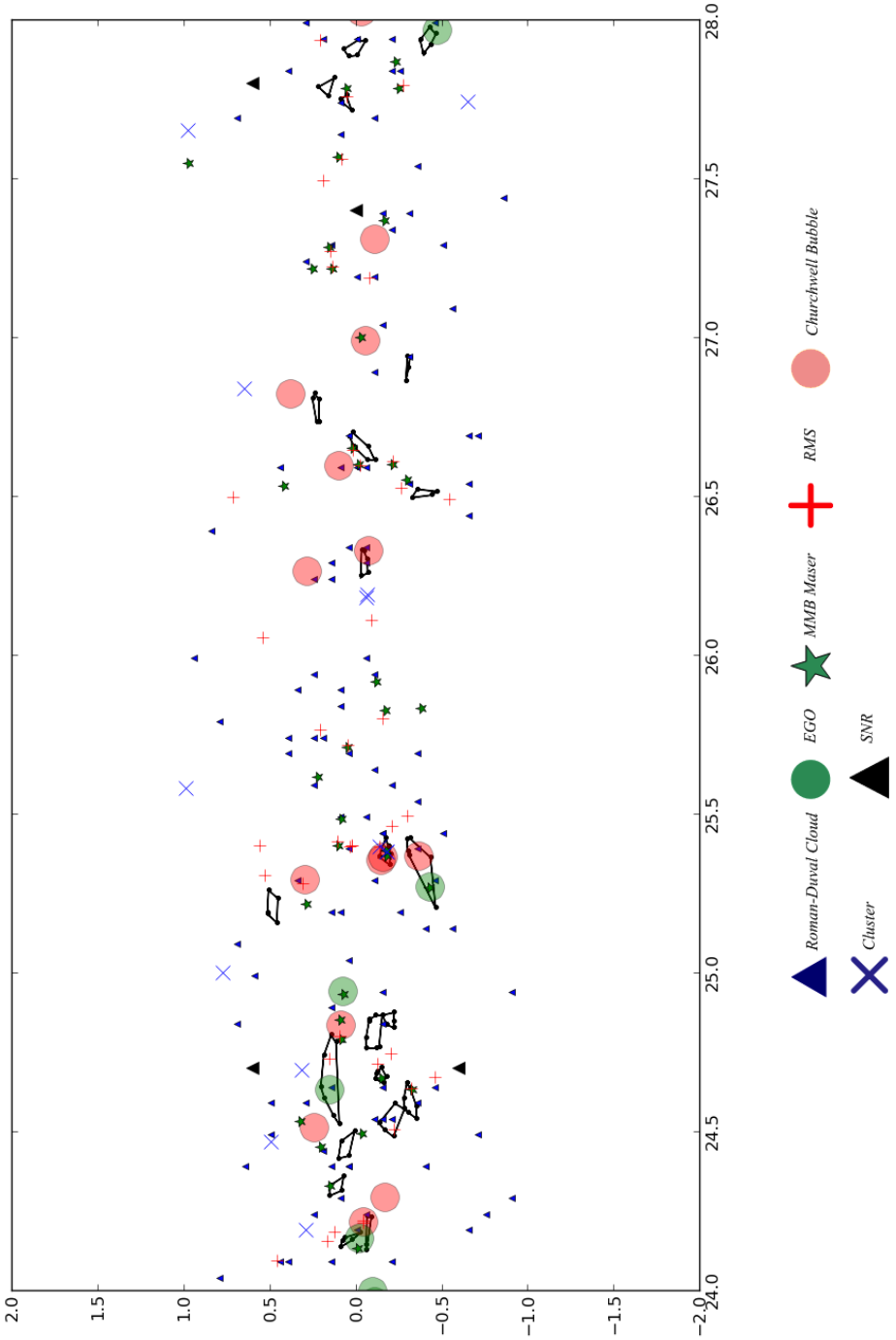


Figure C.3: Convex Hulls for P&F $24^o < l < 28^o$

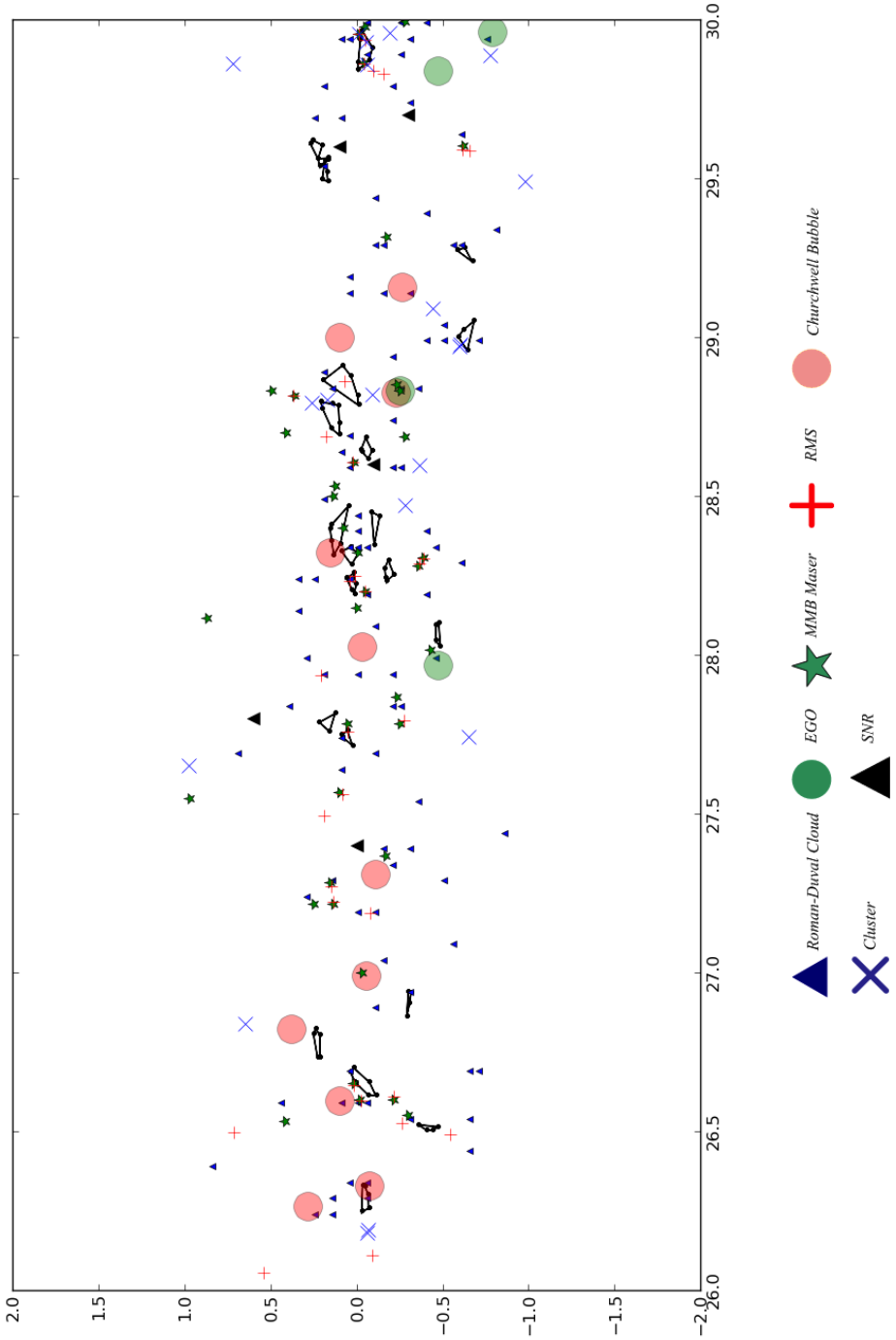


Figure C.4: Convex Hulls for P&F $26^o < l < 30^o$

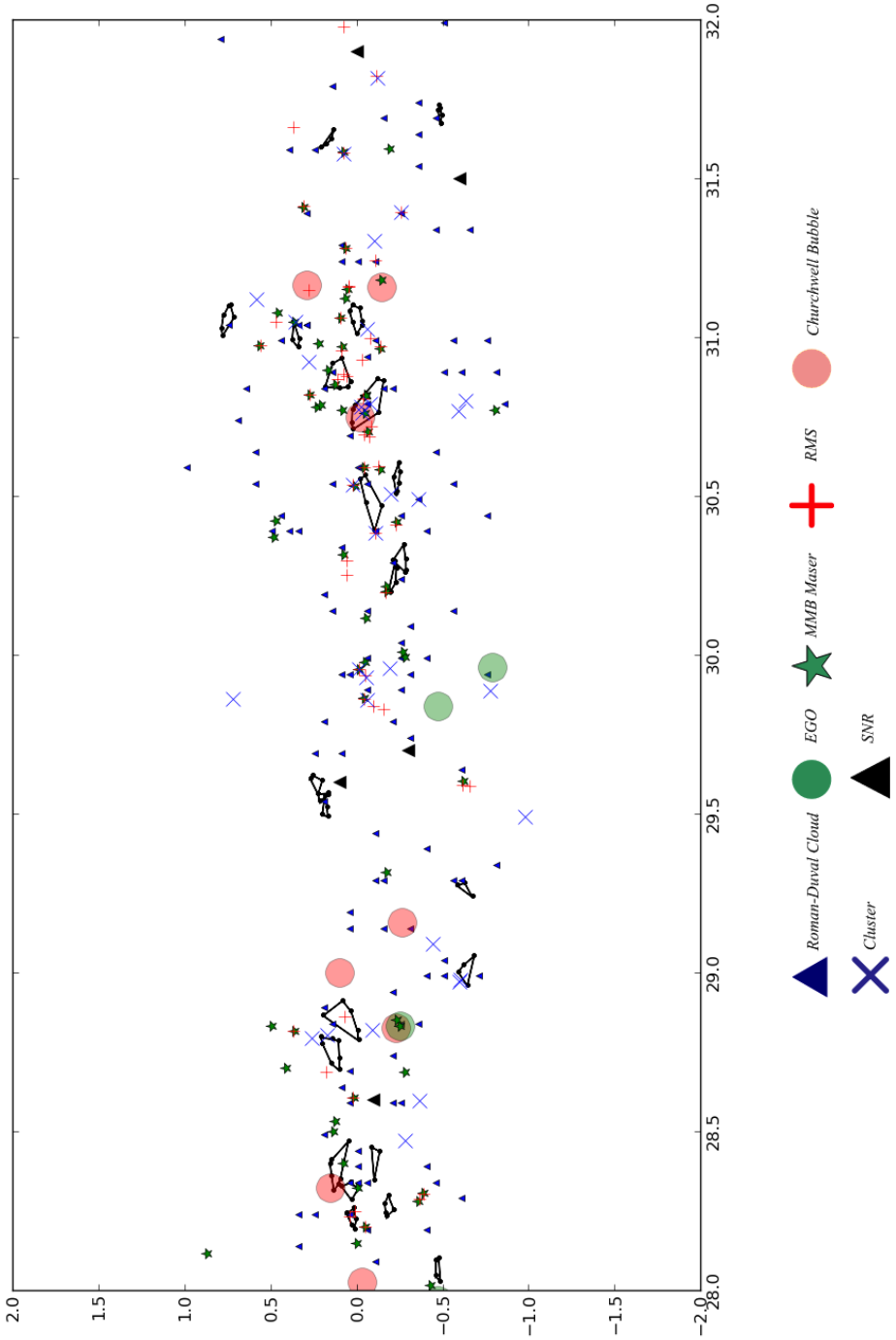


Figure C.5: Convex Hulls for P&F $28^o < l < 32^o$

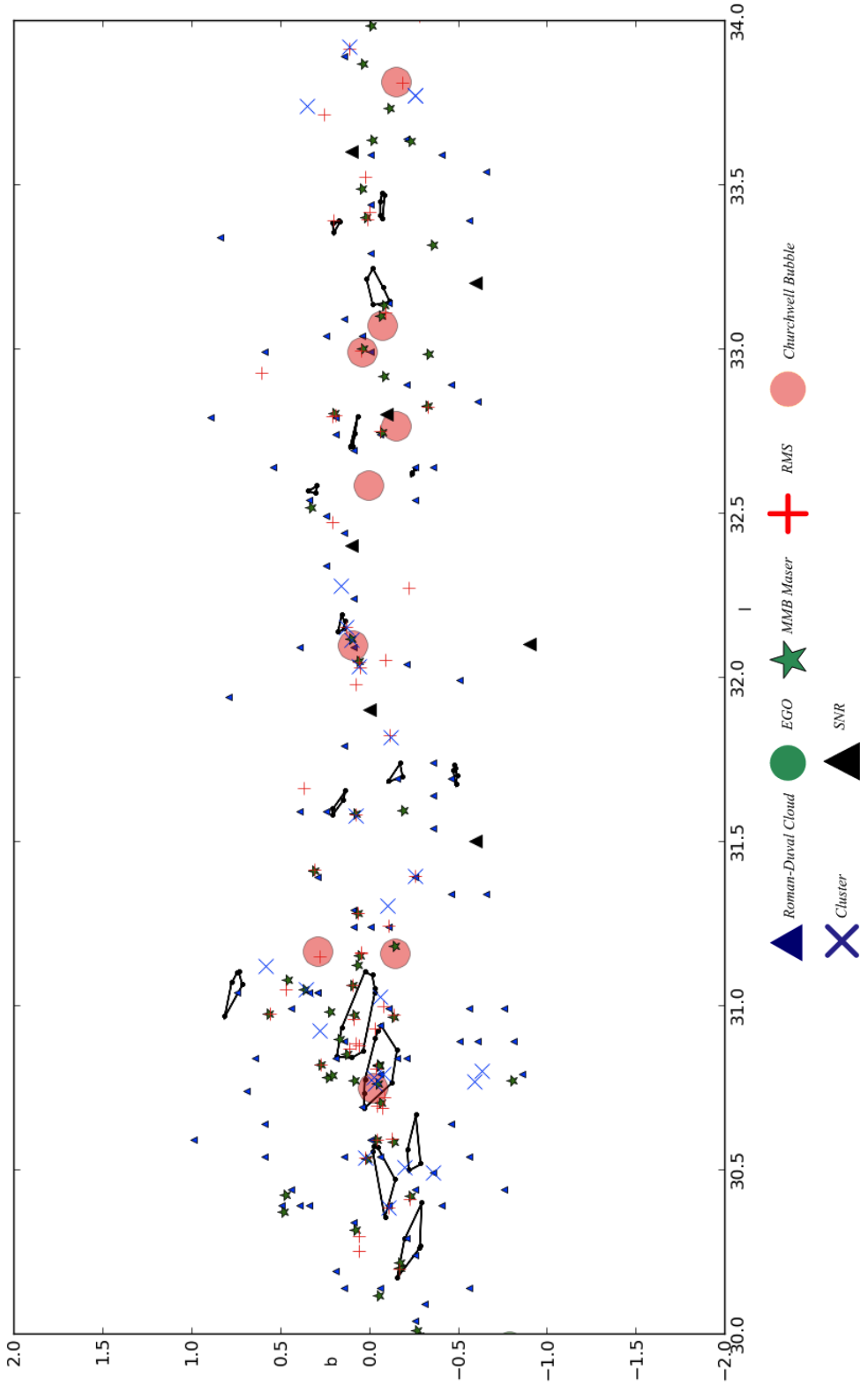


Figure C.6: Convex Hulls for P&F $30^o < l < 34^o$

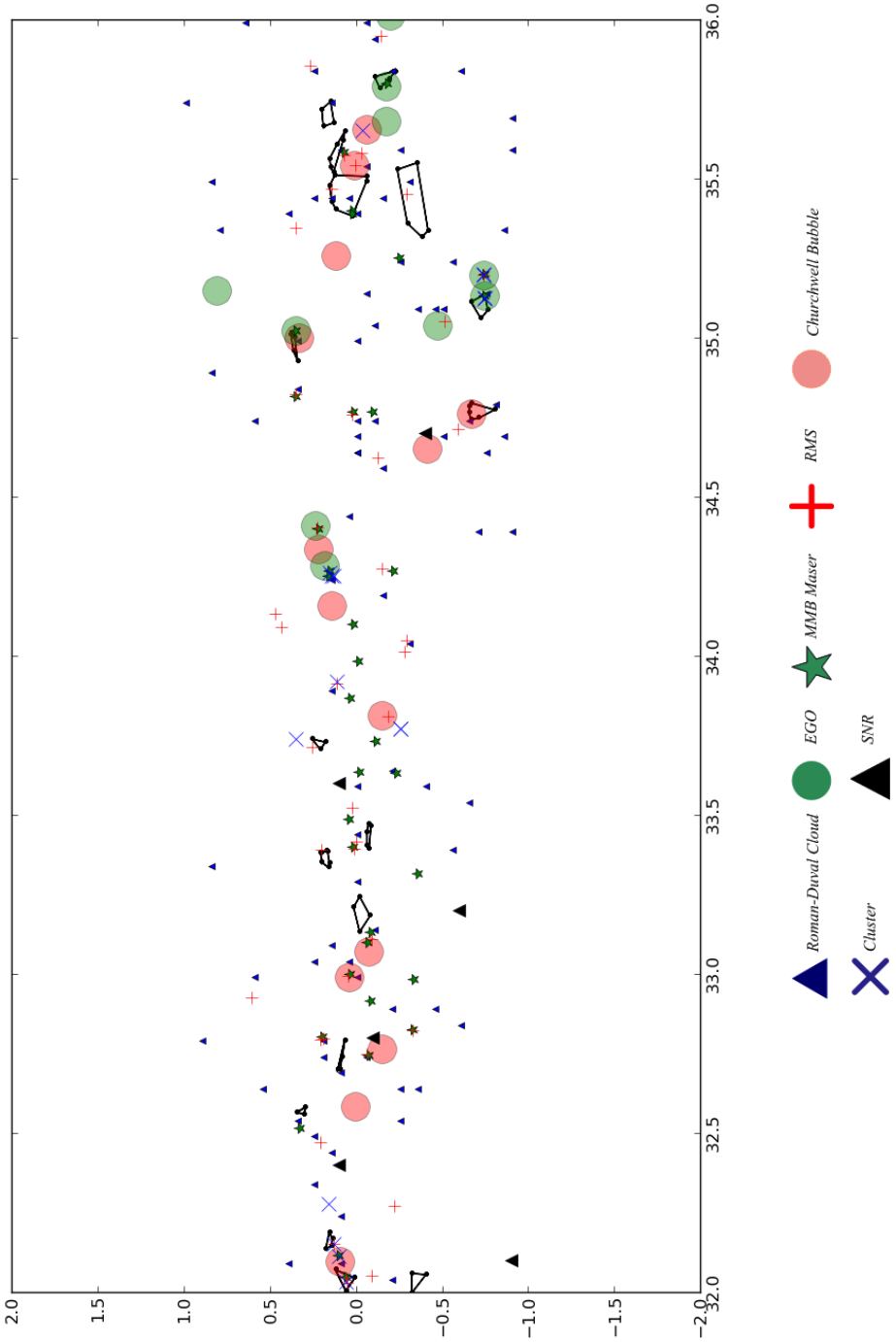


Figure C.7: Convex Hulls for P&F $32^o < l < 34^o$

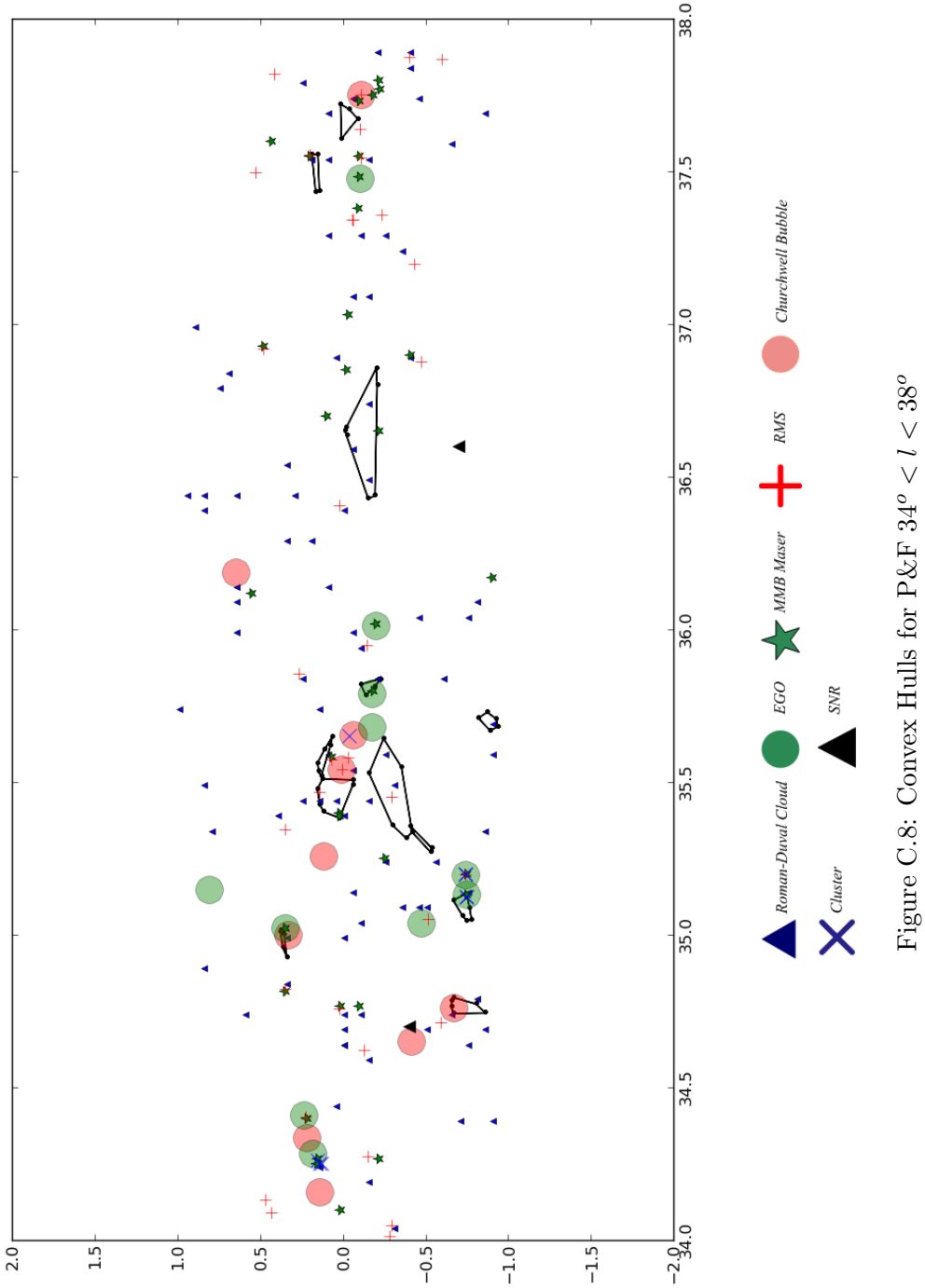


Figure C.8: Convex Hulls for P&F $34^o < l < 38^o$

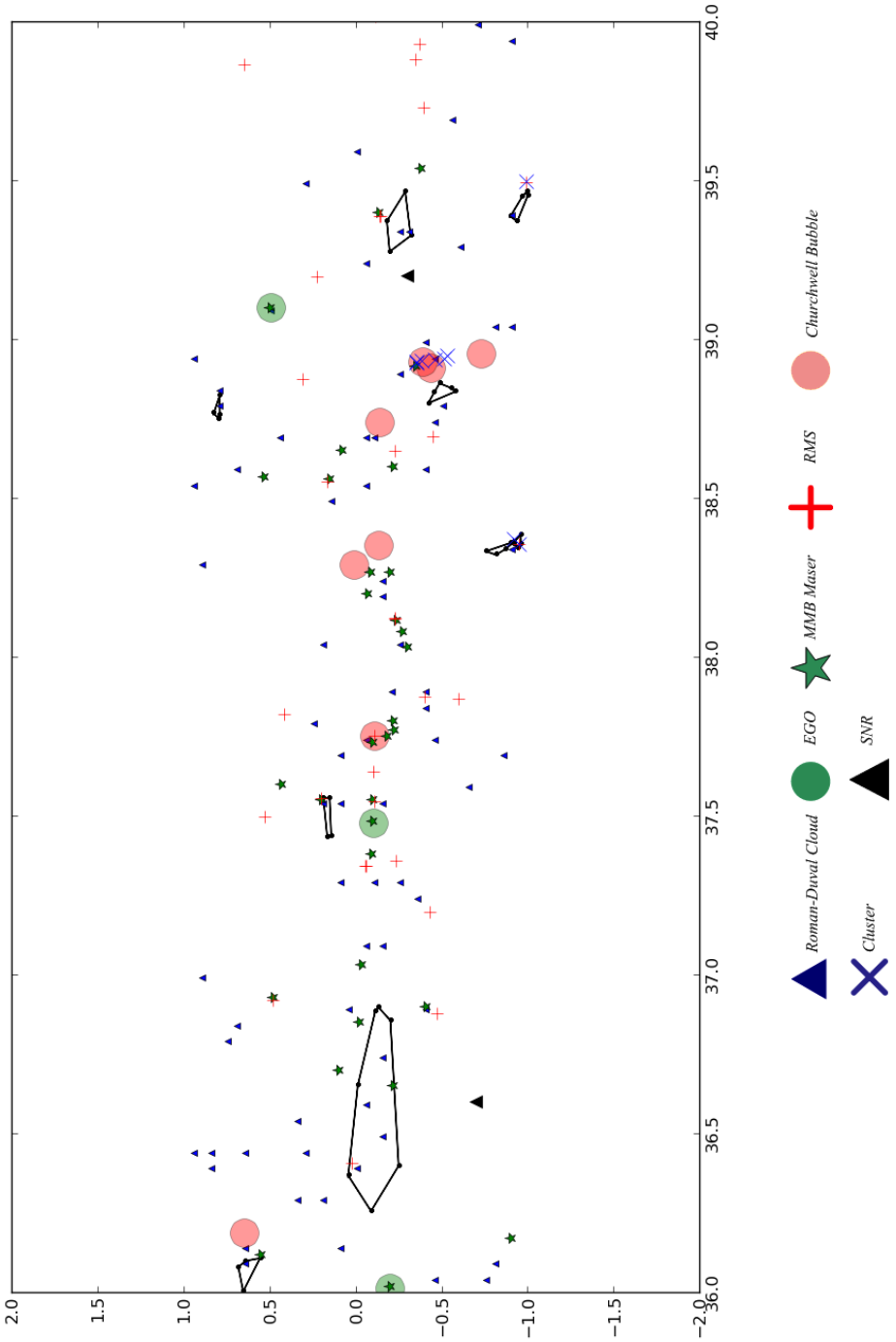


Figure C.9: Convex Hulls for P&F $36^o < l < 40^o$

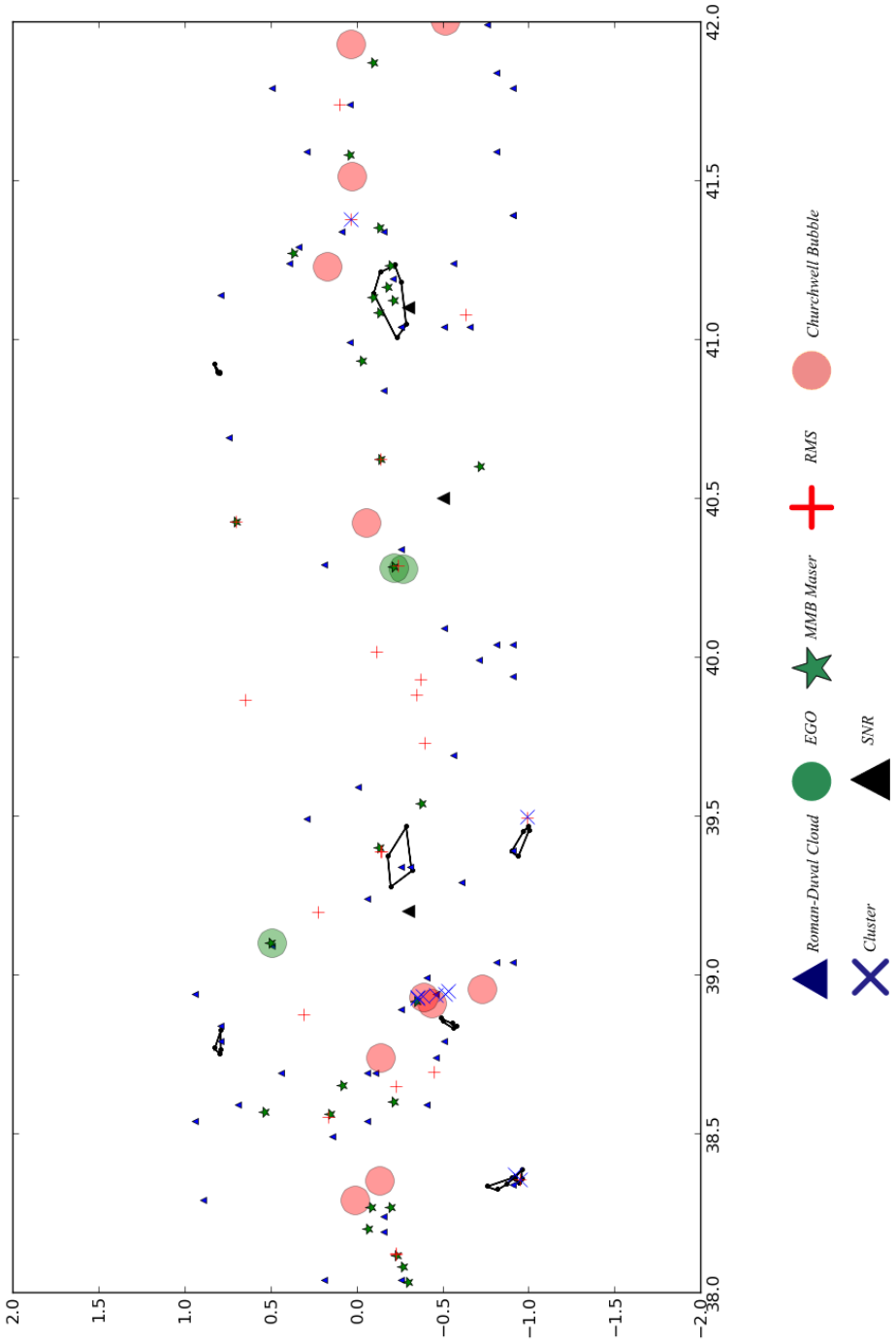


Figure C.10: Convex Hulls for P&F $38^o < l < 42^o$

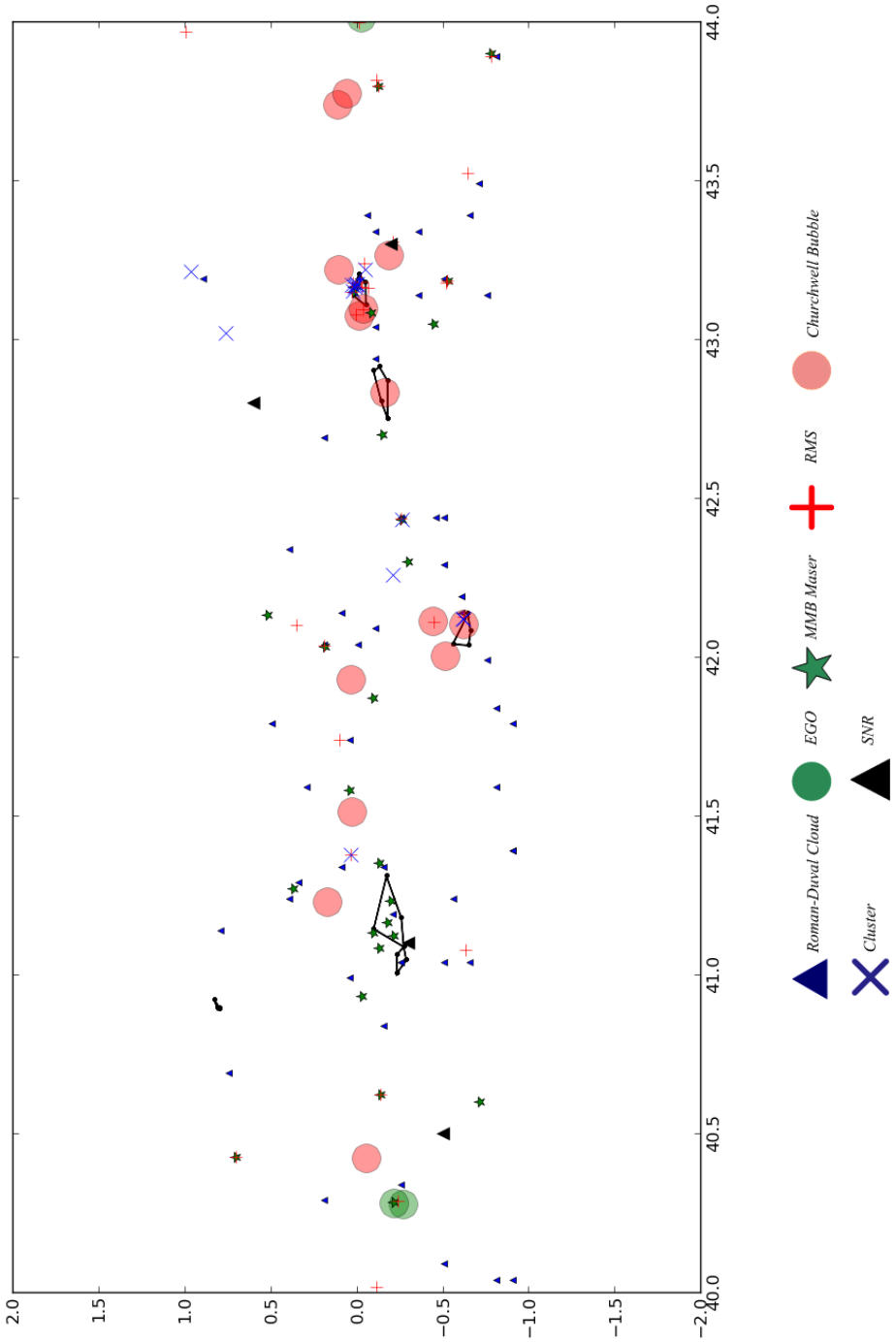


Figure C.11: Convex Hulls for P&F $40^o < l < 44^o$

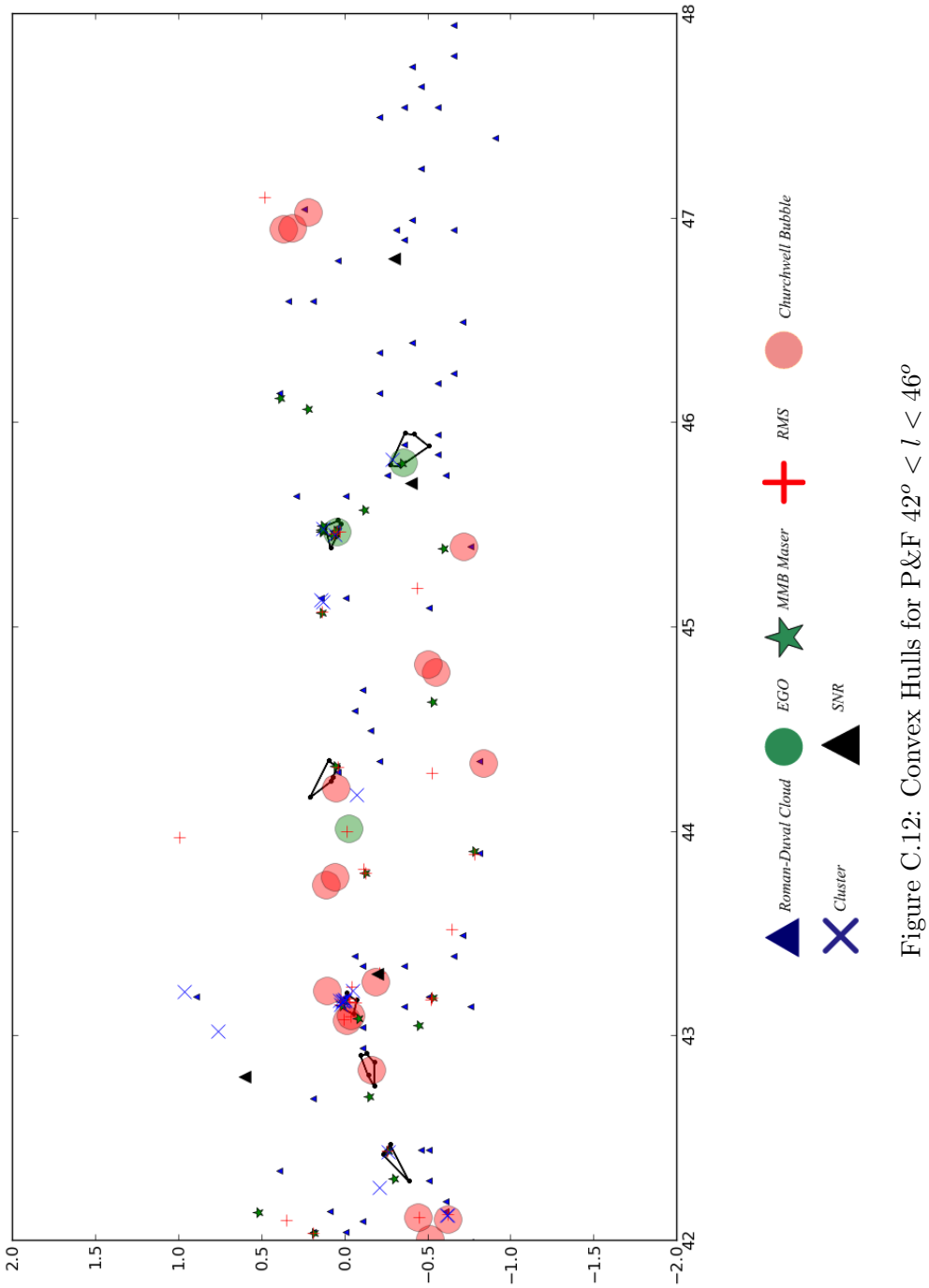


Figure C.12: Convex Hulls for P&F $42^o < l < 46^o$

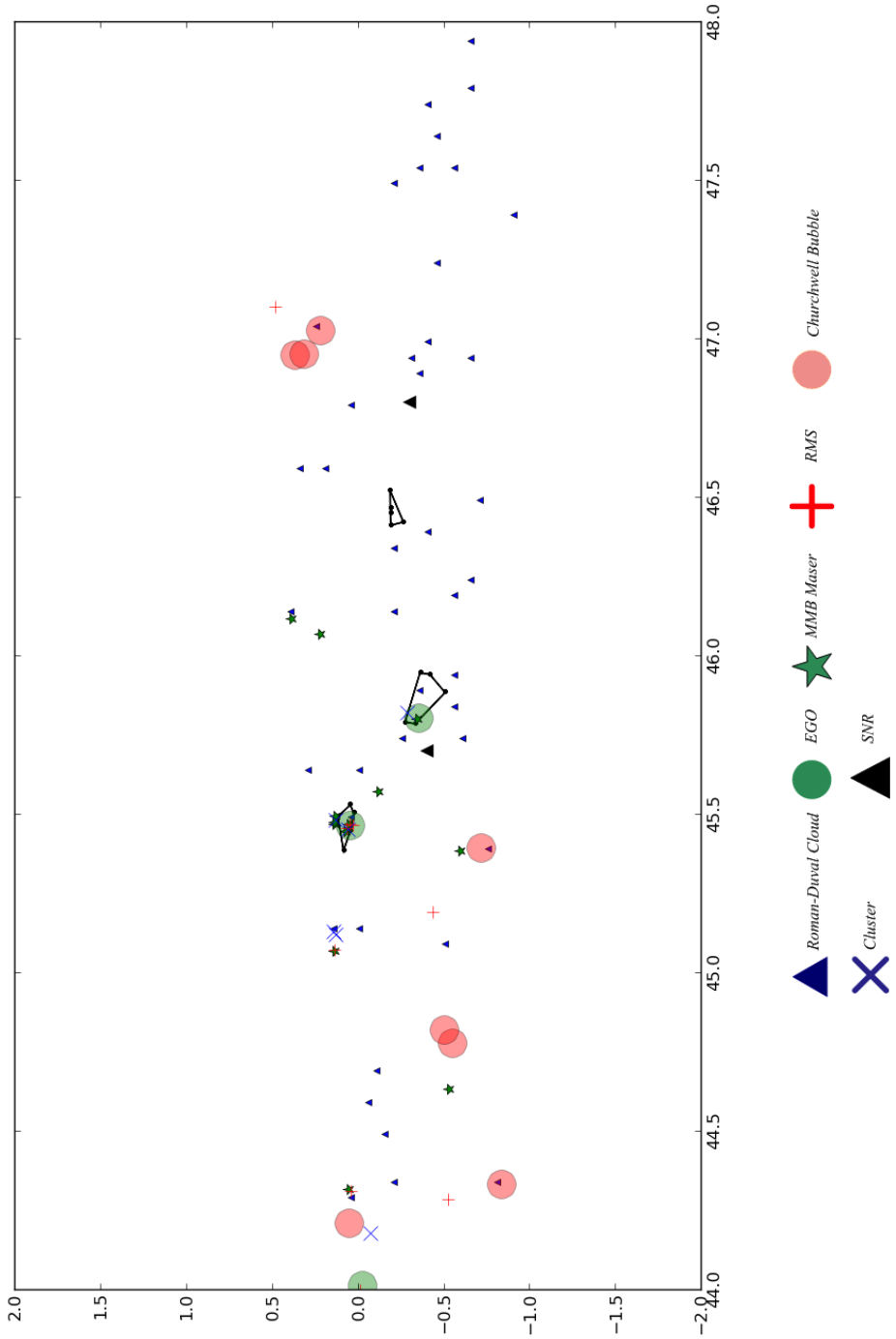


Figure C.13: Convex Hulls for P&F $44^o < l < 38^o$

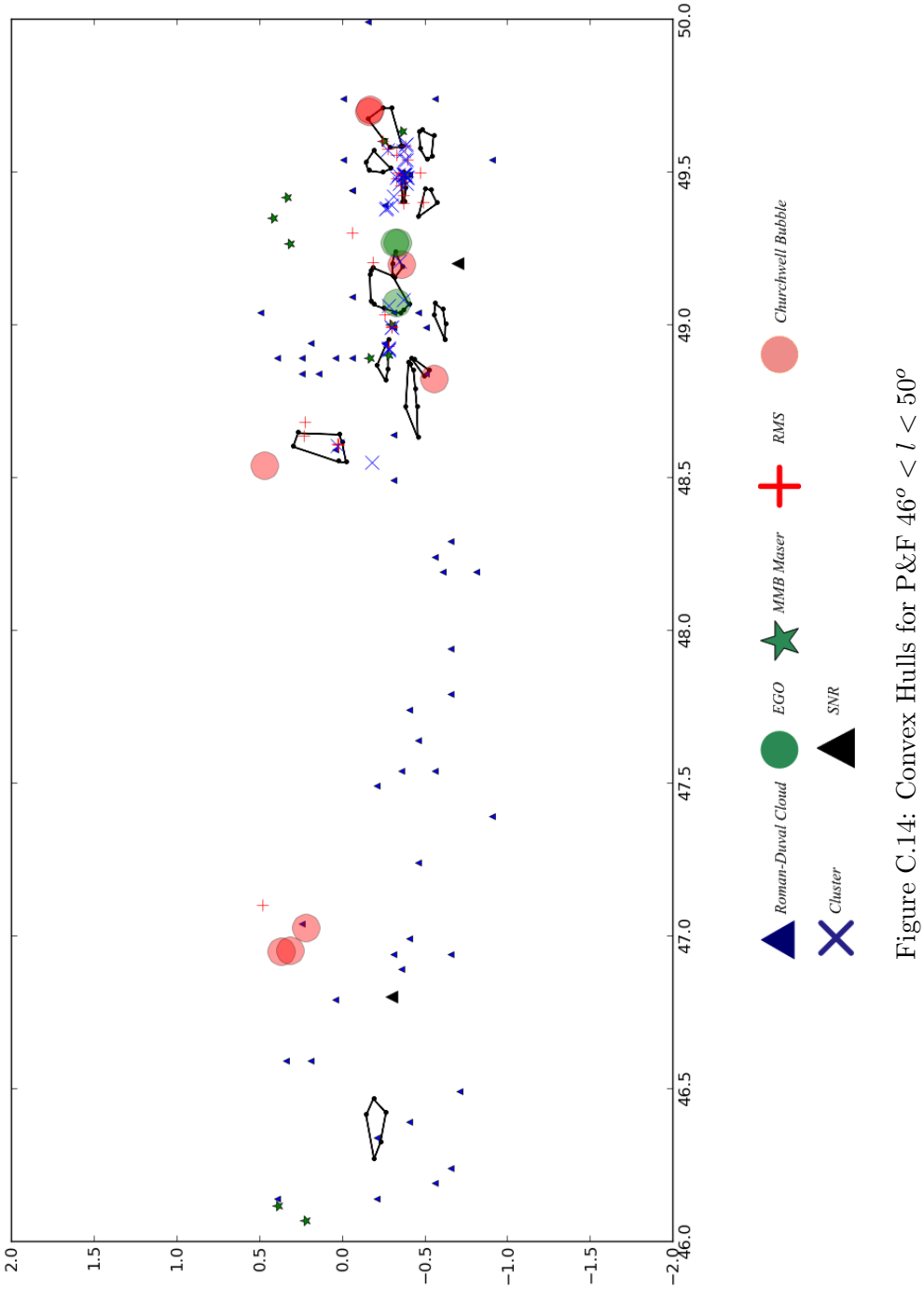


Figure C.14: Convex Hulls for P&F $46^o < l < 50^o$

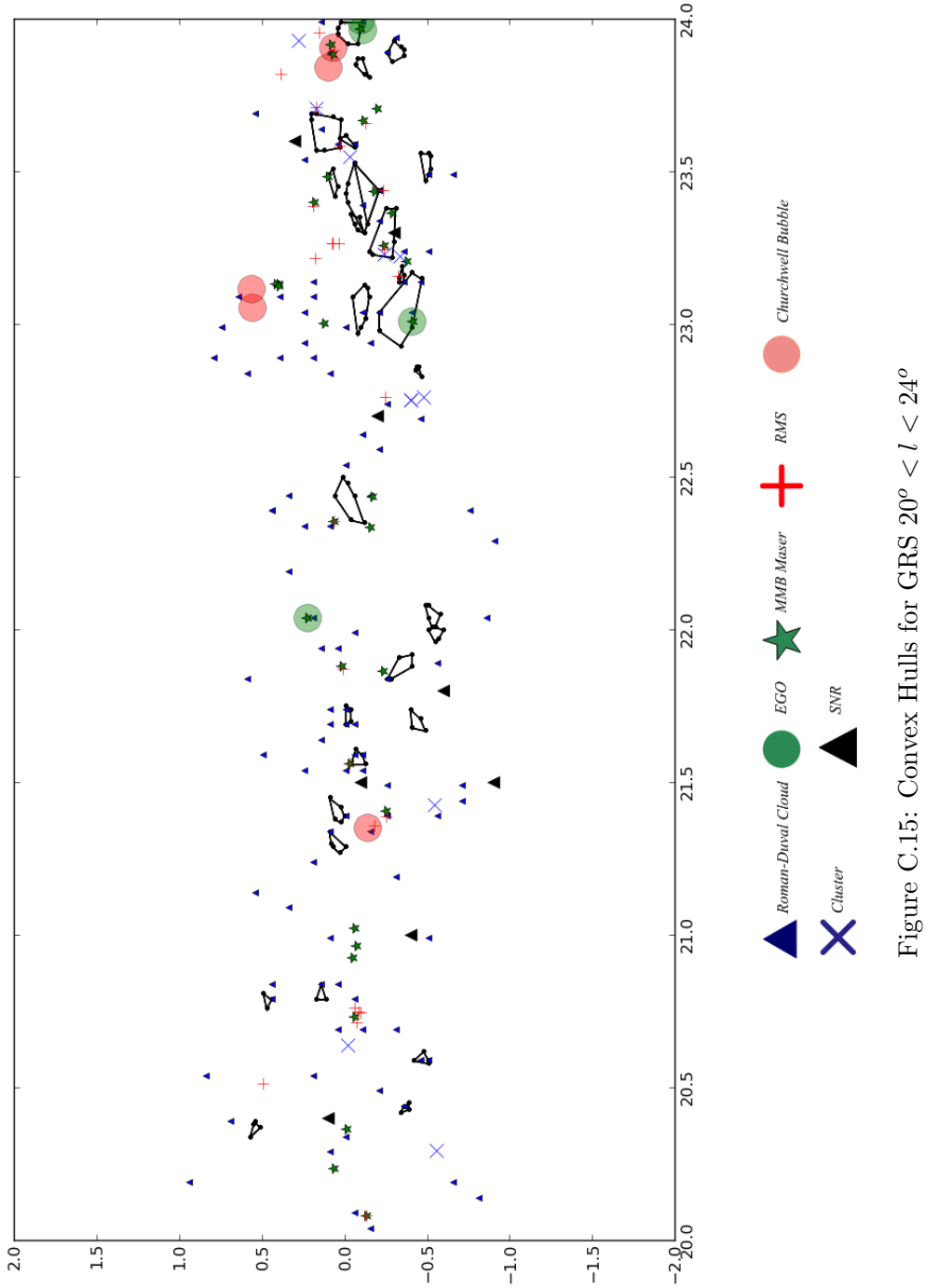


Figure C.15: Convex Hulls for GRS $20^o < l < 24^o$

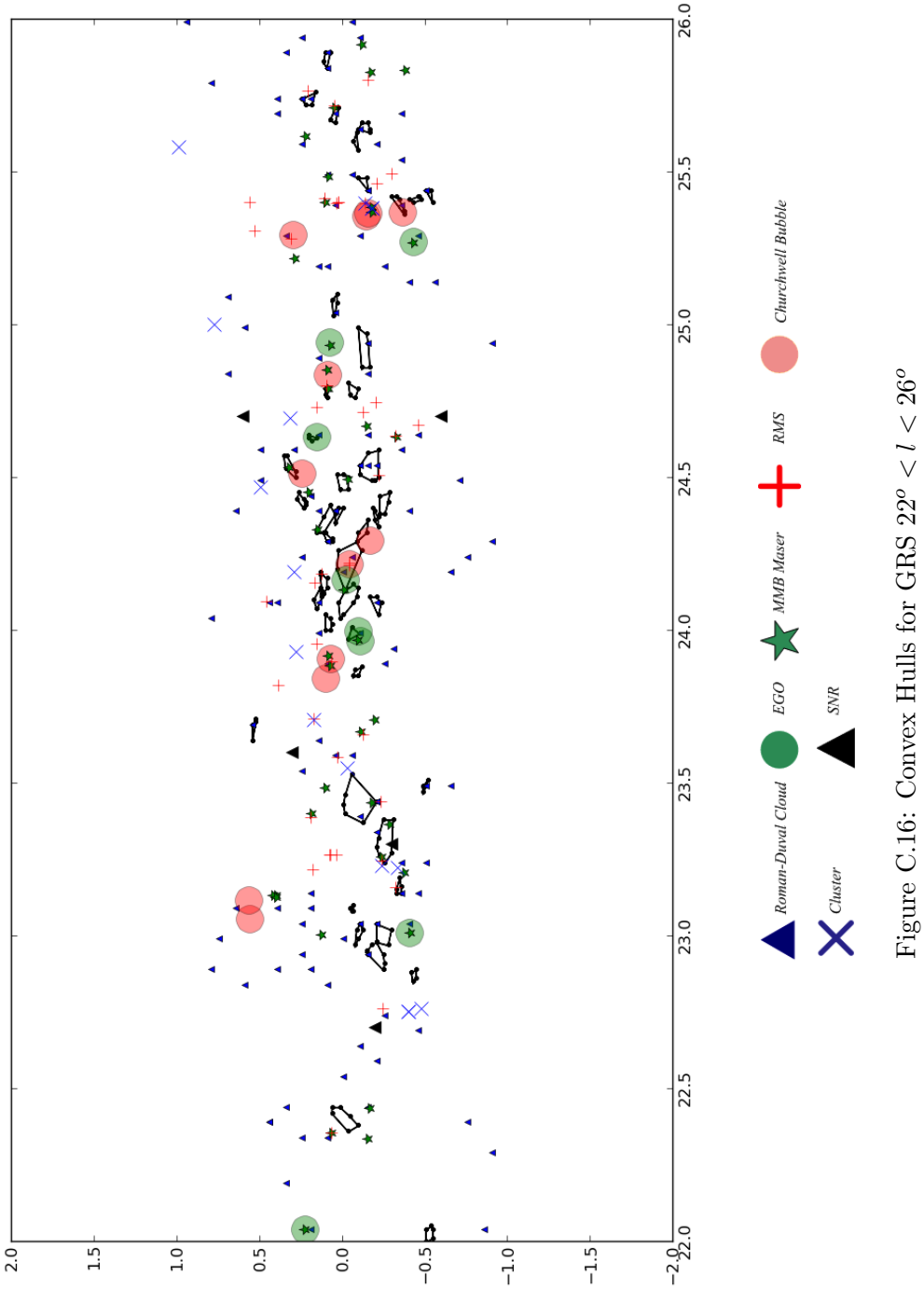


Figure C.16: Convex Hulls for GRS $22^o < l < 26^o$

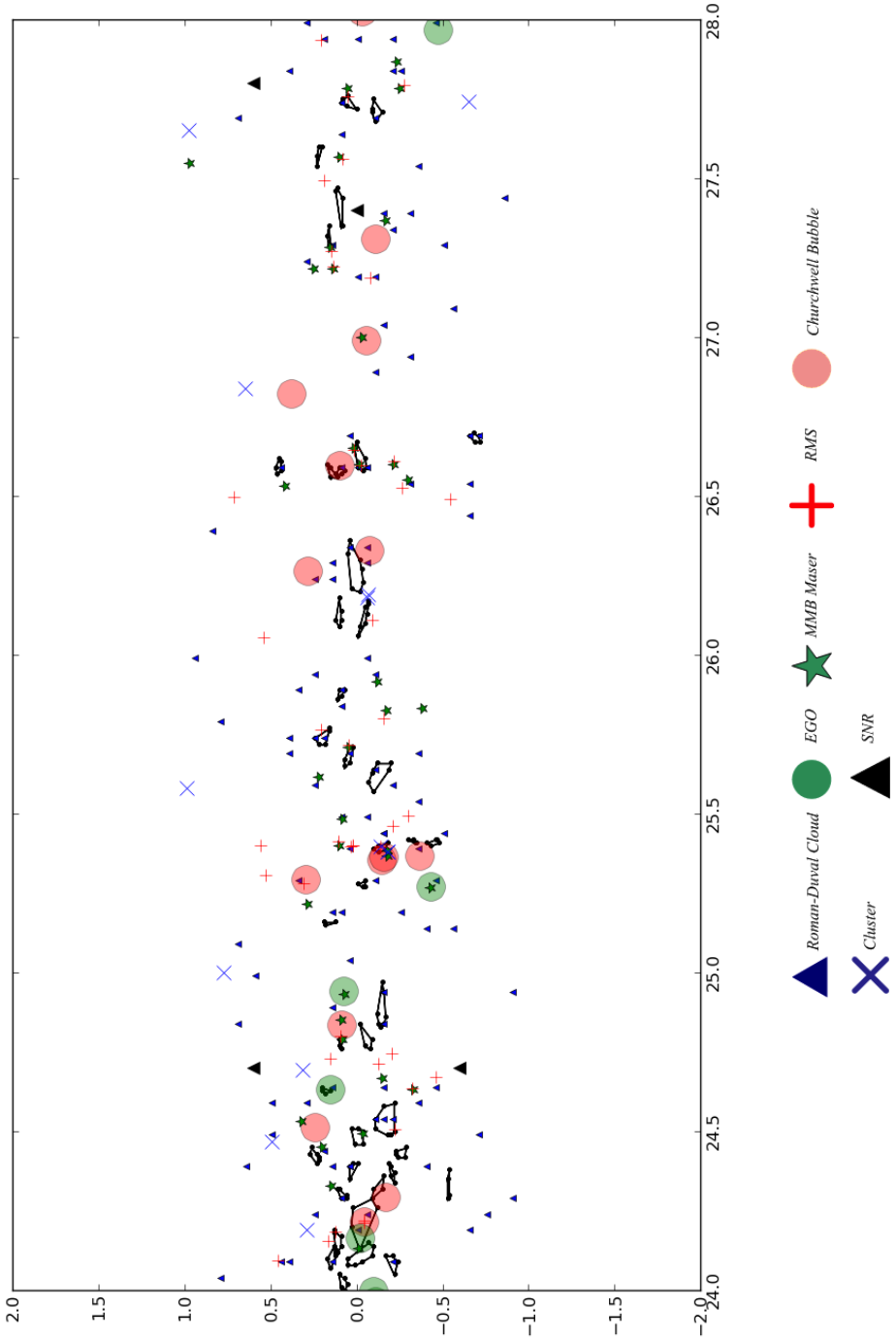


Figure C.17: Convex Hulls for GRS $24^o < l < 28^o$

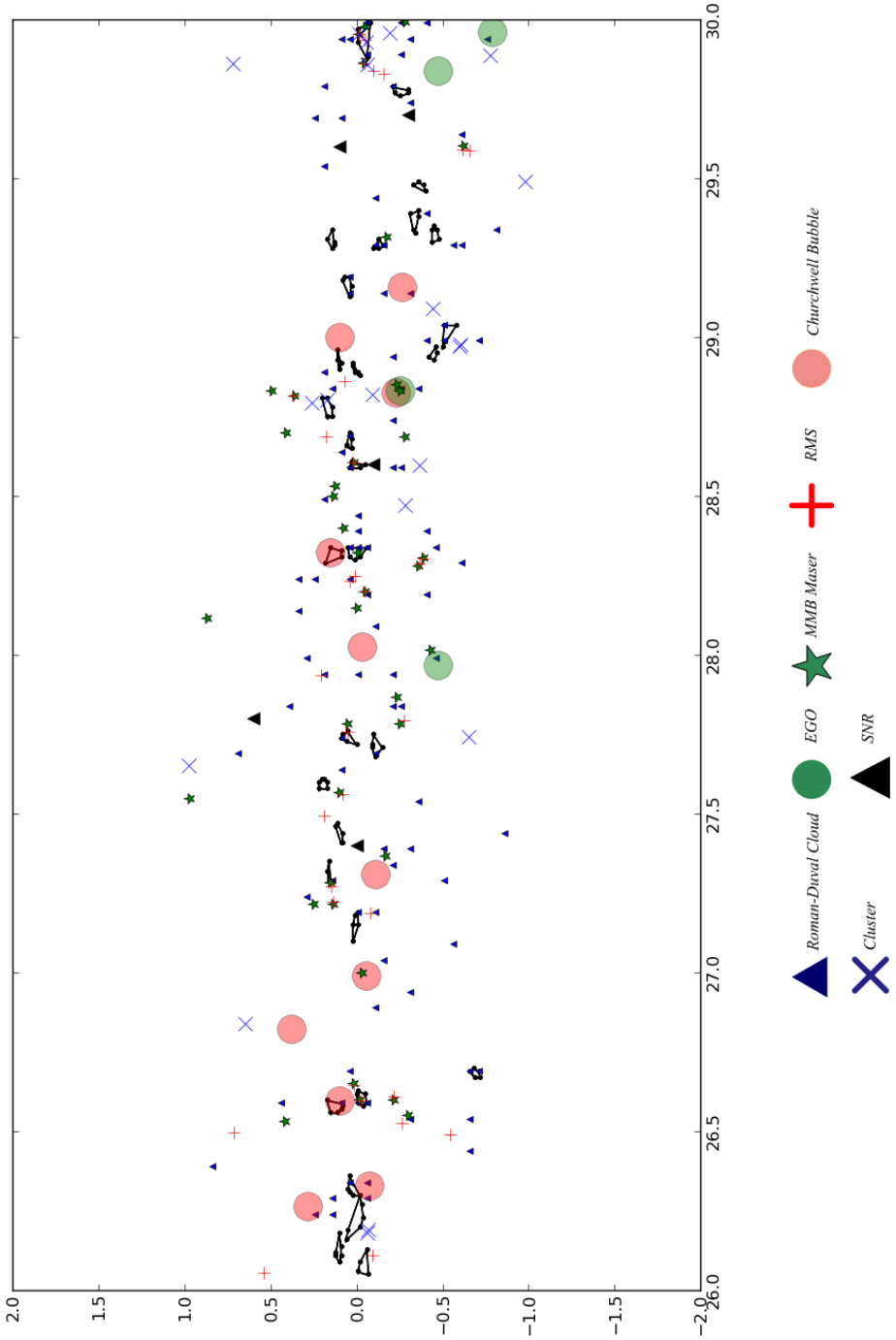


Figure C.18: Convex Hulls for GRS $26^o < l < 30^o$

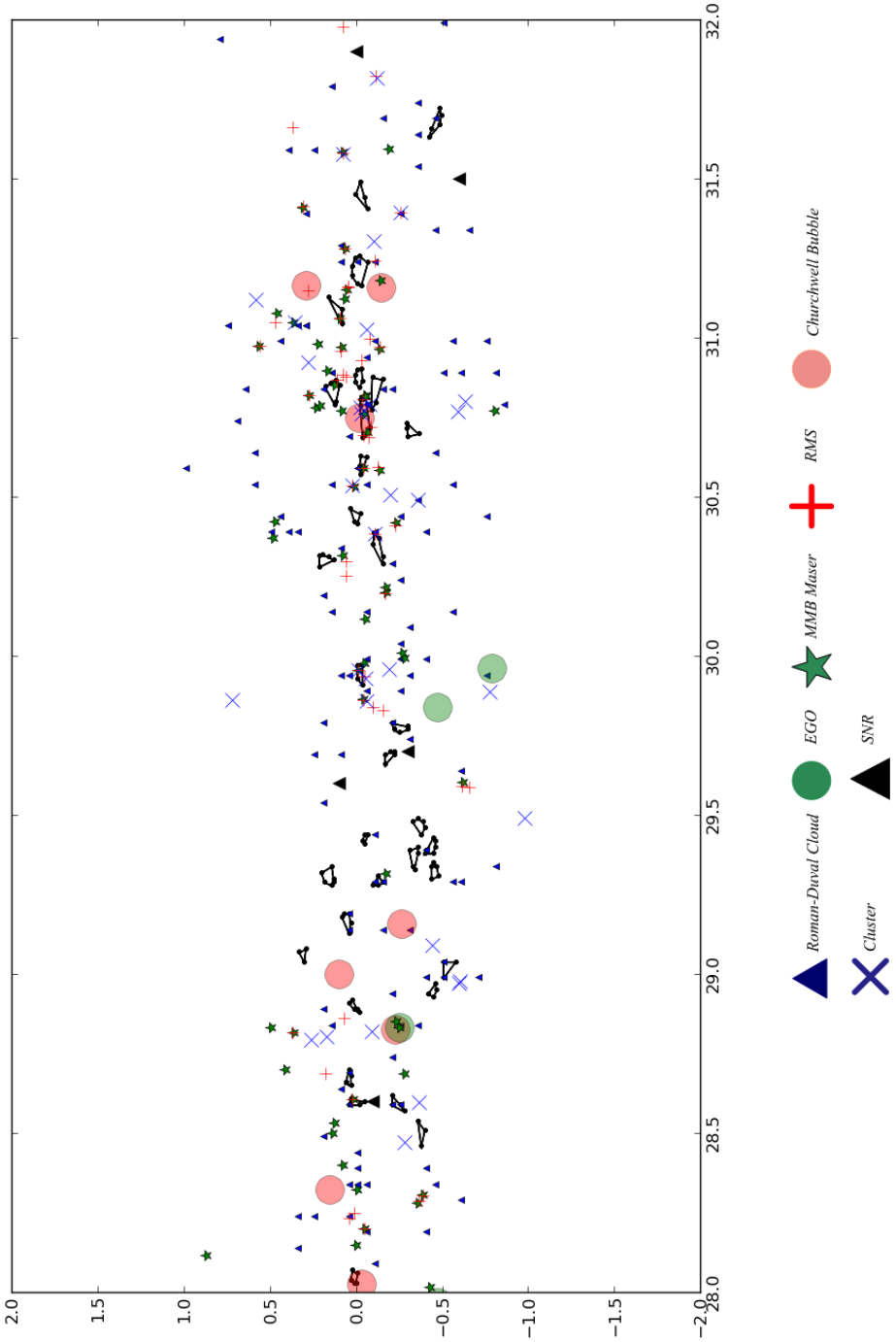


Figure C.19: Convex Hulls for GRS $28^o < l < 32^o$

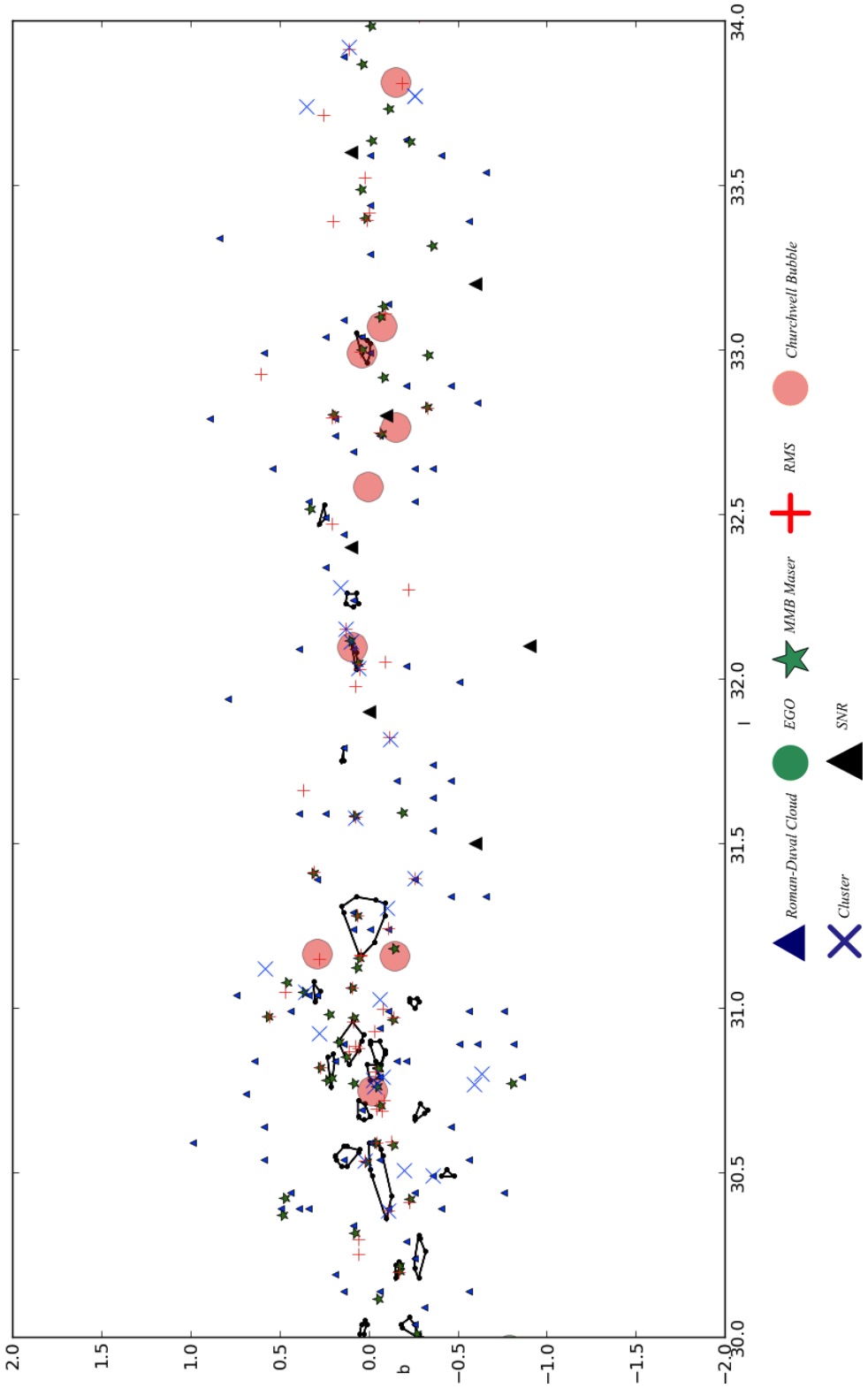


Figure C.20: Convex Hulls for GRS $30^o < l < 34^o$

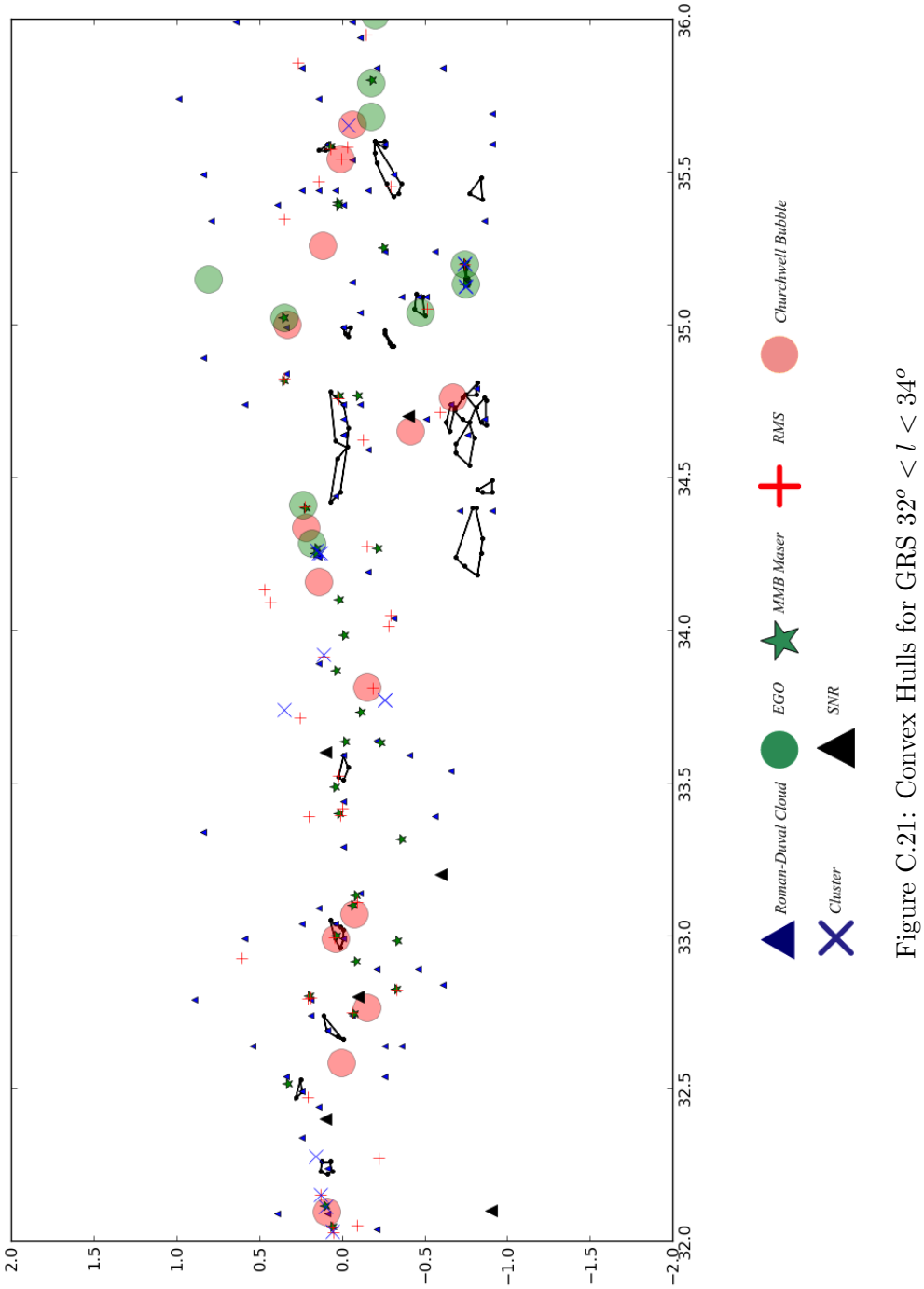


Figure C.21: Convex Hulls for GRS $32^o < l < 34^o$

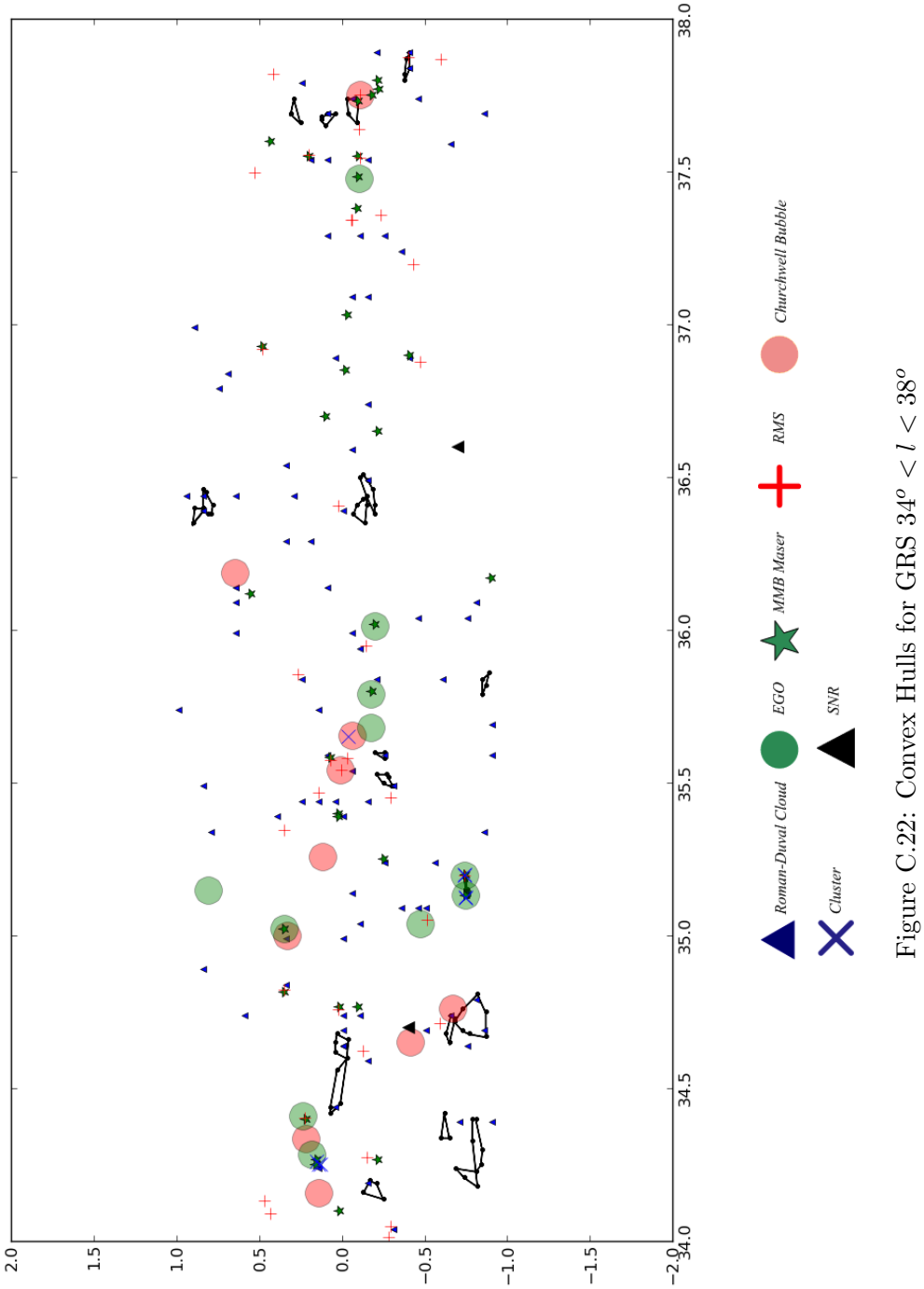


Figure C.22: Convex Hulls for GRS $34^o < l < 38^o$

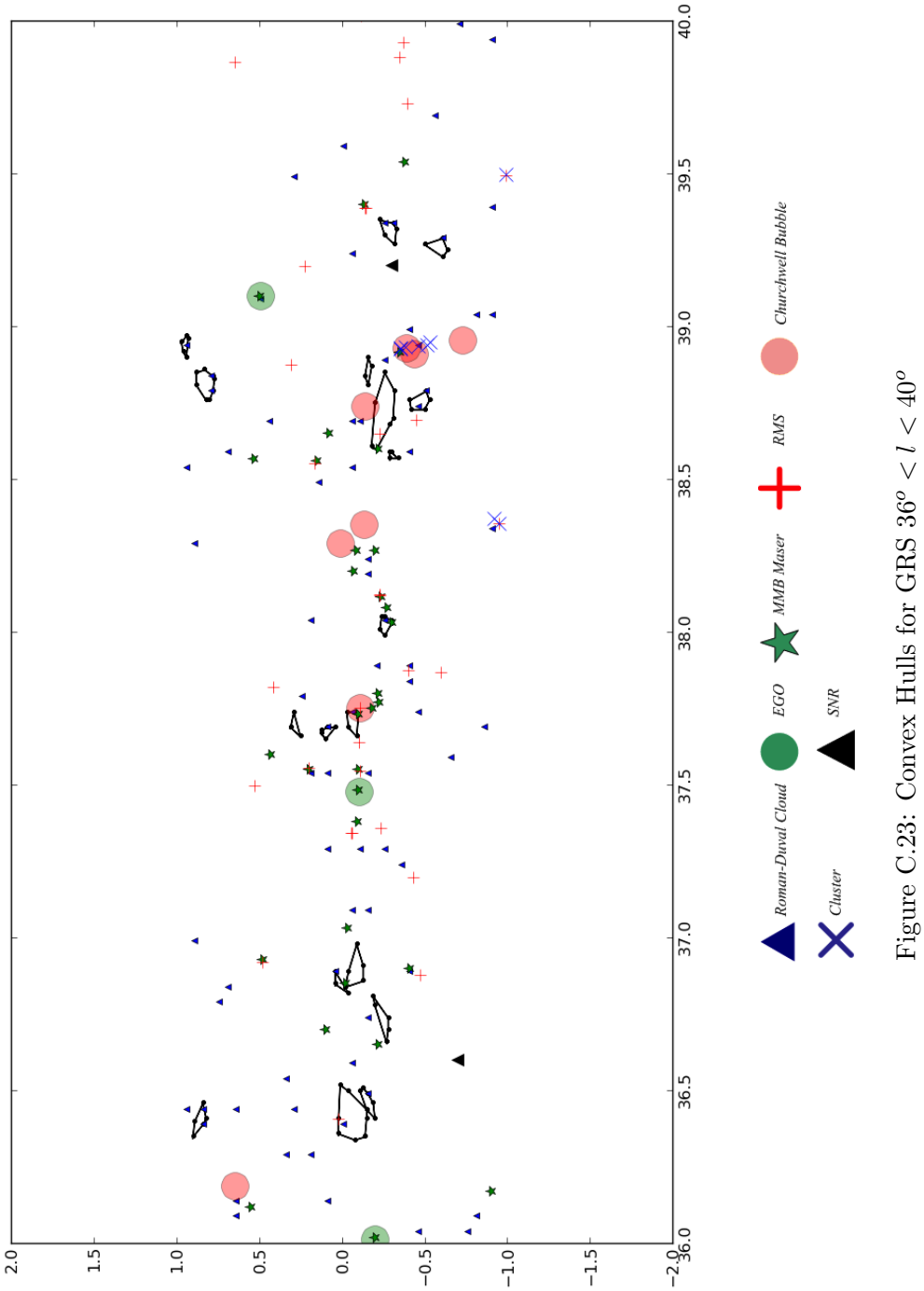


Figure C.23: Convex Hulls for GRS $36^o < l < 40^o$

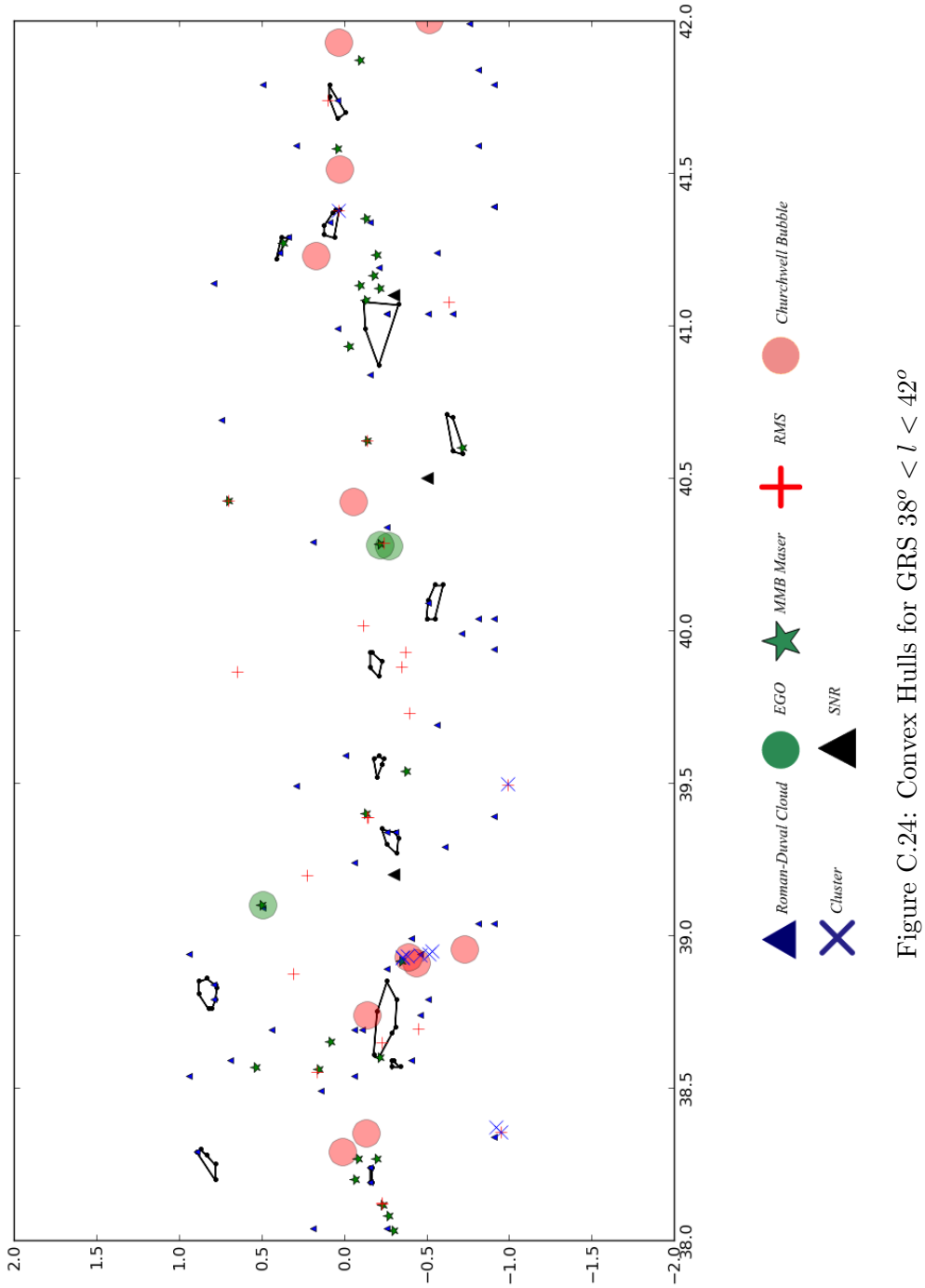


Figure C.24: Convex Hulls for GRS $38^o < l < 42^o$

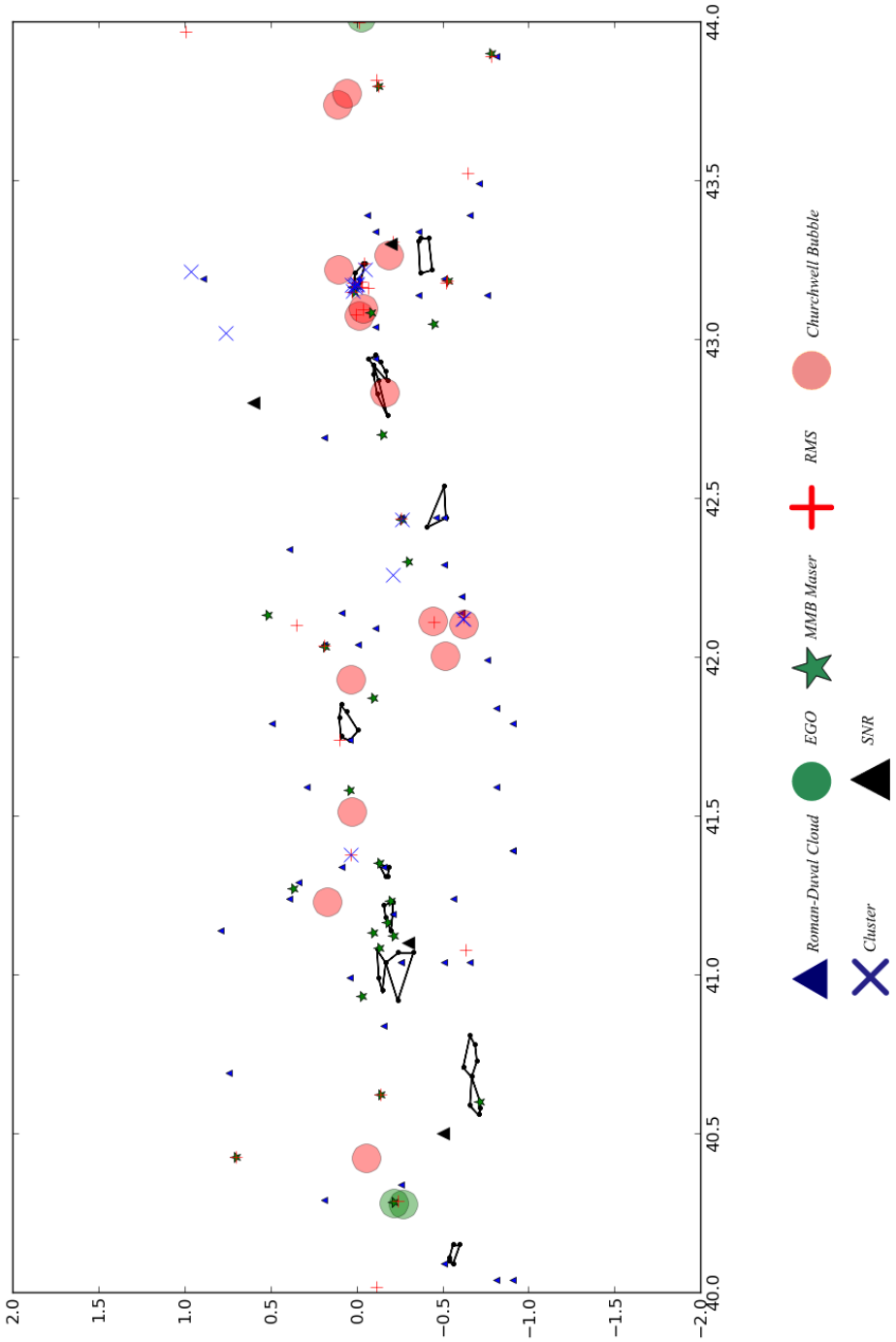


Figure C.25: Convex Hulls for GRS $40^o < l < 44^o$

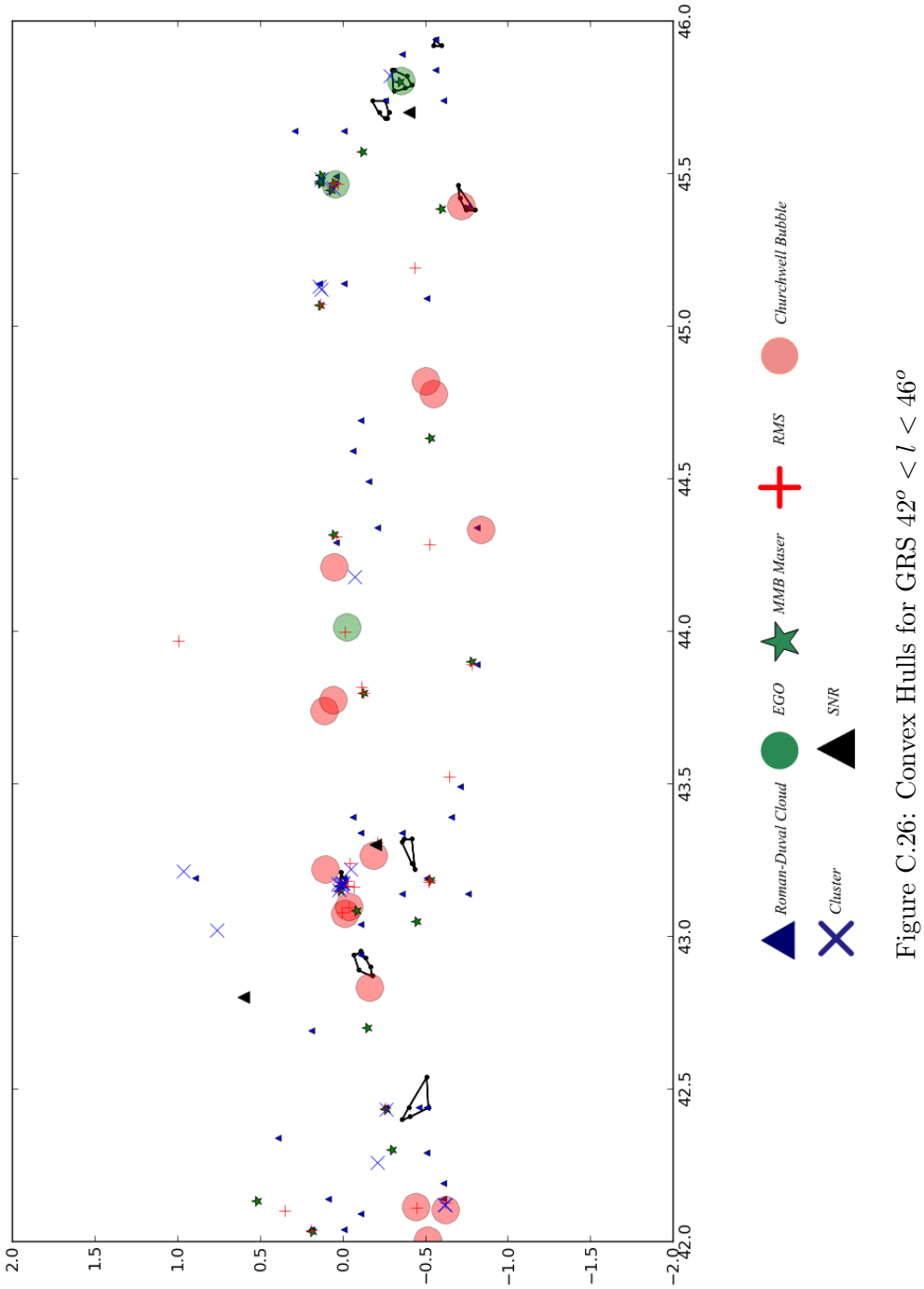


Figure C.26: Convex Hulls for GRS $42^o < l < 46^o$

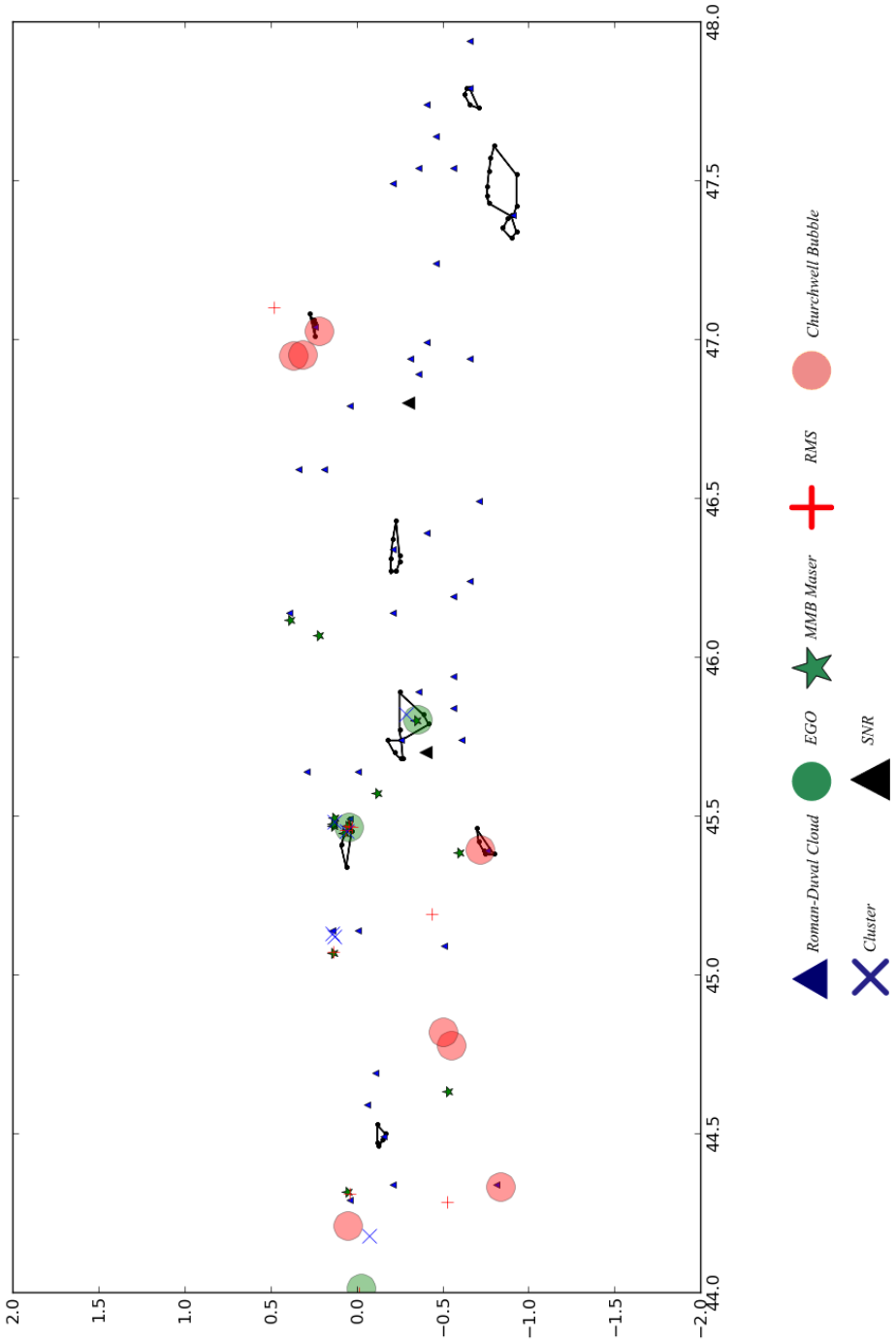


Figure C.27: Convex Hulls for GRS $44^o < l < 38^o$

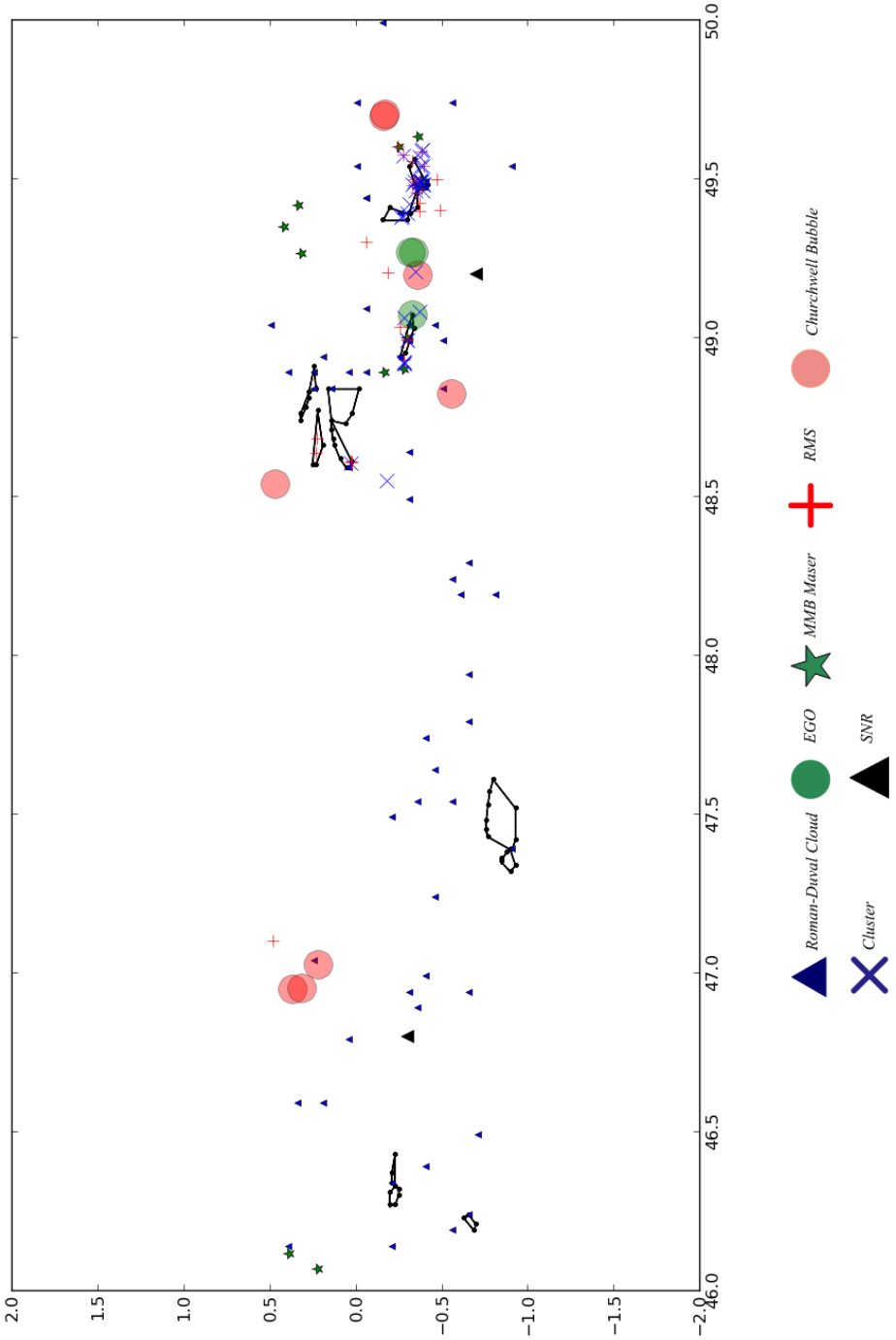


Figure C.28: Convex Hulls for GRS $46^o < l < 50^o$

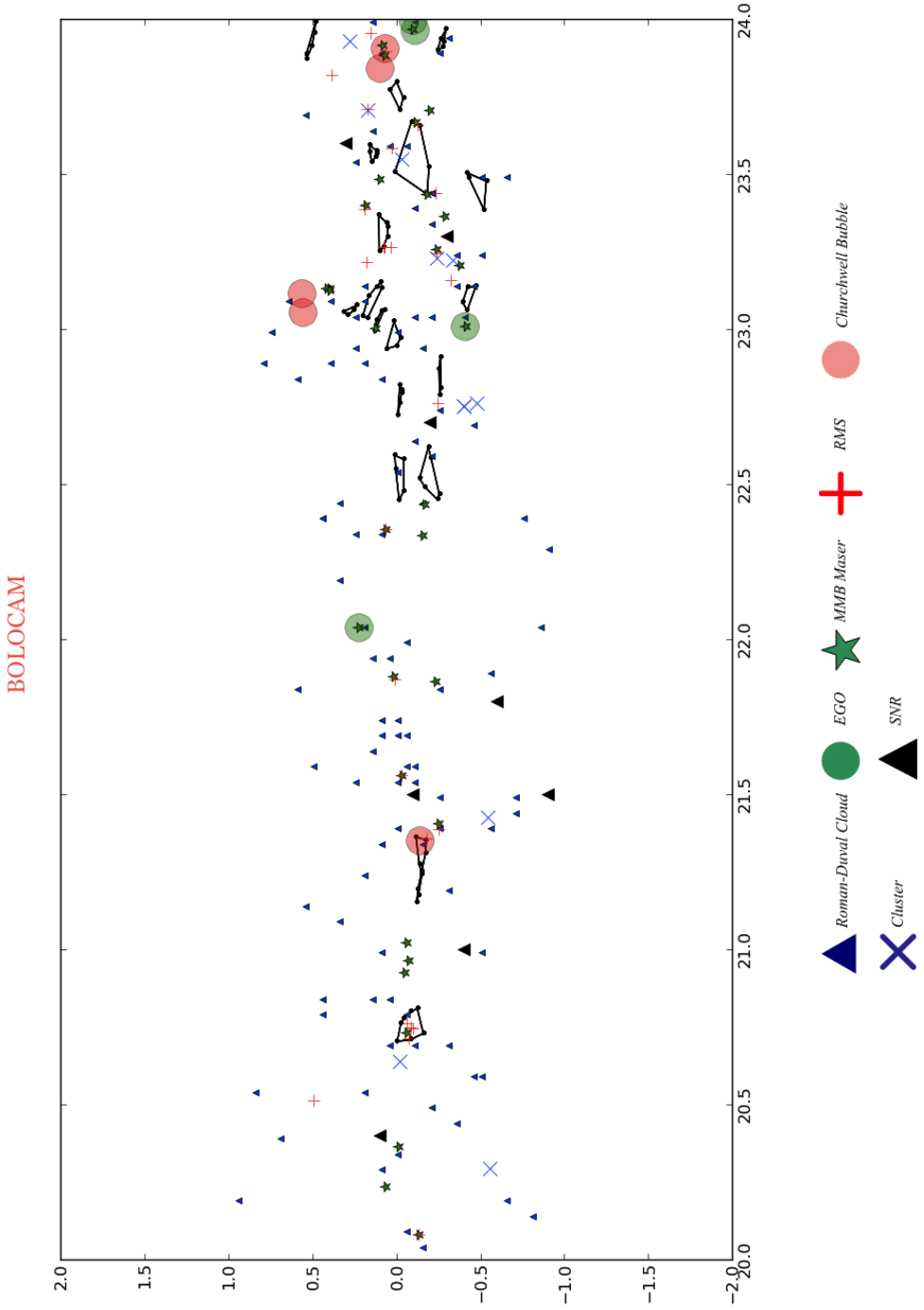


Figure C.29: Convex Hulls for BGPS $20^o < l < 24^o$

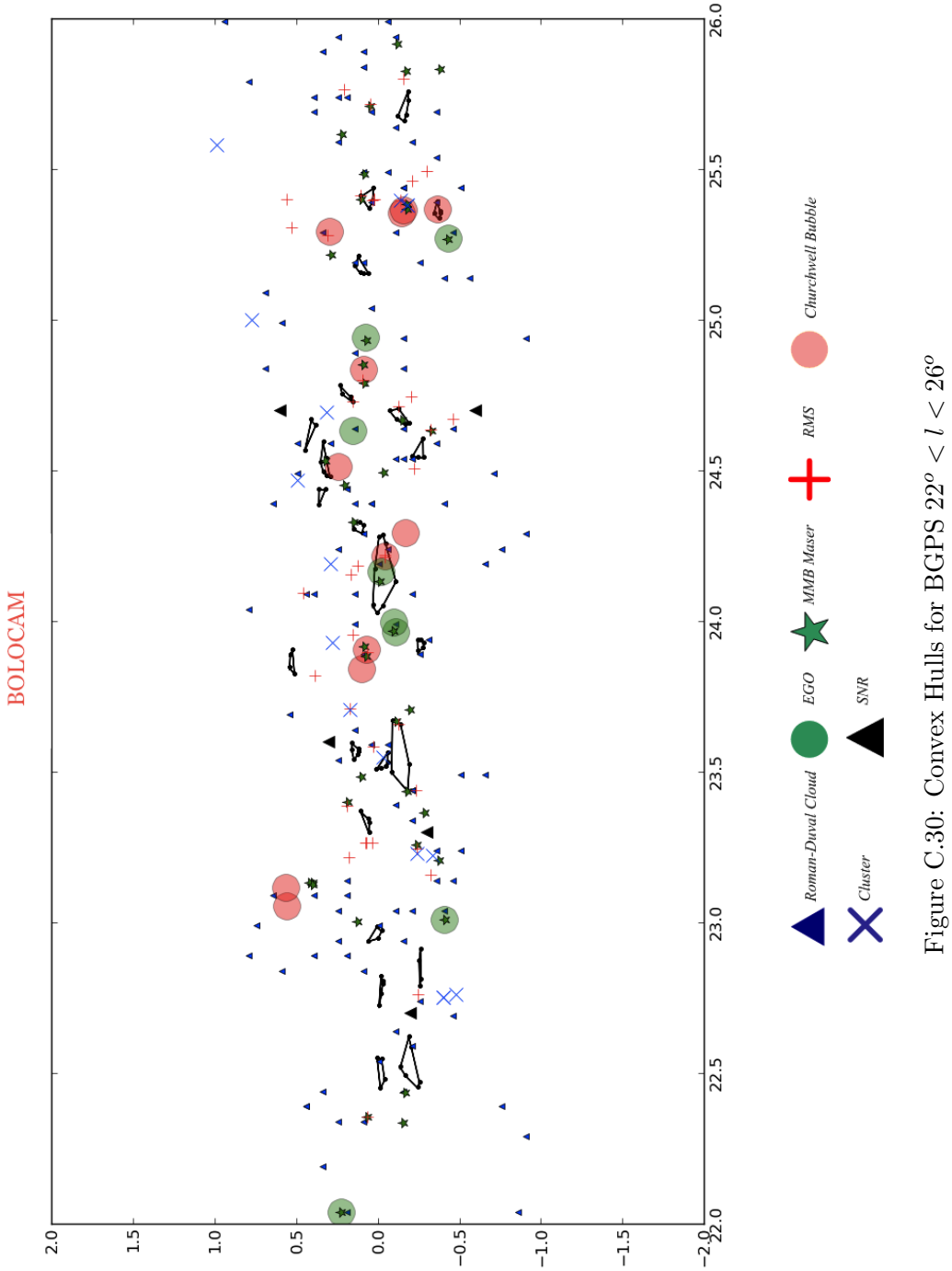


Figure C.30: Convex Hulls for BGPS $22^o < l < 26^o$

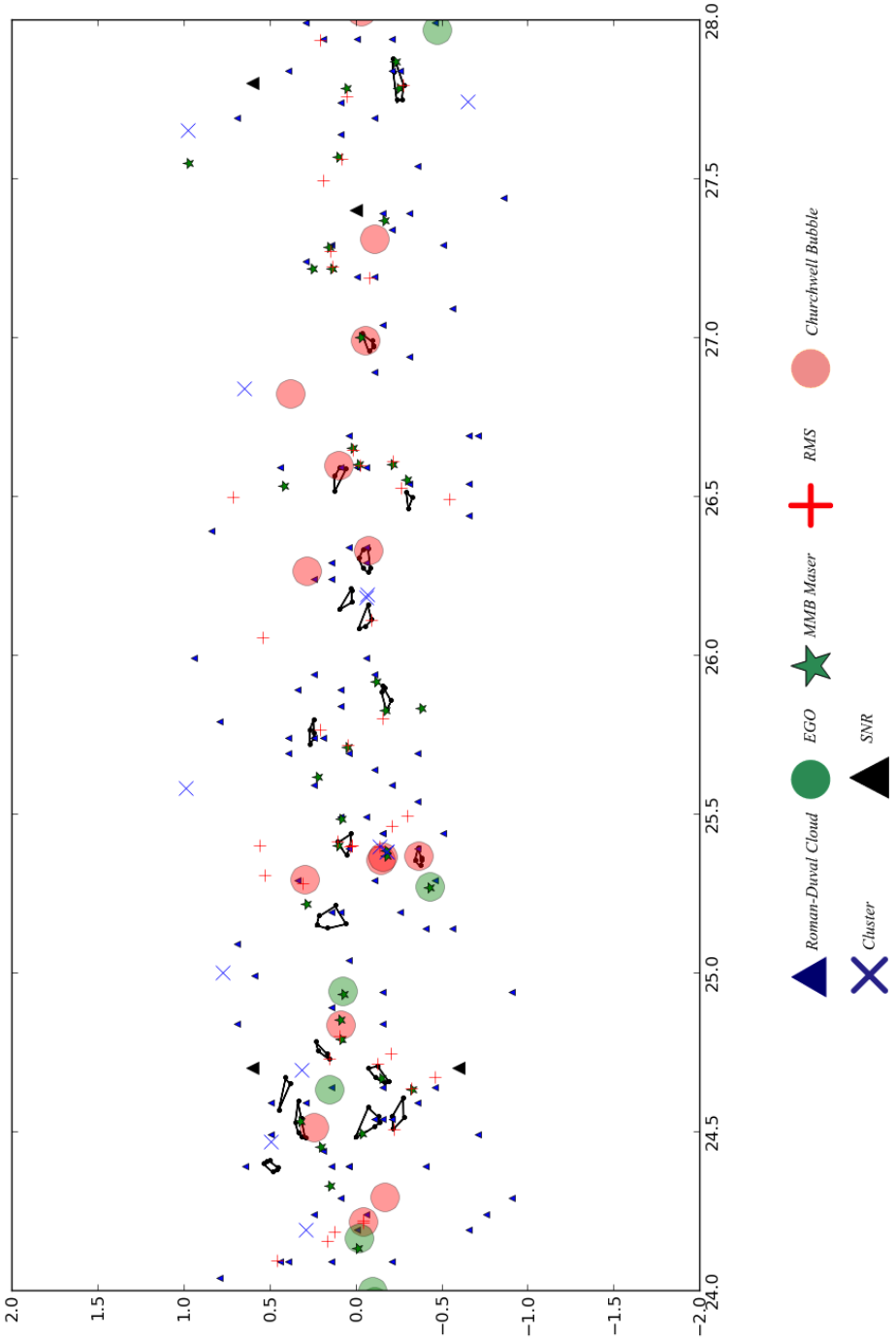


Figure C.31: Convex Hulls for BGPS $24^o < l < 28^o$

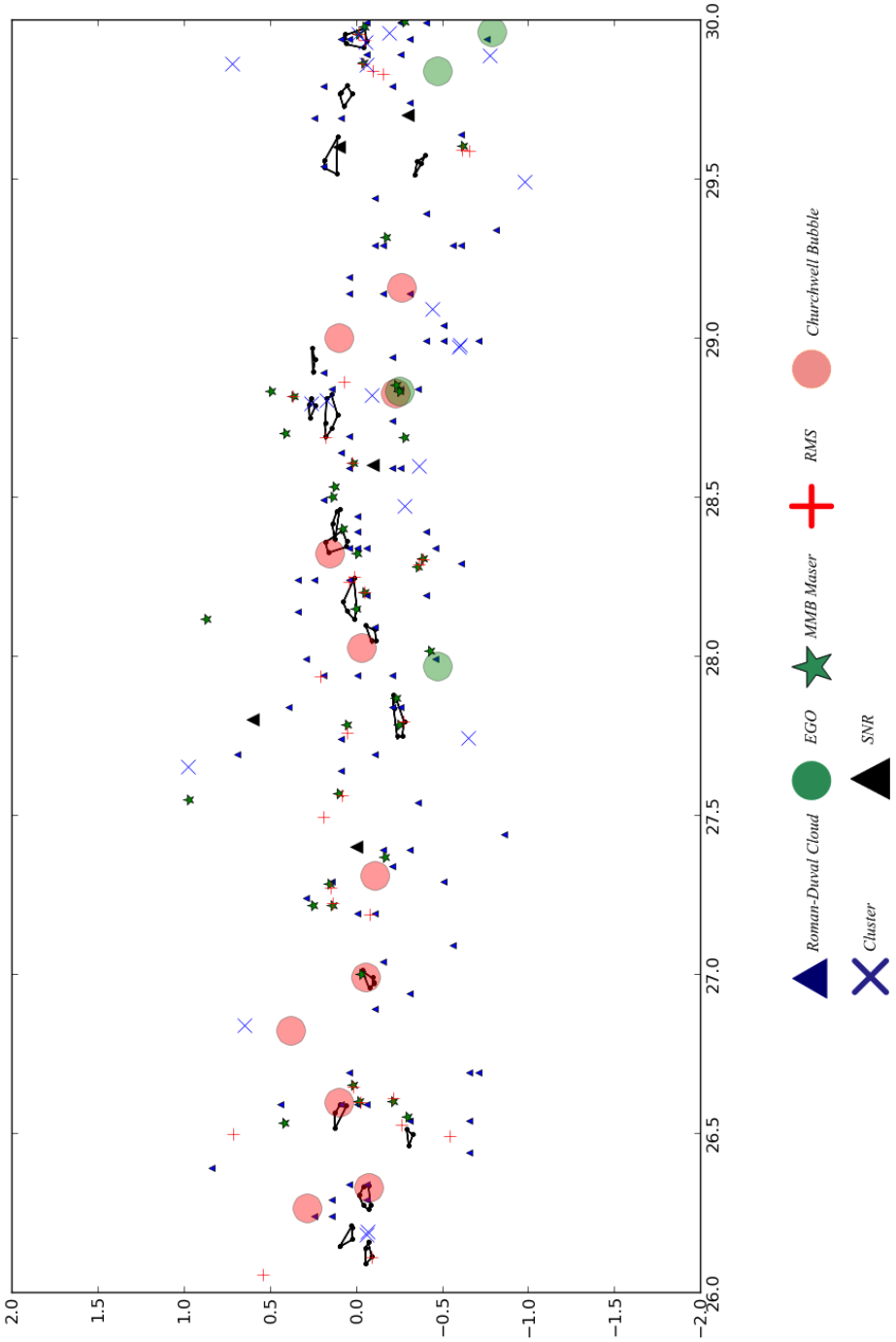


Figure C.32: Convex Hulls for BGPS $26^o < l < 30^o$

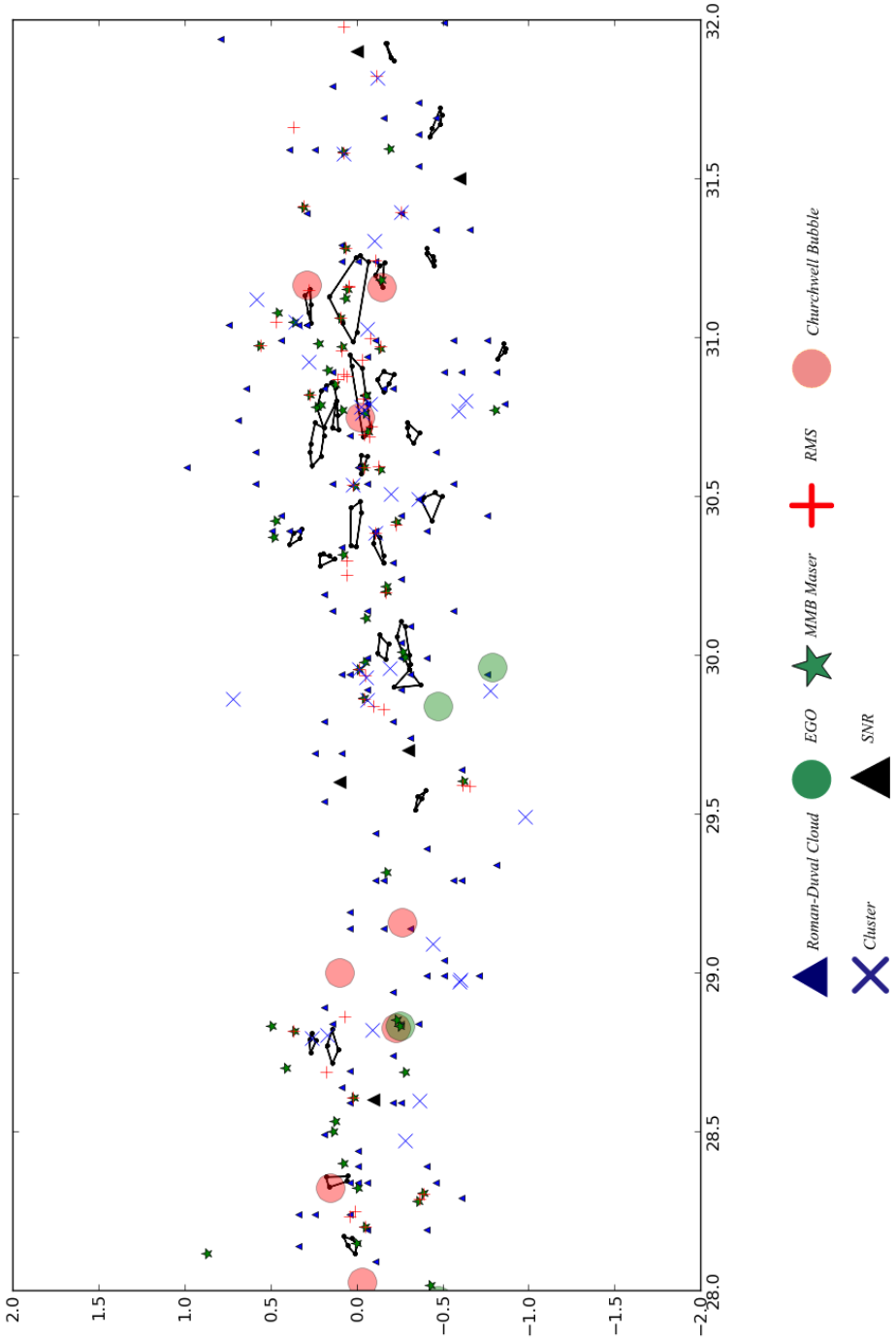


Figure C.33: Convex Hulls for BGPS $28^o < l < 32^o$

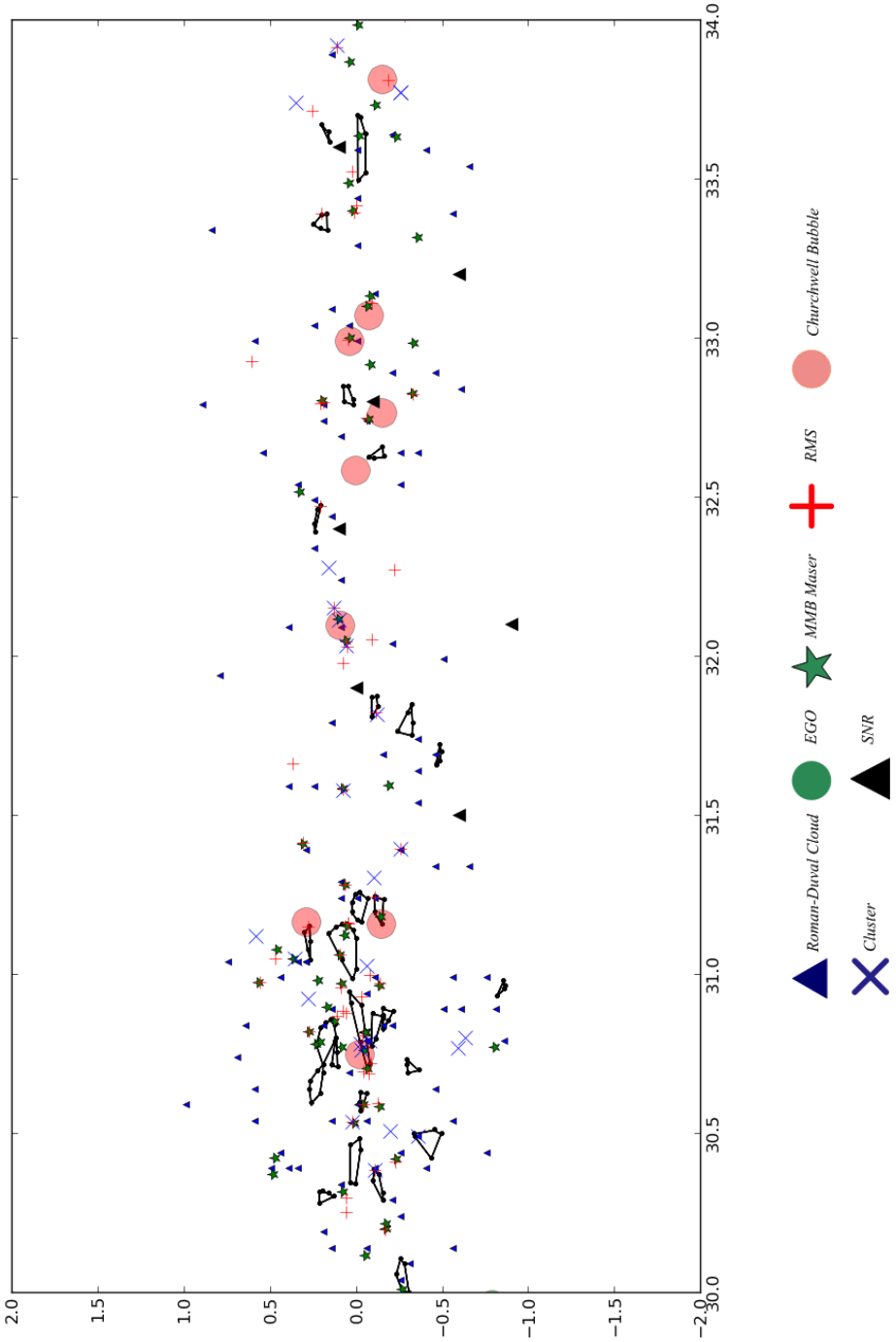


Figure C.34: Convex Hulls for BGPS $30^o < l < 34^o$

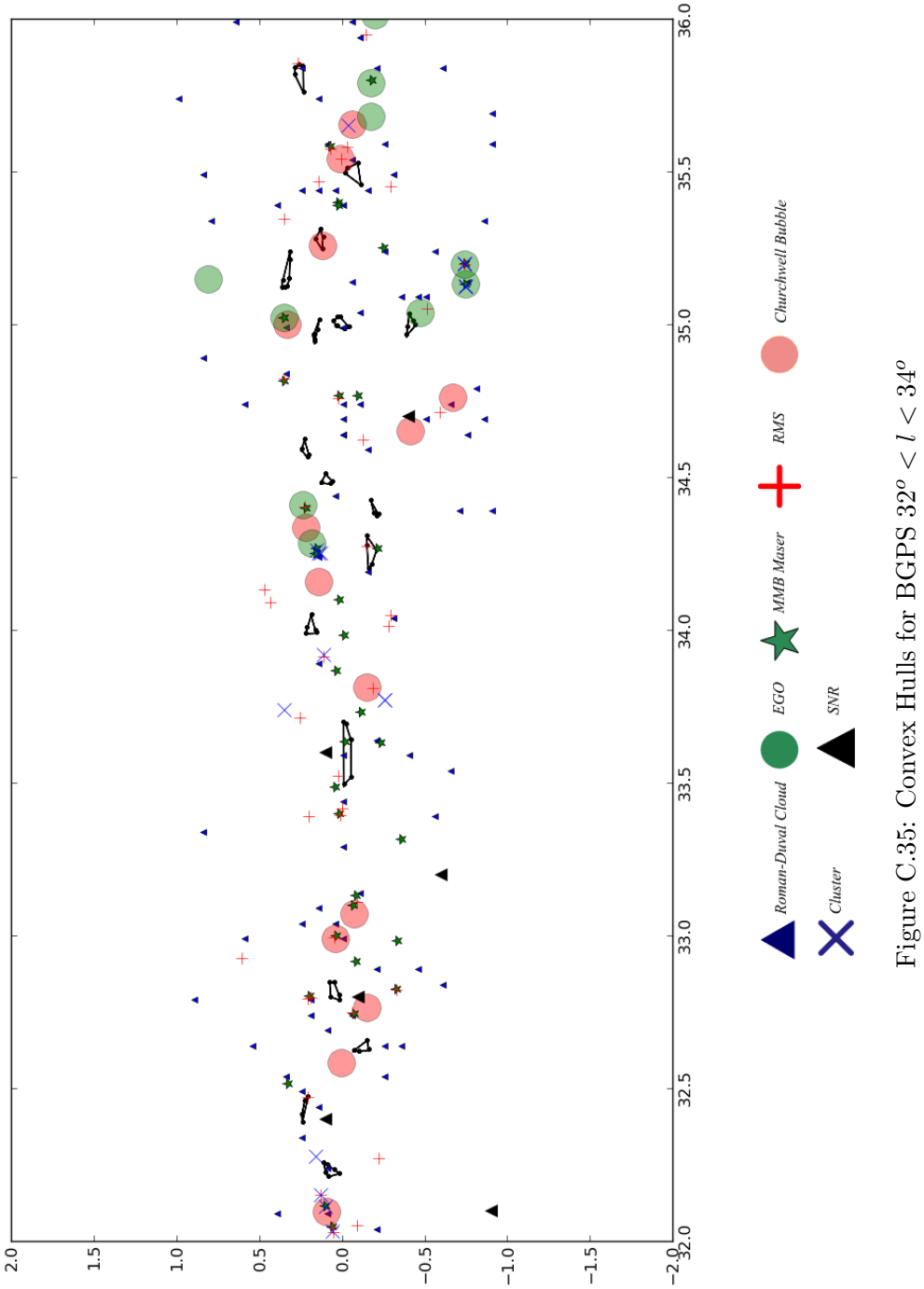


Figure C.35: Convex Hulls for BGPS $32^o < l < 34^o$

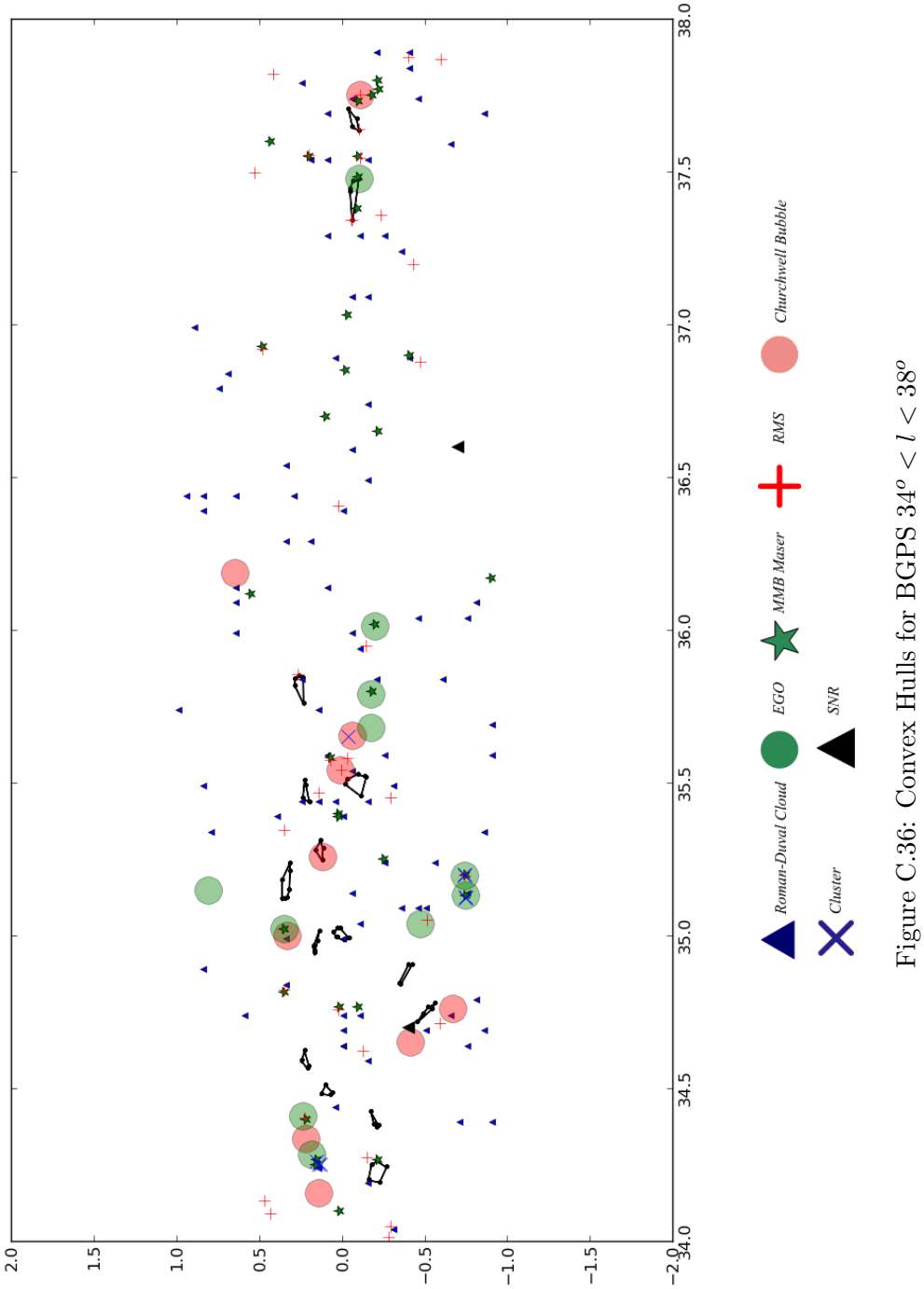


Figure C.36: Convex Hulls for BGPS $34^o < l < 38^o$

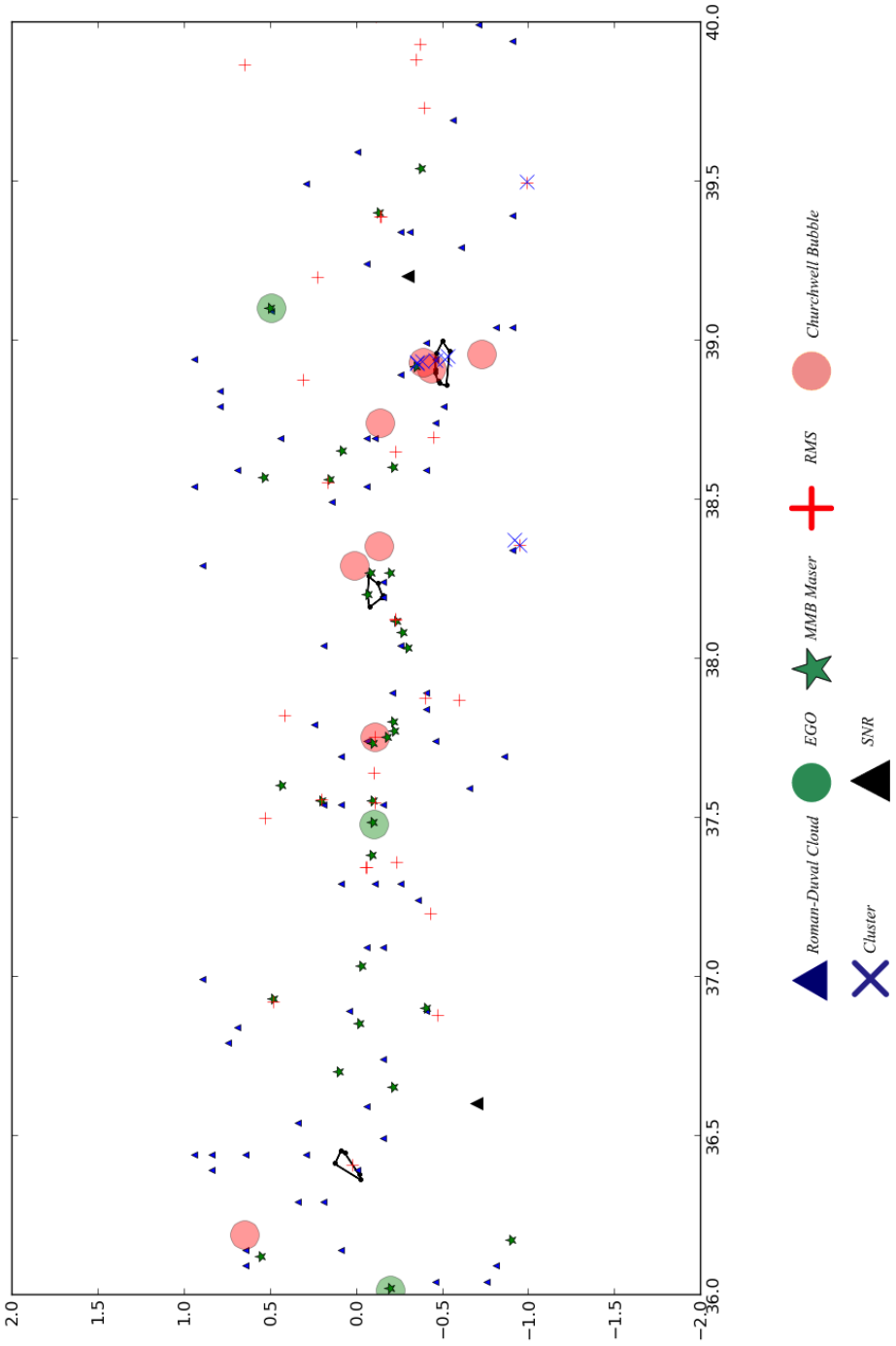


Figure C.37: Convex Hulls for BGPS $36^o < l < 40^o$

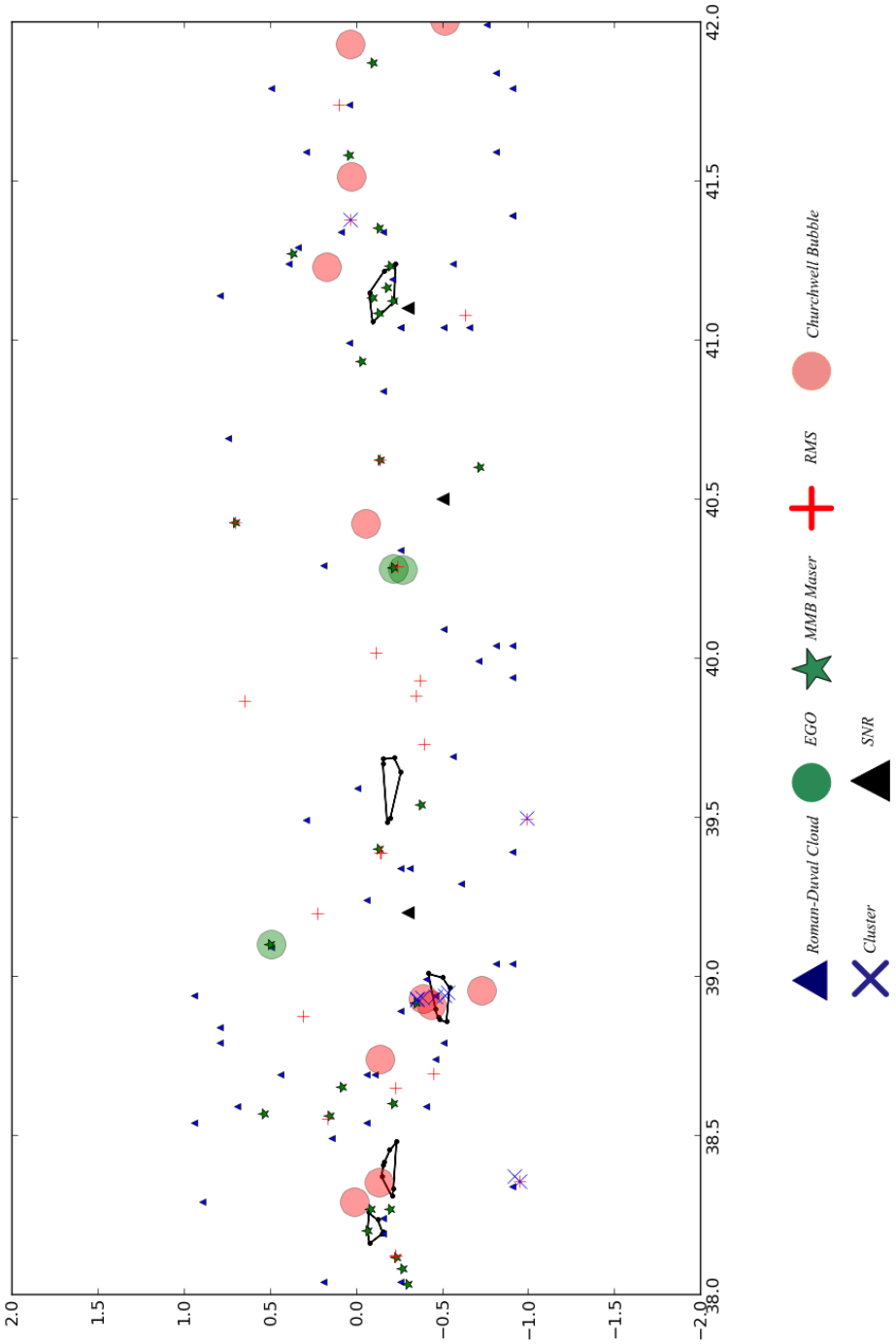


Figure C.38: Convex Hulls for BGPS $38^o < l < 42^o$

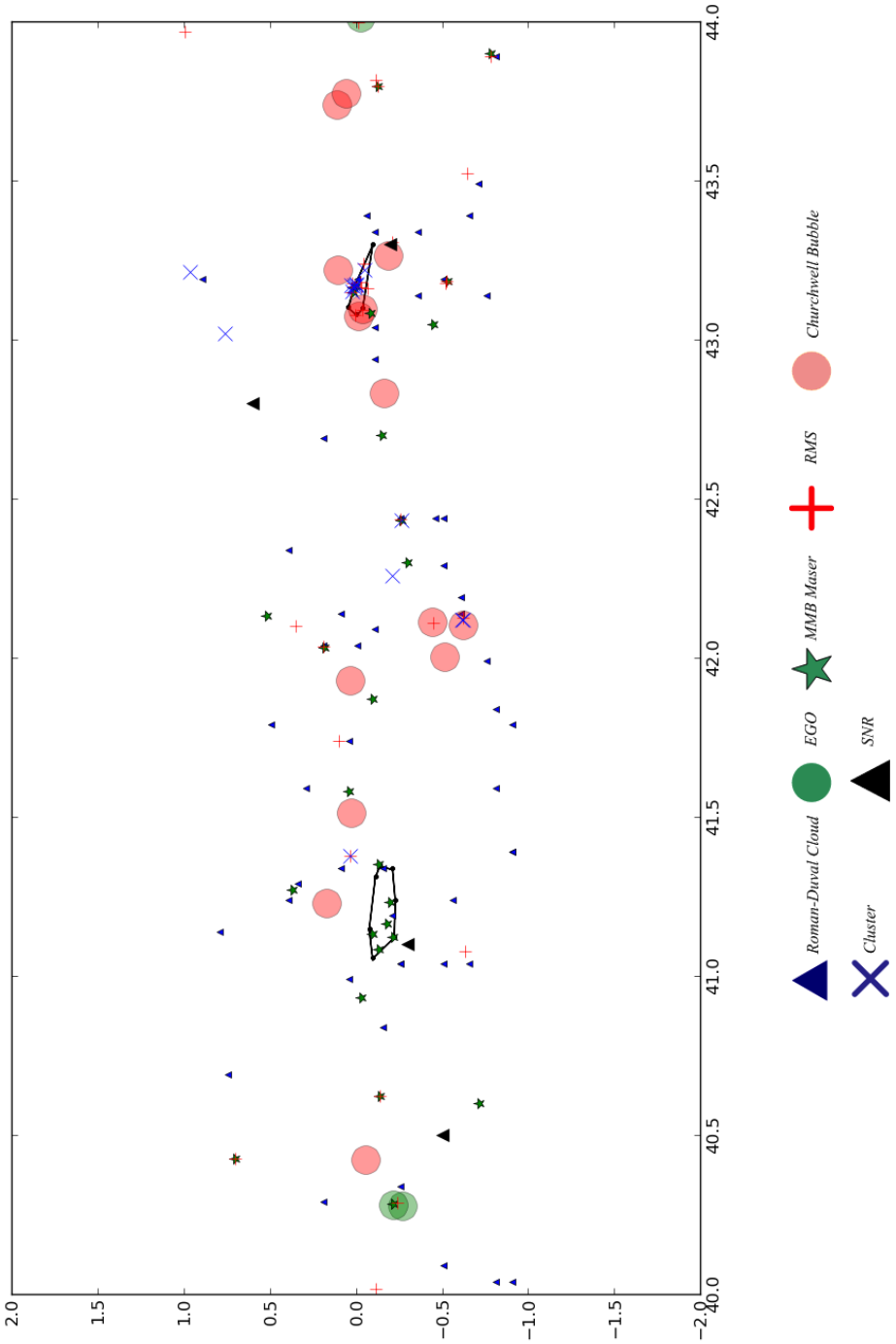


Figure C.39: Convex Hulls for BGPS $40^o < l < 44^o$

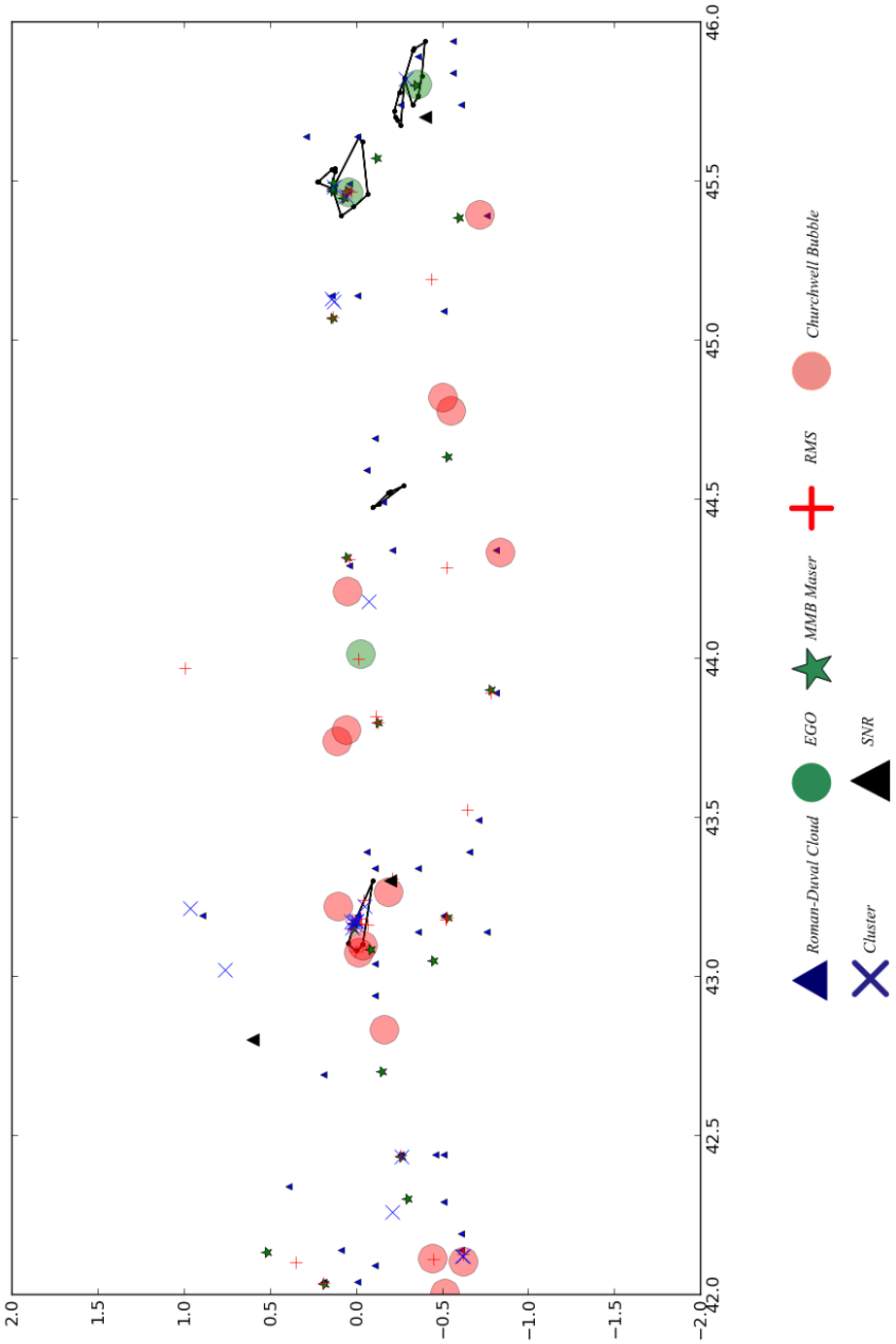


Figure C.40: Convex Hulls for BGPS $42^o < l < 46^o$

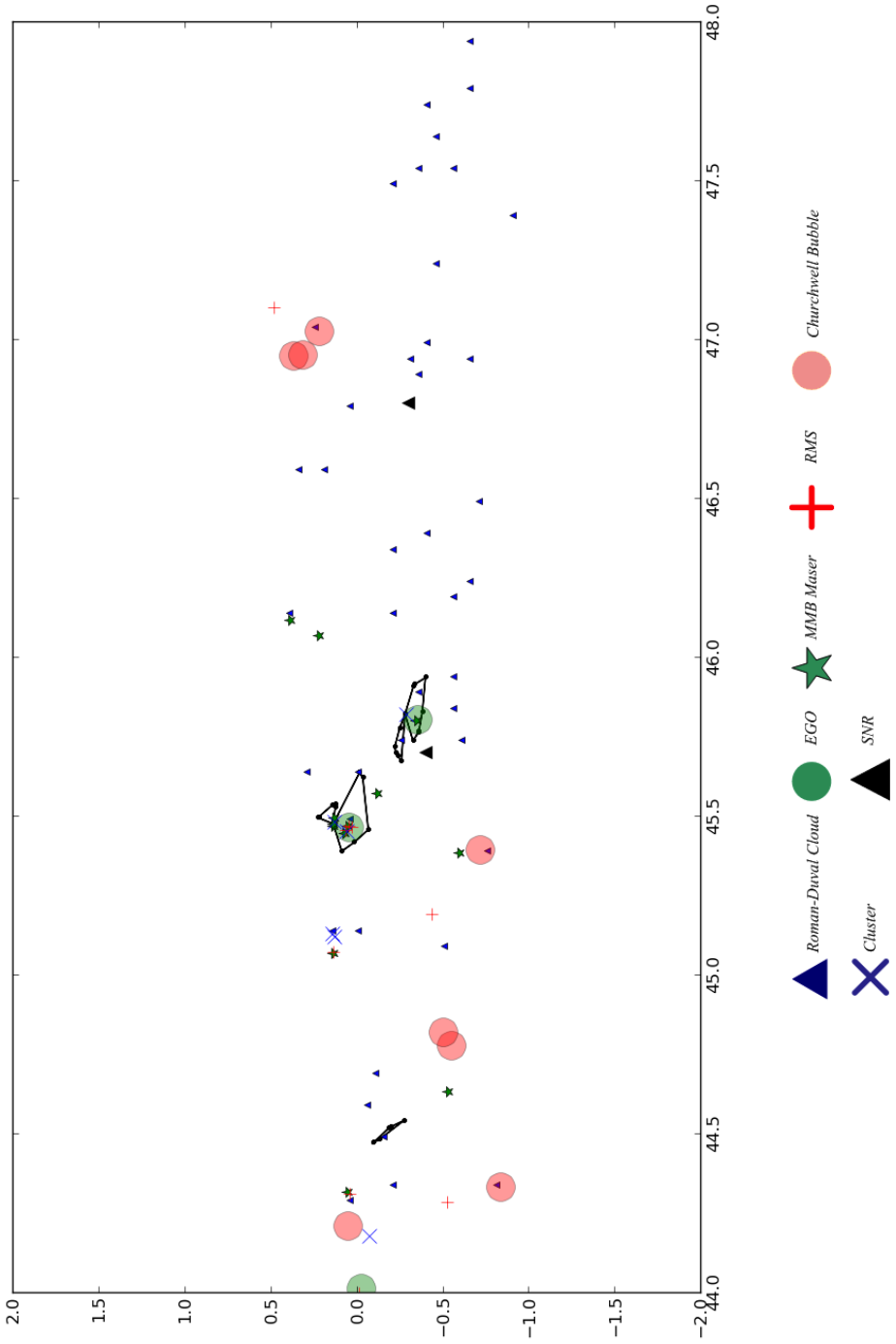


Figure C.41: Convex Hulls for BGPS $44^o < l < 38^o$

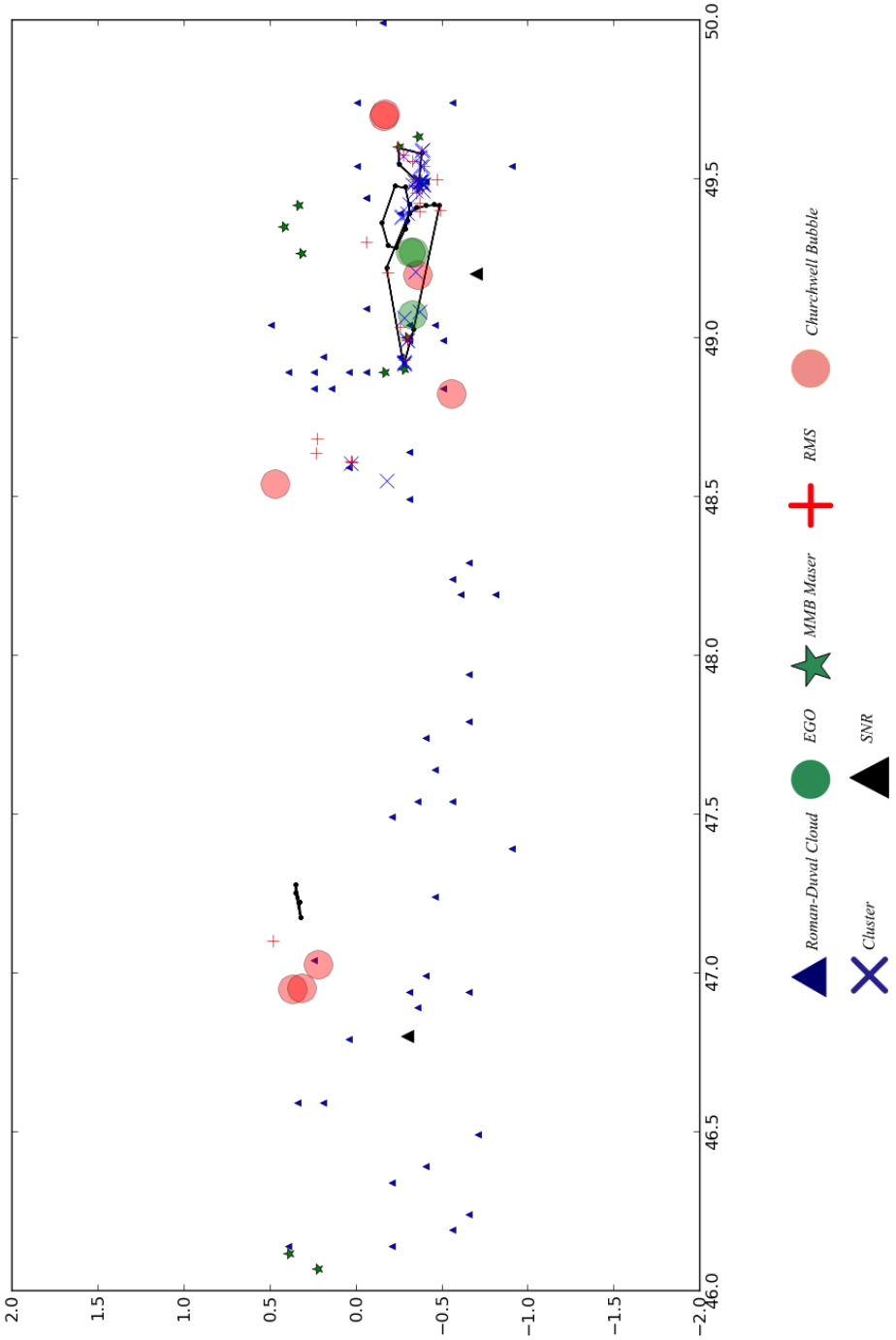
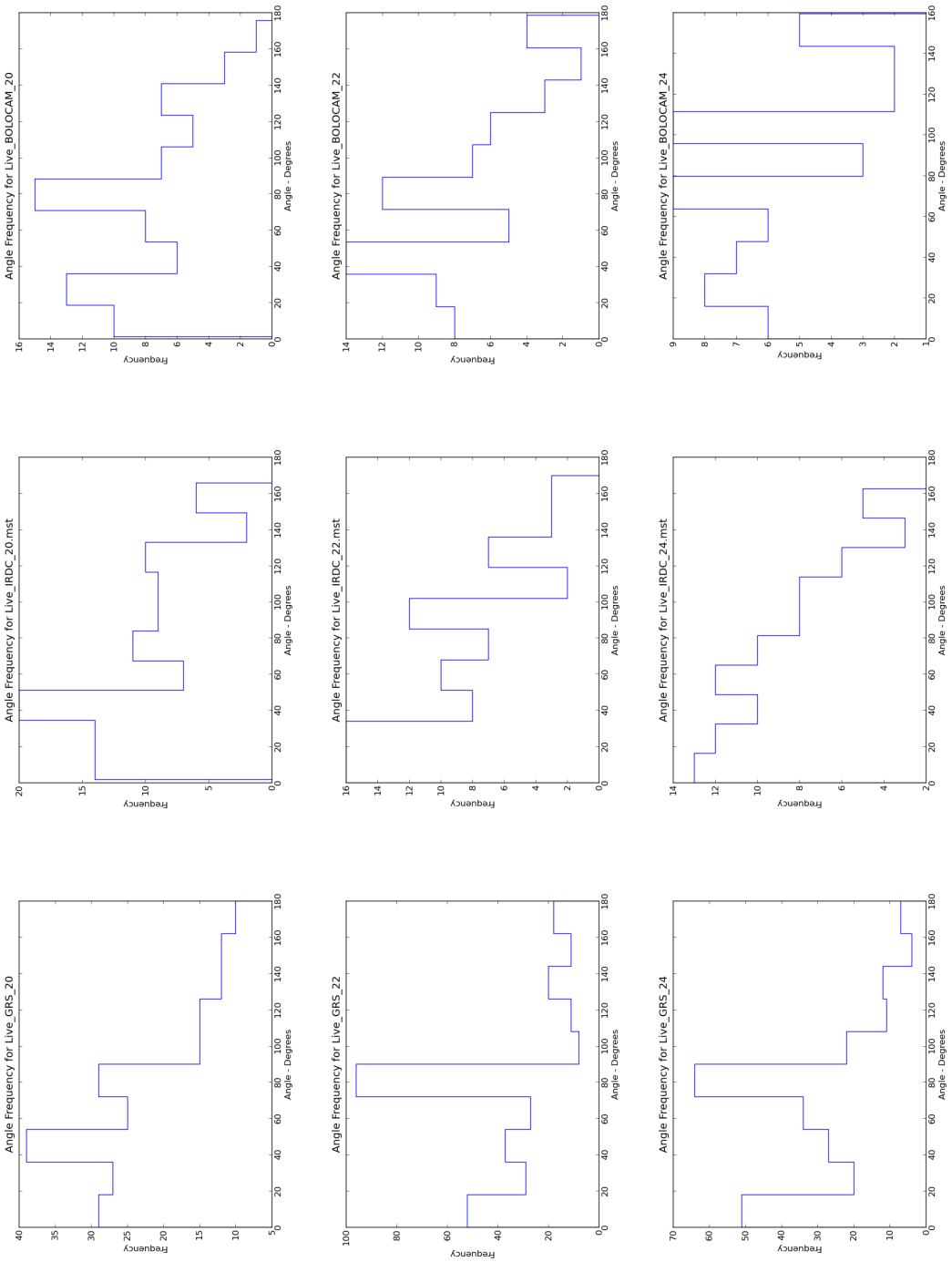
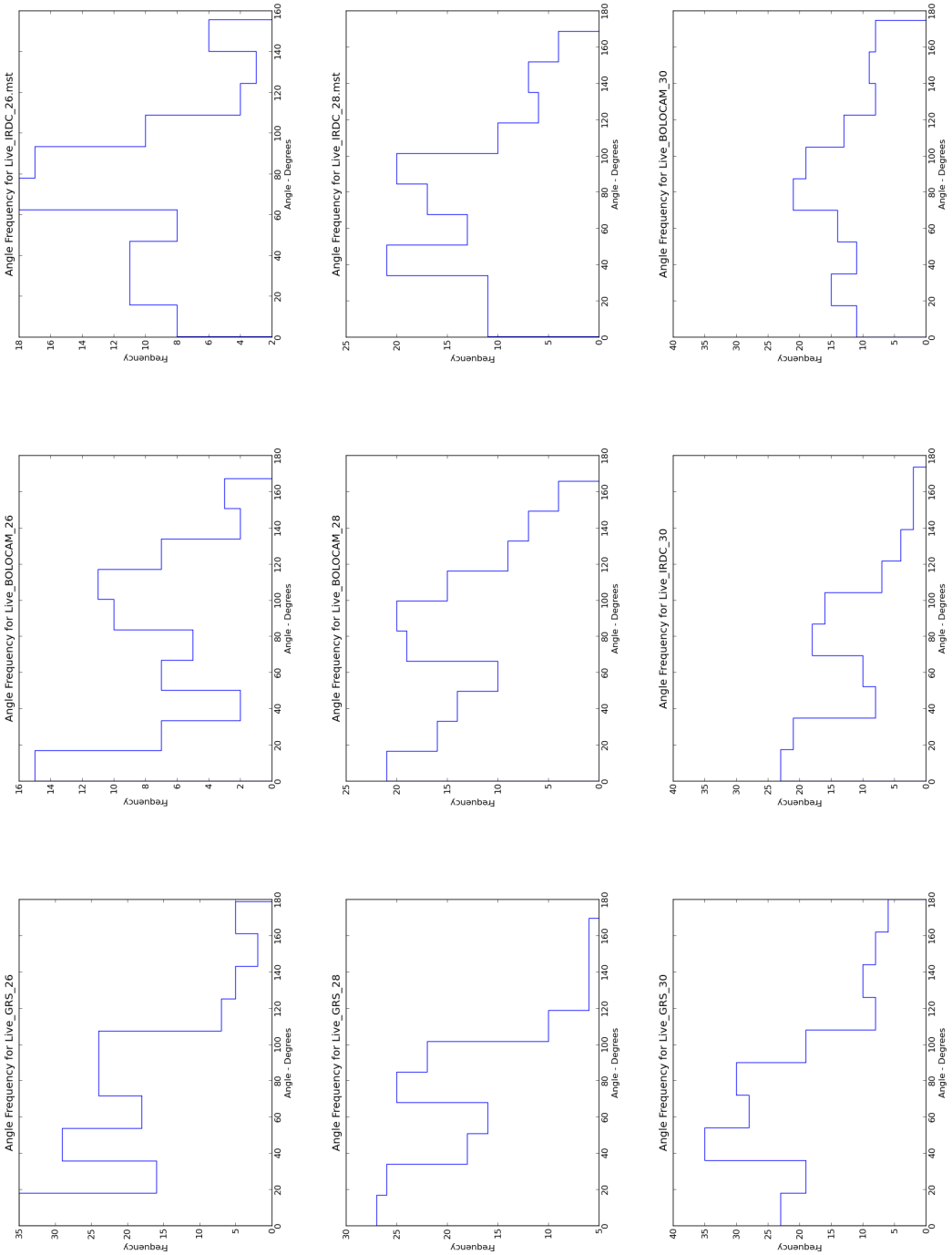


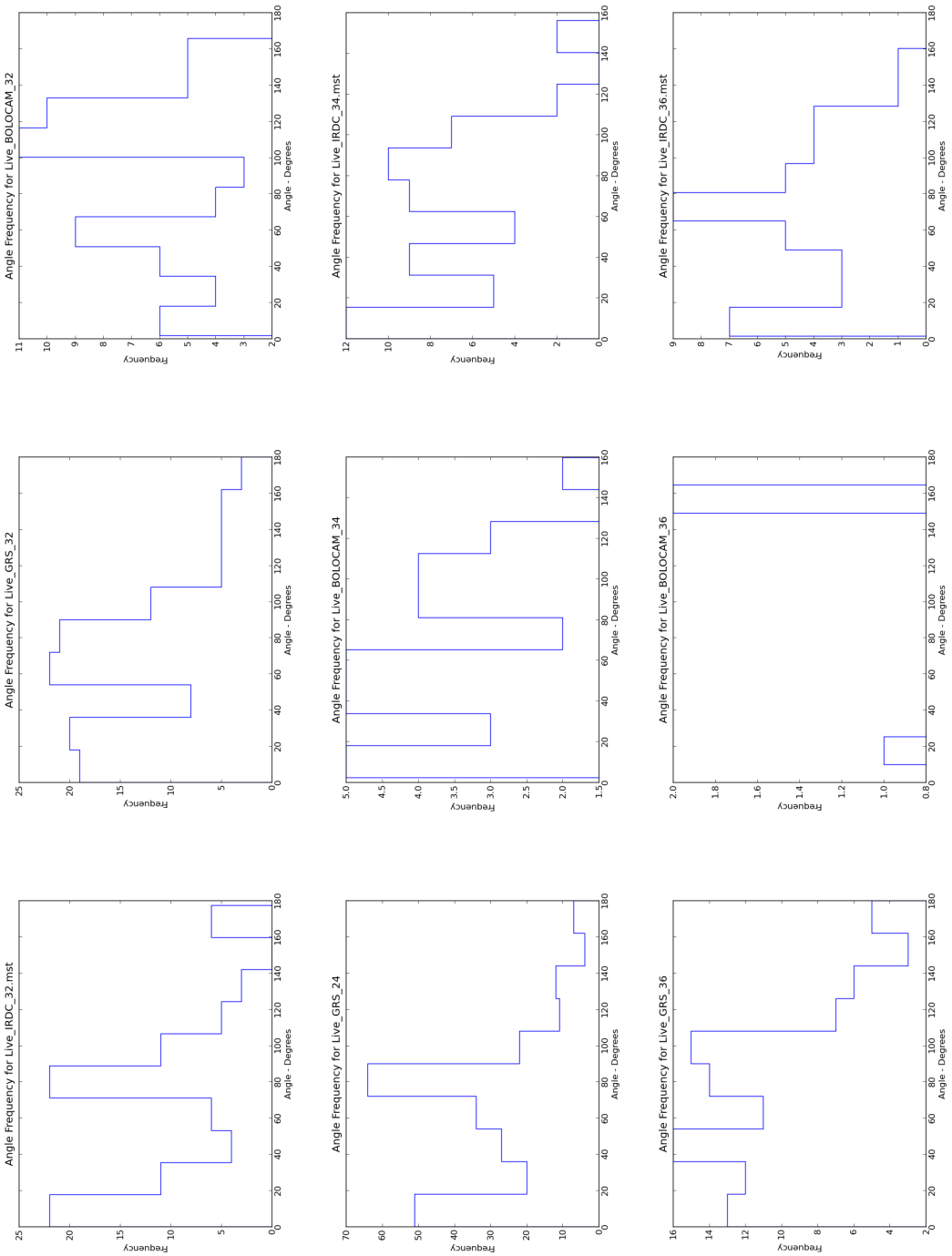
Figure C.42: Convex Hulls for BGPS $46^o < l < 50^o$

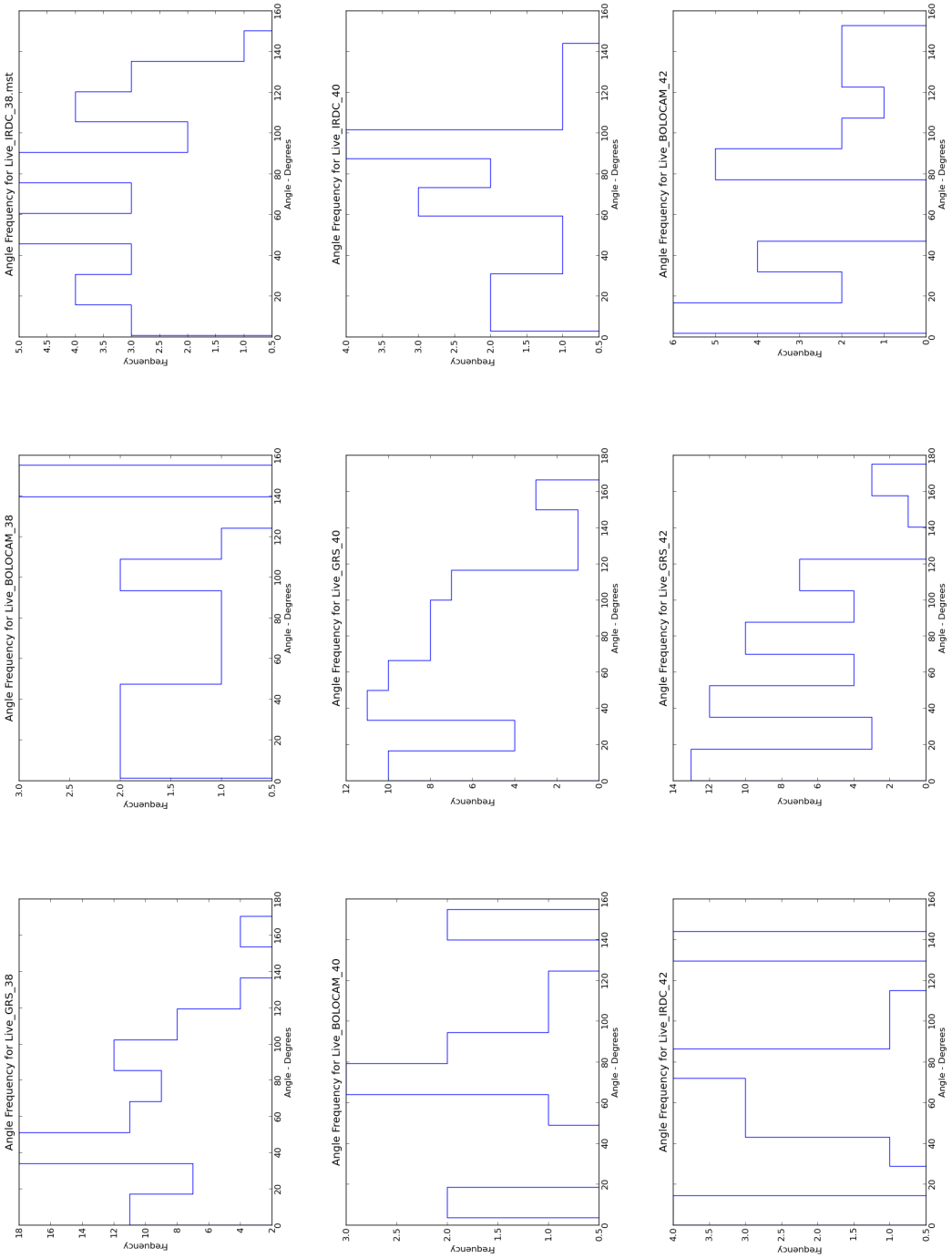
APPENDIX A

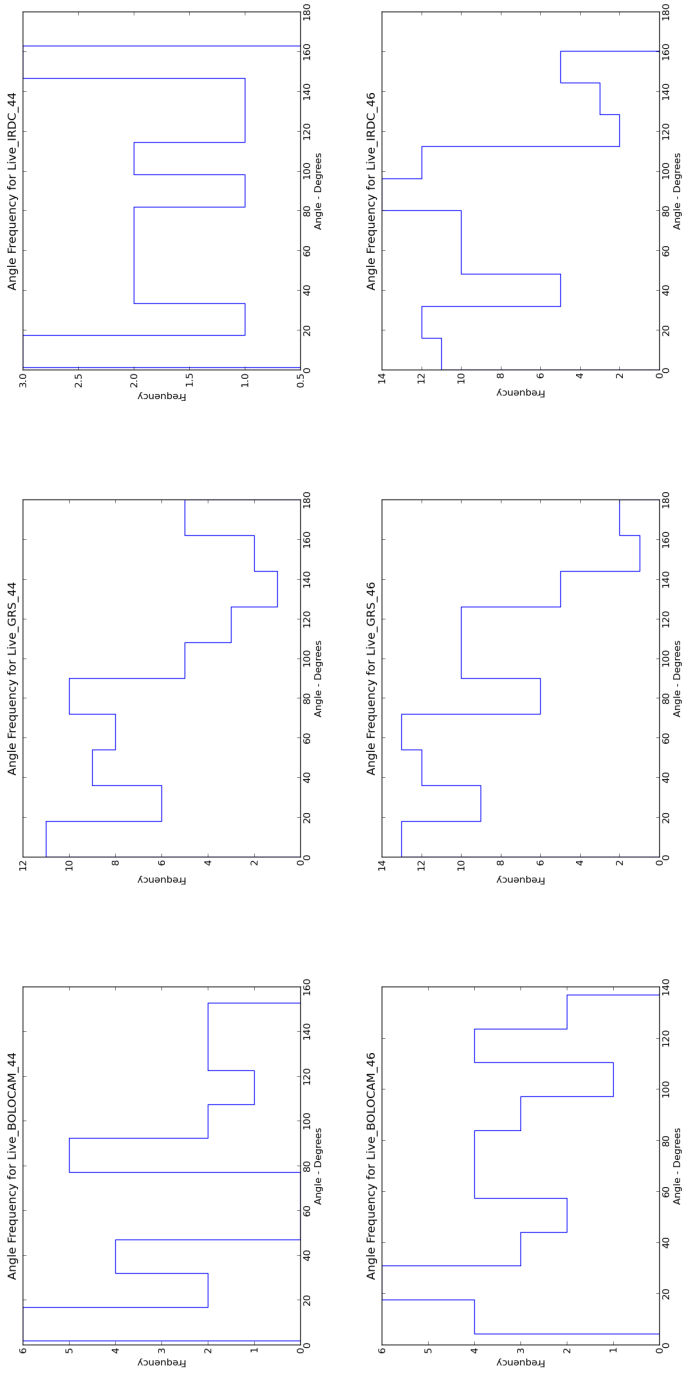
Edge Angle Distribution


Angle Frequency for Live_BOLOCAM_20
Angle Frequency for Live_BOLOCAM_22
Angle Frequency for Live_BOLOCAM_24
Angle Frequency for Live_IRDC_20.mst
Angle Frequency for Live_IRDC_22.mst
Angle Frequency for Live_IRDC_24.mst
Angle Frequency for Live_GRS_20
Angle Frequency for Live_GRS_22
Angle Frequency for Live_GRS_24
Frequency
Angle - Degrees


Angle Frequency for Live_IRDC_26.mst
Angle Frequency for Live_IRDC_28.mst
Angle Frequency for Live_BOLOCAM_30
Angle Frequency for Live_BOLOCAM_26
Angle Frequency for Live_BOLOCAM_28
Angle Frequency for Live_IRDC_30
Angle Frequency for Live_GRS_26
Angle Frequency for Live_GRS_28
Angle Frequency for Live_GRS_30
Frequency
Angle - Degrees


Angle Frequency for Live_BOLOCAM_32
Angle Frequency for Live_IRDC_34.mst
Angle Frequency for Live_IRDC_36.mst
Angle Frequency for Live_GRS_32
Angle Frequency for Live_BOLOCAM_34
Angle Frequency for Live_BOLOCAM_36
Angle Frequency for Live_IRDC_32.mst
Angle Frequency for Live_GRS_24
Angle Frequency for Live_GRS_36
Frequency
Angle - Degrees


Angle Frequency for Live_IRDC_38.mst
Angle Frequency for Live_IRDC_40
Angle Frequency for Live_BOLOCAM_42
Angle Frequency for Live_BOLOCAM_38
Angle Frequency for Live_GRS_40
Angle Frequency for Live_GRS_42
Angle Frequency for Live_GRS_38
Angle Frequency for Live_BOLOCAM_40
Angle Frequency for Live_IRDC_42
Frequency
Angle - Degrees


Angle Frequency for Live_IRDC_44
Angle Frequency for Live_IRDC_46
Angle Frequency for Live_GRS_44
Angle Frequency for Live_GRS_46
Angle Frequency for Live_BOLOCAM_44
Angle Frequency for Live_BOLOCAM_46
Frequency
Angle - Degrees

APPENDIX B

GRS Convex Hull data

GRS Convex hull statistics. Col 1: Hull ID. Col 2: Minimum distance in Kpc from the Roman Duval Cloud association. Col 3: Maximum distance in Kpc from the RomanDuval Cloud association. Col 4: Cell in which the hull is identified. Col 5: Number of Nodes within hull. Col 6: Intercluster mean edge cost. Col 7: Field mean edge cost. Col 8: The Cartwright Clustering measure Q. Col 9: Area of hull in square degrees. Col:10 Semi Major Axis of hull in degrees. Col 11 & 12: Position of centre of hull in galactic coordinates. Col 13: Billot Elongation measure. Col 14 Schmeja elongation measure. Col 15: Q^*. Col 16:Distance to Associated Murray SFR in kpc

1	0.82	0.75	20	10	0.0213	0.0432	0.4937	0.0013	0.0335	20.37	0.5425	1.6488	0.6065	0.6198	-
2	12.57	0.77	20	12	0.0137	0.0432	0.3178	0.0008	0.0292	20.435	-0.37	1.827	0.5473	0.4316	-
3	12.57	0.82	20	10	0.0258	0.0432	0.5964	0.0006	0.0453	20.595	-0.48	3.2762	0.3052	0.6598	-
4	10.62	0.82	20	8	0.0222	0.0432	0.5141	0.0011	0.0269	20.7867	0.4667	1.4728	0.679	0.6554	-
5	14.02	0.82	20	8	0.0276	0.0432	0.6381	0.0015	0.03	20.8067	0.14	1.3729	0.7284	0.7896	-
6	14.02	3.85	21	16	0.0215	0.0432	0.4963	0.0039	0.0559	21.298	0.052	1.5866	0.6303	0.6274	-
7	14.02	0.9	21	12	0.0311	0.0432	0.7202	0.0033	0.0541	21.402	0.038	1.6815	0.5947	0.8439	-
8	12.57	3.85	21	12	0.0268	0.0432	0.6206	0.0018	0.045	21.58	-0.0875	1.9066	0.5245	0.7297	-
9	11.07	3.85	21	12	0.0268	0.0432	0.6196	0.0022	0.057	21.7	-0.44	2.1543	0.4642	0.7161	-
10	13.88	0.9	21	10	0.0226	0.0432	0.5228	0.0012	0.03	21.714	-0.022	1.535	0.6515	0.6583	-
11	11.07	3.8	21	16	0.0302	0.0432	0.6975	0.0074	0.085	21.8671	-0.3214	1.7573	0.569	0.8158	-
12	11.2	4.03	21	14	0.0225	0.0432	0.5211	0.0024	0.045	21.988	-0.556	1.6453	0.6078	0.6475	-
13	11.2	3.85	22	14	0.017	0.033	0.5154	0.001	0.0292	22.028	-0.532	1.6766	0.5965	0.6394	-
14	11.2	3.85	22	20	0.0191	0.0432	0.441	0.0027	0.0474	22.045	-0.5283	1.6332	0.6123	0.5684	-
15	15.07	4.3	22	24	0.0221	0.033	0.6704	0.0028	0.0854	22.4083	-0.01	2.8619	0.3494	0.7431	-
16	15.07	4.3	22	38	0.0223	0.0432	0.5157	0.0055	0.1006	22.4283	-0.0283	2.4049	0.4158	0.6022	-
17	6.88	3.88	22	10	0.0213	0.0432	0.4919	0.0006	0.0224	22.85	-0.444	1.69	0.5917	0.615	10.8
18	6.88	3.88	22	12	0.0148	0.033	0.4497	0.0007	0.0224	22.866	-0.436	1.4472	0.691	0.5934	10.8
19	13.3	3.88	22	22	0.0204	0.033	0.6181	0.0019	0.0585	22.9338	-0.2138	2.4117	0.4146	0.7043	10.8
20	6.88	3.88	23	14	0.0263	0.033	0.797	0.0052	0.0495	23.0025	-0.25	1.2225	0.818	0.9671	10.8
21	13.93	3.88	23	12	0.0209	0.033	0.6332	0.0014	0.0381	23.008	-0.102	1.8038	0.5544	0.7485	10.8
22	6.88	3.88	23	54	0.0234	0.0432	0.5403	0.0212	0.1553	23.0433	-0.3417	1.8908	0.5289	0.6503	10.8
23	13.93	3.88	23	38	0.02	0.0432	0.4617	0.008	0.0825	23.0538	-0.1125	1.6393	0.61	0.5886	10.8
24	13.93	3.88	23	10	0.0094	0.033	0.2842	0.0005	0.0112	23.085	-0.0617	0.9342	1.0705	0.5069	-
25	11.77	3.88	23	10	0.0184	0.033	0.5568	0.001	0.0269	23.1514	-0.3414	1.5092	0.6626	0.6946	10.8
26	11.77	3.88	23	44	0.022	0.0432	0.5083	0.0099	0.1063	23.2867	-0.245	1.8936	0.5281	0.6182	-
27	11.77	3.88	23	34	0.0207	0.033	0.6289	0.006	0.0743	23.3057	-0.2543	1.708	0.5855	0.7507	-

28	13.93	3.88	23	12	0.0245	0.0432	0.5675	0.0026	0.05	23.33	-0.078	1.738	0.5754	0.6872	-
29	13.93	3.88	23	88	0.0154	0.0432	0.3557	0.034	0.1188	23.4109	-0.1027	1.1433	0.8747	0.5377	-
30	13.93	3.88	23	82	0.0146	0.033	0.4426	0.0165	0.1001	23.4286	-0.0829	1.3837	0.7227	0.593	-
31	13.93	3.88	23	12	0.0239	0.0432	0.5525	0.0023	0.0453	23.458	0.066	1.6918	0.5911	0.6755	-
32	11.57	0.65	23	8	0.0156	0.033	0.4729	0.0005	0.025	23.49	-0.5025	1.9817	0.5046	0.5778	-
33	11.57	0.65	23	16	0.0208	0.0432	0.4814	0.0041	0.0474	23.515	-0.5038	1.313	0.7616	0.6398	-
34	13.93	0.88	23	12	0.0193	0.0432	0.4462	0.0007	0.0474	23.6	-0.025	3.1777	0.3147	0.5116	-
35	13.93	0.88	23	50	0.0236	0.0432	0.5454	0.0058	0.1012	23.64	0.1225	2.3563	0.4244	0.6337	-
36	13.93	0.88	23	8	0.0216	0.033	0.6547	0.0002	0.0364	23.676	0.532	4.5621	0.2192	0.7003	-
37	11.7	2.95	23	12	0.0234	0.0432	0.5403	0.0019	0.0461	23.844	-0.106	1.8745	0.5335	0.6513	10.3
38	11.7	2.95	23	14	0.0124	0.033	0.3772	0.0009	0.0292	23.8625	-0.0925	1.7725	0.5642	0.4946	10.3
39	11.57	3.78	23	14	0.0285	0.0432	0.6589	0.0021	0.0502	23.895	-0.3183	1.9436	0.5145	0.7659	10.3
40	11.7	0.88	23	26	0.0248	0.0432	0.574	0.0067	0.0776	23.9567	-0.0183	1.6871	0.5927	0.6973	10.3
41	12.77	2.95	23	8	0.0266	0.033	0.8083	0.0012	0.0364	23.99	-0.07	1.8625	0.5369	0.92	10.3
42	12.77	0.88	24	20	0.0136	0.033	0.4111	0.0012	0.0292	24.022	0.08	1.4917	0.6704	0.5505	-
43	11.7	3.2	24	16	0.0168	0.033	0.5086	0.0012	0.0391	24.09	-0.21	1.9981	0.5005	0.6127	10.3
44	12.77	2.95	24	40	0.0173	0.033	0.5244	0.0035	0.0743	24.1	-0.0413	2.2269	0.449	0.6178	10.3
45	12.77	0.88	24	18	0.0145	0.0331	0.4368	0.0008	0.0364	24.1033	0.14	2.2811	0.4384	0.5279	-
46	12.77	0.88	24	12	0.02	0.033	0.6076	0.0015	0.0364	24.145	0.1113	1.6943	0.5902	0.7304	-
47	12.77	0.88	24	64	0.0155	0.033	0.4698	0.0087	0.0873	24.228	-0.036	1.6593	0.6026	0.5952	10.3
48	12.77	0.88	24	10	0.0166	0.033	0.5034	0.0006	0.0292	24.3	0.0857	2.1096	0.474	0.602	-
49	11.57	3.2	24	10	0.0199	0.0331	0.6005	0.0008	0.0453	24.322	-0.534	2.8373	0.3524	0.6738	10.3
50	12.77	3.25	24	14	0.0203	0.0331	0.6138	0.0017	0.0495	24.3225	-0.125	2.1278	0.47	0.7116	10.3
51	12.77	0.88	24	28	0.0165	0.033	0.4997	0.0034	0.0626	24.355	0.105	1.9185	0.5212	0.6081	-
52	11.7	3.2	24	20	0.0093	0.033	0.282	0.0006	0.0335	24.372	-0.206	2.427	0.412	0.3677	10.3
53	12.77	0.88	24	16	0.0117	0.033	0.3544	0.0009	0.0354	24.385	0.0225	2.0889	0.4787	0.454	-
54	11.25	2.97	24	26	0.0135	0.033	0.408	0.0036	0.0461	24.414	-0.258	1.3713	0.7292	0.5597	-
55	12.77	0.88	24	14	0.0172	0.0331	0.5185	0.0011	0.0292	24.4175	0.2363	1.5581	0.6418	0.652	-
56	11.25	2.97	24	16	0.0127	0.0331	0.3834	0.0012	0.0316	24.432	-0.254	1.618	0.618	0.5119	-
57	12.77	0.88	24	22	0.0171	0.033	0.5189	0.0018	0.043	24.488	-0.01	1.7969	0.5565	0.6346	-
58	11.25	2.97	24	36	0.0214	0.033	0.6505	0.0066	0.068	24.52	-0.1778	1.4894	0.6714	0.7902	6.1
59	12.77	0.88	24	12	0.024	0.033	0.7271	0.0016	0.0495	24.538	0.316	2.1933	0.4559	0.822	6.1
60	12.77	0.88	24	12	0.0122	0.033	0.3695	0.0007	0.0255	24.63	0.1825	1.7725	0.5642	0.4869	6.1

61	12.77	0.88	24	12	0.0122	0.0331	0.3676	0.0007	0.0255	24.63	0.1825	1.7725	0.5642	0.4849	6.1
62	12.77	2.97	24	10	0.011	0.033	0.3324	0.0005	0.0158	24.7775	0.0925	1.2533	0.7979	0.4984	6.1
63	12.77	2.97	24	10	0.011	0.0331	0.3306	0.0005	0.0158	24.7775	0.0925	1.2533	0.7979	0.4966	6.1
64	12.77	2.97	24	16	0.0187	0.0331	0.5637	0.003	0.05	24.79	-0.06	1.618	0.618	0.6923	6.1
65	11.25	0.62	24	26	0.0177	0.0331	0.5328	0.0029	0.0702	24.87	-0.15	2.3098	0.4329	0.6228	6.1
66	11.25	0.62	24	28	0.0191	0.033	0.5781	0.0029	0.0738	24.924	-0.14	2.4298	0.4116	0.6637	6.1
67	9.5	0.62	25	14	0.0157	0.033	0.4768	0.0018	0.0364	25.0571	0.0414	1.5423	0.6484	0.6116	6.1
68	9.5	0.62	25	10	0.0167	0.0331	0.5053	0.0007	0.035	25.154	0.17	2.3447	0.4265	0.594	6.1
69	11.5	0.62	25	8	0.0156	0.0331	0.4703	0.0009	0.0206	25.275	-0.0383	1.2533	0.7979	0.6363	6.1
70	11.5	0.62	25	12	0.0188	0.0331	0.5661	0.0012	0.0412	25.3929	-0.15	2.1096	0.474	0.6647	-
71	5.95	0.62	25	12	0.0141	0.033	0.4287	0.0004	0.035	25.412	-0.45	3.1018	0.3224	0.4957	-
72	7.68	0.62	25	8	0.0146	0.0331	0.4408	0.0002	0.0255	25.415	-0.33	3.1953	0.313	0.5059	4.1
73	4.22	1.1	25	14	0.0104	0.033	0.3147	0.0006	0.0283	25.4267	-0.5333	2.0467	0.4886	0.4163	-
74	11.5	0.62	25	14	0.0139	0.033	0.4224	0.001	0.0361	25.47	-0.1275	2.0209	0.4948	0.5253	-
75	11.5	0.62	25	8	0.0204	0.033	0.6189	0.0008	0.0304	25.6	-0.0867	1.9684	0.508	0.7246	-
76	11.5	0.62	25	30	0.0171	0.0331	0.5155	0.0024	0.0716	25.6286	-0.1243	2.5901	0.3861	0.5958	-
77	11.5	0.62	25	16	0.0147	0.033	0.446	0.0008	0.0354	25.646	-0.144	2.2156	0.4514	0.5399	-
78	11.5	0.62	25	22	0.0124	0.0331	0.3726	0.0012	0.0391	25.68	0.048	1.9981	0.5005	0.4767	-
79	11.5	2.3	25	14	0.0193	0.033	0.5843	0.0024	0.0412	25.735	0.2	1.4917	0.6704	0.7237	-
80	11.5	2.3	25	14	0.0155	0.033	0.4712	0.0012	0.0269	25.87	0.0925	1.3777	0.7258	0.6222	8.8
81	11.5	1.1	26	10	0.0314	0.0339	0.928	0.0042	0.043	26.0825	-0.04	1.1834	0.845	1.1038	8.8
82	10.85	2.3	26	20	0.0132	0.0331	0.3991	0.0011	0.045	26.12	0.1	2.4615	0.4063	0.4836	8.8
83	10.85	2.3	26	24	0.0127	0.0339	0.3748	0.0011	0.045	26.12	0.1029	2.4049	0.4158	0.4613	8.8
84	10.85	1.1	26	32	0.018	0.0339	0.5322	0.0027	0.0806	26.225	0	2.7501	0.3636	0.6079	-
85	10.85	3.8	26	40	0.0176	0.0331	0.5304	0.0037	0.0854	26.27	0.0014	2.5066	0.3989	0.6134	-
86	10.85	3.8	26	18	0.0157	0.0339	0.4647	0.003	0.0424	26.3129	0.0243	1.3729	0.7284	0.6162	-
87	10.85	1.9	26	20	0.0183	0.0339	0.5407	0.0028	0.0461	26.5757	0.1114	1.5581	0.6418	0.6742	-
88	10.85	1.9	26	10	0.0219	0.0331	0.6603	0.0003	0.0361	26.576	0.14	3.6897	0.271	0.7167	-
89	9.8	1.9	26	12	0.015	0.0331	0.4535	0.0011	0.0255	26.5871	0.4543	1.3946	0.7171	0.6027	-
90	10.85	1.9	26	12	0.0225	0.0339	0.6652	0.0018	0.0335	26.5988	-0.0313	1.4211	0.7037	0.8116	-
91	10.85	1.9	26	16	0.0232	0.0331	0.7009	0.0029	0.05	26.6075	-0.0288	1.6457	0.6077	0.8273	-
92	8.02	0.8	26	8	0.0266	0.0331	0.8027	0.0008	0.0316	26.684	-0.692	1.9817	0.5046	0.9077	-
93	8.02	0.8	26	8	0.0266	0.0339	0.7858	0.0008	0.0316	26.684	-0.692	1.9817	0.5046	0.8907	-

94	13.23	3.33	27	12	0.0237	0.0339	0.7004	0.0013	0.0474	27.145	0.0083	2.378	0.4205	0.7879	-
95	13.23	5.25	27	10	0.0151	0.0331	0.4549	0.0011	0.0304	27.305	0.16	1.6636	0.6011	0.5799	-
96	13.23	0.73	27	28	0.0135	0.0331	0.4064	0.0021	0.0618	27.414	0.096	2.3921	0.418	0.4934	-
97	13.23	0.73	27	12	0.0164	0.0339	0.4851	0.0009	0.0335	27.438	0.096	1.9817	0.5046	0.59	-
98	13.23	0.73	27	8	0.0213	0.0331	0.6435	0.0009	0.0335	27.5775	0.22	1.9817	0.5046	0.7485	-
99	13.23	0.73	27	14	0.0146	0.0339	0.432	0.0008	0.0269	27.5943	0.1986	1.6873	0.5927	0.5552	-
100	13.23	0.73	27	16	0.0176	0.0331	0.5304	0.0026	0.0354	27.714	-0.108	1.241	0.8058	0.698	-
101	13.23	0.73	27	18	0.0151	0.0331	0.4543	0.0019	0.0461	27.74	0.056	1.8745	0.5335	0.5652	-
102	12.52	0.73	28	10	0.0192	0.0349	0.5492	0.0014	0.0224	28.0414	0.0071	1.0592	0.9441	0.7456	-
103	12.52	2.95	28	10	0.0284	0.0339	0.8392	0.0015	0.0492	28.3175	0.1275	2.2536	0.4437	0.9315	-
104	12.52	3.03	28	18	0.0212	0.0339	0.6262	0.0037	0.055	28.32	0.004	1.6026	0.624	0.756	-
105	10.07	0.62	28	20	0.0109	0.0349	0.3131	0.0013	0.0412	28.5033	-0.38	2.0269	0.4934	0.4157	-
106	12.52	0.62	28	10	0.0226	0.0349	0.6476	0.0007	0.043	28.5925	-0.24	2.8815	0.347	0.7197	-
107	12.52	3.03	28	14	0.0178	0.0339	0.5258	0.0014	0.0453	28.5975	0	2.1448	0.4662	0.6228	-
108	12.52	3.03	28	22	0.0077	0.0349	0.2191	0.0017	0.0255	28.6657	0.0443	1.096	0.9124	0.4089	-
109	12.52	3.03	28	22	0.0077	0.0339	0.2261	0.0017	0.0255	28.6657	0.0443	1.096	0.9124	0.4159	-
110	12.52	3.47	28	14	0.02	0.0339	0.591	0.0014	0.0424	28.775	0.165	2.0467	0.4886	0.6926	-
111	11.55	3.47	28	12	0.0145	0.0349	0.4136	0.0006	0.0335	28.898	0.008	2.535	0.3945	0.4957	-
112	11.55	3.47	28	14	0.0111	0.0339	0.3274	0.0007	0.0304	28.922	0.102	2.0375	0.4908	0.4295	-
113	11.55	0.62	28	8	0.021	0.0349	0.6009	0.0009	0.0255	28.9475	-0.45	1.55	0.6452	0.7351	-
114	11.55	0.62	28	8	0.021	0.0339	0.6201	0.0009	0.0255	28.9475	-0.45	1.55	0.6452	0.7543	-
115	11.55	0.85	29	20	0.0153	0.0339	0.453	0.0025	0.0532	29.0167	-0.53	1.9033	0.5254	0.5623	-
116	11.55	0.85	29	18	0.0145	0.0349	0.4139	0.0018	0.043	29.0233	-0.5333	1.8224	0.5487	0.5281	-
117	6.25	3.47	29	8	0.0187	0.0349	0.5346	0.0008	0.0212	29.0633	0.3067	1.3729	0.7284	0.6861	-
118	11.55	3.22	29	10	0.0216	0.0349	0.6189	0.0009	0.0335	29.1633	0.05	1.9817	0.5046	0.7238	-
119	11.55	1.58	29	10	0.0193	0.0349	0.5525	0.0012	0.0304	29.29	-0.13	1.5562	0.6426	0.6862	-
120	11.55	5.05	29	22	0.0146	0.0349	0.4191	0.0025	0.0381	29.3	0.1514	1.3499	0.7408	0.5732	-
121	11.55	0.85	29	16	0.0144	0.0349	0.4111	0.0011	0.0255	29.328	-0.456	1.3946	0.7171	0.5603	-
122	11.55	1.58	29	16	0.0173	0.0349	0.4963	0.0021	0.0364	29.3617	-0.34	1.425	0.7018	0.6423	-
123	9.05	0.85	29	14	0.018	0.0349	0.5163	0.0006	0.0361	29.3971	-0.44	2.609	0.3833	0.596	-
124	9.05	3.22	29	10	0.012	0.0349	0.3445	0.0005	0.018	29.4275	-0.0525	1.429	0.6998	0.49	-
125	14	0.85	29	12	0.0195	0.0349	0.5591	0.0008	0.0364	29.465	-0.3733	2.213	0.4519	0.6531	-
126	14	0.85	29	10	0.0178	0.0339	0.525	0.0008	0.0364	29.4775	-0.37	2.213	0.4519	0.619	-

127	14	4.55	29	12	0.0165	0.0349	0.4719	0.0004	0.032	29.6833	-0.1917	2.675	0.3738	0.5497	-
128	14	3.25	29	18	0.0141	0.0349	0.4025	0.0016	0.0461	29.774	-0.256	2.0753	0.4818	0.5027	-
129	14	3.25	29	18	0.0141	0.0339	0.4154	0.0016	0.0461	29.774	-0.256	2.0753	0.4818	0.5156	-
130	14	2.65	29	12	0.0186	0.0349	0.5324	0.0018	0.0335	29.945	-0.025	1.4012	0.7136	0.6808	8.7
131	14	2.65	29	30	0.0147	0.0339	0.4351	0.0023	0.0559	29.952	-0.044	2.066	0.484	0.5357	8.7
132	14	3.25	30	26	0.0153	0.037	0.4144	0.0016	0.0539	30.03	-0.2367	2.3862	0.4191	0.5015	-
133	14	2.65	30	10	0.0191	0.037	0.5166	0.0007	0.025	30.03	0.03	1.738	0.5754	0.6362	-
134	14	1	30	12	0.0146	0.037	0.3947	0.0014	0.0269	30.2111	-0.1678	1.2989	0.7699	0.5548	-
135	14	1	30	16	0.024	0.037	0.6496	0.0028	0.065	30.25	-0.2817	2.1773	0.4593	0.7451	-
136	14	2.65	30	10	0.0241	0.0349	0.689	0.0008	0.0434	30.3063	0.18	2.6711	0.3744	0.7669	6.3
137	14	1	30	12	0.0265	0.0349	0.7576	0.0015	0.0537	30.3341	-0.134	2.4894	0.4017	0.8412	6.3
138	14	2.65	30	8	0.0263	0.0349	0.7518	0.0021	0.0306	30.4375	0.0016	1.1966	0.8357	0.9256	6.3
139	11.48	1	30	12	0.0179	0.037	0.4826	0.0007	0.035	30.4967	-0.4433	2.3447	0.4265	0.5713	-
140	11.75	2.65	30	68	0.0174	0.037	0.4706	0.0066	0.1254	30.5113	-0.0575	2.7463	0.3641	0.5463	6.3
141	11.75	2.65	30	26	0.0197	0.037	0.5334	0.0016	0.0707	30.5525	0.1263	3.1834	0.3141	0.5988	6.3
142	11.75	2.65	30	8	0.0291	0.0349	0.832	0.0014	0.033	30.6024	-0.0361	1.5584	0.6417	0.9655	6.3
143	9.88	1	30	10	0.0233	0.037	0.6296	0.0017	0.0381	30.682	-0.29	1.6616	0.6018	0.7547	6.3
144	9.88	1	30	10	0.0226	0.0349	0.6455	0.0019	0.038	30.7106	-0.3149	1.5441	0.6476	0.7802	6.3
145	11.75	2.65	30	16	0.024	0.0349	0.6879	0.0045	0.0665	30.7416	-0.0534	1.7597	0.5683	0.8061	6.3
146	11.75	2.65	30	24	0.0135	0.037	0.3646	0.0023	0.0403	30.802	-0.024	1.5063	0.6639	0.5027	6.3
147	11.55	2.65	30	16	0.0208	0.037	0.5625	0.0009	0.0502	30.8275	0.2075	2.8896	0.3461	0.6344	6.3
148	11.75	3.5	30	14	0.0264	0.0349	0.7548	0.0055	0.0574	30.8296	-0.1134	1.377	0.7262	0.9059	6.3
149	11.75	2.65	30	30	0.0142	0.037	0.3848	0.002	0.0461	30.8592	-0.0617	1.827	0.5473	0.4986	6.3
150	11.75	2.65	30	12	0.0237	0.0349	0.6789	0.0013	0.0308	30.8763	-0.0156	1.5358	0.6511	0.8144	6.3
151	11.75	2.65	30	50	0.0164	0.037	0.4432	0.0118	0.0707	30.8967	0.0833	1.1562	0.8649	0.6231	6.3
152	11.55	3.5	31	26	0.008	0.037	0.2165	0.0006	0.0255	31.02	-0.254	1.8448	0.5421	0.3293	-
153	7.28	2.65	31	10	0.0183	0.037	0.4938	0.0011	0.0304	31.05	0.2933	1.6636	0.6011	0.6188	-
154	11.55	2.65	31	10	0.0287	0.0349	0.822	0.0019	0.0573	31.0766	0.1006	2.348	0.4259	0.9106	-
155	12.73	1.1	31	22	0.0245	0.0349	0.7003	0.0037	0.0506	31.2147	-0.0115	1.4699	0.6803	0.8418	-
156	12.73	1.1	31	102	0.017	0.037	0.4605	0.0064	0.1209	31.259	0.017	2.6794	0.3732	0.5381	-
157	12.73	1.1	31	12	0.0222	0.0349	0.634	0.0009	0.0465	31.4481	-0.035	2.7393	0.3651	0.7099	-
158	8.88	5.18	31	12	0.0224	0.0349	0.6405	0.0017	0.0546	31.6749	-0.4724	2.3587	0.424	0.7287	-
159	13.48	7.07	31	8	0.0125	0.037	0.338	0.0002	0.0206	31.76	0.145	2.5838	0.387	0.4185	-

160	13.48	6.85	32	8	0.027	0.037	0.7298	0.0005	0.0474	32.062	0.074	3.7599	0.266	0.7851	-
161	13.48	7.07	32	10	0.0295	0.037	0.798	0.002	0.035	32.24	0.094	1.3872	0.7209	0.948	8.2
162	13.48	4.62	32	8	0.0227	0.037	0.6136	0.0009	0.0335	32.4967	0.2567	1.9817	0.5046	0.7185	-
163	13.48	4.62	32	16	0.0214	0.0453	0.4712	0.0032	0.0721	32.69	0.055	2.2594	0.4426	0.5632	-
164	13.48	4.62	32	14	0.0205	0.0349	0.5863	0.0012	0.041	32.8191	0.0427	2.143	0.4666	0.6834	-
165	13.43	4.62	32	18	0.0233	0.0453	0.5135	0.0026	0.0541	32.999	0.016	1.88	0.5319	0.6242	-
166	13.12	3.8	33	12	0.0221	0.0453	0.4883	0.0029	0.04	33.5425	-0.01	1.328	0.753	0.6449	-
167	13.12	2.92	34	10	0.0332	0.0474	0.6991	0.0034	0.0608	34.1725	-0.19	1.849	0.5408	0.8116	-
168	10.15	0.95	34	16	0.0235	0.0474	0.4963	0.0048	0.0716	34.215	-0.765	1.8315	0.546	0.6098	-
169	1.05	0.95	34	32	0.0275	0.0453	0.6071	0.0101	0.111	34.2829	-0.7914	1.9628	0.5095	0.7131	-
170	5.6	0.95	34	16	0.0315	0.0474	0.6648	0.0039	0.0856	34.3183	-0.815	2.4291	0.4117	0.7504	-
171	13.12	0.95	34	8	0.0325	0.0474	0.6859	0.002	0.0427	34.3667	-0.6233	1.6931	0.5906	0.8087	-
172	12.8	0.95	34	8	0.0333	0.0453	0.7345	0.003	0.0474	34.4625	-0.8725	1.535	0.6515	0.87	-
173	13.12	1.9	34	24	0.0231	0.0474	0.4881	0.0069	0.103	34.494	0.03	2.1969	0.4552	0.5828	-
174	13.12	1.9	34	22	0.024	0.0453	0.5286	0.0071	0.103	34.5075	0.02	2.1657	0.4617	0.6246	-
175	12.8	0.95	34	20	0.0289	0.0453	0.6382	0.0064	0.07	34.608	-0.744	1.557	0.6423	0.7718	-
176	12.8	1.9	34	24	0.0191	0.0474	0.4033	0.0035	0.05	34.642	0.008	1.498	0.6676	0.5421	-
177	12.8	1.9	34	52	0.0173	0.0453	0.3813	0.0097	0.103	34.68	0.006	1.8529	0.5397	0.4935	-
178	12.8	0.95	34	12	0.0222	0.0474	0.4686	0.001	0.0453	34.695	-0.6583	2.5378	0.394	0.5505	-
179	12.8	0.95	34	12	0.0222	0.0453	0.49	0.001	0.0453	34.695	-0.6583	2.5378	0.394	0.572	-
180	12.8	0.95	34	18	0.0207	0.0453	0.4576	0.002	0.0453	34.718	-0.85	1.7945	0.5573	0.5735	-
181	12.8	0.95	34	20	0.0266	0.0453	0.586	0.0049	0.0652	34.725	-0.745	1.6507	0.6058	0.712	-
182	12.8	2.45	34	12	0.0187	0.0453	0.4132	0.0012	0.0403	34.7833	-0.7933	2.0626	0.4848	0.5141	-
183	12.8	1.15	34	8	0.0191	0.0453	0.4222	0.0006	0.0354	34.95	-0.284	2.458	0.4068	0.5068	-
184	12.8	1.9	34	10	0.017	0.0453	0.374	0.001	0.0212	34.974	-0.032	1.189	0.841	0.549	11
185	12.8	1.05	35	16	0.0207	0.0453	0.4568	0.0029	0.043	35.0675	-0.47	1.428	0.7003	0.6024	3
186	12.8	1.05	35	10	0.0148	0.0474	0.3127	0.0003	0.0316	35.16	-0.752	3.236	0.309	0.377	-
187	12.8	2.25	35	8	0.0357	0.0453	0.7876	0.0027	0.043	35.44	-0.82	1.4672	0.6816	0.9294	-
188	12.8	1.12	35	32	0.0271	0.0453	0.5974	0.008	0.1101	35.4943	-0.27	2.1889	0.4568	0.6925	11
189	12.8	1.12	35	12	0.0245	0.0474	0.5166	0.0014	0.0539	35.51	-0.27	2.5978	0.3849	0.5967	11
190	12.8	1.12	35	10	0.019	0.0453	0.4187	0.0009	0.0403	35.5775	0.0975	2.3817	0.4199	0.5061	11
191	12.8	1.12	35	12	0.0134	0.0474	0.2816	0.0006	0.0316	35.5933	-0.24	2.2882	0.437	0.3725	11
192	3.75	1.02	35	10	0.0225	0.0474	0.4744	0.0016	0.0403	35.8275	-0.865	1.7862	0.5598	0.5908	-

193	10.88	0.93	36	14	0.0176	0.0474	0.3714	0.0031	0.0391	36.3757	0.8729	1.2432	0.8044	0.5387	-
194	10.88	0.93	36	20	0.0202	0.0564	0.3582	0.0048	0.0626	36.3986	0.8529	1.6028	0.6239	0.4879	-
195	8.38	2.05	36	38	0.03	0.0564	0.5326	0.0078	0.1134	36.4089	-0.0722	2.275	0.4396	0.624	-
196	8.38	2.05	36	18	0.0263	0.0474	0.5538	0.002	0.075	36.4557	-0.1629	2.9725	0.3364	0.6238	-
197	8.38	2.05	36	16	0.0258	0.0564	0.4573	0.0022	0.0636	36.4683	-0.1567	2.4049	0.4158	0.5438	-
198	12.43	1.35	36	14	0.0297	0.0564	0.5277	0.003	0.085	36.725	-0.2483	2.7506	0.3636	0.6033	-
199	12.43	1.35	36	8	0.0334	0.0564	0.5919	0.0006	0.0532	36.85	0.005	3.846	0.26	0.646	-
200	12.43	1.35	36	18	0.0298	0.0564	0.5294	0.0073	0.0783	36.896	-0.082	1.6236	0.6159	0.6576	-
201	10.48	1.05	37	14	0.0177	0.0474	0.3741	0.0014	0.0412	37.676	0.094	1.9531	0.512	0.4806	-
202	10.48	1.05	37	14	0.0177	0.0564	0.3147	0.0014	0.0412	37.676	0.094	1.9531	0.512	0.4212	-
203	10.48	1.05	37	12	0.0208	0.0474	0.4394	0.0018	0.0447	37.695	0.29	1.8683	0.5352	0.5507	-
204	10.48	1.05	37	12	0.0208	0.0564	0.3696	0.0018	0.0447	37.695	0.29	1.8683	0.5352	0.481	-
205	10.48	1.05	37	18	0.0243	0.0564	0.4315	0.0023	0.05	37.7033	-0.0717	1.8683	0.5352	0.5429	10.5
206	10.48	1.02	37	12	0.02	0.0474	0.422	0.0017	0.0474	37.8367	-0.3917	2.0698	0.4831	0.5225	10.5
207	9.93	1.02	38	12	0.0272	0.0564	0.4824	0.0009	0.0381	38.0283	-0.265	2.2498	0.4445	0.5749	-
208	9.93	1.02	38	8	0.0175	0.057	0.307	0.0005	0.0255	38.215	-0.165	2.0209	0.4948	0.4099	3.5
209	12.5	0.93	38	10	0.0358	0.057	0.6282	0.0041	0.0711	38.264	0.83	1.9671	0.5084	0.7339	-
210	9.93	1.35	38	8	0.0197	0.0564	0.3492	0.0006	0.0269	38.58	-0.305	1.9484	0.5133	0.456	-
211	9.93	1.35	38	8	0.0197	0.057	0.3454	0.0006	0.0269	38.58	-0.305	1.9484	0.5133	0.4521	-
212	10.5	1.25	38	28	0.0344	0.0564	0.6103	0.0148	0.1275	38.7114	-0.2529	1.8573	0.5384	0.7223	-
213	10.5	1.25	38	28	0.0344	0.057	0.6036	0.0148	0.1275	38.7114	-0.2529	1.8573	0.5384	0.7156	-
214	10.5	1.35	38	10	0.0386	0.0564	0.6845	0.0054	0.06	38.7471	-0.4814	1.4472	0.691	0.8282	-
215	2.17	0.93	38	40	0.0165	0.0564	0.2918	0.0039	0.0602	38.8125	0.8175	1.7199	0.5814	0.4128	-
216	9.93	1.25	38	8	0.0307	0.0564	0.5442	0.0018	0.045	38.855	-0.16	1.88	0.5319	0.6548	-
217	2.17	0.93	38	12	0.0208	0.0564	0.369	0.0012	0.035	38.9333	0.9467	1.8293	0.5466	0.4827	-
218	9.93	3.53	39	12	0.0337	0.0564	0.5982	0.0034	0.0707	39.26	-0.59	2.1494	0.4652	0.695	-
219	9.93	2.85	39	16	0.0273	0.0564	0.4838	0.0032	0.0602	39.316	-0.29	1.9014	0.5259	0.5932	-
220	9.93	2.85	39	16	0.0273	0.057	0.4785	0.0032	0.0602	39.316	-0.29	1.9014	0.5259	0.5879	-
221	9.93	2.85	39	8	0.0339	0.057	0.5956	0.003	0.0361	39.5643	-0.2171	1.1766	0.8499	0.7723	-
222	9.1	2.85	39	10	0.0347	0.057	0.6082	0.0022	0.0472	39.898	-0.186	1.7825	0.561	0.7249	-
223	9.1	1.05	40	30	0.0185	0.057	0.3238	0.005	0.0743	40.0867	-0.5433	1.8726	0.534	0.4349	-
224	9.1	1.05	40	12	0.0212	0.0595	0.3559	0.0021	0.0391	40.115	-0.56	1.5104	0.6621	0.4936	-
225	-	-	40	14	0.0275	0.0595	0.4615	0.0033	0.0632	40.595	-0.6983	1.9664	0.5086	0.5673	-

226	-	-	40	16	0.0292	0.057	0.5121	0.0029	0.082	40.636	-0.676	2.6991	0.3705	0.5892	-
227	10.23	6.38	40	10	0.0414	0.0595	0.6957	0.0037	0.0652	40.7317	-0.6683	1.8996	0.5264	0.8052	-
228	11.65	1.65	41	40	0.0284	0.057	0.4988	0.031	0.1166	41.0025	-0.1975	1.174	0.8518	0.676	-
229	11.65	4.72	41	16	0.0236	0.0595	0.3972	0.0047	0.0667	41.015	-0.1425	1.7247	0.5798	0.5178	-
230	10.23	4.72	41	26	0.029	0.0595	0.4874	0.0081	0.0875	41.025	-0.245	1.7225	0.5805	0.6081	-
231	11.7	4.72	41	10	0.0365	0.0595	0.6136	0.0013	0.0453	41.19	-0.1929	2.2258	0.4493	0.707	-
232	11.65	2.33	41	8	0.03	0.057	0.5257	0.0014	0.0495	41.2667	0.3767	2.3447	0.4265	0.6144	-
233	11.7	4.72	41	14	0.015	0.0595	0.2519	0.002	0.032	41.324	-0.17	1.285	0.7782	0.4138	-
234	11.65	4.72	41	16	0.0288	0.057	0.5046	0.0019	0.0602	41.3417	0.075	2.4482	0.4085	0.5896	-
235	11.65	4.1	41	12	0.0344	0.057	0.6034	0.0053	0.0673	41.73	0.0525	1.6455	0.6077	0.7298	-
236	11.65	1.17	41	16	0.0313	0.0595	0.5265	0.004	0.064	41.7917	0.0617	1.8058	0.5538	0.6416	-
237	8.6	0.93	42	20	0.031	0.0581	0.5345	0.0119	0.1026	42.446	-0.44	1.6704	0.5987	0.659	-
238	6.28	0.93	42	14	0.0325	0.0595	0.5453	0.0057	0.082	42.4633	-0.48	1.9337	0.5171	0.6529	-
239	11.43	2.8	42	12	0.034	0.0595	0.571	0.0024	0.0894	42.854	-0.126	3.236	0.309	0.6352	-
240	11.43	2.8	42	16	0.0269	0.0581	0.464	0.0021	0.0652	42.9133	-0.1283	2.5215	0.3966	0.5465	-
241	11.43	2.8	43	14	0.0109	0.0581	0.1874	0.003	0.0212	43.1825	0.0025	0.6865	1.4567	0.4904	-
242	11.43	2.8	43	18	0.0161	0.0595	0.2698	0.0015	0.0461	43.208	-0.018	2.1457	0.466	0.3667	-
243	11.43	2.78	43	16	0.0292	0.0595	0.4906	0.0063	0.061	43.276	-0.392	1.3575	0.7366	0.6438	-
244	8.02	2.83	44	14	0.0183	0.0592	0.3085	0.0026	0.0354	44.488	-0.138	1.229	0.8137	0.4777	-
245	7.18	3.62	45	12	0.0259	0.0592	0.4376	0.0041	0.064	45.406	-0.74	1.7834	0.5607	0.5542	-
246	7.18	3.62	45	12	0.0259	0.0581	0.4462	0.0041	0.064	45.406	-0.74	1.7834	0.5607	0.5628	-
247	11.12	1.88	45	14	0.0282	0.0592	0.4764	0.0026	0.0757	45.4225	0.055	2.6558	0.3765	0.5547	-
248	11.12	1.8	45	24	0.016	0.0581	0.2758	0.004	0.0541	45.7067	-0.245	1.5157	0.6598	0.413	-
249	11.12	1.8	45	26	0.031	0.0592	0.5227	0.0126	0.0986	45.802	-0.314	1.5572	0.6422	0.6563	-
250	7.28	1.8	45	10	0.0145	0.0581	0.2495	0.0005	0.025	45.9267	-0.57	1.9817	0.5046	0.3544	-
251	6.72	1.8	46	8	0.0276	0.0545	0.5068	0.0012	0.0364	46.2175	-0.67	1.8625	0.5369	0.6185	-
252	5.6	1.8	46	14	0.0229	0.0545	0.4197	0.002	0.0354	46.2925	-0.2275	1.4191	0.7047	0.5662	-
253	5.6	1.8	46	24	0.0237	0.0592	0.4007	0.0032	0.0814	46.3122	-0.2256	2.5503	0.3921	0.4823	-
254	5.6	1.8	46	10	0.0249	0.0545	0.4579	0.002	0.05	46.3675	-0.22	1.9817	0.5046	0.5629	-
255	6.95	0.65	47	10	0.0165	0.0592	0.2784	0.0008	0.0381	47.05	0.25	2.3862	0.4191	0.3656	-
256	5.75	0.75	47	12	0.0258	0.0592	0.436	0.0019	0.0403	47.356	-0.892	1.6612	0.602	0.5612	-
257	5.75	0.75	47	52	0.0275	0.0545	0.5051	0.0129	0.123	47.4642	-0.8442	1.923	0.52	0.6132	-
258	5.75	0.73	47	8	0.0341	0.0592	0.5761	0.0021	0.0461	47.764	-0.66	1.8046	0.5541	0.6913	-

259	10.38	3.22	48	18	0.0271	0.0545	0.4983	0.005	0.0901	48.65	0.0913	2.2708	0.4404	0.5899	-
260	10.38	3.22	48	24	0.0206	0.0545	0.3777	0.0073	0.0863	48.6575	0.2225	1.7906	0.5585	0.4938	-
261	10.38	3.22	48	26	0.0307	0.0545	0.5641	0.0201	0.0943	48.782	0.072	1.1794	0.8479	0.7405	-
262	10.38	3.22	48	10	0.0237	0.0545	0.4347	0.0015	0.0515	48.7829	0.2929	2.3559	0.4245	0.523	-
263	10.38	3.22	48	10	0.0224	0.0545	0.4115	0.0015	0.0427	48.8525	0.2525	1.9885	0.5029	0.5161	-
264	6.55	3.67	49	18	0.0216	0.0545	0.3961	0.0022	0.0738	49.006	-0.306	2.7897	0.3585	0.4706	5.6
265	6.82	3.62	49	12	0.0325	0.0545	0.5973	0.0038	0.0757	49.385	-0.2425	2.1756	0.4597	0.6929	5.7
266	6.82	5.47	49	10	0.0258	0.0545	0.4744	0.0027	0.0472	49.415	-0.3375	1.609	0.6215	0.6037	5.7
267	6.82	5.47	49	34	0.0138	0.0545	0.2525	0.0051	0.0626	49.4933	-0.3867	1.5549	0.6431	0.3863	5.7

APPENDIX C

IRDC Convex Hull data

IRDC Convex hull statistics. Col 1: Hull ID. Col 2: Minimum distance in Kpc from the Roman Duval Cloud association. Col 3: Maximum distance in Kpc from the RomanDuval Cloud association. Col 4: Cell in which the hull is identified. Col 5: Number of Nodes within hull. Col 6: Intercluster mean edge cost. Col 7: Field mean edge cost. Col 8: The Cartwright Clustering measure Q. Col 9: Area of hull in square degrees. Col:10 Semi Major Axis of hull in degrees. Col 11 & 12: Position of centre of hull in galactic coordinates. Col 13: Billot Elongation measure. Col 14 Schmeja elongation measure. Col 15: Q^*. Col 16:Distance to Associated Murray SFR in kpc

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
1	20	11.38	3.33	20	0.0276	0.0551	0.5001	0.0057	0.0878	20.1188	-0.797	2.0608	0.4852	0.601	-
2	20	14.02	0.82	16	0.0234	0.0551	0.4242	0.0026	0.0443	20.7414	-0.0608	1.5479	0.646	0.5586	-
3	22	11.07	1.45	14	0.0184	0.0551	0.3339	0.0008	0.0411	22.0493	0.2225	2.5971	0.3851	0.414	-
4	22	11.2	4.03	10	0.0182	0.047	0.3874	0.0008	0.0299	22.1644	-0.6583	1.8465	0.5416	0.5	-
5	22	14.27	0.93	10	0.0274	0.0551	0.4967	0.0006	0.061	22.4054	0.3222	4.5244	0.221	0.5427	-
6	22	11.2	11.05	10	0.0301	0.0551	0.546	0.0004	0.0397	22.4385	-0.7326	3.6697	0.2725	0.6027	-
7	22	15.07	4.47	8	0.0298	0.0551	0.54	0.0013	0.0463	22.4879	-0.2175	2.2956	0.4356	0.6306	-
8	22	6.88	3.88	8	0.0248	0.0551	0.4492	0.0011	0.0368	22.7043	-0.448	1.9697	0.5077	0.5548	-
9	22	6.88	3.88	12	0.0322	0.0551	0.5851	0.0042	0.0461	22.8753	-0.4309	1.2667	0.7894	0.7493	10.8
10	22	13.93	3.88	30	0.0351	0.0551	0.6376	0.0055	0.1437	22.9636	-0.1985	3.4465	0.2902	0.698	10.8
11	23	6.88	3.88	22	0.0293	0.0551	0.5309	0.0083	0.0856	23.0075	-0.4232	1.6692	0.5991	0.6555	10.8
12	23	13.93	3.88	10	0.0233	0.047	0.4946	0.0008	0.0462	23.1939	0.0058	2.9731	0.3363	0.5646	-
13	23	6.72	3.88	10	0.0295	0.047	0.6284	0.0022	0.046	23.2072	-0.3924	1.7332	0.577	0.7484	10.8
14	23	11.77	3.88	10	0.0337	0.047	0.7175	0.0023	0.0391	23.2435	-0.2779	1.4425	0.6932	0.8617	10.8
15	23	11.77	3.88	14	0.0347	0.0551	0.6292	0.0048	0.0655	23.2567	-0.253	1.6815	0.5947	0.7529	-
16	23	13.93	3.88	28	0.0317	0.047	0.6733	0.004	0.126	23.2934	-0.1808	3.5338	0.283	0.7321	-
17	23	11.77	3.88	12	0.0246	0.047	0.5224	0.0025	0.0578	23.3877	-0.2995	2.0292	0.4928	0.6249	-
18	23	11.77	3.88	36	0.0272	0.0551	0.4932	0.0072	0.1323	23.4107	-0.2648	2.7565	0.3628	0.5687	-
19	23	13.93	3.88	12	0.0341	0.047	0.7247	0.0041	0.0459	23.4516	-0.0791	1.2663	0.7897	0.889	-
20	23	11.77	3.88	12	0.0242	0.047	0.515	0.0017	0.0404	23.4746	-0.262	1.7458	0.5728	0.6341	-
21	23	13.93	0.88	12	0.0317	0.047	0.6738	0.0032	0.0713	23.4846	0.0625	2.2159	0.4513	0.7676	-
22	23	13.93	0.88	14	0.0346	0.0551	0.6283	0.0055	0.0713	23.4949	0.0507	1.7047	0.5866	0.7503	-
23	23	13.93	0.88	10	0.0264	0.047	0.562	0.0014	0.0344	23.5818	0.1532	1.6351	0.6116	0.6893	-
24	23	11.7	0.88	12	0.029	0.0551	0.5255	0.0019	0.0537	23.8466	0.0453	2.1805	0.4586	0.6209	5.8
25	24	12.77	0.88	12	0.0241	0.047	0.5125	0.0012	0.0601	24.1618	0.0439	3.0976	0.3228	0.5796	-
26	24	12.77	2.95	10	0.0268	0.047	0.5707	0.0024	0.053	24.1738	-0.0697	1.9069	0.5244	0.6798	10.3

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
27	24	12.77	0.88	8	0.0378	0.0506	0.7476	0.0029	0.0504	24.3271	0.1131	1.6678	0.5996	0.8723	-
28	24	12.77	0.88	12	0.033	0.047	0.7016	0.0043	0.0637	24.4546	0.0571	1.7151	0.583	0.8229	-
29	24	11.25	2.97	20	0.0325	0.047	0.691	0.0075	0.0773	24.5371	-0.2082	1.5808	0.6326	0.8226	6.1
30	24	11.25	3.2	18	0.0273	0.047	0.5808	0.0029	0.0627	24.594	-0.3141	2.0459	0.4888	0.6825	-
31	24	12.77	0.88	30	0.0271	0.047	0.5762	0.0037	0.124	24.6411	0.1507	3.5951	0.2782	0.634	6.1
32	24	12.77	0.88	36	0.0302	0.0506	0.5962	0.0118	0.1423	24.6648	0.1491	2.3179	0.4314	0.686	6.1
33	24	11.25	2.97	12	0.0228	0.047	0.4849	0.0011	0.0319	24.6792	-0.1427	1.6955	0.5898	0.6075	6.1
34	24	12.77	2.97	24	0.0236	0.047	0.5011	0.0036	0.0703	24.8163	-0.1011	2.0619	0.485	0.602	6.1
35	24	11.25	2.97	8	0.0325	0.047	0.6903	0.0026	0.039	24.8454	-0.2074	1.3508	0.7403	0.8443	6.1
36	25	9.5	2.62	12	0.0327	0.0506	0.647	0.0044	0.0565	25.2069	0.4874	1.5132	0.6609	0.7845	6.1
37	25	7.68	0.62	32	0.0277	0.0506	0.5483	0.0194	0.1369	25.362	-0.3556	1.7432	0.5737	0.6676	4.1
38	25	11.5	0.62	10	0.0425	0.0506	0.8389	0.0013	0.043	25.3752	-0.1841	2.103	0.4755	0.9378	4.1
39	25	7.68	0.62	12	0.0194	0.047	0.4131	0.002	0.0398	25.3918	-0.3154	1.586	0.6305	0.5443	4.1
40	26	10.85	3.8	16	0.0185	0.0506	0.3659	0.0015	0.0412	26.2957	-0.0505	1.9031	0.5254	0.4752	-
41	26	8.02	0.8	10	0.0325	0.0506	0.6419	0.0023	0.0721	26.5113	-0.4006	2.6734	0.3741	0.7197	-
42	26	13.23	1.9	18	0.0277	0.0506	0.5477	0.0035	0.0831	26.6563	-0.0281	2.473	0.4044	0.6318	-
43	26	13.23	1.9	10	0.0269	0.0506	0.5319	0.0023	0.0456	26.7694	0.2238	1.6768	0.5964	0.6559	-
44	26	10.65	0.8	10	0.0192	0.0506	0.3786	0.0013	0.0394	26.9103	-0.3062	1.9153	0.5221	0.4872	-
45	27	13.23	0.73	10	0.0194	0.0506	0.3842	0.0019	0.0352	27.7438	0.0601	1.4423	0.6933	0.5284	-
46	27	13.23	0.73	8	0.0328	0.0506	0.6486	0.0021	0.0471	27.7899	0.1657	1.8093	0.5527	0.7636	-
47	27	12.7	0.73	14	0.0274	0.0506	0.5415	0.0031	0.0637	27.9116	0.0057	2.0241	0.4941	0.6442	-
48	27	10.07	0.62	10	0.0388	0.0506	0.7661	0.005	0.0476	27.9341	-0.4251	1.1899	0.8404	0.9409	-
49	28	10.07	0.62	8	0.0262	0.0486	0.5386	0.0016	0.0363	28.0692	-0.4718	1.6123	0.6202	0.6676	-
50	28	12.52	0.73	14	0.0198	0.0486	0.4079	0.0023	0.0338	28.226	0.0229	1.2596	0.7939	0.573	-
51	28	12.52	0.62	10	0.0284	0.0486	0.5842	0.0026	0.0335	28.262	-0.1818	1.1556	0.8654	0.7642	-
52	28	12.52	3.03	12	0.0327	0.0486	0.6717	0.0019	0.0485	28.3201	0.035	1.969	0.5079	0.7774	-
53	28	12.52	3.03	36	0.0189	0.0486	0.3893	0.0058	0.088	28.3785	0.1175	2.048	0.4883	0.4908	-
54	28	12.52	3.03	8	0.0362	0.0486	0.7443	0.0024	0.0528	28.4137	-0.1086	1.9175	0.5215	0.8528	-
55	28	12.52	3.03	12	0.0226	0.0495	0.4562	0.0025	0.0335	28.6481	-0.0546	1.1896	0.8406	0.6311	-
56	28	12.52	3.47	22	0.0292	0.0486	0.5999	0.002	0.0741	28.7497	0.1371	2.9392	0.3402	0.6706	-
57	28	7.43	3.47	32	0.0227	0.0486	0.4673	0.0041	0.1094	28.8536	0.0575	3.0374	0.3292	0.5357	-
58	29	5.1	0.85	18	0.0233	0.0486	0.4792	0.0025	0.0531	29.0117	-0.6359	1.8759	0.5331	0.5901	-
59	29	5.1	0.85	8	0.0304	0.0486	0.625	0.0012	0.0478	29.2671	-0.6294	2.4685	0.4051	0.7093	-

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
60	29	14	2.65	10	0.0219	0.0486	0.4504	0.0006	0.0348	29.5188	0.18	2.5045	0.3993	0.5334	-
61	29	14	2.65	10	0.0175	0.0486	0.3608	0.0009	0.0307	29.5587	0.1916	1.8512	0.5402	0.4731	-
62	29	14	2.65	10	0.0246	0.0486	0.5051	0.0019	0.0349	29.6019	0.2365	1.4235	0.7025	0.6512	-
63	29	14	2.65	16	0.0333	0.0495	0.6739	0.0034	0.0629	29.9012	-0.04	1.9271	0.5189	0.7819	8.7
64	30	14	1	12	0.0217	0.0486	0.447	0.0016	0.0497	30.2385	-0.2178	2.2322	0.448	0.5402	-
65	30	14	1	32	0.0267	0.0595	0.4494	0.0147	0.1334	30.2785	-0.2446	1.948	0.5133	0.5562	6.3
66	30	14	1	14	0.0243	0.0486	0.4993	0.002	0.0437	30.2912	-0.271	1.7515	0.5709	0.618	6.3
67	30	11.75	2.65	30	0.023	0.0486	0.4728	0.0092	0.092	30.4935	-0.0758	1.7043	0.5867	0.5948	6.3
68	30	11.75	2.65	34	0.0235	0.0595	0.3943	0.0097	0.1117	30.5041	-0.067	2.0114	0.4972	0.4977	6.3
69	30	11.48	1	12	0.0229	0.0486	0.4709	0.0021	0.0484	30.5492	-0.24	1.868	0.5353	0.5823	6.3
70	30	11.48	1	20	0.0272	0.0595	0.4577	0.0102	0.0863	30.5613	-0.2482	1.5165	0.6594	0.5949	6.3
71	30	11.75	2.65	32	0.028	0.0486	0.577	0.0044	0.1175	30.7871	-0.047	3.1537	0.3171	0.643	6.3
72	30	11.75	2.65	50	0.0289	0.0595	0.4857	0.0085	0.1354	30.8232	-0.0447	2.6098	0.3832	0.5654	6.3
73	30	11.75	2.65	26	0.0285	0.0486	0.5873	0.0084	0.0741	30.8713	0.0923	1.4349	0.6969	0.7323	6.3
74	30	11.55	2.65	42	0.0327	0.0595	0.5505	0.0181	0.1588	30.9851	0.0347	2.0905	0.4784	0.65	-
75	30	7.28	2.65	8	0.0307	0.0486	0.6307	0.0018	0.042	30.9963	0.3499	1.7582	0.5688	0.749	-
76	31	11.55	2.65	12	0.0347	0.0486	0.7138	0.0028	0.0477	31.0567	-0.0068	1.5956	0.6267	0.8441	-
77	31	12.73	6.55	10	0.0231	0.0595	0.3883	0.0013	0.0512	31.6156	0.1723	2.5106	0.3983	0.4712	-
78	31	8.88	4.62	10	0.0166	0.0486	0.3411	0.0003	0.0302	31.7036	-0.4859	3.0174	0.3314	0.41	-
79	32	13.48	4.62	24	0.0317	0.0647	0.4898	0.0051	0.0492	32.0248	-0.3413	1.2168	0.8218	0.6607	-
80	32	13.48	6.85	16	0.0528	0.0647	0.8162	0.0117	0.0578	32.0376	0.0614	0.947	1.056	1.0359	-
81	32	13.48	7.07	16	0.0381	0.0595	0.6411	0.0017	0.0271	32.1572	0.1488	1.1637	0.8593	0.8199	-
82	32	13.48	0.8	16	0.0218	0.0595	0.3673	0.0019	0.0251	32.5702	0.3106	1.0197	0.9807	0.5713	-
83	32	13.48	5.03	12	0.0113	0.0595	0.1894	0.0001	0.0121	32.6245	-0.2446	1.936	0.5165	0.2969	-
84	32	13.48	4.62	24	0.0366	0.0595	0.616	0.0029	0.0508	32.7384	0.0848	1.6627	0.6014	0.7411	-
85	33	13.43	3.8	14	0.039	0.0595	0.6562	0.0043	0.0732	33.1852	-0.046	1.9719	0.5071	0.7616	-
86	33	13.43	3.8	12	0.0365	0.0647	0.5634	0.0054	0.0554	33.195	-0.0283	1.3348	0.7491	0.7192	-
87	33	13.43	4.62	16	0.0164	0.0647	0.2529	0.0004	0.031	33.37	0.1789	2.6659	0.3751	0.331	-
88	33	13.43	4.62	10	0.0145	0.0595	0.2441	0.0002	0.0242	33.3797	0.1875	2.9892	0.3345	0.3137	-
89	33	13.43	3.8	12	0.0195	0.0595	0.328	0.001	0.0391	33.4319	-0.0733	2.2317	0.4481	0.4212	-
90	33	10.15	4.62	8	0.0245	0.0647	0.3782	0.001	0.0394	33.7273	0.2129	2.1651	0.4619	0.4742	-
91	34	12.8	2.45	24	0.0237	0.0684	0.3469	0.0052	0.1043	34.7701	-0.7204	2.5644	0.39	0.428	-
92	34	12.8	2.85	8	0.0274	0.0647	0.4228	0.0013	0.0455	34.9768	0.3581	2.2296	0.4485	0.5161	-

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
93	35	12.8	1.05	22	0.0232	0.0684	0.339	0.0036	0.0644	35.0913	-0.7399	1.9115	0.5231	0.4478	-
94	35	12.8	1.05	18	0.0215	0.0647	0.3325	0.0015	0.0512	35.1077	-0.7303	2.3262	0.4299	0.4219	-
95	35	12.8	1.05	8	0.0424	0.0684	0.6193	0.0021	0.0755	35.3063	-0.4862	2.9453	0.3395	0.69	-
96	35	12.8	1.05	32	0.0335	0.0647	0.5173	0.0277	0.1314	35.4211	-0.3393	1.3999	0.7144	0.6658	11
97	35	12.8	1.05	44	0.0356	0.0684	0.5205	0.0386	0.1767	35.4586	-0.3103	1.5942	0.6273	0.651	11
98	35	12.8	1.12	10	0.0369	0.0647	0.5696	0.0015	0.0741	35.5845	0.1129	3.3733	0.2964	0.6312	11
99	35	3.75	1.02	8	0.0525	0.0684	0.7667	0.0023	0.0607	35.6971	-0.8901	2.2304	0.4483	0.86	-
100	35	5.95	1.92	10	0.0318	0.0647	0.4915	0.0024	0.0604	35.8212	-0.177	2.1732	0.4602	0.5872	-
101	36	8.18	0.93	8	0.0529	0.094	0.5623	0.0047	0.0718	36.0747	0.6323	1.8534	0.5395	0.6745	-
102	36	12.43	1.35	48	0.0304	0.0684	0.4443	0.0072	0.2148	36.6162	-0.1265	4.4793	0.2233	0.4907	-
103	37	11.05	1.35	8	0.0454	0.0684	0.6639	0.0036	0.0639	37.4858	0.1615	1.9002	0.5263	0.7734	-
104	37	10.48	1.05	10	0.0459	0.0684	0.6715	0.0037	0.059	37.679	-0.0262	1.7205	0.5812	0.7924	10.5
105	38	5.25	1.25	10	0.0369	0.094	0.3924	0.0018	0.0818	38.3453	-0.8536	3.3792	0.2959	0.454	-
106	38	1.25	1.25	10	0.0171	0.094	0.1815	0.001	0.0234	38.3573	-0.9425	1.2992	0.7697	0.3416	-
107	38	2.17	0.93	8	0.0274	0.094	0.2916	0.0024	0.0361	38.7695	0.7983	1.2963	0.7715	0.452	-
108	38	10.5	1.35	16	0.0268	0.094	0.2845	0.0028	0.0798	38.8374	-0.5008	2.6812	0.373	0.3621	-
109	38	10.5	1.35	10	0.0219	0.0872	0.2514	0.0006	0.0469	38.8478	-0.5374	3.3793	0.2959	0.313	-
110	39	9.93	2.85	14	0.0522	0.094	0.555	0.0182	0.1046	39.362	-0.2478	1.3738	0.7279	0.7064	-
111	39	6.57	3.53	16	0.0257	0.094	0.2732	0.0011	0.0617	39.4186	-0.9585	3.3079	0.3023	0.3361	-
112	40	11.65	2.33	8	0.0134	0.0913	0.1469	0.0003	0.0209	40.9015	0.8054	2.1442	0.4664	0.2439	-
113	41	10.23	4.72	12	0.0327	0.0913	0.3577	0.0025	0.0472	41.0429	-0.2529	1.6581	0.6031	0.4831	-
114	41	11.7	4.72	26	0.0415	0.0872	0.4756	0.0113	0.1153	41.1194	-0.2108	1.9203	0.5208	0.5839	-
115	41	11.7	4.72	20	0.0497	0.0913	0.5446	0.0106	0.1219	41.1824	-0.2025	2.1004	0.4761	0.6437	-
116	42	6.28	0.93	16	0.0384	0.0913	0.4202	0.0038	0.0649	42.0758	-0.6309	1.8681	0.5353	0.5316	-
117	42	8.6	0.93	8	0.0717	0.1097	0.6534	0.0066	0.1061	42.393	-0.3005	2.3065	0.4336	0.7436	-
118	42	11.43	0.93	10	0.0517	0.1097	0.4714	0.0058	0.0855	42.8335	-0.1536	1.9881	0.503	0.576	-
119	43	11.43	2.8	16	0.038	0.1097	0.3465	0.005	0.0534	43.1577	-0.0314	1.3357	0.7487	0.5022	-
120	44	8.02	2.83	12	0.0535	0.1097	0.4879	0.0161	0.1127	44.2677	0.0911	1.5763	0.6344	0.6198	-
121	45	11.12	1.88	12	0.0495	0.1097	0.4512	0.009	0.0705	45.4642	0.0559	1.3134	0.7614	0.6096	-
122	45	11.12	1.8	14	0.0608	0.1167	0.5208	0.0155	0.1261	45.8712	-0.3812	1.798	0.5562	0.6364	-
123	46	5.6	1.8	24	0.029	0.061	0.4756	0.0119	0.0984	46.3806	-0.2064	1.5952	0.6269	0.606	-
124	46	5.6	0.8	14	0.0273	0.1167	0.2338	0.0062	0.063	46.4563	-0.2058	1.4199	0.7043	0.3803	-
125	48	10.38	4	24	0.0428	0.061	0.7015	0.0108	0.1624	48.6018	0.095	2.7675	0.3613	0.7767	-

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
126	48	8.77	5.55	20	0.0374	0.061	0.6132	0.0042	0.1269	48.783	-0.4276	3.4714	0.2881	0.6731	-
127	48	8.77	5.55	10	0.0323	0.061	0.5284	0.0015	0.0563	48.8664	-0.4719	2.5523	0.3918	0.6098	-
128	49	6.55	3.67	16	0.026	0.061	0.4258	0.0023	0.0658	49.0105	-0.6012	2.4191	0.4134	0.5118	-
129	49	9.9	3.22	54	0.0249	0.061	0.4078	0.009	0.1274	49.1048	-0.259	2.3859	0.4191	0.495	5.6
130	49	5.57	5.5	14	0.0345	0.061	0.5645	0.0038	0.0423	49.1896	-0.3243	1.2203	0.8195	0.735	5.6
131	49	6.82	5.47	12	0.0391	0.061	0.6412	0.004	0.0614	49.4104	-0.5198	1.7161	0.5827	0.7624	5.7
132	49	6.82	5.47	10	0.0259	0.061	0.4239	0.0017	0.0491	49.4424	-0.3736	2.0926	0.4779	0.5233	5.7
133	49	6.82	3.62	12	0.0395	0.061	0.6471	0.0064	0.0766	49.5241	-0.2083	1.6912	0.5913	0.7701	5.7
134	49	5.53	5.47	18	0.0309	0.061	0.5057	0.0042	0.0564	49.593	-0.5056	1.5416	0.6487	0.6406	5.7
135	49	5.53	3.95	20	0.0413	0.061	0.6763	0.0102	0.1135	49.6352	-0.2814	1.9909	0.5023	0.7808	5.7

APPENDIX D

BGPS Convex Hull data

IRDC Convex hull statistics. Col 1: Hull ID. Col 2: Minimum distance in Kpc from the Roman Duval Cloud association. Col 3: Maximum distance in Kpc from the RomanDuval Cloud association. Col 4: Cell in which the hull is identified. Col 5: Number of Nodes within hull. Col 6: Intercluster mean edge cost. Col 7: Field mean edge cost. Col 8: The Cartwright Clustering measure Q. Col 9: Area of hull in square degrees. Col:10 Semi Major Axis of hull in degrees. Col 11 & 12: Position of centre of hull in galactic coordinates. Col 13: Billot Elongation measure. Col 14 Schmeja elongation measure. Col 15: Q^*. Col 16:Distance to Associated Murray SFR in kpc

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
14.02	0.82	26	1	20	0.029	0.0455	0.6383	0.0061	0.0801	20.7582	-0.0757	1.8222	0.5488	0.7524	-
12.93	2.45	10	2	21	0.0213	0.0455	0.4696	0.0006	0.0512	21.197	-0.1355	3.6128	0.2768	0.5272	-
12.93	3.85	12	3	21	0.0318	0.0455	0.6994	0.0042	0.0585	21.3127	-0.1508	1.5994	0.6252	0.8295	-
15.07	4.3	10	4	22	0.0307	0.0393	0.7814	0.0021	0.051	22.5076	-0.0201	1.9951	0.5012	0.8857	-
15.07	3.88	24	5	22	0.0246	0.0393	0.6263	0.0052	0.0882	22.5248	-0.2013	2.1657	0.4617	0.7223	-
15.07	0.6	10	6	22	0.0205	0.0393	0.5213	0.0011	0.0484	22.777	-0.0224	2.5734	0.3886	0.6021	-
15.07	3.88	10	7	22	0.0246	0.0393	0.6256	0.0008	0.06	22.8366	-0.2596	3.8698	0.2584	0.6793	10.8
13.93	0.6	10	8	22	0.0275	0.0393	0.7005	0.0016	0.0453	22.9626	0.0097	2.0003	0.4999	0.8045	-
13.93	0.6	12	9	22	0.0297	0.0455	0.654	0.004	0.0496	22.9726	0.0117	1.3977	0.7155	0.8028	-
13.93	0.6	8	10	23	0.0244	0.0455	0.5372	0.0006	0.0455	23.0385	0.1017	3.2251	0.3101	0.6017	-
15.15	0.9	8	11	23	0.0256	0.0455	0.5643	0.0012	0.0403	23.0629	0.2732	2.1007	0.476	0.6633	-
13.93	0.6	14	12	23	0.0306	0.0455	0.6734	0.0024	0.0757	23.1035	0.1391	2.7341	0.3658	0.7494	-
13.93	3.88	14	13	23	0.0246	0.0455	0.5415	0.0047	0.057	23.3085	0.0691	1.4742	0.6783	0.6826	-
13.93	3.88	10	14	23	0.0211	0.0393	0.537	0.0004	0.043	23.3291	0.0639	3.9849	0.2509	0.5892	-
11.57	0.65	18	15	23	0.0258	0.0455	0.5677	0.0063	0.0787	23.4668	-0.4775	1.7596	0.5683	0.6859	-
13.93	0.88	8	16	23	0.0268	0.0393	0.6832	0.0019	0.044	23.5259	-0.0391	1.7958	0.5569	0.799	-
13.93	2.95	42	17	23	0.023	0.0393	0.5845	0.0105	0.1261	23.5584	-0.1392	2.1793	0.4589	0.6799	-
13.93	0.88	10	18	23	0.0205	0.0393	0.5211	0.0013	0.0265	23.5692	0.1364	1.3263	0.754	0.6779	-
13.93	0.88	12	19	23	0.0269	0.0455	0.5919	0.0028	0.0461	23.7586	-0.0069	1.5496	0.6453	0.7261	10.3
10.3	0.88	8	20	23	0.025	0.0393	0.6371	0.0017	0.0395	23.8673	0.5272	1.7003	0.5881	0.7594	-
11.57	3.78	8	21	23	0.0227	0.0393	0.5773	0.0013	0.0222	23.9205	-0.2664	1.094	0.914	0.7675	10.3
12.77	0.88	46	22	24	0.0254	0.0393	0.6458	0.0121	0.1304	24.1585	-0.0224	2.1038	0.4753	0.7447	10.3
12.77	0.88	8	23	24	0.027	0.0393	0.6885	0.0009	0.0316	24.3121	0.1124	1.8974	0.527	0.7981	-
12.77	0.88	10	24	24	0.023	0.0418	0.549	0.001	0.0446	24.393	0.4894	2.5251	0.396	0.6314	-
12.77	0.88	8	25	24	0.0216	0.0393	0.5492	0.0009	0.0314	24.4214	0.3471	1.8172	0.5503	0.6637	-
12.77	0.88	14	26	24	0.0268	0.0393	0.6826	0.0026	0.0617	24.5214	0.3235	2.1255	0.4705	0.7804	6.1

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
11.25	2.97	12	27	24	0.0291	0.0418	0.6965	0.0021	0.0569	24.5543	-0.2472	2.1836	0.458	0.7917	6.1
11.25	2.97	10	28	24	0.0269	0.0393	0.6845	0.0023	0.0432	24.5633	-0.2552	1.6045	0.6233	0.8141	6.1
12.77	0.88	10	29	24	0.0277	0.0393	0.7061	0.0017	0.0539	24.6149	0.42	2.3427	0.4269	0.7949	6.1
11.25	2.97	12	30	24	0.0299	0.0393	0.7618	0.0034	0.0658	24.6784	-0.1363	1.9905	0.5024	0.8663	6.1
12.77	2.97	8	31	24	0.0262	0.0393	0.6667	0.0007	0.0474	24.7534	0.191	3.2826	0.3046	0.73	6.1
9.5	0.62	22	32	25	0.0255	0.0418	0.6102	0.0083	0.086	25.1678	0.1555	1.676	0.5966	0.7343	6.1
7.68	0.62	8	33	25	0.0249	0.0393	0.6341	0.0015	0.0286	25.3541	-0.3719	1.321	0.757	0.7916	4.1
11.5	0.62	10	34	25	0.0302	0.0393	0.7681	0.0039	0.0414	25.4045	0.0532	1.1815	0.8464	0.9441	-
11.5	0.62	14	35	25	0.024	0.0393	0.6106	0.004	0.0514	25.6949	-0.1684	1.4405	0.6942	0.755	-
11.5	4.62	8	36	25	0.0276	0.0418	0.6584	0.0016	0.0405	25.7497	0.2508	1.8023	0.5548	0.7738	-
11.5	1.1	10	37	25	0.0256	0.0418	0.6112	0.0013	0.0396	25.8661	-0.1735	1.9629	0.5095	0.7172	-
9	1.1	10	38	26	0.0293	0.0418	0.7006	0.0032	0.0458	26.1119	-0.0592	1.4258	0.7013	0.8465	8.8
10.85	1.1	8	39	26	0.027	0.0437	0.6187	0.0018	0.0349	26.1249	-0.0687	1.4765	0.6773	0.7596	-
10.85	2.3	10	40	26	0.0268	0.0418	0.6398	0.0013	0.0474	26.1814	0.0413	2.3031	0.4342	0.7301	-
10.85	3.8	12	41	26	0.0297	0.0418	0.71	0.0024	0.0385	26.2974	-0.0558	1.3803	0.7245	0.8607	-
8.02	0.8	8	42	26	0.0223	0.0418	0.5319	0.0008	0.0267	26.4908	-0.3111	1.6318	0.6128	0.6594	-
10.85	1.9	10	43	26	0.025	0.0418	0.5974	0.002	0.0481	26.5644	0.0984	1.8999	0.5264	0.7069	-
13.23	3.33	10	44	26	0.0223	0.0418	0.5324	0.0025	0.0374	26.9771	-0.0883	1.3403	0.7461	0.6876	-
10.07	0.62	16	45	27	0.0271	0.0418	0.6484	0.0026	0.0695	27.8119	-0.2429	2.4098	0.415	0.7347	-
12.52	0.73	8	46	28	0.0294	0.0437	0.6719	0.0014	0.0391	28.0684	-0.0943	1.8606	0.5375	0.7837	-
12.52	0.73	10	47	28	0.0272	0.0387	0.7032	0.0014	0.0419	28.143	0.0274	1.9502	0.5128	0.8099	-
12.52	0.73	14	48	28	0.0314	0.0437	0.7178	0.0064	0.064	28.1566	0.025	1.4137	0.7074	0.8649	-
12.52	3.03	16	49	28	0.0233	0.0387	0.6026	0.003	0.062	28.348	0.1121	2.01	0.4975	0.7061	-
12.52	3.03	18	50	28	0.0255	0.0437	0.5835	0.0045	0.062	28.3578	0.1052	1.6454	0.6078	0.7099	-
12.52	3.03	12	51	28	0.0207	0.0437	0.4739	0.0009	0.0488	28.425	0.1166	2.8811	0.3471	0.5461	-
12.52	3.47	18	52	28	0.0312	0.0437	0.7138	0.0044	0.0682	28.7543	0.1522	1.8219	0.5489	0.828	-
12.52	3.47	12	53	28	0.0297	0.0387	0.7674	0.0038	0.053	28.7665	0.1405	1.5239	0.6562	0.9039	-
9.45	3.47	8	54	28	0.0259	0.0387	0.6678	0.0012	0.0301	28.784	0.2595	1.535	0.6515	0.8033	-
11.55	3.47	8	55	28	0.0199	0.0437	0.4556	0.0005	0.0371	28.9215	0.2474	2.8691	0.3485	0.5281	-
14	0.85	8	56	29	0.0255	0.0387	0.6595	0.0011	0.0424	29.548	-0.3692	2.239	0.4466	0.7524	-
14	2.65	14	57	29	0.0298	0.0437	0.6819	0.0052	0.06	29.5605	0.1458	1.4793	0.676	0.8225	-
14	2.65	10	58	29	0.0289	0.0437	0.6624	0.0022	0.037	29.7655	0.0654	1.4001	0.7142	0.811	-
14	3.25	16	59	29	0.0263	0.0387	0.6781	0.004	0.0781	29.9208	-0.2989	2.1889	0.4568	0.7732	-

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
14	2.65	18	60	29	0.028	0.0437	0.6409	0.0037	0.0682	29.9408	0.0051	1.9869	0.5033	0.7456	8.7
14	3.25	18	61	30	0.0284	0.0387	0.7332	0.0034	0.0791	30.0179	-0.2853	2.3918	0.4181	0.8202	-
14	3.25	14	62	30	0.0298	0.0383	0.7789	0.0033	0.0577	30.0522	-0.2707	1.7802	0.5617	0.8958	-
14	2.65	10	63	30	0.0241	0.0383	0.6288	0.0008	0.0434	30.3063	0.18	2.6711	0.3744	0.7067	6.3
14	1	12	64	30	0.0265	0.0383	0.6915	0.0015	0.0537	30.3341	-0.134	2.4894	0.4017	0.775	6.3
14	2.65	10	65	30	0.0234	0.0387	0.6051	0.0015	0.0438	30.3741	0.3516	2.0112	0.4972	0.7085	-
14	2.65	20	66	30	0.0274	0.0383	0.7159	0.0069	0.0758	30.3952	0.0036	1.6224	0.6164	0.8441	6.3
11.48	1	16	67	30	0.0256	0.0387	0.6616	0.0028	0.0583	30.4838	-0.4309	1.9428	0.5147	0.7687	-
11.75	2.65	8	68	30	0.0291	0.0383	0.7594	0.0014	0.033	30.6024	-0.0361	1.5584	0.6417	0.8928	6.3
11.55	2.65	18	69	30	0.029	0.0383	0.7566	0.0025	0.0698	30.661	0.2305	2.4677	0.4052	0.8409	6.3
9.88	1	12	70	30	0.0252	0.0387	0.6508	0.0022	0.038	30.702	-0.3181	1.423	0.7027	0.797	6.3
9.88	1	10	71	30	0.0226	0.0383	0.5891	0.0019	0.038	30.7106	-0.3149	1.5441	0.6476	0.7238	6.3
11.75	2.65	10	72	30	0.0261	0.0383	0.6806	0.001	0.0462	30.7476	0.1224	2.5803	0.3876	0.7612	6.3
11.75	2.65	34	73	30	0.0261	0.0387	0.6742	0.0085	0.1345	30.7885	-0.038	2.5893	0.3862	0.7546	6.3
11.55	2.65	26	74	30	0.0274	0.0383	0.7151	0.0033	0.0734	30.811	0.1702	2.2722	0.4401	0.8066	6.3
11.75	3.5	14	75	30	0.0264	0.0383	0.6889	0.0055	0.0574	30.8296	-0.1134	1.377	0.7262	0.84	6.3
11.75	3.5	10	76	30	0.0278	0.0383	0.7259	0.0017	0.0396	30.8653	-0.1713	1.6993	0.5885	0.8483	-
9.88	5.07	8	77	30	0.0193	0.0383	0.5046	0.0011	0.0298	30.9562	-0.8529	1.5828	0.6318	0.636	-
11.55	1.1	36	78	31	0.0303	0.0383	0.7903	0.0046	0.0997	31.0989	0.058	2.5977	0.385	0.8704	-
12.73	1.1	60	79	31	0.0284	0.0387	0.7337	0.0192	0.1384	31.1316	0.0248	1.7712	0.5646	0.8511	-
11.55	1.1	10	80	31	0.0253	0.0387	0.6535	0.0012	0.0383	31.2041	-0.1389	1.9626	0.5095	0.7595	-
9.88	1.1	8	81	31	0.02	0.0387	0.5162	0.0004	0.0336	31.2482	-0.4343	2.9484	0.3392	0.5867	-
8.88	5.18	12	82	31	0.0224	0.0387	0.5781	0.0017	0.0546	31.6749	-0.4724	2.3587	0.424	0.6662	-
13.48	4.62	14	83	31	0.0296	0.0383	0.7721	0.0009	0.0587	31.7955	-0.3038	3.386	0.2953	0.8336	-
13.48	6.85	8	84	31	0.0277	0.0383	0.7234	0.0012	0.0353	31.8411	-0.1042	1.8407	0.5433	0.8364	-
13.48	4.62	8	85	31	0.0195	0.0387	0.5029	0.0003	0.0369	31.9005	-0.1905	3.5449	0.2821	0.5615	-
13.48	6.85	12	86	32	0.0245	0.0396	0.6193	0.0009	0.05	32.2346	0.0722	2.8772	0.3476	0.6915	8.2
13.48	4.62	8	87	32	0.0255	0.0383	0.6649	0.001	0.0444	32.4361	0.2273	2.539	0.3939	0.7469	-
13.48	5.03	10	88	32	0.0255	0.0383	0.6657	0.0031	0.046	32.6305	-0.1162	1.4676	0.6814	0.8075	-
13.48	4.62	14	89	32	0.0205	0.0383	0.5351	0.0012	0.041	32.8191	0.0427	2.143	0.4666	0.6322	-
13.43	4.62	12	90	33	0.0261	0.0383	0.6808	0.0023	0.0431	33.364	0.1985	1.5999	0.625	0.8108	-
13.12	3.8	22	91	33	0.0263	0.0383	0.6867	0.0038	0.1024	33.6098	-0.0321	2.9519	0.3388	0.7572	-
10.15	4.62	8	92	33	0.0214	0.0383	0.5596	0.0005	0.0375	33.6456	0.1713	2.9176	0.3427	0.6308	-

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
10.15	3.78	10	93	34	0.0242	0.0396	0.6098	0.0014	0.034	34.0083	0.1854	1.6095	0.6213	0.7391	-
13.12	2.92	16	94	34	0.0274	0.044	0.6227	0.0033	0.0589	34.2312	-0.2097	1.816	0.5507	0.7372	-
13.12	2.92	12	95	34	0.0277	0.0396	0.6986	0.0032	0.0542	34.254	-0.1709	1.7083	0.5854	0.8204	-
13.12	2.85	8	96	34	0.0202	0.0396	0.5093	0.001	0.0318	34.3909	-0.2004	1.8113	0.5521	0.6241	-
12.8	1.9	8	97	34	0.0247	0.0396	0.6224	0.0013	0.0351	34.4913	0.0871	1.7505	0.5713	0.7412	-
12.8	1.9	10	98	34	0.02	0.0396	0.504	0.0013	0.0304	34.5846	0.2165	1.5162	0.6595	0.6412	-
12.8	3.08	8	99	34	0.0329	0.044	0.7482	0.0006	0.0623	34.7549	-0.5155	4.5437	0.2201	0.794	-
12.8	1.05	10	100	34	0.0231	0.044	0.5242	0.0014	0.0492	34.8754	-0.3823	2.3614	0.4235	0.6123	-
12.8	1.9	10	101	34	0.0179	0.0396	0.4521	0.0009	0.0386	34.9723	0.157	2.3079	0.4333	0.5422	11
12.8	1.05	10	102	35	0.0236	0.0396	0.5948	0.0013	0.0359	35.0016	-0.4146	1.7462	0.5727	0.7139	-
12.8	3	18	103	35	0.0194	0.0396	0.4907	0.0013	0.0637	35.1575	0.3305	3.1121	0.3213	0.5575	-
12.8	1.1	12	104	35	0.0191	0.0396	0.4811	0.001	0.0324	35.2755	0.1286	1.8045	0.5542	0.5964	11
12.8	1.12	8	105	35	0.0266	0.044	0.604	0.001	0.0381	35.4727	0.2184	2.0829	0.4801	0.7039	11
12.8	1.12	12	106	35	0.0296	0.0396	0.747	0.0056	0.0529	35.5003	-0.0656	1.2545	0.7972	0.9128	11
5.95	1.98	14	107	35	0.0258	0.0396	0.6511	0.0034	0.0485	35.8238	0.2583	1.4745	0.6782	0.7921	-
8.38	1.3	10	108	36	0.0407	0.0603	0.6757	0.0019	0.0791	36.4098	0.0448	3.237	0.3089	0.7399	-
11.05	1.35	12	109	37	0.0303	0.044	0.6887	0.0022	0.0699	37.4255	-0.0681	2.6255	0.3809	0.7679	10.5
10.48	1.05	8	110	37	0.0348	0.044	0.7899	0.0014	0.0482	37.6664	-0.0747	2.3032	0.4342	0.8802	10.5
9.93	1.02	10	111	38	0.0421	0.0603	0.6979	0.0028	0.0506	38.2051	-0.115	1.6837	0.5939	0.8215	3.5
9.93	1.25	14	112	38	0.0364	0.0699	0.5198	0.0044	0.087	38.396	-0.1903	2.324	0.4303	0.6093	-
10.5	1.35	16	113	38	0.0311	0.0603	0.5155	0.0043	0.0709	38.9175	-0.4878	1.9108	0.5234	0.6243	-
11.6	2.85	20	114	39	0.0336	0.0699	0.4802	0.0063	0.1032	39.611	-0.1963	2.2999	0.4348	0.5706	-
11.7	4.72	22	115	41	0.037	0.0699	0.5294	0.0097	0.1124	41.1572	-0.1578	2.0235	0.4942	0.6322	-
11.7	4.72	32	116	41	0.0397	0.0928	0.4283	0.011	0.1515	41.2238	-0.1558	2.5642	0.39	0.5094	-
11.43	2.8	14	117	43	0.052	0.0949	0.5477	0.0101	0.121	43.1499	-0.0175	2.1375	0.4678	0.645	-
8.02	2.83	12	118	44	0.0323	0.0879	0.3672	0.001	0.0956	44.5085	-0.1773	5.3444	0.1871	0.4061	-
11.12	1.88	24	119	45	0.0444	0.0879	0.5055	0.022	0.1332	45.5012	0.0187	1.5919	0.6282	0.6361	-
11.12	1.8	10	120	45	0.0361	0.0879	0.4109	0.0027	0.0748	45.7309	-0.2479	2.57	0.3891	0.4918	-
11.12	1.8	24	121	45	0.0334	0.0879	0.3802	0.0094	0.1053	45.8355	-0.3476	1.9292	0.5183	0.4881	-
6.95	2.9	16	122	47	0.0348	0.0879	0.3957	0.0032	0.1332	47.1592	0.3034	4.1476	0.2411	0.4458	-
6.95	6.95	8	123	47	0.0278	0.0695	0.3996	0.0008	0.0529	47.2317	0.3349	3.2937	0.3036	0.4628	-
6.82	5.5	72	124	49	0.0329	0.0695	0.4727	0.0377	0.2708	49.2764	-0.3521	2.4726	0.4044	0.5568	-
6.82	5.5	36	125	49	0.0288	0.0695	0.4149	0.0137	0.1038	49.3804	-0.2662	1.5722	0.6361	0.5472	5.7

ID	D min	D max	Cell	N	$\bar{s}$	$\bar{m}$	Q	Area	SMA	l	b	BE	ξ	Q^*	MSFR
6.82	3.95	14	126	49	0.0376	0.0695	0.5408	0.0097	0.0814	49.5536	-0.3143	1.462	0.684	0.6831	5.7