

HII Regions in the Pilot Ku-band Galactic Reconnaissance Survey

Author:

Mubela MUTALE

Supervised by:

Dr. M.A THOMPSON

Dr. L HINDSON

Centre for Astrophysics Research
School of Physics, Astronomy and Mathematics
University of Hertfordshire

*Submitted to the University of Hertfordshire in partial fulfilment of the requirements of
the degree of Master of Science by Research.*

October 2017

Abstract

This thesis presents data from the Ku-band Galactic Reconnaissance Survey (KuGARS), a large area survey observed on the Jansky Very Large Array (JVLA). KuGARS is the first galactic plane survey to explore the sub-arcsecond and sub-mJy regime at 14 GHz. We have described the development of the data reduction procedures in *CASA* for the On the Fly (OTF) mapping method used for KuGARS. Included in our discussion are the calibration and imaging processes: as the OTF mapping method has not been extensively explored we have gone through several steps to present quality assurance of the calibration. Due to the large volume of data, we focused on a pilot study of the area surrounding the W49A massive star forming region and a subset of the final frequency range.

Using an automated source finding algorithm, AEGEAN (Hancock et al., 2012), we have detected 35 islands with 57 components. We compared the brightest of these sources to the literature to determine their nature. In order to carry out a robust analysis, we compared our detections to observations over multiple frequencies. Comparing to GLIMPSE and Herschel showed our sources deeply embedded in $8\mu\text{m}$ and $70\mu\text{m}$ emission, respectively, which indicate heated dust and therefore infer star formation. We further, through the SIMBAD database, found several star formation tracers, including: Masers, YSOs, and HII regions, associated with our KuGARS detections. Our comparison with CORNISH and De Pree et al. (1997) revealed previous detections of these sources as HII regions with similar morphologies to KuGARS. We have detected and identified 2 spherical or unresolved HII regions, and 2 irregular HII regions in the ring of HII regions (W49 North), 1 isolated spherical/unresolved HII region, and 1 cometary HII region in W49 South. We find that one of the objects (source B) is recovered by KuGARS and De Pree et al. (1997) but not in CORNISH. This source, which is undetected at 5GHz and below has a steep positive spectrum ($\alpha = 1.1$) and has broad radio recombination lines with line widths ($> 45\text{kms}^{-1}$) (De Pree et al., 1997). It, therefore, must be a Hypercompact HII region and thus demonstrates KuGARS capability to detect Hypercompact HII regions.

Declaration

I declare that no part of this work is being submitted concurrently for another award of the University or any other awarding body or institution. This thesis contains a substantial body of work that has not previously been submitted successfully for an award of the University or any other awarding body or institution.

Except where indicated otherwise in the submission, the submission is my own work and has not previously been submitted successfully for any award.

Mubela Mutale

October 2017

Acknowledgements

I owe a world of gratitude to my principal supervisor Dr Mark Thompson, for not only giving me the opportunity to delve into a career of astronomy under his tutelage but also for being an amazing mentor, teaching me the basics, and giving me unending guidance throughout my MSc. In much the same way, I thank my co-supervisor Dr Luke Hindson for “holding my hand” and making sure I learnt and mastered all the techniques necessary for my Master’s.

I would like to thank the Development in Africa with Radio Astronomy (DARA) Newton Fund Project for awarding me the scholarship that enabled me to take on a career in astronomy in the most effective way possible.

I am deeply indebted to my parents, Emmanuel and Peggy, for giving me unconditional love and support from the moment I developed an inquisitive mind to this day. For not only tolerating my desire to discuss “unusual” topics, but actually enjoying it. I would also like to thank my older brother Cupa, for being my best mate, for being the person that understands me the most, and for having my back in everything I do. Finally, I would like to thank my friends for their support, particularly those in the Centre for Astrophysics Research for helping me out when I needed it, and for making my MSc memorable.

Contents

Abstract	i
Acknowledgements	iii
Contents	iv
List of Figures	vii
List of Tables	xi
List of Abbreviations	xii
1 Introduction	1
1.1 Massive Star Formation	1
1.1.1 Molecular Clouds	2
1.1.2 Cores	3
1.1.3 Gravitational Collapse	5
1.1.3.1 Virial theorem and Jeans mass	5
1.1.3.2 Fragmentation and accretion discs	6
1.1.3.3 Time scales	6
1.1.4 HII regions	7
1.2 Massive Star Formation Models	7
1.3 Large Area Surveys	9
1.4 The Ku-band Galactic Reconnaissance Survey	10
1.4.1 Description	10
1.4.2 Observation	15
1.5 Interferometry	16
1.5.1 The Very Large Array	18
1.5.2 The KuGARS Observation Technique	20
1.6 Project Goals	21
2 Calibration and Imaging in CASA	22
2.1 CASA	22
2.2 Data Examination and Editing	23
2.2.1 Initial Inspection	24
2.2.2 Flagging	25
2.3 Synthesis Calibration	26

2.3.1	Preparing for Calibration; “A Priori” Corrections	27
2.3.1.1	Correction of Antenna Position Errors	27
2.3.1.2	Gain Curves	28
2.3.1.3	Atmospheric Optical Depth	28
2.3.1.4	Requantizer Gain Corrections	29
2.3.1.5	Setting Flux Density	29
2.3.2	Preliminary Phase Calibration	30
2.3.3	Delay Calibration	30
2.3.4	Bandpass Calibration	30
2.3.5	Gain Calibration	31
2.3.6	Applying the Calibration	32
2.4	Splitting the Data	32
2.5	Deconvolution and Synthesis Imaging	32
2.5.1	Imaging (clean)	33
2.6	Self-Calibration	34
2.7	Data and Image Analysis	35
3	Data Reduction on KuGARS Block 5	36
3.1	Initial Data Examination	36
3.1.1	Observation report (VLA Operator Logs)	37
3.1.2	Data Layout - listobs	37
3.1.3	Antenna Layout and (\mathbf{u}, \mathbf{v}) Coverage	37
3.1.4	Data layout plots (plotms)	40
3.2	Initial Flagging (flagdata)	41
3.3	Calibration	44
3.3.1	Preparing for Calibration with “a priori” Corrections	45
3.3.1.1	Antenna Position Corrections	45
3.3.1.2	Gain Curves	45
3.3.1.3	Atmospheric Optical Depth	45
3.3.1.4	Requantizer Gain Corrections	46
3.3.1.5	Setting the Flux Density Scale	46
3.3.2	Initial Phase Calibration	47
3.3.3	Delay Calibration	47
3.3.4	Bandpass Calibration	49
3.3.5	Gain Calibration	50
3.3.6	Scaling the Amplitude Gains	52
3.4	Apply Calibration	52
3.5	Deconvolution and Imaging	55
3.5.1	Imaging W49A	57
3.5.1.1	Unusual Features in On-The-Fly Mapping (OTFM)	57
3.5.1.2	Imaging the Target	59
4	Data Analysis	64
4.1	Calibration Quality Assurance	64
4.1.1	Calibrator Parameters	64
4.1.1.1	Flux Calibrator	65
4.1.1.2	Phase Calibrator	67

4.2	Target-Noise Evaluation and Statistics	70
4.2.1	Spectral Window 2 (13.99 GHz)	70
4.3	Source Catalogue	70
5	Discussion	74
5.1	GLIMPSE Three Colour Maps	74
5.1.1	Description of GLIMPSE	74
5.1.2	KuGARS Sources Overlaid On GLIMPSE	75
5.2	Comparison of KuGARS with the De Pree et al. (1997) 3.6 cm observations. . .	78
5.2.1	Overlapping Sources	79
5.3	Comparison of KuGARS with CORNISH	79
5.4	KuGARS Detections Compared with the Literature	81
5.5	Comparison of KuGARS With Herschel 70 μ m Data	82
6	Conclusion	85
7	Future Work	87
7.1	Spectral Lines	87
7.2	Data Reduction and Future Analyses	88
A	LISTOBS Output	90
	Bibliography	98

List of Figures

1.1	Herschel Hi-GAL 250 μm image of the $l = 42^\circ - 46^\circ$ region. CORNISH UC HII regions are indicated by circles and AMGPS CH ₃ OH masers by triangles over 250 μm Hi-GAL colour scale. The box shows the proposed $4^\circ \times 2^\circ$ region to be mapped. Regions of interest are indicated by text labels.	13
1.2	A close-up of the W49A complex. UC HII regions and CH 3 OH masers are indicated as before. The large circle represents the primary beam from De Pree et al's 1997 W49A study. Note that the colour-scale stretch is different between Figs. 1.1 and 1.2.	13
1.3	Theoretical spectra of homogenous isothermal HII regions at a distance of 20 kpc. Solid, dashed and dot-dashed lines show spectra for hypercompact, ultra-compact and compact HII regions, respectively (with diameters 0.005, 0.05 & 0.5 pc and typical emission measures as given in Kurtz 2005). 5σ continuum sensitivities for MeerGAL (both for 1 and after 3 epochs), KuGARS, CORNISH, GLOSTAR (5 & 8 GHz bands) and VLASS-Galactic are indicated (figure from Thompson and Goedhart, 2016.	14
1.4	The Galactic longitude distribution of ATLASGAL clumps (shaded histogram) and CORNISH UC HII regions (blue histogram), showing the high surface density of UC HII regions between $l = 42^\circ - 46^\circ$	14
1.5	Block diagram of the simplest two-element interferometer observing in a very narrow frequency range.	17
1.6	The VLA is a collection of 27 radio telescopes, each antenna is 25 metres in diameter and weighs 230 tons. Images courtesy of NRAO/AUI and NRAO.	19
1.7	VLA telescopes positioned in a Y shaped configuration.	20
1.8	Antenna positions of the VLA in the B configuration. Image Courtesy of the NRAO.	21
2.1	Flow chart describing the general path of data reduction in CASA.	23
3.1	Correlator Setup during the KuGARS Observations. Total BI. BPs Used: 48 of 64. Total Data Rate: 59.97 MB/s or 215.89 GB/h. Total Spectral Points: 5184. Total Bandwidth: 4.232 GHz. Capability Mode: Shared risk.	36
3.2	Output file of the PLOTANTS task. The image shows the setup of the array during the observation.	38
3.3	(u, v) coverage plots of the calibrators. From top to bottom: the flux calibrator 3C48, the bandpass calibrator J2253+1608, and the phase calibrator J1922+1530.	39
3.4	Combined plot of the (u, v) coverage of the calibrators and target in plotms with baseline as colour axis. The x-axis represents the U coordinates, ranging from -10,000 to 15,000 while the y-axis represents the V coordinates, ranging from -15,000 to 10,000. Both axes are in units of metres.	40

3.5	Data stream plot with antenna2 in the y-axis ranging from 1 to 25 and time in the x-axis, from 16:06:40 to 18:53:20 with ea01 as reference point and field as colour axis.	41
3.6	A plot of amplitude as a function of time of the raw data in spw 2 showing all the calibrators and the target prior to RFI flagging. The y-axis shows amplitudes from -0.5 to 2.5 while the x-axis covers time from 10:36:40 to 18.33.20.	41
3.7	Amp-time plot after flagging configuration scans with field as colour axis. The y-axis shows amplitudes from -0.2 to 1 and the x-axis shows time form 16:04:40 to 18:53:20.	42
3.8	Inspection plots of the phase calibrator – amp v time (left) and amp v channel (right). Both y-axes show amplitudes from -0.02 to 0.14, left x-axis shows time from 16:40:00 to 18:53:20 and the right x-axis shows channels 0 to 61.	43
3.9	Amp-Time plot of the target showing distinct RFI features. The y-axis shows amplitudes from -0.05 to 0.2 and the x-axis shows time form 16:40:00 to 18:53:20.	43
3.10	Amp-time plot of the calibrators and target after flagging. The y-axis shows amplitudes from -0.2 to 1 and the x-axis shows time form 16:06:40 to 18:53:20.	44
3.11	Plot weather output showing the weather summary for the measurement set during correction for atmospheric optical depth.	46
3.12	Terminal output of SETJY task with parameter ‘listmodels’ set to ‘True’. Each calibrator source has a model listed in multiple bands.	47
3.13	Terminal output of SETJY task showing the flux density value assigned to the flux calibrator.	47
3.14	Plots of some of the antennas showing the variation of phase with time on the initial phase calibration table. The y-axes show Gain phases over an 80 degree range while the x-axes show time from 16:36:28.8 to 16:39:56.1.	48
3.15	Time delay plot on delay calibration table. The y-axes show the delays in nano seconds while the x-axes show time.	48
3.16	Antenna delay plot on delay calibration table. The y-axis shows the delay in nano seconds from -5 to 2 and the x-axis shows the antenna index from 0 to 26.	49
3.17	Plot of bandpass prior to calibration (with antenna ea01as reference) showing right circularly polarised data for the bandpass calibrator source (J2253+1608). Colour coded by spw and averaged over all baselines. In the y-axis, the amplitudes range from 0.1 to 0.6, in the x-axis are the channel numbers from 0 to 63.	49
3.18	Top right and left show bandpass amplitudes for J2253+1608 Right and Left Polarisation, respectively while bottom right and left show bandpass phases for J2253+1608 Right and Left polarisations, respectively. The x-axis range is 0-64 and the y-axis range is -300 to 300.	50
3.19	Amp-time plot of complex gain calibration table. The y-axis shows amplitudes from 0 to 1 and the axis shows time from 16:04:40 to 18:43:12.	51
3.20	Phase-time plots of complex gain calibration table. The y-axes show gain phases over 300 degrees and the x-axes show time from 16:04:40 to 18:43:12.	51
3.21	Dictionary results generated while running the task FLUXSCALE.	52
3.22	Amp-time plots of the fluxscale table. The y-axis shows amplitudes from 0.12 to 0.24 and the x-axis shows time from 16:04:40 to 18:43:12.	53
3.23	Plot of amplitude against time for the entire spectral window after calibration - with field as colour axis. In the y-axis, the amplitudes range from 0 to 35, in the x-axis is time from 16:05:40 to 18:53:20.	54

3.24	Amp-time plot of the target after calibration with field as colour axis. N.B the spikes in the first 3 blocks are caused by the target (see § 3.2). In the y-axis, the amplitudes range from -0.5 to 4, in the x-axis is time from 16:40:00 to 18:53:20.	54
3.25	Plot of bandpass after calibration (with antenna ea01 as reference) showing right circularly polarised data for the bandpass calibrator source (J2253+1608). Colour coded by spw and averaged over all baselines. In the y-axis, the amplitudes range from 28 to 31, in the x-axis are the channel numbers from 0 to 63.	55
3.26	Power distribution across baseline length. In the y-axis, the amplitudes range from -0.5 to 4, in the x-axis are the baselines lengths in units of uv wave (lambdas) from 0 to 500,000.	56
3.27	Comparison between the Primary beam response of the calibrator J1922+1530 -standard pointing (left) and the primary beam response of the target - OTFM pointing (right). Note the beam of the OTFM pointing has been truncated because the process is computationally heavy at this resolution.	58
3.28	Left image shows the beam of a single scan in an OTFM stripe – the image to the right shows the beam of several scans morphed together.	58
3.29	An offset phase-centre (left) compared with a correctly set phase-centre (right).	59
3.30	Dirty image of the target, created by running CLEAN with zero iterations.	61
3.31	PSF image of the target.	61
3.32	Raster image of the final clean image- stokes I mosaic of 15A-043 block 5, spw2 in the region surrounding W49A. Below are close ups of the groups of sources.	62
4.1	VLA calibrator manual image of 0137+331= 3C48 at 14.0 GHz in 1995.	65
4.2	KuGARS block 5pm flux calibrator, 0137+331= 3C48 at 13.99 GHz.	66
4.3	Flux of 3C48 at different frequencies on the VLA calibrator manual (blue dot) and the Ku-band (green cross).	67
4.4	VLA Calibrator Manual image of J1922+1530 at 8.46 GHz in 1997	68
4.5	KuGARS block 5pm phase calibrator, J1922+1530 at 13.99 GHz.	68
4.6	Flux of J1922+1530 at different frequencies on the VLA calibrator manual (blue dots) and the Ku-band (green cross).	69
4.7	Source overlay plot of the imaged mosaic in spectral window 2, plotted in inverted grey scale with each source represented by a circle. We have followed the naming convention of Dreher et al. (1984). The sources overlaid on this plot are presented in table 4.3.	73
5.1	Contours overlaid on an inverted grey scale image of the KuGARS map of W49A. The contours have been set at the 3, 6, 9, and 18 σ detection limits. We have followed the naming convention used in Dreher et al. (1984).	75
5.2	Three-colour GLIMPSE image at 99.5% colour scale using the IRAC 4.5, 5.8, and 8.0 μ m bands in blue, green and red, respectively. The green rectangle is an overlay of the region presented in Fig. 3.32	76
5.3	Contours of the brightest KuGARS detections in the region surrounding W49A overlaid on a three-colour GLIMPSE map with the min-max colour scale using the IRAC 4.5, 5.8, and 8.0 μ m bands in blue, green and red, respectively.	77

-
- 5.4 Contours from KuGARS sources in the region around W49A overlaid on the 3.6 cm *De Pree et al. 1997* image. The contours are set at the 3, 6, 9 and 18σ detection limits. The top shows the image in full view. The lower images are close-ups of the sources with the highest emission overlaid with KuGARS contours. 78
- 5.5 Contours from KuGARS sources in the region around W49A overlaid on a CORNISH image. The contours are set at the 3, 6, 9 and 18σ detection limits. The top image shows a full view of the CORNISH sources around W49A. The lower images are close-ups of the sources with the highest emission overlaid with KuGARS contours. 80
- 5.6 The sources in the region around W49A overlaid with KuGARS contours and SIMBAD sources on three colour GLIMPSE images set to the min-max colour scale using the IRAC 4.5, 5.8, and $8.0\mu\text{m}$ bands in blue, green and red, respectively. 81
- 5.7 W49A in Herschel at $70\mu\text{m}$. The top image shows the complete field of view under discussion. The bottom images show closeups of the three regions of interest overlaid with KuGARS contours. Note: the colour scales in the closeup images have been adjusted to highlight the brightest pixels in each image. Of the three, the brightest pixels were found in the region containing W49A (right), the least bright were in the isolated source (middle). 83

List of Tables

1.1	Physical Parameters of HII regions (Kurtz, 2005a).	9
3.1	Important parameters from the LISTOBS output.	38
3.2	Peak and integrated fluxes of the individual sources within the imaged mosaic on spectral window 2.	63
4.1	Fluxes of 3C48 across multiple frequencies from the VLA calibrator manual and KUGARS.	66
4.2	Fluxes of J1922+1530 across multiple frequencies from the VLA calibrator manual and KUGARS.	69
4.3	Source positions and fluxes in the region surrounding W49A.	71
4.3	Continued	72
A.1	listobs of the entire block	91
A.1	Continued	92
A.1	Continued	93
A.1	Continued	94
A.1	Continued	95
A.1	Continued	96
A.1	Continued	97

List of Abbreviations

ALMA	A tacama L arge M illimetre A rray
CASA	C ommon A stronomy S oftware A pplications
CORNISH	C Oordinated R adio a ND I nfrared S urvey for H igh-Mass S tar F ormation
FITS	F lexible I mage T ransport S ystem
GLIMPSE	T he G alactic L egacy I nfrared M id- P lane S urvey E xtraordinaire
GMC	G iant M olecular C loud
GUI	G raphical U ser I nterface
IRAC	I nfra R ed A rray C amera
ISM	I nter S tellar M edium
JVLA	J ansky V ery L arge A rray
KuGARS	Ku -band G alactic R econnaisance S urvey
MC	M olecular C loud
MFS	M ulti- F requency- S ynthesis
MS	M easurement S et
NRAO	N ational R adio A stronomy O bservatory
OTFM	O n- T he- F ly M ap/ M apping
OTF	O n- T he- F ly
PAH	P olycyclic A romatic H ydrocarbon
PDRs	P hoton- D ominated R egions or P hoto- D issociation R egions
RFI	R adio F requency I nterference
RRL	R adio R ecombination L ine
SED	S pectral E nergy D istribution
SPW	S Pectral W indow
YSO	Y oung S tellar O bject

Chapter 1

Introduction

Large area surveys of the Milky Way in the radio regime are continually building our understanding of massive star formation and stellar evolution. They cast a wide net and therefore tend to pick up discoveries far beyond the original expectations. This project is based on a large area survey observed on the Jansky Very Large Array (JVLA): the Ku-band Galactic Reconnaissance Survey (KuGARS). The survey is aimed at uncovering high frequency (> 10 GHz) radio emitters, which have gone undiscovered in current low frequency radio surveys.

1.1 Massive Star Formation

The overview of the process of massive star formation given in this section is based on more extensive reviews by Krumholz 2016, Kurtz et al. 2000, Kurtz 2000, Kurtz 2005a, Kurtz 2005b, Motte et al. 2017, Sewilo et al. 2004, and Wood and Churchwell 1989.

The study of the formation of massive stars is an important part of astrophysics. It however is also very poorly understood due to a low observational record. In its most basic and summarised form, our understanding of massive star formation can be considered a scaled up version of low mass star formation. Thus, the massive star can be considered to form via in fall through an accretion disk. The formation begins when gravitational instabilities in a molecular cloud initiate a collapse of the cloud material. This collapse proceeds on the free fall time scale of 10^5 years. When sufficient material has fallen onto the instability, a hydrostatic core forms. This evolves on the Kelvin Helmholtz scale of 10^4 years. As the molecular density continues to increase, the collapsing core eventually becomes optically thick and the energy released by the

gravitational collapse begins to heat the core, producing the so-called hot molecular core phase. Eventually, nuclear reactions begin and a star is born, and the resulting ultraviolet photon flux ionizes the surrounding material producing an HII region.

In the subsequent sections we will discuss the stages involved in massive star formation in slightly more detail and end by considering a few models proposed for how massive stars are formed.

1.1.1 Molecular Clouds

Massive stars form in giant molecular clouds (GMCs). These are made up of filaments and clumps, therefore, their mass is not evenly distributed. Most of their mass lies in the low-density structures while only a small portion lies in the very dense parts. However, in dividing the total mass by a rough volume occupied by ^{12}CO , a mean density of $\sim 100 \text{ cm}^{-3}$ is found. GMCs have typical sizes of tens of parsecs.

The most abundant element in the universe is hydrogen. In the ISM, hydrogen is found to exist in three particular forms: free atomic hydrogen (HI), Molecular hydrogen (H_2), and interstellar atomic hydrogen that has been ionised (HII). The most prevalent of these is H_2 , hence the dense molecular clouds in which stars are born. It is, however, easiest to observe hydrogen in its free atomic form, HI. This is because HI emits radio waves at a wavelength of 21 cm (1.4 GHz), associated with a hyperfine transition from a state in which the spins of the electron and proton are parallel to a state in which the spins are anti-parallel. Because the energy difference between these two states is really low corresponding to a very low temperature ($\ll 1$ K), HI can be excited in cold regions. Thus, it can be seen in the Milky way and in many nearby galaxies. Hydrogen is much harder to observe when in molecular form, H_2 . This is due to two major reasons: firstly, diatomic molecules have three excitation types: electronic, vibrational, and rotational - from highest to lowest in excitation scale, respectively. H_2 , in the first rotational excited state is 175 K above the ground state and the temperature at which molecules form in dense ISM is ~ 10 K. Ergo, molecules are unlikely to be in this excited state. Secondly, H_2 is a homonuclear molecule and therefore a transition to the first excited state is forbidden. This means that there are effectively no H_2 molecules in a molecular cloud in states capable of emitting (cf. Krumholz 2016).

1.1.2 Cores

Cores are smaller and denser and less massive than GMCs. They typically have densities of $\sim 10^7 \text{ cm}^{-3}$. Cores are only a few M_{\odot} , tend to be rounder than the lower density material around them, and are normally laid out across filaments of lower density gas. Cores are thought to be the forebears of single stars or star systems.

The prevailing thought is that star formation occurs within a core that is purely gas. The evolution is presumed to begin with gaseous sources with no traces of a star, followed by a core within which a protostar forms and evolves into a star which eventually ends up on the main sequence. Therefore, it follows that this would begin with a cold cloud without a central point source. Within this cloud forms a protostar which gradually heats up the cloud while the gas in the cloud collapses onto the protostar, reducing the opacity. With time, after enough material has been accreted from the envelope onto the protostar, it progressively becomes transparent at shorter wavelengths, first in the near infrared (IR) and then the optical – allowing direct observation of the star. Finally, after the accretion disk has fully been accreted, the star eventually contracts onto the main sequence.

At its earliest stage the star forming in the core will have a very low mass and hence a very low luminosity. Thus, it will only heat up the dust nearest to the centre of the core and by a very small amount. This will inherently make it difficult to directly trace the presence of the nascent star because the observed light output will be that of the thermal emission of the dust at its equilibrium temperature and thus the spectral energy distribution (SED) will appear much the same as before the star formed. However, the presence of an outflow or a sharp unresolved peak in the density flow can vicariously indicate the birth of a star. This type of object represents the earliest stage of star formation and at this stage, the object would fall into the so-called class 0. It will remain a class 0 object until it starts to heat up the dust around it beyond the dust's equilibrium temperature, consequently, emitting significantly in the IR; at this point, it evolves into a class I object. The cutting point for this is said to be when the source emits more than 0.5% of the bolometric output at wavelengths longer than $350 \mu\text{m}$, that is, when the ratio of the sub mm luminosity to the bolometric luminosity ($L_{\text{submm}}/L_{\text{bol}}$) is greater than 0.5% - where the sub mm luminosity considers all wavelengths of $350 \mu\text{m}$ or longer. The evolution of the protostar from this point on is defined in terms of the IR SED. Therefore, the wavelength at which the SED peaks infers how evolved the protostar is. During the earliest stages of the evolution, the envelope of dust around the protostar is optically thick at

shorter wavelengths (visible and near IR). Ideally, this envelope absorbs all the radiation from the protostar. Nonetheless, the dust is in thermal equilibrium, and therefore re-radiates this energy. This emission will, however, be at longer wavelengths since the radius of the dust is much larger than the star and thus for the luminosity output of the dust to equal that of the star, this emission must be at a lower temperature; hence the longer wavelengths. We are able to observe the photosphere because the radiation shifts to longer wavelengths as it gets transmitted through the dust and at wavelengths longer than the characteristic sizes of the dust grains, the opacity roughly decreases as $\kappa_\lambda \propto \lambda^{-2}$, therefore the radiation eventually shifts to wavelengths at which the dust is optically thin, and consequently escapes. This essentially means we vicariously observe a stellar photosphere through a “dust photosphere”. Thus, the greater the column density of the envelope of dust, the further the radiation from the star will have to diffuse in wavelength in order to escape. Evidently, the wavelength at which the emission peaks or the slope of the spectrum at a fixed wavelength provides a good estimate to the amount of circumstellar dust. Objects whose SED peaks at longer wavelengths are early in their evolution while objects whose SED peaks closer to the optical are presumed to be more evolved having lost more of their envelopes (cf. Krumholz, 2016).

The infrared classification scheme used for classes I, II, and III was based on fluxes measured by the *Infrared Astronomical Satellite (IRAS)*. The infrared spectral index (based on two points from the *IRAS* SED: 2.2 μm and 10 – 25 μm) is therefore defined as,

$$\alpha_{IR} = \frac{d \log(\lambda F_\lambda)}{d \log \lambda} \quad (1.1)$$

Where,

- α_{IR} is the infrared spectral index,

- λ is wavelength,

- F_λ is flux density as a function of wavelength.

This essentially means that positive values of α_{IR} indicate SEDs that peak in the longer wavelengths (further into the IR) while negative values of α_{IR} indicate SEDs that peak at shorter wavelengths (closer to the optical). By this classification scheme, class I objects would therefore be sources with $\alpha_{IR} \geq 0.0$.

The more the envelope accretes, the more optically thin it becomes at the peak emitting wavelengths of the stellar photosphere, thus we see the stellar blackbody spectrum mixed in with the residual IR emission from the remnants of the warm, dusty gas surrounding the star. Stars in this class are known as class II objects (also known as T Tauri stars). Here, the SED appears as a stellar blackbody with additional emission in the near/mid IR. Without the additional IR emission, these stars would have $\alpha \sim -1.6$, however, they have indices in the range $-1.6 < \alpha_{IR} < 0$. This is the last phase in which there is a disk of any significant mass and presumably the phase during which planet formation occurs.

The final group are the class III objects (weak line T Tauri stars). At this stage, the SEDs have $\alpha_{IR} < -1.6$. During this phase, there is no strong IR emission as the disk has begun to dissipate and is either optically thin in the IR, or has in fact completely dissipated. Therefore, any residual IR emission will be in the very far IR indicating cold circumstellar material distant from the star. Hence, the SEDs of these stars will appear as bare photospheres in the optical and mid-IR.

1.1.3 Gravitational Collapse

1.1.3.1 Virial theorem and Jeans mass

A cloud's gravitational stability against gravitational contraction determines whether or not it collapses and goes through the process of star formation. The virial theorem states that; for a self-gravitating system in equilibrium,

$$\begin{aligned}
 2U + \Omega &= 0 \\
 \text{Or} & \\
 \Omega &= -\frac{1}{2}U
 \end{aligned}
 \tag{1.2}$$

Where,

- U is the total thermal energy,
- Ω is the total potential gravitational energy.

In order for contraction to occur, the molecular cloud's mass must exceed a critical mass known as the Jeans mass. For this, we need the gravitational term in the virial theorem to overcome the

pressure term. Therefore, for a molecular cloud to gravitationally collapse Eq. 1.2 becomes,

$$-\Omega > 2U \quad (1.3)$$

For a constant density cloud, this gives,

$$M_c > \left(\frac{5kT}{G\mu m_H} \right)^{\frac{3}{2}} \left(\frac{3}{2\pi\rho_c} \right)^{\frac{1}{2}} \quad (1.4)$$

This is the critical mass, known as Jeans mass, required for contraction in a molecular cloud.

1.1.3.2 Fragmentation and accretion discs

Initially when the cloud collapses, its temperature remains constant so that the Jeans mass drops as the density increases. Consequently, smaller fragments become unstable to collapse leading to star clusters. With time, the fragments will heat up and a hydrostatic core will form and if the cloud has a uniform rotation rate, conservation of momentum results in the cloud contracting parallel to the rotation axis. This leads to flattened rotating structures known as accretion discs.

1.1.3.3 Time scales

The timescale for contraction from the protostellar hydrostatic core to the hydrogen burning main sequence is,

$$\tau_{K-H} \sim \frac{GM^2}{R} / L \quad (1.5)$$

and since for massive stars, $L \propto M^{3.3}$,

$$\tau_{K-H} \propto M^{-1.3}. \quad (1.6)$$

By this we see that for massive stars $\tau_{K-H} \ll \tau_{ff}$ and thus they arrive on the main sequence while still deeply embedded in their molecular clouds. Therefore, because the massive star turns on its UV illumination while still accreting and embedded in a molecular cloud, it heats

and eventually ionizes the circumstellar gas cloud thus creating an HII region. The HII region develops by expanding within the cloud.

1.1.4 HII regions

An HII region is a region of interstellar atomic hydrogen that has been ionized primarily by the ionizing radiation of young stars. However, only stars of $> 10 M_{\odot}$ produce photons with energies above the ionization potential of hydrogen, 13.6eV. It, therefore, follows that HII regions are clouds of partially ionized gas in which massive star formation has recently taken place. HII regions are over pressured with respect to their surroundings, hence they expand over time. Therefore, the smallest of these regions are most likely to be the youngest and thus trace recent star formation. The three classes of H II regions that most closely trace recent massive star formation are the compact, ultracompact (UC), and hypercompact (HC) H II regions (see table 1.1). Even though there is no universal agreement on the process of massive star formation, throughout the 1970s and 80s, the chief tracers of massive star formation were the compact and UC HII regions (Kurtz, 2005b) and more recently the HC HII region (Kurtz et al. 2000; Hoare et al. 2007).

HII regions are important to the study of massive star formation because they are associated with optical emission from transitions between energy levels of hydrogen and other atoms, and free-free (bremsstrahlung) emission in the radio. Bremsstrahlung is electromagnetic radiation produced by free electrons scattering ions. It is produced when a charged particle is deflected by another charged particle causing it to decelerate and therefore lose energy and produce a photon. A typical example is an electron and an atomic nucleus. Evidently, HII regions are important because they allow us to see where recent massive star formation has taken place and since we can calculate their electron density, and Lyman photon flux, we can also determine the photon flux of the stars.

1.2 Massive Star Formation Models

The study of massive stars is important for several reasons. One such reason is the effect they have on the inter stellar medium (ISM) around them. For instance, throughout their short main sequence lifetimes, they run through processes that inherently affect the composition of star

formation within a galaxy, and create H II regions as well as photon dominated regions through the ultraviolet photon flux they emit (Kurtz, 2005b).

As mentioned earlier, massive star formation is one of astrophysics' least understood problems. It is particularly difficult to study massive star formation (MSF) at near-IR wavelengths and more so optical wavelengths because this typically occurs within dense condensations of molecular clouds, with visual extinctions of order a thousand (Kurtz, 2005b). Even though no dominating theory has arisen yet, a number of scenarios have been proposed. One such scenario is coalescence, which was proposed by Bonnell et al. (1998) and Stahler et al. (2000). In this scenario stars of greater than 10 solar masses are assumed to form when stars of intermediate masses merge. The less massive stars would prior to this have formed through accretion. One would think that this would directly apply to more massive stars as well, but instead the coalescence scenario was developed due to the difficulty of accreting mass onto very luminous stars. This difficulty is caused by massive stars exerting a radiation pressure on their surroundings that is greater than their gravitational attraction. Ergo, coalescence was in contrast with a scenario formulated two decades earlier when (Stahler et al. 1980a; Stahler et al. 1980b) proposed that stars formed through accretion. This went on to be developed for low and intermediate mass stars (Palla and Stahler, 1993). It was later investigated for massive star formation and despite its successes, accretion was deemed impossible for massive stars due to their high luminosities, which are able to reverse collapse (Bonnell et al., 1998 ; Stahler et al., 2000). The basic implication here was that the more massive the star, the higher the luminosity and the greater the effect of radiation pressure on the accreting material. Therefore, the question of how massive stars can sustain accretion was open for decades. However, several of these models show that radiation pressure can be overcome by high accretion rates, or Rayleigh Taylor instabilities in the accretion flow. Kuiper et al. (2012) pose a solution to the radiation pressure problem. The solution states that an anisotropy of the thermal radiation field can overcome the problem. In addition, they argue that Radiative Rayleigh-Taylor instabilities in the cavity regions are neither required nor do they occur in massive star formation.

Other scenarios describe massive star formation through HII regions. Among these is the (Keto, 2003) gravitationally trapped model of HII regions. This model assumes the formation of a massive star by accretion onto Young stellar objects (YSOs) through gravitationally trapped HII regions surrounding the star. In 2004 the conventional model assumed that UC H II regions represented the earliest manifestation of massive stars after forming via rapid accretion of ambient gas onto a protostellar hydrostatic core. In this model massive protostars evolve from prestellar

cores, to hot cores, to UC H II regions to compact H II regions, to classical H II regions (Sewilo et al., 2004). The compact H II region is larger and has lower density than the UC H II region. However, Gaume et al. (1995) discovered a smaller and denser region than the UC H II region. The Hypercompact (HC) H II region, which they discovered near the NGC 7538 UC H II region, is 10 times smaller and 100 times denser than the UC H II region (Kurtz and Franco, 2002 ; Kurtz, 2002). Clearly, the physical parameters of H II regions span orders of magnitude in scale, but the classes most closely linked to star formation are the smallest, densest and presumably the youngest stages (i.e. compact, UC and HC H II regions (Kurtz, 2005b)). The bright radio emission from these nebulae serves as a signpost of recent massive star formation.

Evidently, all of the aforementioned theories have merit. However, a unified theory can only be developed after our understanding has significantly improved. This will only happen when the underlying problem in a lack of observations is solved.

TABLE 1.1: Physical Parameters of HII regions (Kurtz, 2005a).

Class of Region	Size (pc)	Density (cm^{-3})	Emis. Meas. ($pc\ cm^{-6}$)	Ionized Mass (M_{\odot})
Hypercompact	$\lesssim 0.03$	$\gtrsim 10^6$	$\gtrsim 10^{10}$	$\sim 10^{-3}$
Ultracompact	~ 0.1	$\gtrsim 10^4$	$\gtrsim 10^7$	$\sim 10^{-2}$
Compact	~ 0.5	$\gtrsim 5 \times 10^3$	$\gtrsim 10^7$	~ 1
Classical	~ 10	~ 100	~ 100	$\sim 10^5$
Giant	~ 100	~ 30	$\gtrsim 5 \times 10^5$	$10^3 - 10^6$
Supergiant	> 100	~ 10	$\sim 10^5$	$10^6 - 10^8$

1.3 Large Area Surveys

Radio surveys have an important role in astronomy, one that has changed with technology and scientific requirements. Due to the fact that the galaxy is largely transparent at radio wavelengths, radio surveys are particularly powerful and are able to study objects across the Galaxy even at low Galactic latitude (Thompson and Goedhart, 2016). Most objects studied by radio astronomers today are the unexpected discoveries of early surveys. For instance, Jansky's detection of Galactic radio emission and Reber's 160 MHz maps showing that emission is non-thermal stimulated the survey "discovery". Surveys made just after World War II revealed strong discrete sources which were later identified with: supernova remnants, radio galaxies, and quasars. Pulsars were discovered during a sky survey for scintillating sources, BL Lac objects were recognised in early high-frequency surveys, the first gravitationally lensed quasar

appeared in the extensive Jodrell Bank 960 MHz survey, and the first measurement of gravitational radiation came from the binary pulsar serendipitously found in a pulsar survey (Condon, 1998). Continuum surveys before 1998 spanned 30 MHz to 5 GHz and most known blazars were found in the higher-frequency surveys sensitive to compact sources with flat radio spectra. However, higher frequency surveys had been considered in the hope that they would find a so-called new population of sources with peaked or inverted spectra (Condon, 1998). The problem lay in that these were technically difficult as the beam solid angle of a telescope scales as ν^{-2} and system noise generally increases with frequency, implying that the time needed to survey a fixed area of sky to a given flux limit rises quite rapidly above 5 GHz (Condon, 1998).

As stated earlier, low frequency surveys have a strong selection bias against objects whose spectra rise with frequency (Thompson and Goedhart, 2016). The turnover frequency between the optically thick and thin regimes for thermal bremsstrahlung is essentially a linear function of electron density and so this means that low frequency ($\nu \leq 5\text{GHz}$) surveys preferentially select against dense plasma ($N_e \gtrsim 10^5 \text{cm}^{-3}$) (Mezger and Henderson, 1967). Therefore, the lack of surveys in this frequency regime means we have a poor understanding of the properties and occurrence of steep spectrum emission in the Galaxy.

1.4 The Ku-band Galactic Reconnaissance Survey

1.4.1 Description

The Ku-band Galactic Reconnaissance Survey (KuGARS), is the first sub-arcsecond resolution (12-18 GHz) survey of the Galactic Plane. Our primary aim is to discover and characterise the population of steep positive spectrum objects in the Galaxy. In the full survey we aim to achieve the following major science goals:

- i) a complete and unbiased survey for young HII regions, allowing their numbers, relationship to 12.2 GHz methanol masers, physical properties and statistical lifetimes to be determined
- ii) determine accurate mass loss rates for hundreds of OB stars as a function of their stellar parameters and search for colliding-wind binaries
- iii) measure magnetic fields on small scales in massive YSOs via Zeeman splitting of OH masers

- iv) characterise the population of young steep spectrum PNe
- v) enable the study of anomalous microwave emission and spinning dust in a wide variety of environments at high angular resolution.

KuGARS is the first Galactic Plane survey to explore the sub-arcsecond and sub-mJy regime at 14 GHz, and has significant scope for new and unexpected discoveries. In particular, KuGARS has high legacy value as the highest frequency radio component of a family of Galactic Plane surveys from the optical to 5 GHz, allowing ready cross-correlation with optical/IR/low frequency counterparts. The pilot to KuGARS is being carried out in order to validate our observing approach and fully test On The Fly (OTF) mapping (see § 1.5.2) at Ku-band. We have mapped a $4^\circ \times 2^\circ$ region around W49A (Fig. 1.1) in this self-contained project that will provide the first estimate of the statistical lifetime of HC HII regions.

Through matching volumes between the CORNISH & GLOSTAR UC HII region surveys, we can improve our understanding of the birth of HII regions by determining their statistical lifetimes. Finding these objects requires a large area survey as massive YSOs/UC HII regions themselves are rare (only ~ 1000 in the Milky Way, Lumsden et al., 2013) it is likely that HC HII are even rarer. As Fig. 1.3 shows, there is no survey preceding KuGARS that can reach the same completeness limit for HC HII regions. A Ku-band survey is the only suitable way, in terms of sensitivity to steep spectrum sources and survey speed, to detect the youngest, most dense and earliest stage in the development of an HII region (Sjouwerman and Mills, 2013). The MeerGAL survey with MeerKAT will conduct a multi-epoch Ku-band survey of the southern GLIMPSE region at similar depths to KuGARS.

This pilot project allows us to place an initial constraint on the lifetime & properties of HC HII regions. We, however, will need a much larger survey to place tight constraints on the frequency of these objects. The Galactic longitude range $l = 42^\circ - 46^\circ$ has the highest UC HII region surface density in the Milky Way, mostly due to the highly luminous W49A star forming region (see Fig. 1.4 and also Urquhart et al., 2013) which itself contains almost 10% of the UC HII regions in the CORNISH survey. Also within this region are 20 6.7 GHz CH₃OH masers (Pandian and Goldsmith, 2007). If the range of HC/UC HII and HC/CH₃OH maser ratios lies within the range of recent limited surveys (De Pree et al., 2014, Sanchez-Monge et al., 2013) we expect to detect between 10–26 HC HII regions. We regard this as a lower estimate as we have not counted potential HC HII regions associated with the numerous bright FIR sources from Hi-GAL (Molinari et al., 2010) that have similar characteristics to the HII region and maser-hosting

clumps (see Figs. 1.1 and 1.2). This would more than double the known sample of HC HII regions recovered from unbiased surveys and finally allow an estimate of their statistical lifetime to be determined with better than 3σ precision. In addition, the wealth of multiwavelength data covering this region (UKIDSS, GLIMPSE, MIPS GAL, Hi-GAL, ATLASGAL, SCUBA-2, GLOSTAR, THOR) and kinematic/trigonometric distance determinations (Urquhart et al., 2013, Reid et al., 2014) makes it possible to investigate the luminosity dependence of the HC and UC HII regions.

The science we can carry out with these observations extends beyond HC HII regions and we expect to detect much more than just these objects. As Fig. 1.3 shows, we will also detect UC HII regions (although we will likely resolve out more extended compact HII regions) and so by combining our observations with GLOSTAR 5 GHz continuum we will also be able to accurately measure their spectral indices, which are poorly constrained (Urquhart et al., 2013), as well as the spectral indices of the many radio galaxies we will detect. There are also more than 40 OB stars within the survey region, including the clusters GLIMPSE19 and 20 (Mercer et al., 2005), of which the latter contains 3 OB stars and an early WC star (Messineo et al., 2009). Many lie within ~ 2.5 kpc and so we will be able to constrain their mass loss via stellar winds and also identify colliding wind binaries. Towards bright and compact continuum sources we expect to observe recombination lines in emission (with typical line-continuum ratios towards all detected UC/HC HII regions with fluxes >50 mJy) and also absorption from the 14.4 GHz H_2CO line. For the latter we will combine our data with the GLOSTAR detections of the 4.8 GHz line to obtain the H_2 volume density around these objects (Ginsburg et al., 2011). Finally, we have obtained a simultaneous survey of 12.2 GHz CH_3OH and 13.4 GHz excited OH masers, in a bid to reveal their relationship to the star-forming regions traced by the HII regions. For OH we will be able to measure the Zeeman splitting for all masers brighter than 1 Jy, enabling the measurement of the magnetic field around massive YSOs on sub-arcsec scales, complementing the MAGMO survey (Green et al., 2012).

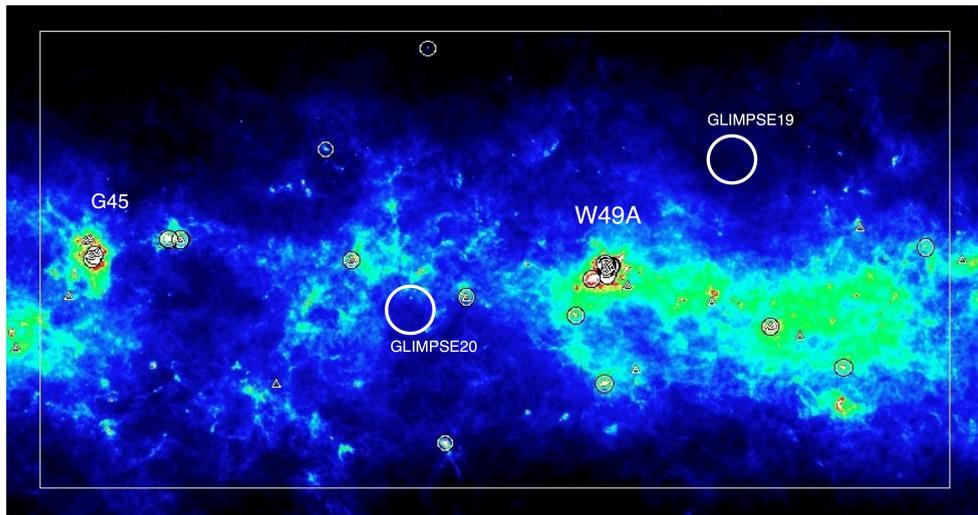


FIGURE 1.1: Herschel Hi-GAL 250 μm image of the $l = 42^\circ - 46^\circ$ region. CORNISH UC HII regions are indicated by circles and AMGPS CH_3OH masers by triangles over 250 μm Hi-GAL colourscale. The box shows the proposed $4^\circ \times 2^\circ$ region to be mapped. Regions of interest are indicated by text labels.

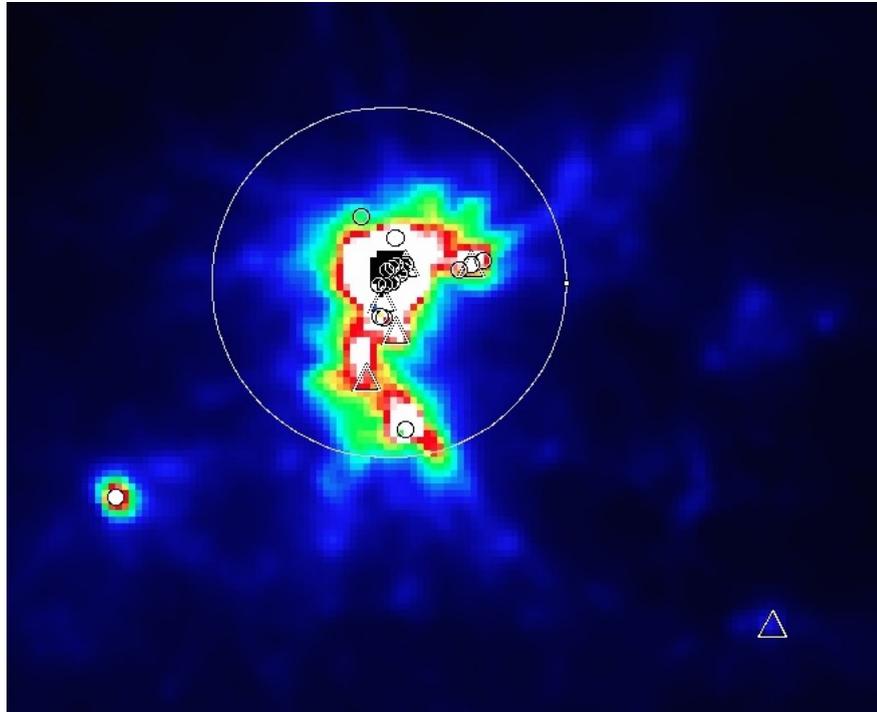


FIGURE 1.2: A close-up of the W49A complex. UC HII regions and CH_3OH masers are indicated as before. The large circle represents the primary beam from De Pree et al's 1997 W49A study. Note that the colour-scale stretch is different between Figs. 1.1 and 1.2.

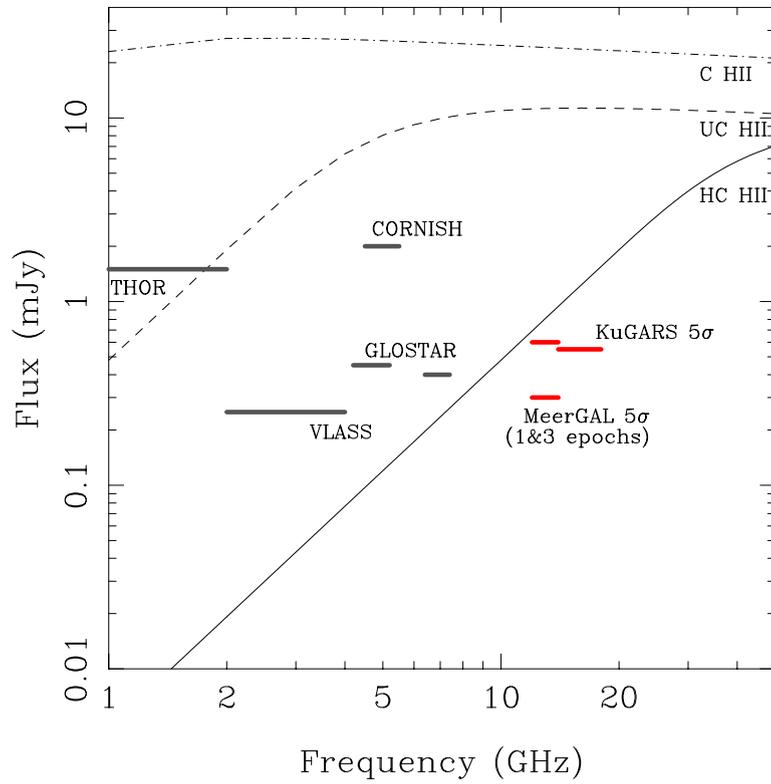


FIGURE 1.3: Theoretical spectra of homogenous isothermal HII regions at a distance of 20 kpc. Solid, dashed and dot-dashed lines show spectra for hypercompact, ultracompact and compact HII regions, respectively (with diameters 0.005, 0.05 & 0.5 pc and typical emission measures as given in Kurtz 2005). 5σ continuum sensitivities for MeerGAL (both for 1 and after 3 epochs), KuGARS, CORNISH, GLOSTAR (5 & 8 GHz bands) and VLASS-Galactic are indicated (figure from Thompson and Goedhart, 2016).

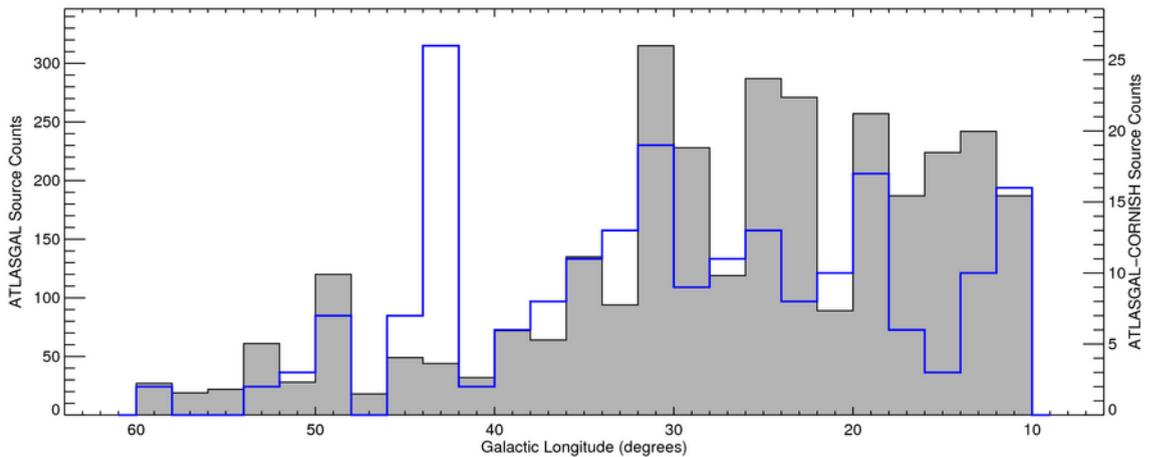


FIGURE 1.4: The Galactic longitude distribution of ATLASGAL clumps (shaded histogram) and CORNISH UC HII regions (blue histogram), showing the high surface density of UC HII regions between $l = 42^\circ - 46^\circ$.

1.4.2 Observation

In KuGARS our primary goal is to detect and identify HC HII regions out to the edge of the Milky Way. To maximise our sensitivity to these steep spectrum objects we have observed in continuum 14-18 GHz, splitting this bandwidth into five 0.7 GHz wide channels (allowing for RFI losses) over the bandpass. We have split the band into 5 in order to determine the spectral index of the emission - most HC HII regions will not have counterparts in the CORNISH or GLOSTER 5 GHz surveys. HC HII regions are faintest at 14 GHz and so this frequency drives the sensitivity requirements; a typical HC HII region at 20 kpc has a 14 GHz flux of 0.9 mJy, which requires an rms of 0.18 mJy for a 5σ detection. In the KuGARS pilot, we have set our sensitivity slightly below this limit (0.11 mJy) to enable us to determine the optimum observing strategy for the larger KuGARS survey.

To reach an rms noise of 0.11 mJy at 14 GHz requires an on-source time of 20s according to the VLA ECT (with 3 bit samplers, dual polarisation & natural weighting). We have mapped a $4^\circ \times 2^\circ$ area in OTF mode. Using a $\theta_{PB}(18GHz)/\sqrt{2}$ (where θ_{PB} is the angular resolution of the primary beam) spacing between 2° long scan rows resulted in 80 rows in total. Using a 1s dump time and a $0.25'/s$ scan rate to maintain 10 samples per $2.5'$ primary beam, each row took 10 minutes to complete (not counting overheads). We split our area into 8 blocks of 20 rows. Each block took ~ 2.5 hours, including 25% overheads. We repeated each block twice within a single 5 hour track to obtain 20s on source time and to improve the uv coverage (a similar strategy was adopted by the CORNISH survey, Hoare et al., 2012). The total time was therefore, 8 blocks x 2 repeats x 2.5 hours, or 40 hours.

We allocated unused baseline board pairs (keeping within the data rate limits of 60 MB/s) to maximise the scientific productivity of the survey with simultaneous spectral line observations of CH₃OH and OH masers, plus 8 Radio Recombination lines (RRLs) and H₂CO in absorption. The OH maser transition was observed with 0.1 km s^{-1} channels in order to measure the Zeeman splitting (possible for all masers $> 1 \text{ Jy}$), whereas the CH₃OH maser and H₂CO lines were observed with 0.6 and 0.3 km s^{-1} wide channels, respectively. RRL line windows are centred on the hydrogen lines from H81 α to H74 α and have 5.7 km s^{-1} channels, which are sufficient to resolve these broad ($>30 \text{ km s}^{-1}$) lines. We have obtained line sensitivities per channel of 41 mJy, 17 mJy, 22 mJy and 6 mJy for OH, CH₃OH, H₂CO and RRLs, respectively. This was suitable to detect the strong maser lines, plus observe H₂CO absorption against strong continuum

and also RRL emission from all compact HII regions brighter than ~ 50 mJy (binning to 10 km s^{-1} channels and stacking the 8 $\text{Hn}\alpha$ lines).

1.5 Interferometry

Interferometry is the method through which the phenomenon of interference of waves is applied to electromagnetic (EM) waves, which are superimposed in order to study the source of the EM waves. In radio astronomy, this is carried out using two or more telescopes. Fig. 1.5 shows the simplest setup of a radio interferometer.

Radio telescopes sample different physical processes than optical ones. Whereas aspects such as light pollution, turbulent air in the atmosphere, atmospheric absorption, bad weather, and interstellar dust clouds limit the effectiveness of optical telescopes, they do not present a problem in the radio regime. These aspects and the Galaxy are largely transparent at radio wavelengths (Thompson and Goedhart, 2016). However, simply observing in the radio is not entirely sufficient as radio wavelengths are longer and require a wider beam-width. As a consequence, the angular resolution of a single antenna is insufficient for most observations. Radio interferometers, which are ensembles of two-element interferometers, improve the angular resolution in observations of radio emission (Thompson et al., 2017)

The voltages V_1 and V_2 of antennas 1 and 2, respectively, are induced by radio emission incident on the antennas. These voltages, as a function of time at the receiving antenna, appear as very short parts of a random rhythm that builds into a waveform with a Gaussian amplitude distribution (Thompson et al., 2017). The correlator multiplies and time averages the signal voltages, effectively filtering out high frequencies.

Essentially, the wavefront incident from a source in the direction of \hat{s} arrives at the right antenna at a time τ_g before it reaches the left antenna. This is called the geometric delay and is given by:

$$\tau_g = \frac{\vec{b} \cdot \hat{s}}{c}, \quad (1.7)$$

where c is the speed of light, and \vec{b} is the vector baseline from antenna 1 to antenna 2. Therefore, if the centre frequency is $\nu = \omega/2\pi$, the output of the multiplier will be:

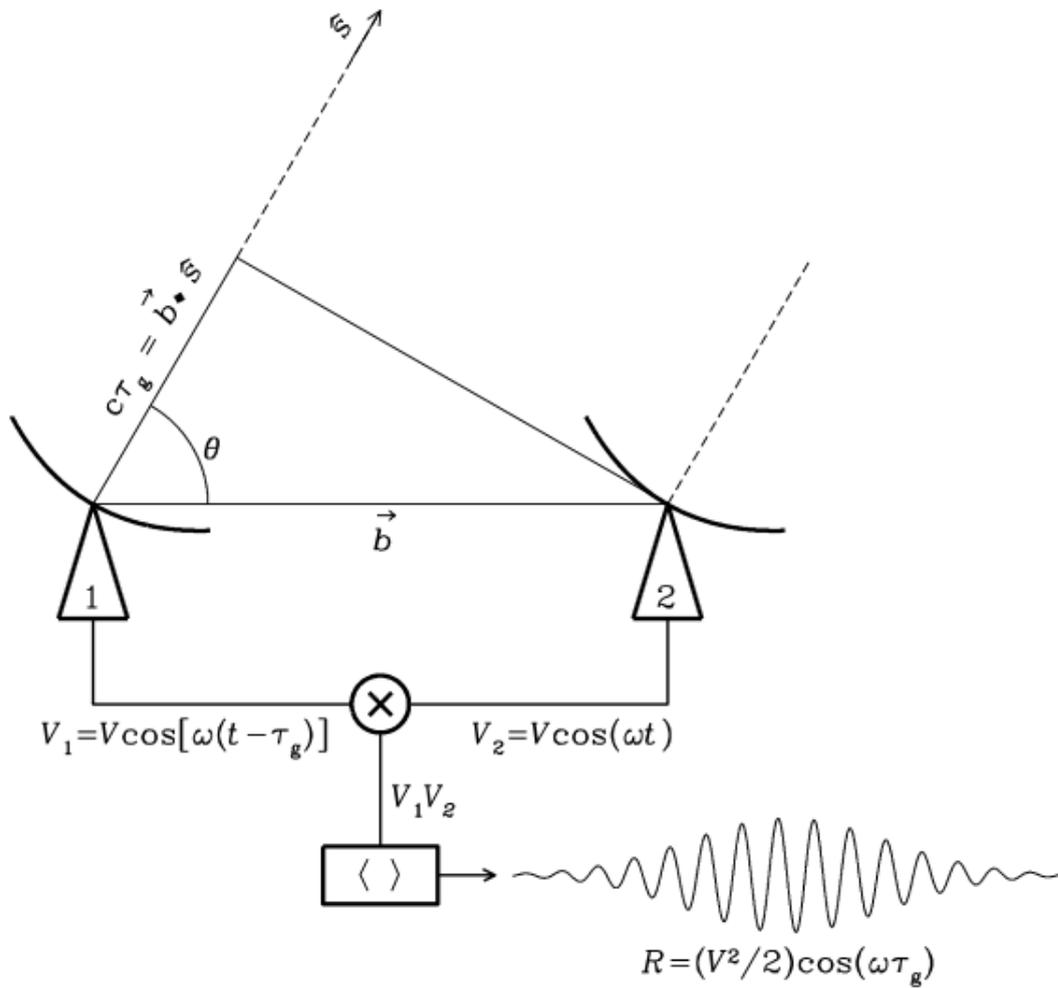


FIGURE 1.5: Block diagram of the simplest two-element interferometer observing in a very narrow frequency range. In this diagram, the centre frequency is $\nu = \omega/2\pi$, the voltage outputs of dish 1 and 2 are V_1 and V_2 , respectively. The voltages are the same except for the delay $\tau_g = \vec{b} \cdot \hat{s}/c$ on V_2 , where \hat{s} is the unit vector in the direction of a distant point source and \vec{b} is the vector baseline from antenna 1 to antenna 2. The correlator multiplies and time averages V_1 and V_2 to generate an output response with a delay and frequency dependent amplitude R that is proportional to the point-source flux density. From <http://www.cv.nrao.edu/course/ast534/Interferometers1.html>

$$\begin{aligned}
 F &= 2 \sin(2\pi\nu t) \sin 2\pi\nu(t - \tau_g) \\
 &= 2 \sin(\omega t) \sin \omega(t - \tau_g) \\
 &= 2 \sin^2(\omega t) \cos(\omega\tau_g) - 2 \sin(\omega t) \cos(\omega t) \sin(\omega\tau_g).
 \end{aligned} \tag{1.8}$$

Averaging this over time, and substituting the values $\omega(2\pi c/\lambda)$ and τ_g leaves us with,

$$F = \cos \omega\tau_g = \cos \left(\frac{2\pi\vec{b} \cdot \hat{s}}{\lambda} \right). \tag{1.9}$$

Alternatively, we consider the interferometer in Fig. 1.5¹, a quasi-monochromatic interferometer centred on the frequency $\nu = \omega/2\pi$, within a narrow band. V_1 and V_2 are the voltages of antennas 1 and 2, respectively, t is time, and τ_g (Eq. 1.7) is the geometric delay. \hat{s} is the unit vector in the direction of a distant point source, and \vec{b} is the vector baseline from antenna 1 to antenna 2. Then, multiplying the two voltages at the correlator we have,

$$\begin{aligned} V_1 V_2 &= V \cos[\omega(t - \tau_g)] V \cos(\omega t) \\ &= V^2 \cos(\omega t) \cos[\omega(t - \tau_g)] \\ &= \left(\frac{V^2}{2}\right) [\cos(2\omega t - \omega\tau_g)]. \end{aligned} \quad (1.10)$$

To remove the high frequencies ($\cos(2\omega t - \omega\tau_g)$) from the final output R , we take a long time average ($\Delta t \gg (2\omega)^{-1}$) we also take $\vec{b} \cdot \hat{s} = b \sin \theta$ to get,

$$\begin{aligned} R = \langle V_1 V_2 \rangle &= \left(\frac{V^2}{2}\right) \cos(\omega\tau_g) \\ &= \left(\frac{V^2}{2}\right) \cos\left(\frac{2\pi\vec{b}\cdot\hat{s}}{\lambda}\right) \\ &\approx \left(\frac{V^2}{2}\right) \cos\left(\frac{2\pi\theta}{\lambda}\right) \end{aligned} \quad (1.11)$$

as the final output of the correlator. The output R multiplied by the beam's strength per angle is the pattern produced on the image.

1.5.1 The Very Large Array

The KuGARS observations were obtained using the Very Large Array (VLA) in B configuration using On-The-Fly (OTF) mode. The VLA is a collection of 27 radio telescopes located at the National Radio Astronomy Observatory (NRAO) site on the plains of San Agustin ~ 80 km from Socorro, New Mexico (Fig. 1.6). With 27 radio antennas (each 25 metres in diameter) in a Y shaped configuration, the VLA is a “next generation” radio interferometer - one of the world's premier observatories. The electronically combined data from the antennas gives the resolution of a 36km wide antenna with the sensitivity of a dish 130 metres in diameter².

The VLA is an interferometer and therefore works by multiplying the data from each baseline pair to form interference patterns. These patterns are transformed into maps using Fourier transforms. The VLA telescopes are set up in a Y shape (see Fig. 1.7). This can be used in one

¹<http://www.cv.nrao.edu/course/ast534/Interferometers1.html>

²<http://www.vla.nrao.edu/>



FIGURE 1.6: The VLA is a collection of 27 radio telescopes, each antenna is 25 metres in diameter and weighs 230 tons. Images courtesy of NRAO/AUI and NRAO.

of four configurations: the A array has the maximum antenna separation of 36km, the B and C arrays have 10km and 3.6km maximum antenna separations, respectively, while the D array is the most compact with a maximum antenna separation of 1km. It therefore follows that the resolution of the VLA is set by the size of the array. The resolution will also be dependent on the observation frequency as the VLA has multiple receivers available.

The KuGARS observations were carried out in the B configuration. The VLA changes configuration \sim every 4 months. Fig. 1.8 shows the positions in the B configuration.



FIGURE 1.7: The VLA telescopes are positioned in a Y shaped configuration. Top: centre of the VLA array, each of the arms extends farther in the respective directions. Bottom: Showing the full extension of one of the arms. The VLA antennas can be set up in different configurations, A, B, C, or D. Images courtesy of NRAO/AUI and NRAO.

1.5.2 The KuGARS Observation Technique

The KuGARS data was obtained through the On the Fly (OTF) Mapping observation technique. In the standard methods (step and integrate), the telescope remains fixed while data is collected. In the OTFM method, the telescope scans across a field in Right Ascension (RA), smoothly and rapidly (with a constant slewing speed) while continuously recording observation information. For KuGARS, this technique provides a high observing efficiency. The telescope overhead is significantly reduced, properties of the atmosphere and the system such as antenna pointing and calibration change less, and "dead time" is minimized. On the other hand, sampling the sky properly in an interval shorter than the changes in the instrumental and atmospheric calibration requires a very high data rate (Mangum et al., 2000).

Chapter 2

Calibration and Imaging in CASA

2.1 CASA

The Common Astronomy Software Applications (CASA) package has been developed to calibrate, image, and analyse data obtained from radio telescopes such as the Atacama Large Millimetre Array (ALMA), and the Very Large Array (VLA). The version of CASA used in this study is 4.6.0. Here, the basis of the data reduction follows the National Radio Astronomy Observatory (NRAO) CASA cookbook¹.

CASA, which is written in PYTHON, has a group of astronomical data reduction tools and tasks that can be run in the IPYTHON interface to PYTHON to achieve the various steps of the data reduction process. The user can and will normally interactively run the set of built-in tasks, but because CASA operates under an IPYTHON wrapped interface, the user can also write a script to run these tasks.

Every interferometry dataset is unique and in order to achieve the optimum image the reduction process needs to be tailored to the observation. However, the reduction process itself follows a general route which is outlined in Fig. 2.1 below. CASA manages data through a table system in which the visibility data are stored in a table known as a Measurement Set (MS) and images are handled through special CASA image tables (standard FITS I/O is also supported).

¹https://casa.nrao.edu/casa_cookbook.pdf

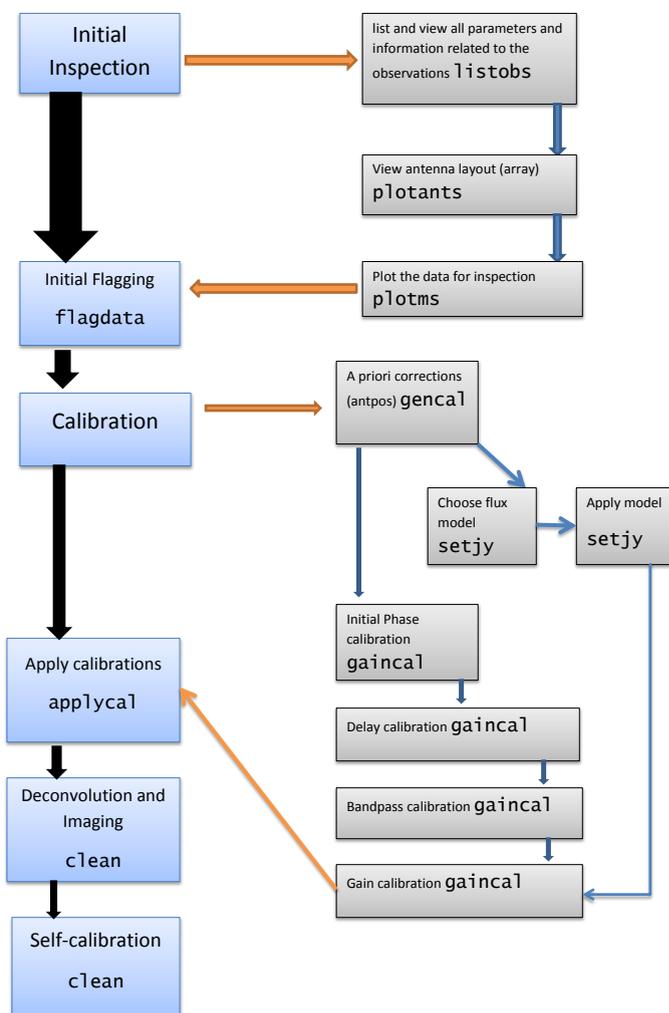


FIGURE 2.1: Flow chart describing the general path of data reduction in CASA.

2.2 Data Examination and Editing

Data examination involves inspecting the data for aspects such as variations of amplitude and phase with time and frequency. The primary aim of the examination is to determine the ‘quality’ of the data and the effectiveness of the reduction process. It is through this inspection that bad data (e.g. RFI, bad baselines) are discovered. The CASA task used to create the plots in which these inspections are carried out is PLOTMS. The PLOTMS GUI allows the user to carry out several actions within it such as: edit the axes, average data, and replot it. It also has options to locate specific portions of the data and flag data.

A more robust and automated way to flag data is through the CASA task FLAGDATA. This task

can be used in various configurations to flag data in the most suitable way. There are options to set parameters based on information of bad data passed on by inspection through PLOTMS. In addition to this or as an alternative, the data can be flagged through automated flagging algorithms (see § 2.2.2). Each time the task is run, CASA creates a ‘flag version’ which can be used to restore the flagged data through the task FLAGMANAGER. The task FLAGMANAGER has options to: list, save, restore, and delete ‘flag versions’. This allows the user to control the data they flag.

Knowing the layout of the antennas in a given observation is important to the subsequent data reduction. This is done in the CASA task PLOTANTS. This task is useful to see which antennas were used (a list of the antennas can be seen in LISTOBS) as well as the antenna positions relative to each other. One particularly good application of this task is to determine which antennas are good candidates to be used as the reference antenna (the antenna used as a standard), ideally one close to the centre of the array. The choice of reference antenna is made based on this and aspects such as a ‘datastream’ plot in PLOTMS to see which antennas are present throughout the observations. It is also necessary to plot amplitude as a function of frequency iterating over baseline to make sure the chosen reference antenna has good data on all the baselines. Even though this is not the case for the VLA which has antennas of equal size, it is important to note that in the event that the antenna sizes vary, it would be preferred that the largest was chosen as reference antenna.

2.2.1 Initial Inspection

Each Measurement Set will be unique, possessing different calibrators, scan patterns, and sources of bad data. It is therefore critical to perform an inspection of the data before beginning the data reduction process. The observers log will always have some information about the weather conditions and antenna performances during the observations. Ergo it must be viewed and all necessary information such as antennas that have corrupt data and need immediate flagging should be noted. Following this, the general setup of the Measurement Set should be inspected to examine the data and all the parameters therein as well as the antenna layouts. Thereafter, the data can be viewed through various plots for a more informative inspection. These inspections can be performed in CASA using tasks such as: LISTOBS, PLOTANTS and PLOTMS, respectively. Note that, a log runs parallel to the CASA terminal. This log stores and displays all the tasks run, as well as their outputs and errors.

The LISTOBS task displays information such as: fields, spectral windows, frequency ranges, bandwidth, and number of scans (as well as scan times). It is therefore imperative to run this function prior to the data reduction. On the other hand, the PLOTANTS task simply outputs the antenna layout used during the observation. This is helpful in picking out a reference antenna which ideally should be at the centre of the array. Now PLOTMS, which is used not only at the beginning but throughout the data reduction process, is used to make various plots of the data, each of which supply a different perspective on the data. For instance, during initial inspection, PLOTMS is used to make plots such as phase vs time or amplitude vs time. These can be used to identify bad data – be it: Radio Frequency Interference (RFI), dummy scans, antennas with a bad baseline, antennas with uncharacteristically low amplitudes or something else that would reduce the quality of the data.

2.2.2 Flagging

Following the initial inspection with LISTOBS and PLOTMS, the data noted as bad (pointing scans, RFI, entire antennas noted to have had technical problems... etc.) must be flagged and removed from further processing. This is done by specifying the specific parameters of the data that require flagging in the CASA task FLAGDATA. Some of the flagging modes available in CASA include ‘quack’, ‘manual’ and ‘Tfcrop’. Quacking is used to remove data at scan boundaries and it can apply the same edit to all scans for all baselines. A common application of quack is to set the parameter ‘quackmode’ to ‘beg’ with a ‘quackinterval’ of 5 or 10 seconds ideally, depending on the ‘sampling time’. This flags the first 5 or 10 seconds of every scan. The ‘quack’ option is particularly useful because it is common for the array to ‘settle down’ at the start of a scan. This approach cannot be used for observation techniques such as On-the-fly-mapping (OTFM), as the settle down occurs at the beginning of a stripe of scans rather than each scan. The parameter ‘Tfcrop’^{2 3} on the other hand is an automated flagging algorithm useful for expunging RFI. The algorithm detects outliers in the 2D time-frequency plane and operates on non-bandpass-corrected data. It iterates through lumps of time and applies statistical tests to find and expunge different types of RFI. The flagging process is fairly subjective and so caution must be taken to avoid excessive or deficient data flagging. Just as the inspection is done throughout the reduction process, there may be need for more flagging as the reduction progresses.

²NCRA Technical Report 202 (Oct 2003)

³<https://casa.nrao.edu/Release3.4.0/docs/userman/UserMansu161.html>

2.3 Synthesis Calibration

Calibration is the process of determining the net complex correction factors that must be applied to each visibility in order to make them as close as possible to what an idealized interferometer would measure, resulting in an accurate picture of the sky when the data is imaged. This is not an arbitrary process, and there is a philosophy behind the calibration process.

The purpose of calibration is to correct the antenna based gains as a function of time and frequency. Some of these corrections are known beforehand while others must be determined from observations of well known calibrators, or from observations of the target itself, i.e. “self-calibration”.

The correction factors used to correct for the various instrumental and environmental variations encountered during observing are derived using models of well known calibrator sources. These correction factors are then applied to the science target. This method relies on the fact that the calibrators are well known sources which can therefore be used to derive scaling factors based on the conditions during the time of observation and this in turn gives correction solutions that can be applied to the target.

The calibration process can be described explicitly, if the true visibility is V and the observed visibility is V' , then in general, the relation between the two is given by:

$$V'_{i,j}(u, v, f) = b_{i,j}(t) [B_i(f, t) B_j^*(f, t)] g_i(t) g_j(t) V_{i,j}(u, v, f) e^{i[\theta_i(t) - \theta_j(t)]} \quad (2.1)$$

Where, for generality:

- f is frequency
- u and v are spatial wave numbers
- g_i & θ_i are the amplitude and phase portions of the complex gain
- B_i is the complex bandpass
- $b(t)$ is the baseline term
- t is time

There are several tasks in CASA that can be used to derive calibration solutions. These solutions are stored in calibration tables. The BANDPASS task solves for the complex bandpass, and can produce channel-binned or polynomial solutions when required. The GAINCAL task solves for complex gain (gaintypes G, T) and delay (gaintype K) and has the option for time-binned or spline solutions. POLCAL is used for polarization calibration including leakage and angle. The SETJY task computes model visibilities with the correct flux density for a specified source. Note that the solutions determined by these tasks are not automatically applied to the observation. Therefore, the task APPLYCAL is used to apply calculated calibration solutions. The task creates a new table by dividing the calibration solutions by the data. In the event that the calibration needs to be undone, the task CLEARCAL re-initializes the calibration for the visibility dataset. The tasks used to examine the calibration solutions are LISTCAL and PLOTCAL which list and plot the calibration solutions, respectively.

At its most basic level, calibration in CASA is the process of taking “uncalibrated” data, setting up the operation of calibration tasks using parameters, solving for new calibration tables, and then applying the calibration tables to form “calibrated” data. Iteration can occur as necessary, with the insertion of other non-calibration steps such as imaging to generate improved source models for “self-calibration”.

2.3.1 Preparing for Calibration; “A Priori” Corrections

There are a number of quantities that may need to be initialized or estimated before further calibration is carried out based on “a priori” information. These are then carried along throughout the calibration process as ‘initial conditions’ of sorts for the subsequent calibration steps. The ones we found necessary here were: corrections of antenna position errors, gain-elevation curves, atmospheric optical depth corrections, requantizer gains, and setting the flux density scale.

2.3.1.1 Correction of Antenna Position Errors

Corrections of antenna position errors account for offsets in the positions of antennas assumed during correlation. Therefore, before anything else, it is imperative to correct for antenna elevation discrepancies because if the antenna positions are wrong the delays will be wrong and as a consequence, the signals to the telescopes will not correlate.

Antenna position corrections are derived using the task `GENCAL`. This makes use of the most recent information in the VLA baseline corrections database. Running the task generates a table which is then applied during subsequent tasks (note that the corrections do not necessarily always exist).

2.3.1.2 Gain Curves

This involves computing antenna gain-elevation dependence. This calibration is necessary because large antennas have a forward gain and efficiency that change with elevation. The gain curve calibration compensates for the effects of elevation on the amplitude of the received signals at each antenna. Due to gravitational deformation of the surface of the antenna, the aperture efficiency (effective collecting area) and net surface accuracy varies with elevation. By design, this effect is least for elevations between 45 and 60 degrees; ergo the gain is maximized in this range but reduces at higher and lower elevations. Furthermore, these deformations represent a greater fraction of the observing wavelength at high frequencies thus making it imperative that this calibration is carried out for these frequencies.

2.3.1.3 Atmospheric Optical Depth

This process adds to the corrections for the signal's elevation dependence. Its primary purpose is to correct the attenuation of the signal by the atmosphere, which is necessary because the troposphere is not completely transparent to radio waves. Moreover, at frequencies greater than 15 GHz, water vapour and molecular oxygen begin to have a substantial effect on radio observations. Three factors contribute to this. Firstly, the radio waves from astronomical sources are absorbed and therefore attenuated before reaching the antenna. Secondly, in as much as a good absorber is a good emitter, significant noise-like power is added to the overall system noise. Finally, the optical path length through the troposphere introduces a time-dependent phase error. Each of these escalates with decrease in elevation because the air mass through which the telescope is looking increases as elevation decreases.

Prior to creating the table in `GENCAL`, the `PLOTWEATHER` task is used to display weather information (see § 3.3.1.3) and calculate the zenith opacities for each spectral window. Following this, `GENCAL` re-computes the correct elevation of the data and creates the calibration table.

2.3.1.4 Requantizer Gain Corrections

When observing with the VLA's 8-bit and 3-bit correlator modes, it is necessary to set the requantizer levels. This optimizes the digital power in the correlator which results in improved signal-to-noise ratio within each spectral window. Higher power in spectral windows can also result from: strong RFI, observing very strong isolated sources, or observing part of the sky with high background power. Such high digital power can lead to large correlation errors and reduce the quality of the data. To avoid these errors, optimizing the digital power by using re-quantization setup scans can become important, if not critical.

Requantizer setup scans are performed each time the frequency setup changes. This introduces small gain changes that must be corrected for before proceeding with the rest of the calibration. The values of the corrections are stored in the Measurement Set but must be converted to a gain table to be applied to the data.

2.3.1.5 Setting Flux Density

The flux density scale is set through the SETJY task which calculates absolute fluxes for a Measurement Set based on known calibrator sources. This is then used at a later stage in the calibration to scale the amplitude of the other calibrators sources and science target field. The task SETJY can calculate the flux density as a function of frequency for several standard VLA flux calibrators. Ideally, SETJY will rescale the flux in the models for known sources to match those it would have calculated. It does this by extrapolating the flux from the frequency band of the model image to the frequency band specified in the Measurement Set. In doing this, it rescales the flux but keeps the structure of the source in the model image. The standard models can be viewed by setting the parameter 'LISTMODELS' to True (see Fig. 3.4). For these known sources, SETJY compares the observed DATA column of the Measurement Set with the Fourier transform of the calibrator model. However, if no model is specified a 1Jy point source at the phase centre is assumed. For "unrecognised" sources, SETJY has the option for the flux density to be manually specified. If the source is not well-modelled as a point source, a model image of that source structure can be used with the flux density scaled by the given values or the flux density can be calculated in SETJY.

2.3.2 Preliminary Phase Calibration

A preliminary phase calibration is performed so that we can define the shape of the bandpass response. We carry this out prior to averaging over time in order to correct for the effects of variations from scan to scan in the phase of the bandpass calibrator. These variations are dependent on the frequency and are particularly likely to take place if the scans happen at much different elevations. This step is quite useful in making low level bad data visible and is discarded once used because we have the phase calibrator which more accurately tracks phase variations whilst observing the science target.

2.3.3 Delay Calibration

Antenna based delays are solved for using Fourier transforms of the spectra on baselines to the reference antenna. This is achieved using GAINCAL and setting the ‘gaintype’ to ‘K’. From the initial inspection, particularly in the task PLOTANTS, a reference antenna is chosen. This antenna needs to be placed somewhere close to the centre of the array; if the dish sizes vary it must be the largest or close to it in the event that the largest does not qualify. This antenna must be present for all scans and contain as little RFI as possible. Solving for the antenna based delays introduces a phase ramp against frequency in each spectral window and the K gain type in the task GAINCAL will solve for the relative delays of each antenna with respect to the reference antenna.

2.3.4 Bandpass Calibration

The purpose of the bandpass calibration is to solve for the instrumental response of the receiver. The CASA task BANDPASS calculates this calibration solution and creates a table that will be applied in subsequent steps along with those created in the former steps. Since the bandpass varies more slowly than the overall gain solved for in GAINCAL, it becomes necessary to use a long time scale when solving for the bandpass. It is also necessary to have high Signal to Noise (S/N) to be able to define the bandpass shape accurately because it varies from channel to channel. Furthermore, with the bandwidth split into channels, variations in frequency will arise due to frequency-dependent effects in signal transmission also making it necessary to solve for the gain variations in both frequency and time as these effects will usually vary on timescales

much longer than the time-dependent effects dealt with by the gain types ‘G’ and ‘T’. Therefore, they are solved for as a separate term, ‘B’.

The bandpass calibration process measures and corrects the frequency-dependent part of the gains $B_{i,j}$ in Eq. 2.1. It determines the variations of phase and amplitude with frequency and accounts for slow time-dependency of the bandpass response. Accurate bandpass calibration is crucial to the detection and accurate measurement of spectral features and can also be the limiting factor in dynamic range of continuum observations (Braatz, 2014).

2.3.5 Gain Calibration

Following bandpass calibration, corrections for the complex antenna gains g_i and θ_i are derived. The absolute magnitude of the gain amplitudes g_i is set when the flux density scale is set using SETJY. The best results for the complex gains of the target are achieved by observing a phase calibrator that is in close proximity to the science target ($< 10^\circ$). This condition is made necessary by the fact that the atmosphere is varying with time and direction, thus we want to minimize the differences between the line of sight to the phase calibrator and the line of sight to the target by observing the closest calibrator to the target.

Ideally, this step should be applied after a constant or slowly-varying bandpass has been determined from the frequency channels. The table generated from the bandpass calibration can then be used as an input in the ‘GAINTABLE’ parameter of the GAINCAL task.

The GAINCAL task determines solutions for the time-based complex antenna gains, for each spectral window, from the specified calibration sources. Systematic time-dependent complex gain errors are almost always the dominant calibration effect and a solution for them is almost always necessary before proceeding with any other calibration. The gain type ‘G’ solves for the net effect of relative amplitude and phase gain for each antenna, phase and amplitude drifts in the electronics of each antenna, amplitude response as a function of elevation (gain curve), and tropospheric amplitude and phase effects. It can handle any per-spectral window multiplicative polarization and time-dependent complex gain effects downstream of the polarizers.

2.3.6 Applying the Calibration

Having derived all the required solutions, the last step in the calibration process is collectively applying the corrections in the calibration tables to the DATA column in the visibility using the task APPLYCAL. This writes a new column called CORRECTED_DATA to the observation Measurement Set. The CORRECTED_DATA column can be processed in much the same way as the DATA column functions such as plotting or imaging or can be split out as the DATA column in a new Measurement Set using the task SPLIT. Among its parameters, APPLYCAL has an option to include all the calibration tables, thus allowing them to be collectively applied.

2.4 Splitting the Data

The task SPLIT is useful for extracting specific portions of the data. It can be used at any point prior to, during and after the calibration to split off data (e.g. spectral windows, fields, scans). Its most common use is most likely at this point, to split off the CORRECTED_DATA column into a new Measurement Set. It reduces the data set size which improves processing times and also provides a way of segmenting the calibration and imaging process.

2.5 Deconvolution and Synthesis Imaging

Our observations are convolved with the dirty beam which is the fourier transform of the sampling function. To account for this, we apply CLEAN which deconvolves the dirty beam from our image.

Mathematically, the sky brightness distribution I , and the visibility data V , are Fourier transform pairs, and are related by,

$$I(l, m) = \int V(u, v) e^{2\pi i(ul + vm)} dudv \quad (2.2)$$

Here,

-the u, v coordinates are baselines measured in units of the observing wavelength

-the l, m coordinates are the direction cosines on the sky (rather than right ascension and declination, for generality)

Through the task CLEAN, CASA will Fourier transform and deconvolve the resulting image.

2.5.1 Imaging (clean)

The CLEAN algorithm, published by Jan Högbom in 1974, is a computational algorithm to perform a deconvolution on images created in radio astronomy. It reconstructs a model image based on interferometer data. This is carried out in the CASA task CLEAN. The CLEAN algorithm is an iterative procedure that operates in the image plane, using the dirty beam to distinguish between real structure and sidelobs on the dirty map (Högbom, 1974). The algorithm iteratively removes, at each step, a fraction of the flux in the brightest pixel in a defined region (called clean regions/boxes) of the dirty image and puts this in the model image. The clean task has several parameters that can be set to get the desired result. These will vary based on the source morphology (point or extended sources) the type of weighting, resolution which is based on the longest baseline and array configuration. The CLEAN task has options to do wide-field and wide-band multi-frequency synthesis imaging and can jointly deconvolve mosaics as well as single fields.

Following the original publication by Högbom, there have been several variations of the clean algorithm. The family of CLEAN algorithms has six variations - three of which are basic and the other three are advanced variations to deal with extended structures. The former include: Högbom, Clark, and MX while the latter include SDI, MRC and Multi. We make use of the Clark CLEAN algorithm which essentially applies the same approach as Högbom but is more efficient. The Clark CLEAN introduces the concept of major and minor cycles. It is based on the fact that the Högbom algorithm essentially convolves the CLEAN components with the PSF and therefore the process can be improved by using Fast Fourier Transforms. For the most part the minor cycle proceeds in the same manner as the Högbom algorithm. In this case however, the only residuals considered are those with an intensity value greater than the product of the fractional height of the PSF's worst exterior sidelobe and the peak residual.

An important part of imaging is applying the appropriate weighting scheme to the (u, v) data. Weighting can be used to improve the dynamic range and fidelity of the image because it can account for the differences in the density of sampling in different parts of the uv-plane. It can account for different noise levels in different samples, or improve sensitivity to extended objects.

Some of the more common weighting schemes are: Natural, Uniform, and Robust. Natural weighting spreads out evenly giving constant weights to all visibilities. This weighting scheme maximises sensitivity for point sources in an image but results in a poor beam shape and sidelobe levels compared to other weighting schemes. On the other hand, uniform weighting minimizes the sidelobe levels but has a noise level that can be a factor of 2 worse than natural weighting. Robust weighting is the intermediate weighting scheme providing an optimal compromise between the two. The weighting scheme can be tailored to suit complex imaging situations by using different combinations of parameters.

The CLEAN task can be run in one of four modes, multi-frequency synthesis (mfs), channel, velocity or frequency. The goal of this study is to identify the continuum emission associated with HII regions and so we use the MFS mode. Multi-frequency synthesis is the practice of using visibility data measured over a range of frequencies when forming a continuum image. Because observing frequency is easier to vary than antenna location, it is an effective way of filling the uv-plane for an observation (Sault & Conway 1999).

2.6 Self-Calibration

After standard calibration there are still residual errors caused primarily by atmospheric variation between the phase calibrator and the target. Self-calibration corrects for this by using a model of the target itself, or nearby bright source, to calibrate the complex gain. This last step is carried out using the tasks CLEAN, GAINCAL, and APPLYCAL. If a sufficient amount of visibility data has been obtained, then the process of self-calibration will use a model created during the standard imaging to alleviate the presence of either or both residual phase and amplitude errors in the data.

In this case, Eq. 2.1 becomes:

$$V'_{i,j}(t) = g_i(t)g_j^*(t)V_{i,j}(t) \quad (2.3)$$

Where g_i is the complex gain on the i^{th} antenna and $V_{i,j}$ is the true visibility.

Upon completing the standard clean, a deconvolved model image is computed. This model can be used to begin a self-calibration loop. The model image is passed back for another calibration

to be carried out on the target. The calibration is much the same as the standard calibrations carried out in § 2.3.5. To begin with, the calibration is focused on phase only calibrations. This is repeated until there is no improvement before an amplitude calibration is included. The overall number of iterations of self-calibration is largely dependent on whether the image is still improving and the iterations can continue until there is no further improvement.

2.7 Data and Image Analysis

CASA has a number of image analysis utilities. Image analysis can therefore be managed through the viewer as well as the image analysis utilities and imaging utilities tools, IA and IM respectively. Some of the tasks that are run through these tools are: IMHEAD, IMMATH, IMSTAT, and IMREGRID. The task IMHEAD summarizes (keywords and values) and manipulates (retrieves and changes values) the header information in an image while the task IMMATH performs mathematical operations on or between images and IMSTAT calculates the statistics on an image or part of an image (can be restricted to a box region, specified channels, or stokes of the cube). The purpose of IMREGRID, is to regrid an image onto the coordinate system of another image. CASA viewer is used as the GUI for interactive cleaning and also has useful region statistics and image plotting capabilities. The images can be displayed in raster, contour, or vector form. Lastly, the data can be exported through the task EXPORTFITS which exports a CASA image table to a standard Flexible Image Transport System (FITS) format.

Chapter 3

Data Reduction on KuGARS Block 5

The KuGARS survey was split into 8 scheduling blocks of data (see § 1.4). Here we use block 5 to perform our initial data reduction, pipeline development, and testing. This block covers the area of sky that contains the massive star forming region W49A (see Fig. 1.1 and also Fig. 5.2). Each block of data contains 44 spectral windows, 32 continuum windows and 10 spectral line windows (Fig. 3.1). Here, we focus on spectral window 2, the lowest frequency in the continuum band.

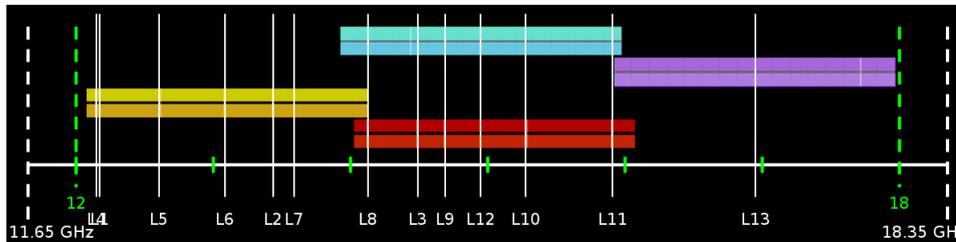


FIGURE 3.1: Correlator Setup during the KuGARS Observations. Total BI. BPs Used: 48 of 64. Total Data Rate: 59.97 MB/s or 215.89 GB/h. Total Spectral Points: 5184. Total Bandwidth: 4.232 GHz. Capability Mode: Shared risk.

3.1 Initial Data Examination

The first step in the data reduction process was examination of the data. This combined looking at figures as well as data plots. The examination is necessary to gain familiarity with the variables used to identify parameters as well as to spot RFI and other forms of bad data.

3.1.1 Observation report (VLA Operator Logs)

VLA operator logs are made available through the NRAO site and can be accessed using the observation date and the project code, these are March 29, 2015 and 15A-043, respectively for block 5-pm. The operator logs start out by showing the array configuration as B, the bands used as X and Ku, and the observation type as Science. They then outline the weather conditions before pointing out the technical events during the observation. Logs of the technical events were particularly important as they pointed out aspects that would need to be corrected for during the data reduction. Antenna 4 had recently updated baseline parameters to correct for errors resulting from its recent relocation and would need to be observed for potential errors. The data on antenna 19 were corrupted because the fringe amplitudes were low due to focus indexing problems. This caused a downtime of 75 minutes (1.9%) of the total time (4049.6 minutes).

3.1.2 Data Layout - listobs

In order to run the calibration process, it is important to become familiar with the variables used to represent the different parameters of the dataset. This was done using the CASA task LISTOBS. The task is run by simply specifying the visibility name. It has options for slightly more specific inputs and outputs to narrow the range of information being looked at or saved.

All of the inputs used in the subsequent steps were specific to the names given in the LISTOBS task. Some of the more common parameters include: 'spw', 'field', 'scan', and 'antenna'. Other useful information included: the number of channels, channel width, total bandwidth, and centre frequency. This process was also useful for selecting the spectral windows to be used for this study.

A summary of these parameters is in table 3.1, and the full LISTOBS output can be viewed in Appendix A. An outstanding feature in these parameters is the large number of fields and sources, this is due to the OTFM technique (see § 1.5.2).

3.1.3 Antenna Layout and (u,v) Coverage

The PLOTANTS task was used to examine the array layout at the time of the observations. This was particularly useful in selecting a reference antenna which aside from being in all the scans

TABLE 3.1: Important parameters from the LISTOBS output.

Parameter	Count	Field Type	Field Name	Field ID
Spectral Windows	44	Flux Calibrator	0137+331=3C48	0
Continuum Spectral Windows	32	Bandpass Calibrator	J2253+1608	1
Channels per Spectral Window	64	Phase Calibrator	J1922+1530	2
Channel width (MHz)	2	Target	specified by "RA+Dec"	3-562
Bandwidth (MHz)	128			
Fields	563			
Sources	23652			

needed to be at the centre of the array. Fig. 3.2 shows exactly how the antennas were placed during the observations.

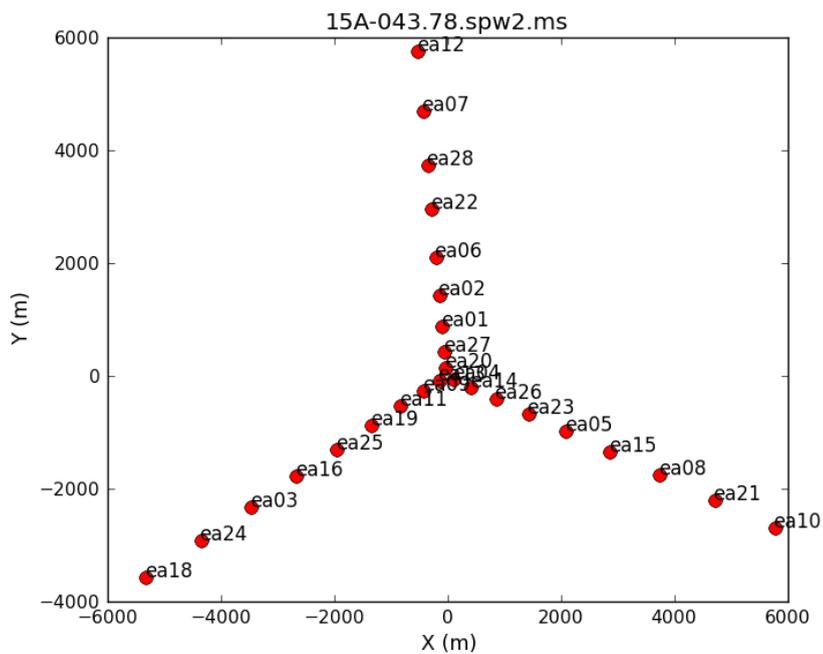


FIGURE 3.2: Output file of the PLOTANTS task. The image shows the setup of the array during the observation.

On this block, ea20 was chosen as reference antenna. It was placed fairly close to the centre and was well behaved during all the scans.

The baseline distribution (to show the (u, v) coverage) can be plotted using two methods. The aptly named CASA task PLOTUV is the first of these methods. However, it was found that this task only plots a single field at a time. Because we are using an OTFM technique, there is only a very short (1s) integration time with 10 integrations per scan. This results in very limited (u, v) coverage per field. Therefore, only the calibrators were plotted using this method. The (u, v)

coverage of the calibrators is shown in Fig. 3.3. Because, the PLOTUV task could not place the entire observation in one plot, this was done with the second method in the task PLOTMS.

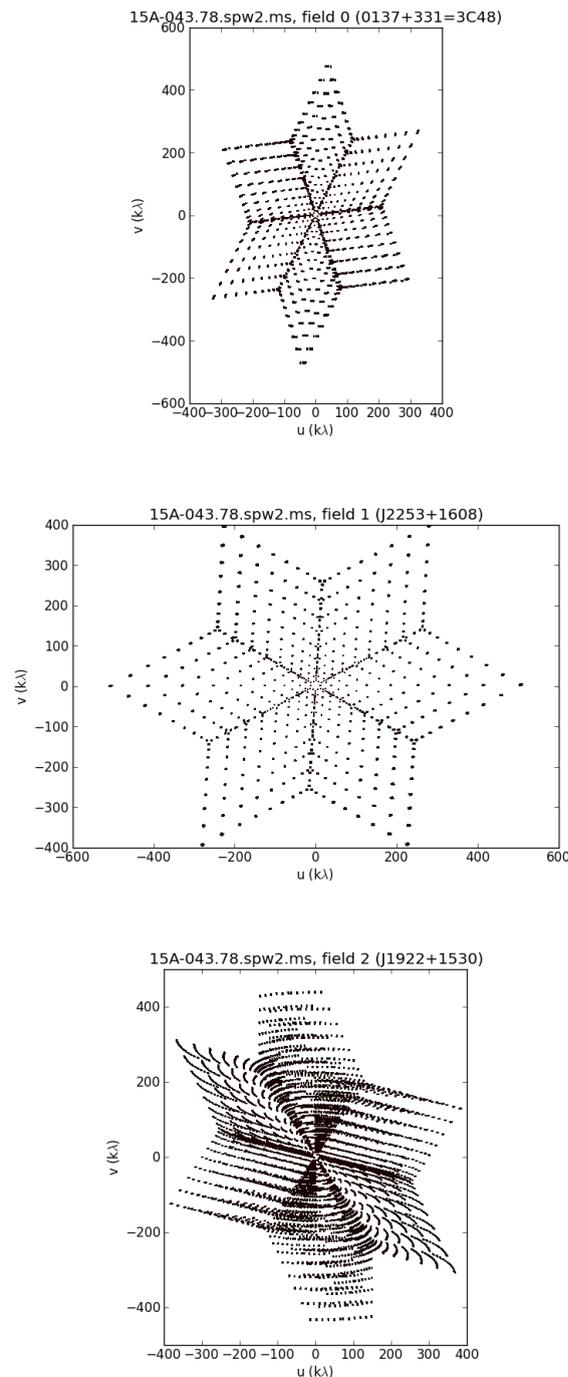


FIGURE 3.3: (u, v) coverage plots of the calibrators. From top to bottom: the flux calibrator 3C48, the bandpass calibrator J2253+1608, and the phase calibrator J1922+1530.

The second method used PLOTMS to plot the (u, v) coverage. Using PLOTMS the range was not restricted to a single field, allowing for the (u, v) coverage of the entire spectral window (or more should it be desired) to be put on a single plot as shown in Fig. 3.4.

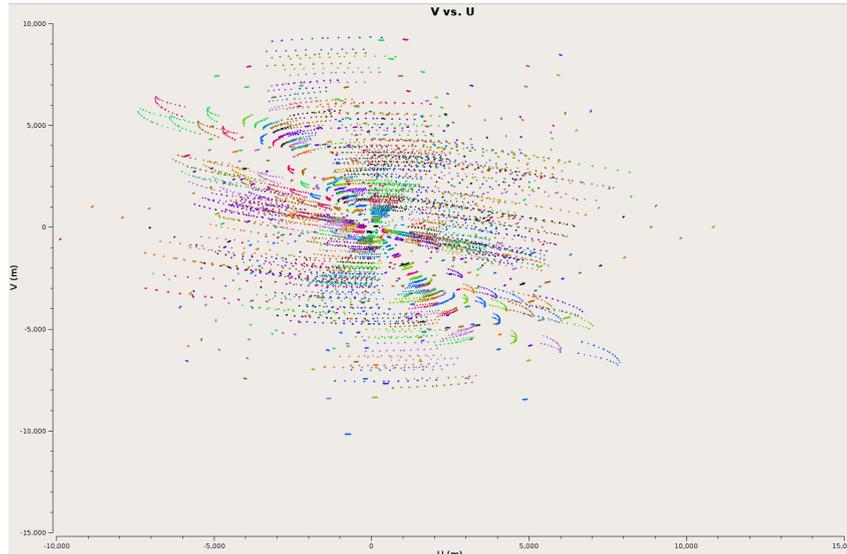


FIGURE 3.4: Combined plot of the (u, v) coverage of the calibrators and target in plots with baseline as colour axis. The x-axis represents the U coordinates, ranging from -10,000 to 15,000 while the y-axis represents the V coordinates, ranging from -15,000 to 10,000. Both axes are in units of metres.

3.1.4 Data layout plots (plotms)

The task PLOTMS was used to view the data in different configurations. While the task LISTOBS gave a general outline of the data and the tasks PLOTANTS and PLOTUV gave the antenna layout and baseline distribution, respectively, PLOTMS gave a graphical outline of the (u, v) data layout. Plotting the (u, v) data in time vs. amplitude allowed the amplitude to be monitored throughout the observations. RFI is often apparent on short timescales which appear as easily recognisable spikes in the time vs. amplitude plots. In a similar manner, plotting (u, v) data in amplitude vs. frequency was used to identify RFI. This is because RFI is often confined to a narrow frequency range. Plotting antenna2 against time allowed antenna drop outs to be identified when the amplitude drops to zero. This was helpful in determining the fidelity of the chosen reference antenna.

The observation was set up to first observe the primary calibrator (Fig. 3.5: pink line), and then move on to the bandpass calibrator (Fig. 3.5: black line) after which the observation would alternate between the phase calibrator (Fig. 3.5: orange line) and the target. While alternating between the phase calibrator and the target, the pointing would first be focused on the phase calibrator before observing a stripe of 54 steps in OTF mode. Following the 54 steps the pointing would focus back on the phase calibrator before proceeding to the next 54 steps. Each of the 8(am and pm) scheduling blocks consisted of a total of 10 stripes of 54 target steps bridged by



FIGURE 3.5: Data stream plot with antenna2 in the y-axis ranging from 1 to 25 and time in the x-axis, from 16:06:40 to 18:53:20 with ea01 as reference point and field as colour axis.

a phase calibrator pointing (see § 1.4). The pattern this represents in a single scheduling block can be seen in Fig. 3.6.

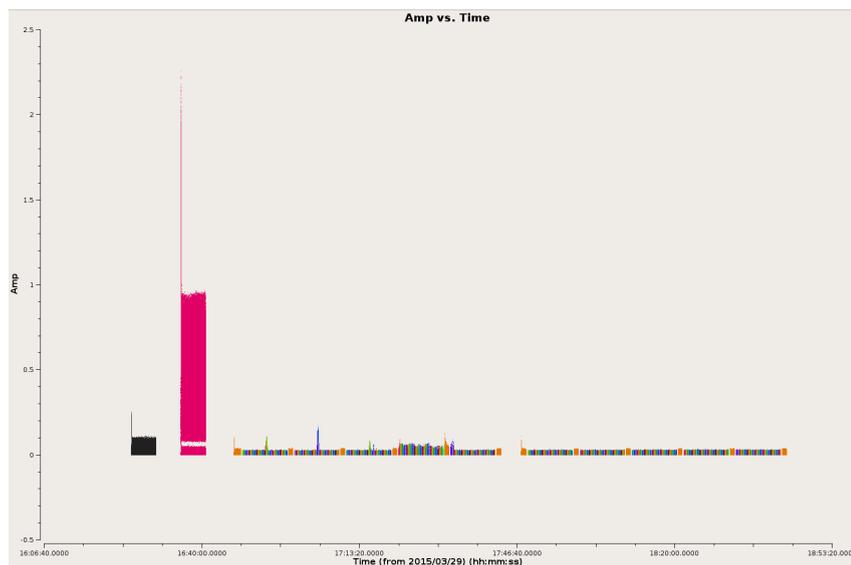


FIGURE 3.6: A plot of amplitude as a function of time of the raw data in spw 2 showing all the calibrators and the target prior to RFI flagging. The y-axis shows amplitudes from -0.5 to 2.5 while the x-axis covers time from 10:36:40 to 18.33.20.

3.2 Initial Flagging (flagdata)

The observation was set up so that configuration scans preceded each calibrator pointing and each observation stripe. The spikes at the edges of the data stripes in Fig. 3.6 appear as a

consequence of the configuration scans that precede each stripe (see § 1.4). These scans were put into FLAGDATA as a python string and flagged. The specifics of the configuration scans were obtained using LISTOBS and the locate function in PLOTMS. Spotting and specifying the parameters of the configuration scans for flagging is fairly straightforward. However, the configuration scans are not the only bad data. Fig. 3.7 which was plotted after flagging the configuration scans, shows evidence of RFI. The bad data is mixed in with both the calibrators and the target. Further inspection in different configurations is necessary to determine exactly where the RFI lies.

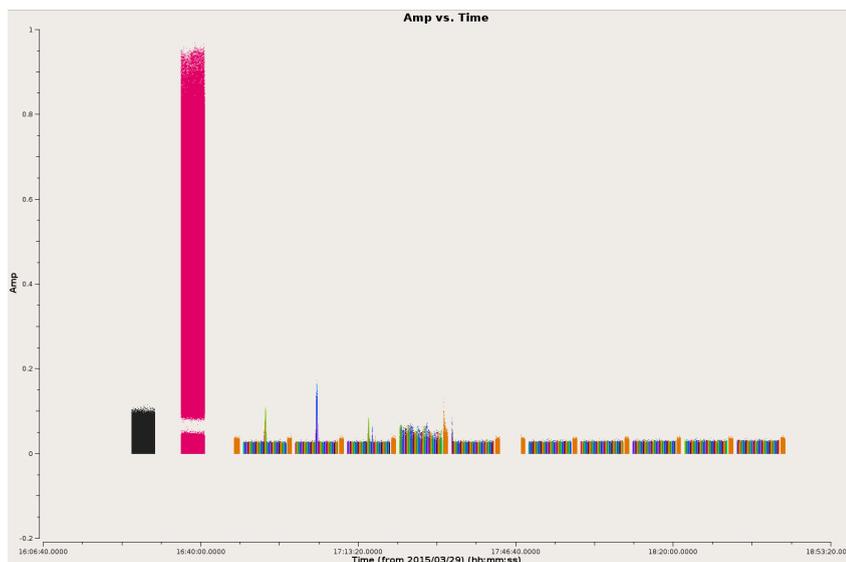


FIGURE 3.7: Amp-time plot after flagging configuration scans with field as colour axis. The y-axis shows amplitudes from -0.2 to 1 and the x-axis shows time from 16:04:40 to 18:53:20.

Fig. 3.7 shows that the remaining RFI is predominantly in the phase calibrator and the target, therefore these will be plotted separately in order to find and expunge the remaining RFI effectively. The plots of the phase calibrator in Fig. 3.8 show RFI in both amplitude plotted as a function of time and as a function of frequency. Here, just as with the configuration scans, the locate task in PLOTMS was used to determine the specifics of the observed RFI which would be used as inputs in the flag commands.

Unlike the phase calibrator which showed RFI in both time and frequency, the target only showed distinct evidence of RFI in the amp-time plot - particularly in the fourth stripe (see Fig. 3.9). Extra caution needs to be taken when flagging the science scans because of the moving beam which means that short spikes of the amplitude in time may be the beam travelling over a real source and not RFI. In traditional pointed observations, we do not expect to see such spikes with time being caused by the science field.

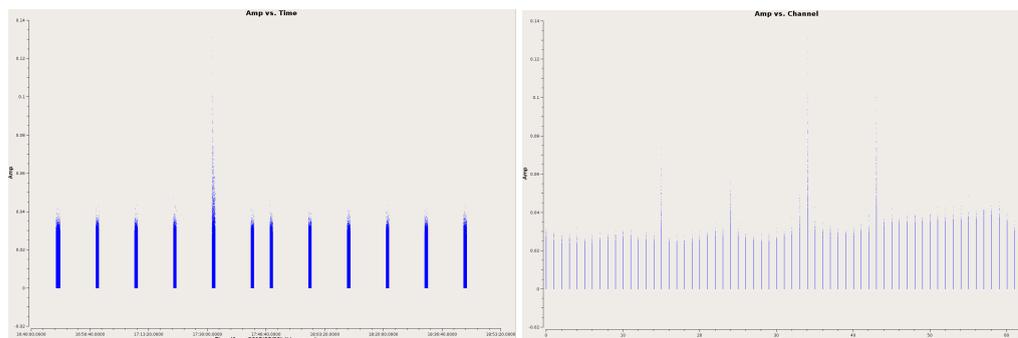


FIGURE 3.8: Inspection plots of the phase calibrator – amp v time (left) and amp v channel (right). Both y-axes show amplitudes from -0.02 to 0.14, left x-axis shows time from 16:40:00 to 18:53:20 and the right x-axis shows channels 0 to 61.

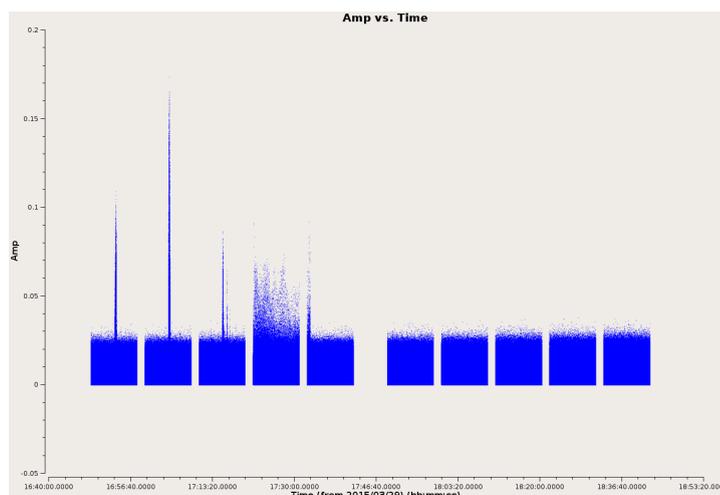


FIGURE 3.9: Amp-Time plot of the target showing distinct RFI features. The y-axis shows amplitudes from -0.05 to 0.2 and the x-axis shows time form 16:40:00 to 18:53:20.

The inspection plots showed no further RFI features in the primary and bandpass calibrators but showed traces in the phase calibrator and a burst of RFI in the fourth stripe of the target as well as smaller traces in some of the other stripes. However, before flagging the bad data, there were a couple of things to note. Firstly, the target has 3 distinct spikes that at first glance appear to be RFI features –further investigation revealed that these are in fact real source features. Attempting to flag these spikes resulted in reduced flux in the brightest sources of the final image. Furthermore, all three spikes appeared over the i th, j th, and k th steps of their respective stripes, that is, within a second of Right Ascension at different Declinations. In the subsequent seven stripes the spikes disappear, implying that the beam had now moved to declinations above the source. Had they been caused by RFI, the spikes would have been present in all 10 stripes. Therefore, the spikes were not flagged. Secondly, to avoid ‘over flagging’, the locate function in PLOTMS was used to determine the specific range of data that was to be flagged. The details of the data selected through the locate function were displayed in the CASALOGGER. These showed

that only some channels in the affected scans actually had RFI. We therefore constrained the flag command by specifying not only the scans but also the channel ranges affected. Therefore, all the data noted as ‘bad’ were specified as strings in various parameters in the FLAGDATA task. To avoid flagging good data by flagging the entire antenna, field or spectral window in which the RFI occurred, the flag command was constrained to the range of affected channels (e.g. channels 15, 24, 34 and 43 on scan range 192 to 245 and 250 to 251). As a final step in the initial flagging phase, the automated flagging algorithm ‘Tfcrop’ was applied (see § 2.2.2). The collective effect of this procedure resulted in flagging $\sim 12\%$ of the data.

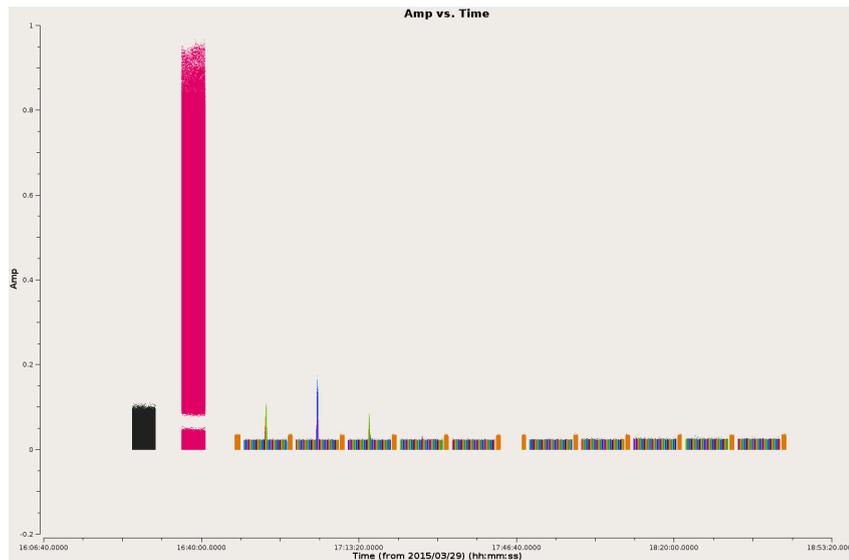


FIGURE 3.10: Amp-time plot of the calibrators and target after flagging. The y-axis shows amplitudes from -0.2 to 1 and the x-axis shows time from 16:06:40 to 18:53:20.

Fig. 3.10 shows the data after the RFI has been expunged. The plot shows 4 distinguishable features; the 3 spikes earlier noted as bright parts of the target rather than RFI and a gap in the brightest peak (pink - which is the bandpass calibrator). The gap appearing between the higher and lower parts of the bandpass calibrator occurs because the RL and LR correlations have lower amplitudes than the RR and LL correlations. The gap is the region over which they intersect.

3.3 Calibration

Having performed some basic editing of the data through these flagging processes based on ‘a priori’ information and analysis of the (u, v) data in PLOTMS, the data was ready for calibration.

3.3.1 Preparing for Calibration with “a priori” Corrections

Prior to calibration, some quantities were initialized based on “a priori” information (see § 2.3.1). These were the antenna positions, gain curves, atmospheric optical depth and requantizer gains.

3.3.1.1 Antenna Position Corrections

In § 3.1.1 the operators log shows that antenna ea04 had recently updated baseline parameters to correct for errors resulting from its recent relocation. In order to avoid calculating u and v in Eq. 2.1 incorrectly (due to incorrect baselines) and as a consequence causing errors in the image, all improved baseline positions would need to be incorporated. The calculations were inserted via the CASA task GENCAL.

3.3.1.2 Gain Curves

Gain curve calibration corrects for the effects of elevation on the amplitude of the received signals at each antenna (see § 2.3.1.2). Antennas are not absolutely rigid, and so their effective collecting area and net surface accuracy vary with elevation as gravity deforms the surface. The task GENCAL was also used to generate the ‘gaincurve’ table. This time the parameter ‘caltype’ was set to ‘gc’ to instruct the task to write gain curve corrections into this table.

3.3.1.3 Atmospheric Optical Depth

The atmospheric optical depth correction is made necessary because the troposphere is not completely transparent and as a consequence attenuates the signal (see § 2.3.1.3). The PLOTWEATHER task displays weather information and calculates the zenith opacities for each spectral window as shown in Fig. 3.11.

Aside from displaying this information, a variable (opac_var) was used to capture the information generated by PLOTWEATHER. Thereafter, opac_var was used to determine the solutions in the opacity table. Yet again, the task GENCAL was used to generate the opacity corrections. This time with ‘caltype’ set to ‘opac’ and parameter set to ‘opac_var’.

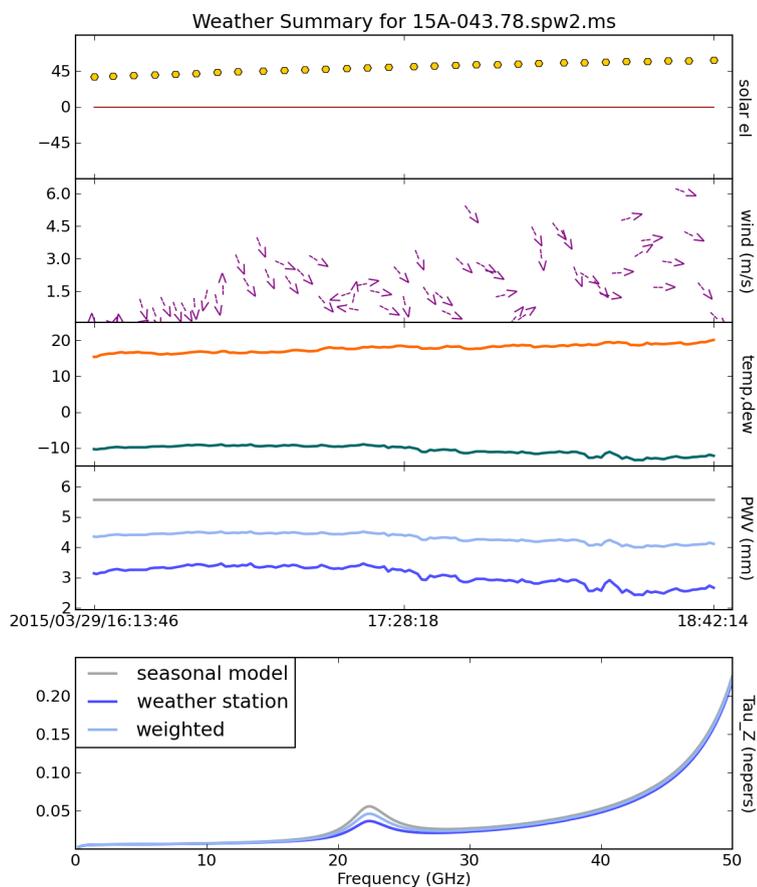


FIGURE 3.11: Plot weather output showing the weather summary for the measurement set during correction for atmospheric optical depth.

3.3.1.4 Requantizer Gain Corrections

Finally, the requantizer gain corrections which are stored in the measurement set were converted into a calibration table. This was done by setting 'caltype' to 'rq' which would instruct the task GENCAL to write the requantizer gain corrections to the table.

3.3.1.5 Setting the Flux Density Scale

To set the flux density scale, the task SETJY'S parameter LISTMODELS was set to 'True'. This brought up a list of standard models, as in Fig. 3.12.

The model was specified to 3C48_K.im. This would scale the flux accordingly and set the flux value (Fig. 3.13) to the value computed for the frequency at which the observations were carried out (see § 2.3.1.5).

```

Candidate modimages (*) in /local/home/mubela/soft/casa-release-4.6.0-el6/data/nrao
/VLA/CalModels:
3C138_A.im 3C138_U.im 3C147_Q.im 3C286_K.im 3C48_A.im 3C48_U.im
3C138_C.im 3C138_X.im 3C147_S.im 3C286_L.im 3C48_C.im 3C48_X.im
3C138_K.im 3C147_A.im 3C147_U.im 3C286_Q.im 3C48_K.im README
3C138_L.im 3C147_C.im 3C147_X.im 3C286_S.im 3C48_L.im
3C138_Q.im 3C147_K.im 3C286_A.im 3C286_U.im 3C48_Q.im
3C138_S.im 3C147_L.im 3C286_C.im 3C286_X.im 3C48_S.im

```

FIGURE 3.12: Terminal output of SETJY task with parameter ‘listmodels’ set to ‘True’. Each calibrator source has a model listed in multiple bands.

```

{'0': {'0': {'fluxd': array([ 1.9317112,  0.          ,  0.          ,  0.          ]),
            'fieldName': '0137+331=3C48'},
      'format': "{field Id: {spw Id: {fluxd: [I,Q,U,V] in Jy}, 'fieldName':field name }}"

```

FIGURE 3.13: Terminal output of SETJY task showing the flux density value assigned to the flux calibrator.

3.3.2 Initial Phase Calibration

If significant phase drift occurs over the duration of the bandpass calibrator scan, this needs to be corrected prior to averaging over time. Therefore, in preparation for the bandpass calibration an initial phase calibration was performed on the bandpass calibrator source. This is a “disposable” calibration carried out before the bandpass calibration in order to moderate the effects of gain variations from scan to scan on the bandpass calibrator. Phase variations unlike amplitude variations would have adverse effects on the bandpass solutions making it necessary to solve for any phase variations with time and prevent decorrelation when vector averaging the data in computing the bandpass solutions. The CASA task used for this was GAINCAL. The task was run with each of the tables generated as “a priori” corrections. The tables were specified as a python list in the parameter ‘GAINFIELD’.

Following the initial phase calibration, the correction tables were inspected via the CASA task PLOTICAL as shown in Fig. 3.14. The ‘caltable’ was set to the table generated from the GAINCAL. The solutions were inspected by viewing a single antenna at a time in plots of phase as a function of time. All of the antennas had fairly smooth phase variations with time, indicating that the data was fine and would behave well during the calibration.

3.3.3 Delay Calibration

Earlier (§3.1.3) ea20 was chosen as reference antenna. Among the checks carried out in the process of choosing it, was making sure it did not drop out in any of the scans in the observation and that it was fairly well behaved on all baselines. This all becomes important when it is used as

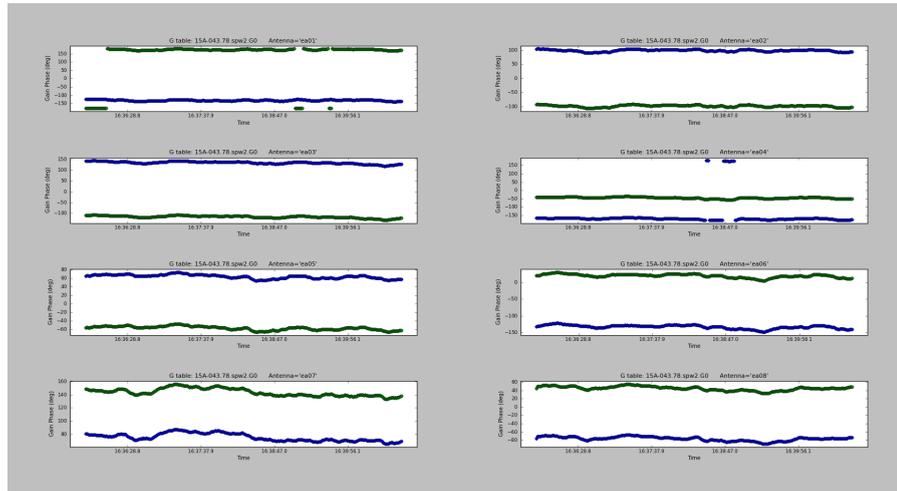


FIGURE 3.14: Plots of some of the antennas showing the variation of phase with time on the initial phase calibration table. The y-axes show Gain phases over an 80 degree range while the x-axes show time from 16:36:28.8 to 16:39:56.1.

reference when solving for delays, as the K gain type in GAINCAL solves for the relative delays of each antenna with respect to the reference antenna (ea20). Rather than a full global delay, this gave one value per spectral window per polarisation. In the same manner as the initial phase calibration, the “a priori” corrections were pre-applied. As an addition, the calibration was set to cut off the edge channels because they have lower sensitivity. Following the calibration, the solutions were inspected through plots of the delays as a function of time (Fig. 3.15) and antenna (Fig. 3.16) respectively.

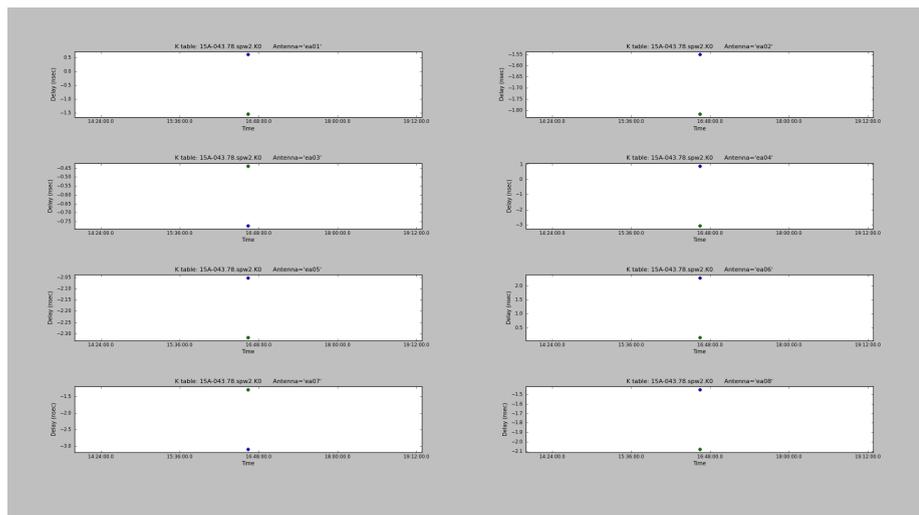


FIGURE 3.15: Time delay plot on delay calibration table. The y-axes show the delays in nano seconds while the x-axes show time.

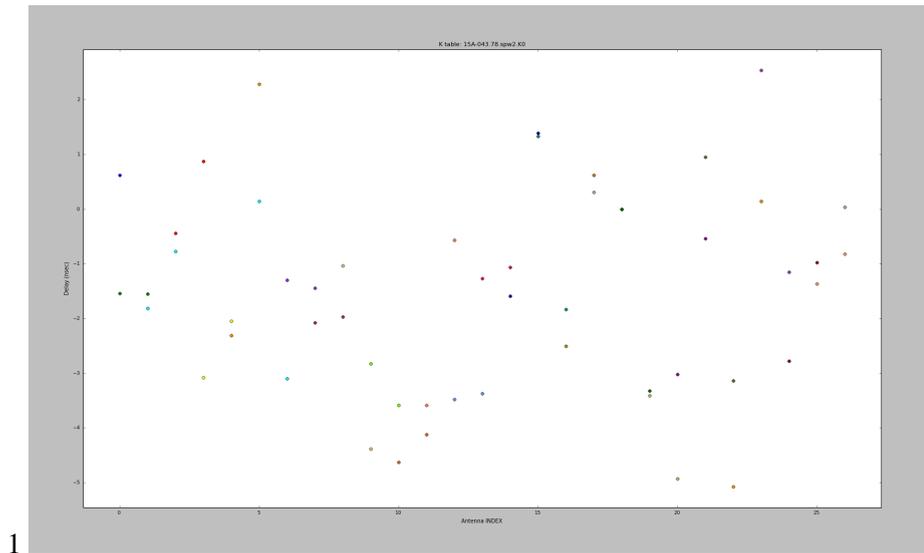


FIGURE 3.16: Antenna delay plot on delay calibration table. The y-axis shows the delay in nano seconds from -5 to 2 and the x-axis shows the antenna index from 0 to 26.

3.3.4 Bandpass Calibration

Before the bandpass calibration, we examine a plot of the data in the pre-calibration state (Fig. 3.17). In an “ideal setting” the visibility data should be constant as a function of frequency. This is not quite the case here. The variations with frequency are a reflection of the slightly different antenna bandpasses. The effect of correcting for the bandpass can be seen by comparing the plot in Fig. 3.17 to the plot in Fig. 3.25.

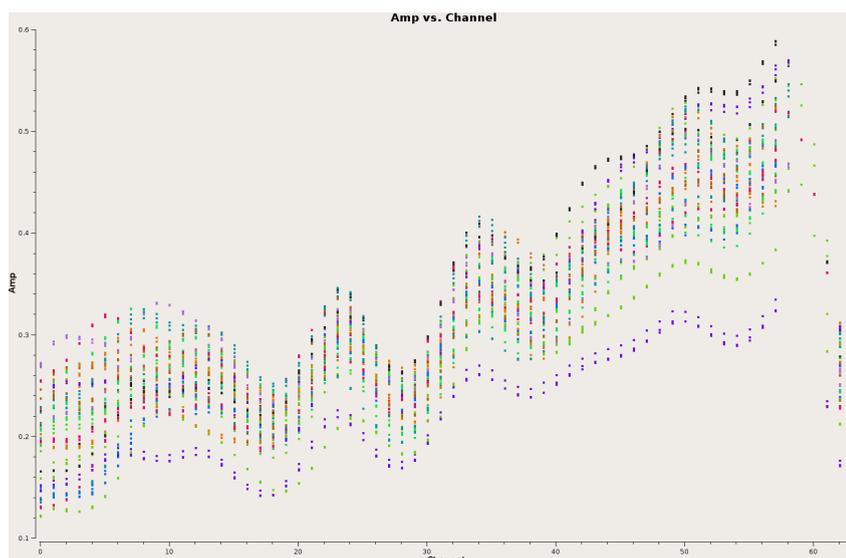


FIGURE 3.17: Plot of bandpass prior to calibration (with antenna ea01as reference) showing right circularly polarised data for the bandpass calibrator source (J2253+1608). Colour coded by spw and averaged over all baselines. In the y-axis, the amplitudes range from 0.1 to 0.6, in the x-axis are the channel numbers from 0 to 63.

Therefore, the bandpass solutions were formed after applying the phase solutions derived in the former steps. The CASA task BANDPASS was used. Again, the “a priori” corrections, phase solutions and delays were pre-applied. PLOTICAL was used to display the solutions after the calibration and as expected the bandpass phases were relatively flat (Fig. 3.18).

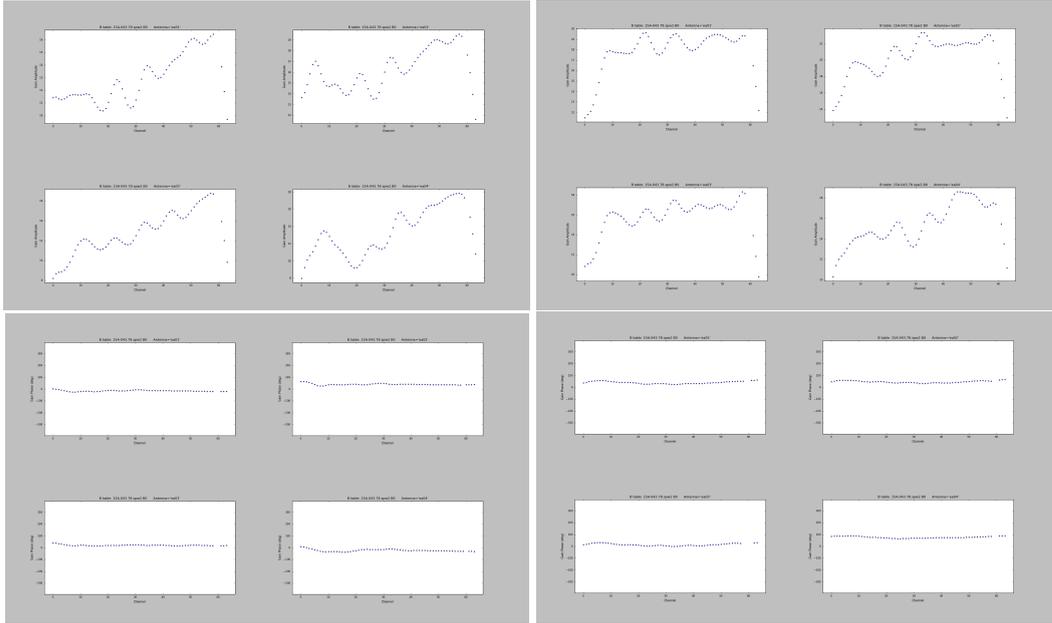


FIGURE 3.18: Top right and left show bandpass amplitudes for J2253+1608 Right and Left Polarisation, respectively while bottom right and left show bandpass phases for J2253+1608 Right and Left polarisations, respectively. The x-axis range is 0-64 and the y-axis range is -300 to 300.

3.3.5 Gain Calibration

When correcting for the complex antenna gains, as discussed in (§ 2.3.5), the absolute magnitude was determined by referring to a standard flux density calibrator. In this case this was the model of 3C48 set in § 3.3.1.5. The gain calibration was carried out in two steps. The first was merely the derivation of the appropriate complex gains g_i and θ_i , for the flux density calibrator 3C48 (0137+331). The field was restricted to only 3C48 at this point. The idea was to relate the measured data values to the presumed flux density of 3C48. Consequently, the complex gain solution included both amplitude and phase. The second step was a run of GAINCAL, in which the phase calibrator was used to determine the appropriate complex gains for a direction of the sky close to the target source. As in the other parts of the reduction process, the solutions from the calibration were inspected. Here, PLOTICAL was used for plots of amplitude and phase as functions of time. The amplitude-time plot (Fig. 3.19), shows most of the points lying around

0.2. One group of points has a much higher value at 1.0. This is because these points all belong to the bandpass calibrator (J2253+1608) which has shown higher amplitudes in previous points.

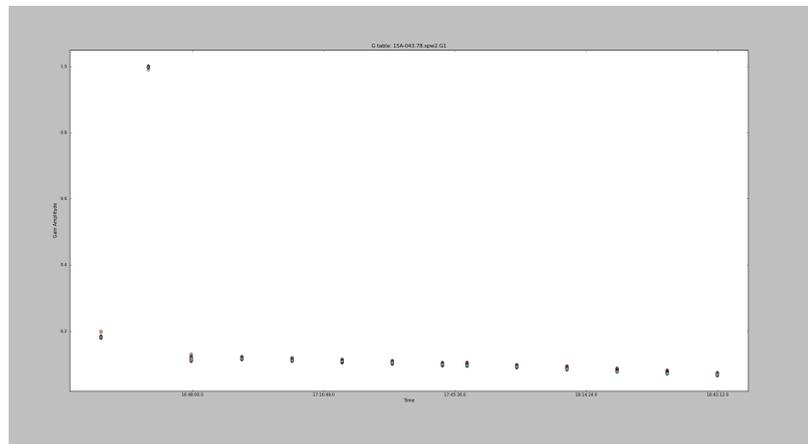


FIGURE 3.19: Amp-time plot of complex gain calibration table. The y-axis shows amplitudes from 0 to 1 and the axis shows time from 16:04:40 to 18:43:12.

Good phase solutions are expected to show a relatively smooth variation in time as the atmosphere varies. Fig. 3.20 shows smooth phase solutions and even though there was a bit of drift in some places, there were no sharp phase jumps indicating that ea20 was in fact, a good choice for reference antenna.

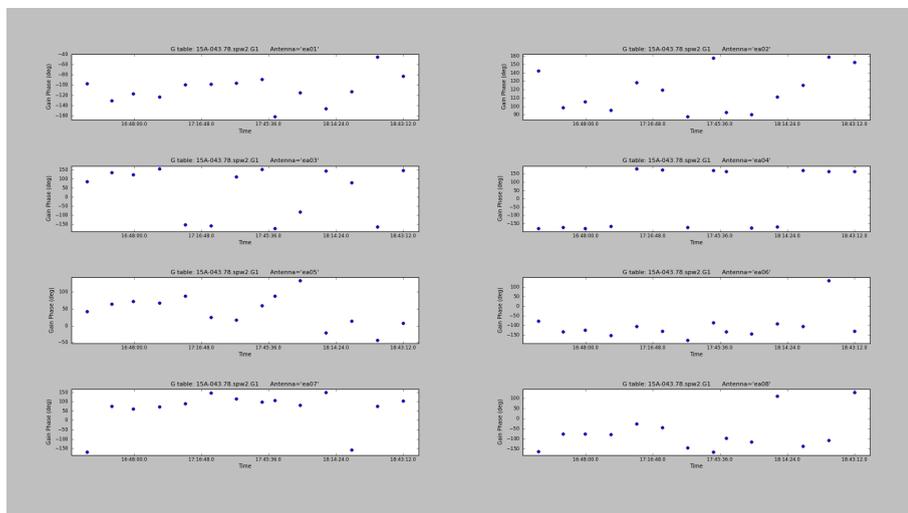
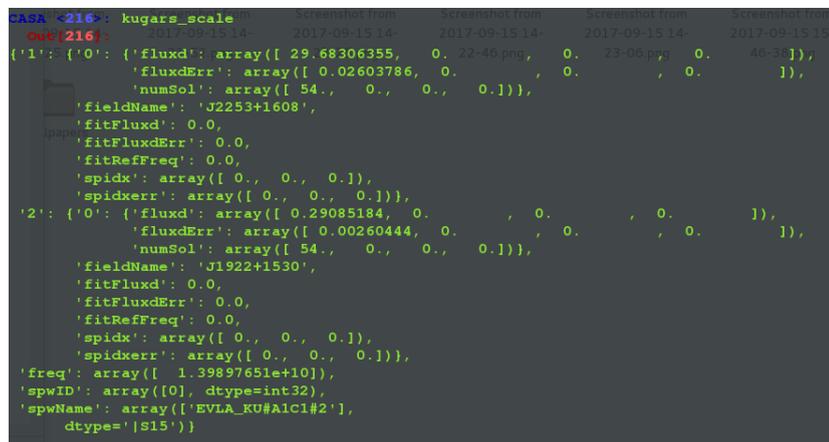


FIGURE 3.20: Phase-time plots of complex gain calibration table. The y-axes show gain phases over 300 degrees and the x-axes show time from 16:04:40 to 18:43:12.

3.3.6 Scaling the Amplitude Gains

We know the flux density of our primary calibrator 3C48, which we set previously using SETJY. However, we do not know the flux density of the secondary calibrators. It is reasonable to assume that the mean gain amplitudes for the primary calibrator are the same for the secondary calibrator. We can, therefore, bootstrap the flux of the primary calibrator to the secondary calibrator through the task FLUXSCALE. Fig. 3.21 shows the output of fluxscale stored in the specified variable. The output shows the result from bootstrapping the flux of 3C48 (field 0) on to J1922+1530 (field 2) and J2253+1608 (field 1).

The task FLUXSCALE takes the complex gain table as the caltable and uses the primary calibrator as a reference. FLUXSCALE returns a dictionary of results which can be captured in a variable (in this case kugars_scale - see Fig. 3.21).



```

CASA ~ 216: kugars_scale
Out(216):
{'1': {'0': {'fluxd': array([ 29.68306355,  0.2246,  0.2306,  0.4638]),
            'fluxdErr': array([ 0.02603786,  0.02603786,  0.02603786,  0.02603786]),
            'numSol': array([ 54.,  0.,  0.,  0.])},
      'fieldName': 'J2253+1608',
      'fitFluxd': 0.0,
      'fitFluxdErr': 0.0,
      'fitRefFreq': 0.0,
      'spidx': array([ 0.,  0.,  0.]),
      'spidxerr': array([ 0.,  0.,  0.])},
      '2': {'0': {'fluxd': array([ 0.29085184,  0.29085184,  0.29085184,  0.29085184]),
            'fluxdErr': array([ 0.00260444,  0.00260444,  0.00260444,  0.00260444]),
            'numSol': array([ 54.,  0.,  0.,  0.])},
      'fieldName': 'J1922+1530',
      'fitFluxd': 0.0,
      'fitFluxdErr': 0.0,
      'fitRefFreq': 0.0,
      'spidx': array([ 0.,  0.,  0.]),
      'spidxerr': array([ 0.,  0.,  0.])},
      'freq': array([ 1.39897651e+10]),
      'spwID': array([0], dtype=int32),
      'spwName': array(['EVLA_KU#A1c1#2'],
                      dtype='|S15')}

```

FIGURE 3.21: Dictionary results generated while running the task FLUXSCALE.

To inspect the scaling of the amplitude gains, a plot of amp against time was made on the 'fluxscale' table.

3.4 Apply Calibration

With all the calibration solutions having been derived and verified, the CASA task APPLYCAL was used to apply the solutions to the data.

Initially, the calibration solutions were applied to each calibrator using the gain solutions derived on that calibrator alone to compute the so-called corrected data. This was done by specifying the calibrator in the field name parameter as well as including the appropriate calibration tables

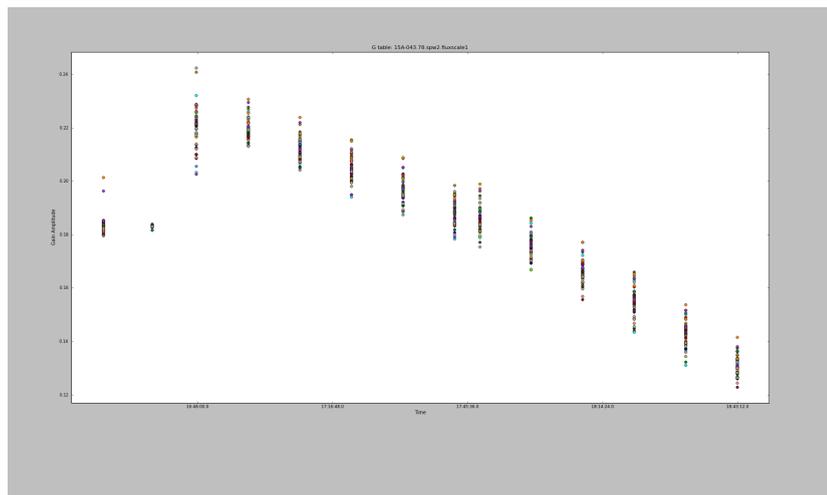


FIGURE 3.22: Amp-time plots of the fluxscale table. The y-axis shows amplitudes from 0.12 to 0.24 and the x-axis shows time from 16:04:40 to 18:43:12.

in the ‘gainfield’ parameter—that is: the fluxscale, delay, bandpass and the “a priori” correction tables. Then the calibration was applied to the target fields in the mosaic, this time by linearly interpolating the gain solutions from the secondary calibrator, J1922+1530. The interpolation was necessary because the gains on J1922+1530 were derived at a different time and in a different position on the sky from the target. Nonetheless, the amplitudes and gains derived from the phase calibrator were applied under the assumption that J1922+1530 was close to the target source in the sky and that the gains applicable to the target source were very similar to those derived in the direction of J1922+1530. Also note that the corrections applied to the source are the interpolated values derived from the calibrators. Thus, although the calibrators are observed at different times, as long as the observations are performed on short enough intervals, the interpolated values accurately track the atmospheric conditions. Therefore, if the phase calibrator is in close proximity to the target, changes induced by atmospheric variations can be accurately modelled and corrected for resulting in good calibration solutions.

Following this was yet another inspection of the data. Again, this was performed using plotms and the most apparent change in the parameters used was the ‘ydatacolumn’ - this was now specified as corrected (i.e. to plot the calibrated data). Examining these plots revealed no signs of residual bad data.

Although the plots revealed no bad data, the gap feature in the bandpass (Fig. 3.23 pink) still remained. It was therefore taken that the explanation given earlier sufficiently explained its existence. The gap is merely being caused by the region over which the RL and LR correlations which have lower amplitudes intersect with the RR and LL correlations which have higher

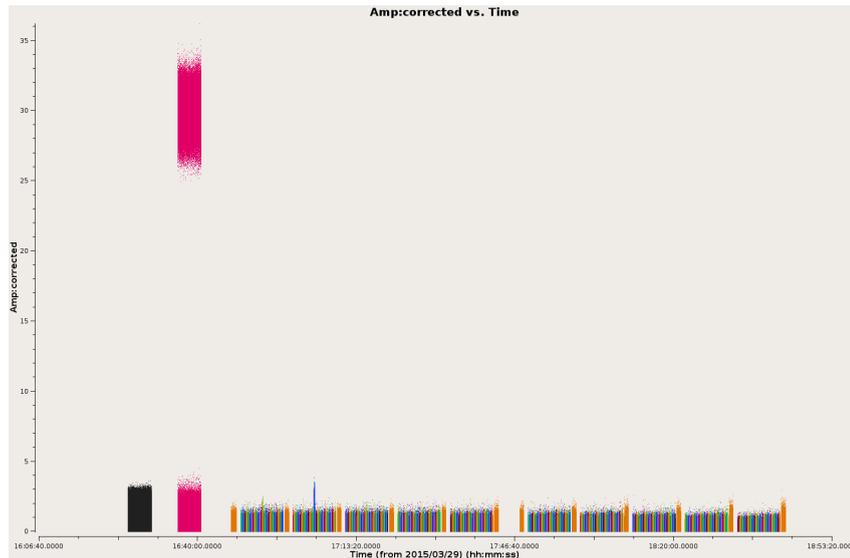


FIGURE 3.23: Plot of amplitude against time for the entire spectral window after calibration - with field as colour axis. In the y-axis, the amplitudes range from 0 to 35, in the x-axis is time from 16:05:40 to 18:53:20.

amplitudes.

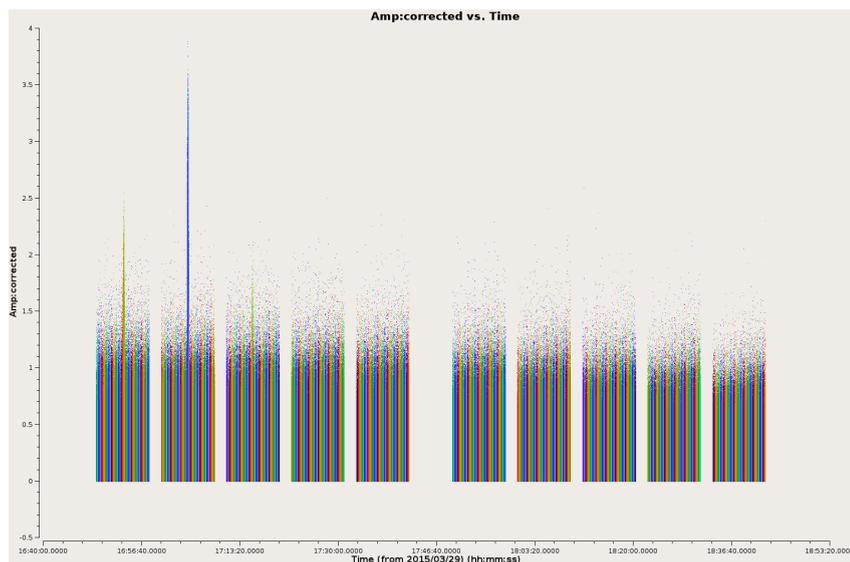


FIGURE 3.24: Amp-time plot of the target after calibration with field as colour axis. N.B the spikes in the first 3 blocks are caused by the target (see § 3.2). In the y-axis, the amplitudes range from -0.5 to 4, in the x-axis is time from 16:40:00 to 18:53:20.

Fig. 3.24 shows a change in the amplitudes of the target when compared to Fig. 3.9 which was a plot of unflagged and uncalibrated data. Aside from the bad data having been expunged it can be seen that the amplitudes in Fig. 3.24 are much higher than those of Fig. 3.9.

Finally, in taking a look at the bandpass, there is a significant difference in Fig. 3.17 and 3.25. The ideal bandpass is expected to have constant amplitude across the entire frequency. We saw

in Fig. 3.17 that this was not quite the case due to the different bandpasses of the individual antennas. As can be seen in Fig. 3.25, there is still some residual variation in the amplitude as a function of frequency for our bandpass calibrator. However, the plot shows nearly constant amplitude over most of the frequency range.

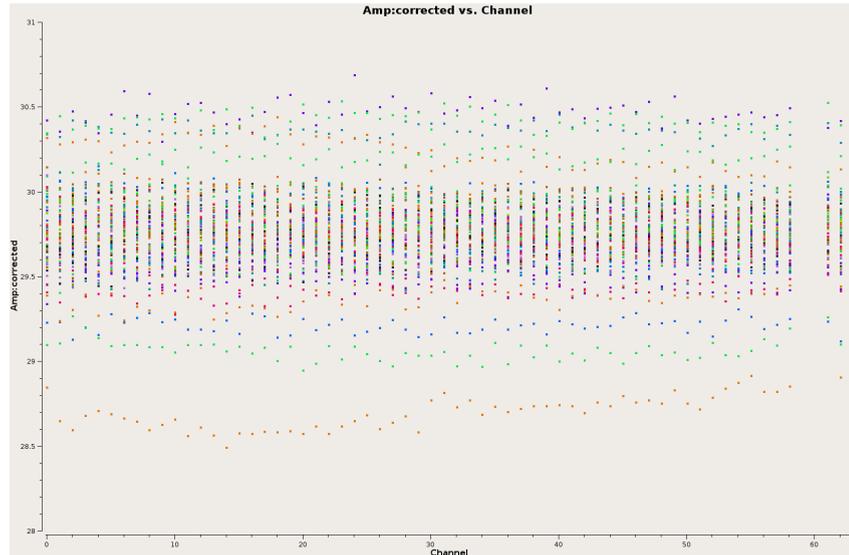


FIGURE 3.25: Plot of bandpass after calibration (with antenna ea01 as reference) showing right circularly polarised data for the bandpass calibrator source (J2253+1608). Colour coded by spw and averaged over all baselines. In the y-axis, the amplitudes range from 28 to 31, in the x-axis are the channel numbers from 0 to 63.

At this point, since the calibration had been applied to the target and the target was seen to no longer contain bad data, the science targets were now split off to create a new measurement set containing all the target fields. This was done using the CASA task SPLIT.

3.5 Deconvolution and Imaging

The process of deconvolution and imaging was carried out using the CLEAN task described in § 2.5. The imaging consisted of a cycle of clean and view stages that are described in the subsequent sections. The clean was performed interactively and the regions to be cleaned were varied. The cleaning was stopped once all the sources had been removed from the residual image. Clean basically subtracts the dirty beam until the residual image no longer shows any signs of real emission. At this point the model clean components are reinserted into the residual map after having been convolved with a gaussian corresponding to the synthesised beam of the observation.

We determine the resolution of our observations using the 'uvwave' plot in Fig. 3.26. The plot shows high amplitudes on the short baselines and lower amplitudes on the longer baselines implying that there is extended structure in the image.

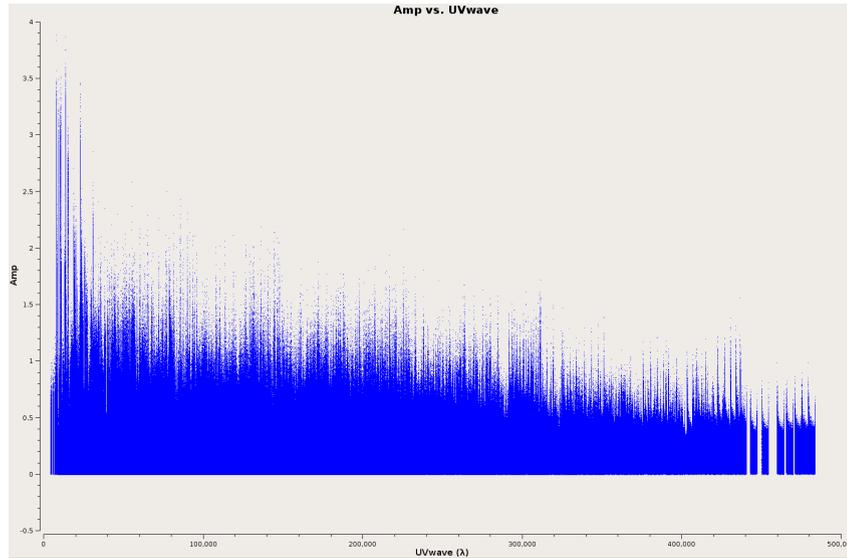


FIGURE 3.26: Power distribution across baseline length. In the y-axis, the amplitudes range from -0.5 to 4, in the x-axis are the baselines lengths in units of uv wave (lambdas) from 0 to 500,000.

The VLA's resolution is diffraction-limited so the resolution is set by the array configuration and the observing frequency. We therefore based our spatial resolution on the longest baseline and sampled it by 5 pixels per beam to obtain the actual value used as the cell size (size of each pixel in arcseconds). Note that, at 5 pixels per beam, we are oversampling our synthesised beam since Nyquist sampling requires 3 pixels per beam. This ensures that we recover all the available information from our observations

Using the OTF observing approach requires significant processing time compared to more traditional single pointing observations. This is because the (u, v) coverage is built up by including large areas of sky. In an effort to reduce processing time we chose to image sections of our field but ensured these were large enough that we did not lose (u, v) coverage.

Looking for a source that may be a few arcseconds across in a wide field is much the same as looking for a star in the night sky without any prior knowledge of where to look or what to look for. To overcome this problem, one of two things can be done. First, the literature can be used to locate known sources within the field. W49A is a well studied region and so there is a significant number of sources identified in a range of studies with published positions. We, therefore, used the centre positions of the published sources as inputs to the 'phasecenter'

parameter in the CLEAN task. The more rigorous way would be to divide it into smaller regions which would make it easier to pick out bright sources. It however remains true that this only works for exceptionally bright sources - since bright sources affect the colour scale and stand out. To find the rest of the sources, an automated source finding algorithm is required.

3.5.1 Imaging W49A

We identify sources in W49A using both of the methods described in §3.5. W49A is a well studied region and has been well documented in the literature. Therefore, the literature was used to obtain its coordinates and image the surrounding region and the sources. Nonetheless, a short script was written to call and instruct CASA to divide the field into 18 smaller regions. These regions were then used to obtain the coordinates of the brightest sources in the field. In retrospect, most of these bright sources happened to be close to W49A. As a consequence, this study's main focus remained centred at W49A.

3.5.1.1 Unusual Features in On-The-Fly Mapping (OTFM)

OTF mapping was introduced to the VLA in 2015 and so is relatively new. There are therefore some unusual features compared to the standard single or multiple pointing observations. Due to the nature of the scan pattern, the primary beam response took a form different from that of the standard single-pointing method. In the standard method, the primary beam response is simply a Gaussian (Fig. 3.27: left) whereas the primary beam response in the OTFM mode produces a rectangular lattice pattern of blended Gaussians (Fig. 3.27: right).

To gain perspective on how the lattice was formed, a single step was imaged and compared to the beam of several scans (Fig. 3.28). The single step produces a primary beam response image in the form that we expect from a standard observation. Comparing this with the stripes reveals that the primary beam response in OTFM mode is formed by morphing together several of these single step Gaussians to form a lattice pattern. This lattice pattern forms because the telescope scans in Right Ascension and changes rows by moving up in Declination before scanning another row, hence creating several rows and columns of intersecting Gaussians.

Because this is a wide field observation and CASA defines the phase-centre based on the first scan rather than working out the middle of a group of fields, it is imperative that a phase-centre be specified during imaging. Not setting the phase-centre results in imaging an area of sky

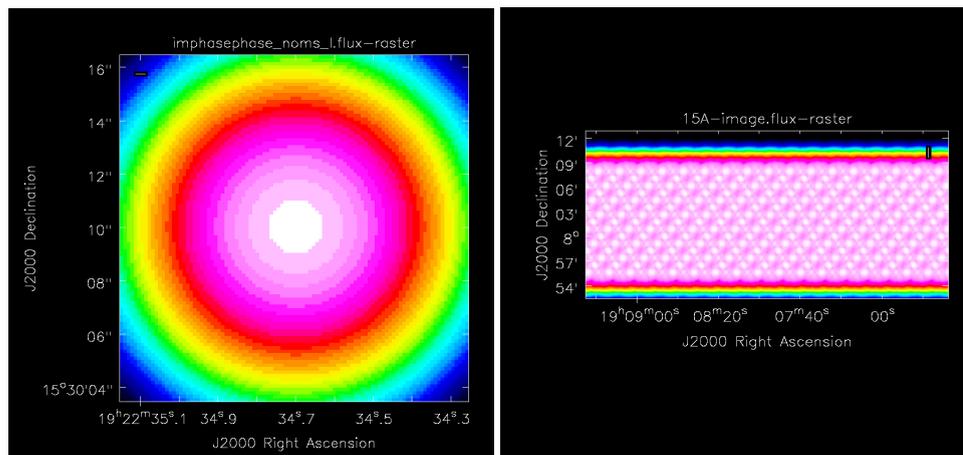


FIGURE 3.27: Comparison between the Primary beam response of the calibrator J1922+1530 -standard pointing (left) and the primary beam response of the target - OTFM pointing (right). Note the beam of the OTFM pointing has been truncated because the process is computationally heavy at this resolution.

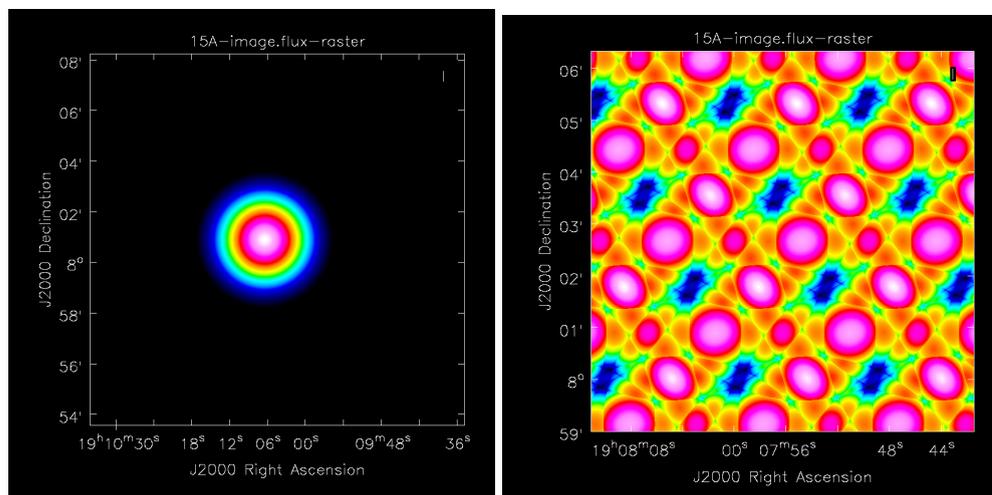


FIGURE 3.28: Left image shows the beam of a single scan in an OTFM stripe – the image to the right shows the beam of several scans morphed together.

centred on different coordinates. The shift has a more drastic effect on an image with several steps (scans), however, we will use a single step to demonstrate its effect.

On the left hand side of Fig. 3.29, the phase-centre parameter was left blank and CASA was left to set the phase-centre. The result was, as seen, a shift to the right. Without constraints, the shift would have been greater. On the right hand side, we have an image in which the phase-centre was set; hence we get a well aligned image.

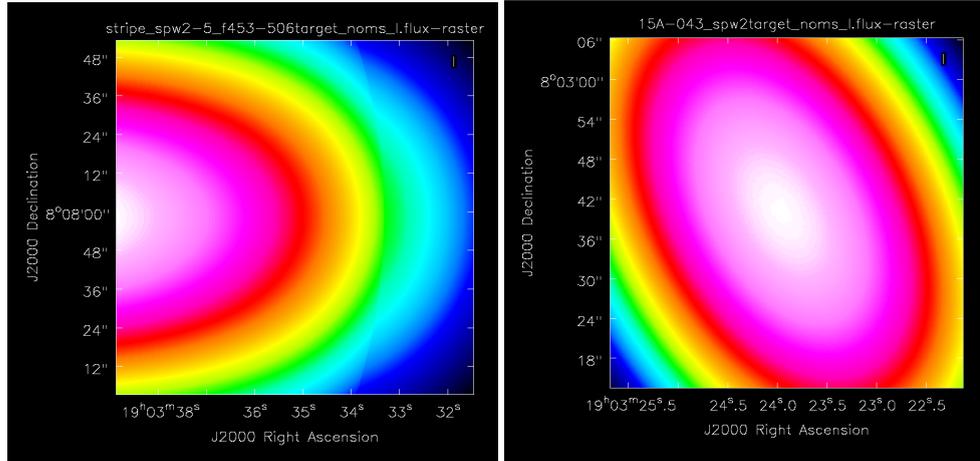


FIGURE 3.29: An offset phase-centre (left) compared with a correctly set phase-centre (right).

3.5.1.2 Imaging the Target

The data from the observation of W49A in spectral window 2 was continuum data. Therefore, to image it we apply mfs (see § 2.5.1). The parameters in the CASA task CLEAN were set considering the discussion in the previous sub-sections. Therefore, the phase-centre parameter in the CASA task was set to the Right Ascension and Declination coordinates of W49A.

W49A is the brightest star forming region in the galaxy (De Pree et al., 1997) and is a very dynamic and complex region. As a consequence, cleaning the image is a rather daunting task. The region contains a combination of point and extended-emission sources with brightness levels varying by many orders of magnitude. Because of the wide range of the types of sources, the weighting used was Briggs robust weighting. Since this was designed to give optimum results for a compromise between natural and uniform weighting, it was expected that this would work best for this highly diverse region. Also note that the array configuration used impacts the chosen weighting scheme. If it is an extended (long baseline) array like A or B, then the observations will only really be sensitive to point sources. Whereas if a compact (short baseline) configuration (C-D array) was used the observations would be sensitive to extended emission. Our observations are subject to significant flux loss on large angular scale, i.e., they are not sensitive to large scale emission above a certain size. At 14 GHz in the B configuration the largest angular structure visible to the array during the KuGARS observations was ~ 12 arcseconds¹.

The process of cleaning these images was done interactively. Clean boxes were used to improve the efficiency and output of the clean process. The clean boxes were set in the CLEAN VIEWER

¹<https://science.nrao.edu/facilities/vla/docs/manuals/oss/performance/resolution>

GUI. These were adjusted after every cycle either to extend the clean regions to cover new emission or reduce them to avoid cleaning too deep. With these regions set, the clean algorithm removes the PSF shape (Fig. 3.31) from the dirty image (Fig. 3.30). The final image (Fig. 3.32) is then made by smoothing the clean components with a restoring beam and putting them back into the residual image.

The angular resolution was calculated using the following equation:

$$\theta = \frac{\lambda}{D} \quad (3.1)$$

Where,

- θ is angular resolution,

- λ is wavelength,

- D is the longest baseline

For spectral window 2, which has a frequency of 13.99 GHz, the wavelength (λ) is 0.0214 metres and the longest baseline is 10311.7 metres. Inserting these values into Eq. 3.1 gives an angular resolution (θ) of 2.08×10^{-6} radians or 0.428 arcseconds. The resolution was then sampled by 5 pixels per beam in order to determine the cell size to be used in CLEAN. The cell size was therefore 0.086 arcseconds. The image size was set to 4000 by 6000 pixels in order to cover a large enough area for the (u, v) coverage to be sufficient. Finally, the number of iterations was set to 10,000, with a threshold of 0mJy. We used interactive CLEAN in order to have more control over the cleaning process and ensure that all significant emission was cleaned without cleaning noise spikes. The noise statistics were determined by running IMSTAT on the dirty image (Fig. 3.30) and predicted a noise threshold of 0.03mJy.

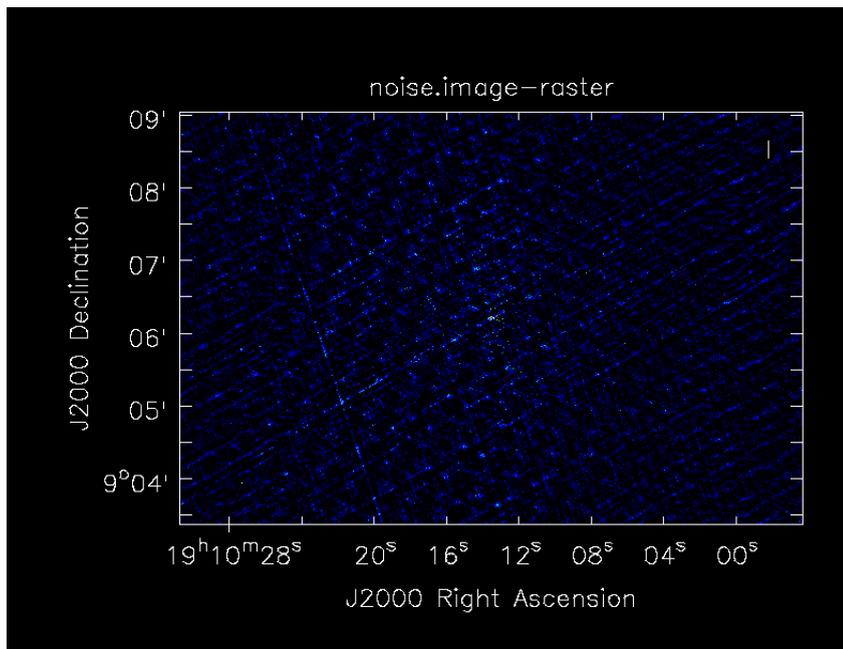


FIGURE 3.30: Dirty image of the target, created by running CLEAN with zero iterations.

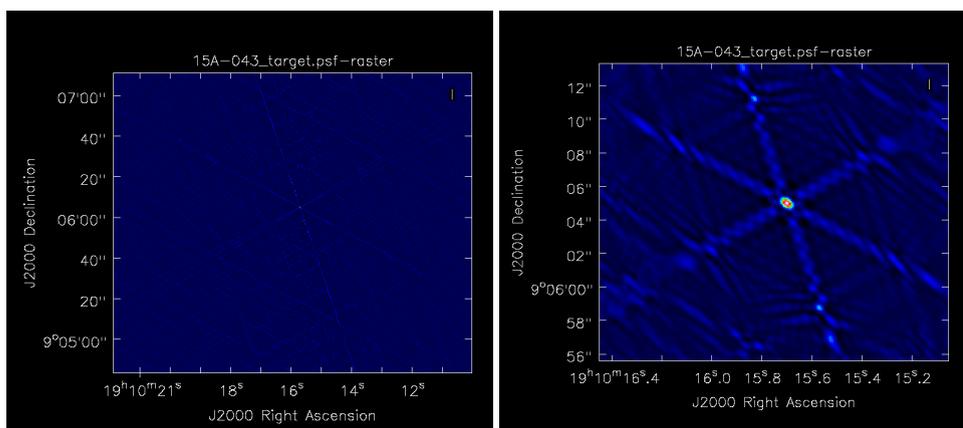


FIGURE 3.31: PSF image of the target.

The clean regions were set prior to the first 100 iterations. They were set using either ellipses or polygon points because of the varying shapes of the sources on the image. At the end of every cycle, the clean boxes were adjusted to include previously hidden source emission. However, after the fourth cycle, some of the emission within the clean regions began to appear noise-like. This time the clean regions were adjusted to erase these areas in order to avoid cleaning too deeply. After 6 cycles, it became difficult to distinguish noise from real emission and after 11 cycles, all the emission in the region looked like noise. At this point, to ensure that any sources that may have been missed were cleaned, the threshold was set at 0.03mJy and clean was left to run until it self-terminated.

The CLEAN task created image files of: the primary beam response, the cleaned and restored

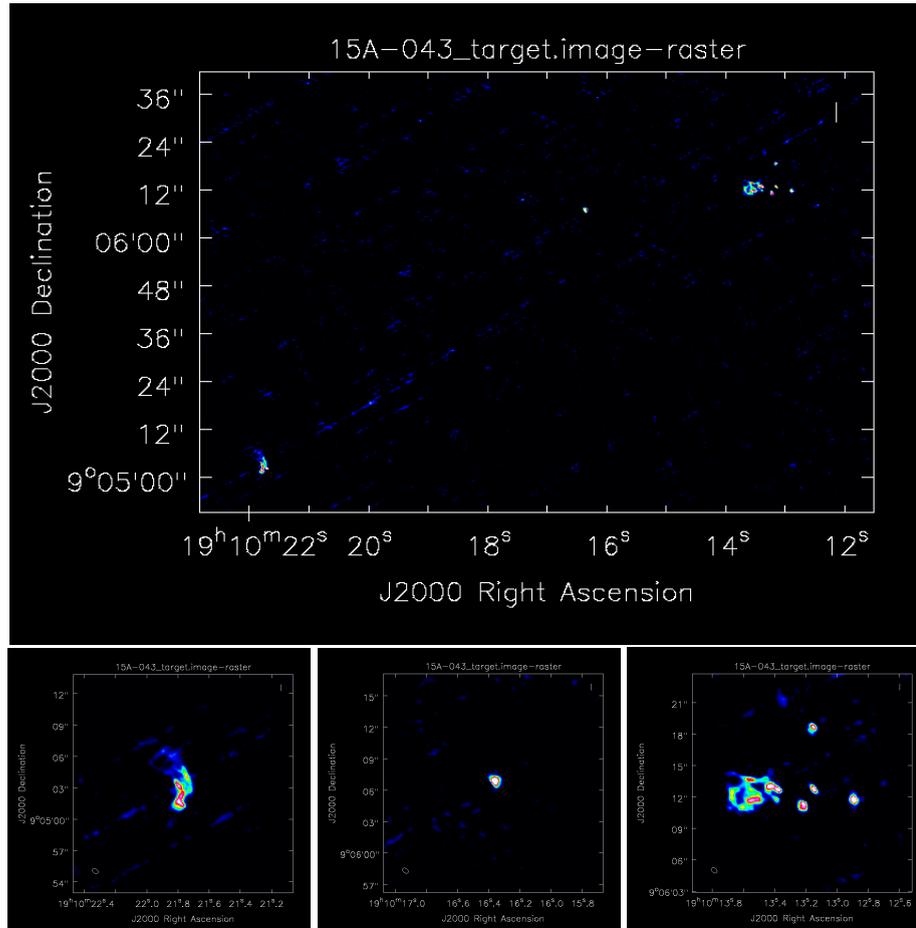


FIGURE 3.32: Raster image of the final clean image- Stokes I mosaic of 15A-043 block 5, spw2 in the region surrounding W49A. Below are close ups of the groups of sources.

image, the residual, the model, the point spread function (psf), and the mask. These image files can be viewed in the CASA task, VIEWER.

The properties and statistics of each source were determined by drawing a region in the same way as during the clean process. Once the region has been completed, the VIEWER GUI displays the statistics of the selected area. For more precise values of the peak and integrated fluxes, a gaussian was fit to the respective regions and again the result was displayed within the GUI. A summary of the properties and statistics is shown in table 3.2.

The results showed 3 immediately evident groups of sources in Fig. 3.32. These are not the only sources in the region. It is expected that W49A being the brightest star forming region in the galaxy, would overshadow the weaker regions close to it. The statistics and positions of these regions were catalogued using an automated source finding algorithm (§ 4.3).

TABLE 3.2: Peak and integrated fluxes of the individual sources within the imaged mosaic on spectral window 2.

Left Group				
	RA	DEC	Peak Flux Density	Integrated Flux Density
	J2000	J2000	(mJy/beam)	(mJy)
1	19:10:21.787	+09:05:01.86	1.51	5.94
2	19:10:21.794	+09:05:03.19	2.25	5.00
3	19:10:21.745	+09:05:04.09	1.15	2.67
4	19:10:21.724	+09:05:02.85	2.26	9.75
5	19:10:21.833	+09:05:06.07	3.98	0.52
6	19:10:21.894	+09:05:06.50	5.95	0.66
Middle Group				
1	19:10:16.359	+09:06:06.89	3.94	9.69
Right Group				
1	19:10:13.671	+09:06:12.17	0.93	1.42
2	19:10:13.605	+09:06:11.02	1.02	1.40
3	19:10:13.536	+09:06:11.72	1.56	3.57
4	19:10:13.552	+09:06:12.68	0.60	0.87
5	19:10:13.558	+09:06:13.63	1.42	1.84
6	19:10:13.427	+09:06:12.99	2.72	6.53
7	19:10:13.380	+09:06:12.71	3.66	4.96
8	19:10:13.220	+09:06:11.14	4.49	23.6
9	19:10:13.148	+09:06:12.74	3.58	4.68
10	19:10:13.155	+09:06:18.60	2.83	5.59
11	19:10:12.892	+09:06:11.74	4.10	5.81

Chapter 4

Data Analysis

4.1 Calibration Quality Assurance

Having been observed in the standard mode, the calibrators made good subjects for calibration quality assurance. The calibrators are also frequently observed, therefore, obtaining results close to those recorded in the literature within reasonable error would validate the calibration. This was achieved by deriving the calibrator properties through the *CASA VIEWER* and *IMSTAT* tasks. The flux was then compared to the literature while the noise was compared to the expected noise from the VLA sensitivity calculator.

This scheduling block was observed in the B configuration with 27 antennas. At 13.99 GHz in the Ku band with a bandwidth of 128 MHz, the VLA exposure calculator predicted an rms noise of 0.12mJy/beam and an approximate beam size of 0.458 arcseconds.

4.1.1 Calibrator Parameters

The *VIEWER* task was used to image the calibrators and derive noise and flux values – these were then used to compute the signal to noise ratio in the respective images. The *IMSTAT* task was then used to verify the noise values obtained by displaying the image statistics.

4.1.1.1 Flux Calibrator

The primary calibrator used here was the compact radio source 3C48 (0137+331) shown in Fig. 4.1.

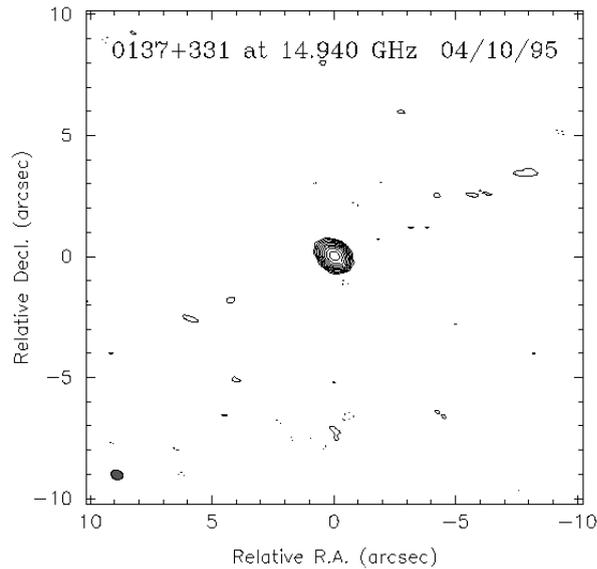


FIGURE 4.1: VLA calibrator manual image of 0137+331= 3C48 at 14.0 GHz in 1995.

The imaging followed the procedure in § 3.5.1.2. To begin with, the resolution required to image 3C48 was determined using Eq. 3.1. The longest baseline was 10172 metres, while the wavelength was 0.0214 metres. Therefore, sampling at 5 pixels per beam the resolution required to image 3C48 was 0.08 arcsecs. Unlike the target which covers a wide area and as a consequence requires a large image size, the calibrators are point sources and require much smaller images. In this case the image size used was 480 square pixels.

As expected, Fig. 4.2 shows 3C48 at 13.99 GHz as a point source. Nonetheless, the fidelity of the image is dependent on how well it agrees with the literature and the VLA exposure calculator for flux and noise respectively. Depending on the spectral index of a source, the flux of the source varies directly, inversely or remains flat. In table 4.1, the flux is seen to reduce as the frequency increases. Here, we see the KuGARS value following the trend. This is expected as the primary calibrator was in fact used to set the scale. Furthermore, Fig. 4.3 shows that the flux in the Ku band lies within 10% error of those in the VLA calibrator manual; given inherent uncertainties in the absolute flux scale, this is expected.

The rms noise of the images was determined by running the CASA task IMSTAT on the final map. Among the statistics that IMSTAT returns, are the peak flux and rms noise of the image. At 13.99

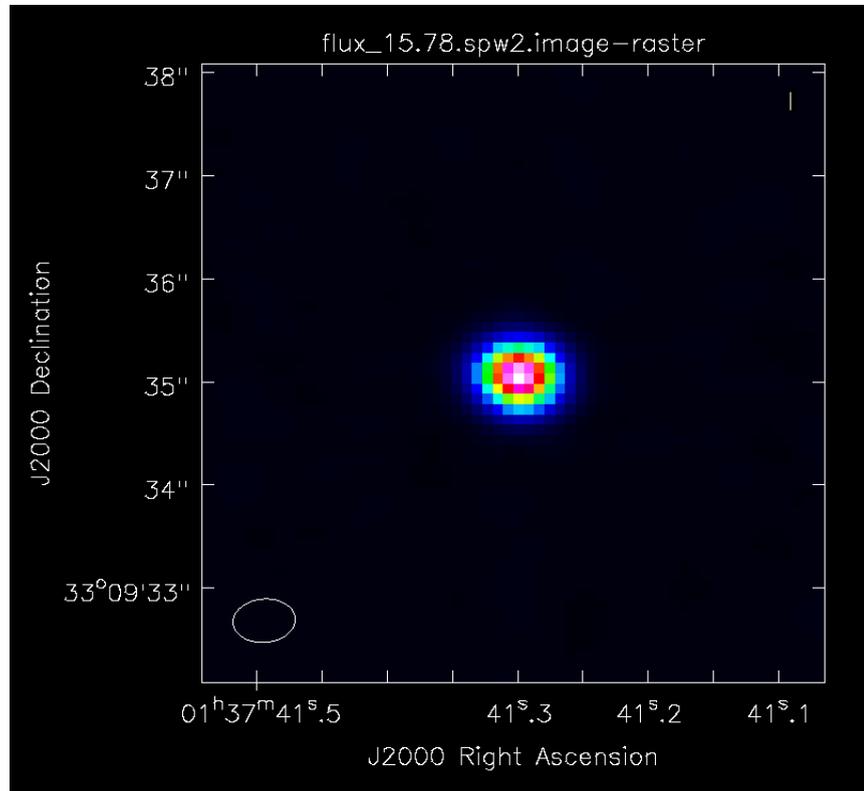


FIGURE 4.2: KuGARS block 5pm flux calibrator, 0137+331= 3C48 at 13.99 GHz.

TABLE 4.1: Fluxes of 3C48 across multiple frequencies from the VLA calibrator manual and KUGARS.

VLA Calibrator Manual			
Band	Wavelength (cm)	Frequency (GHz)	Flux Density(Jy)
P	90.0	0.33	42.0
L	20.0	1.5	16.5
C	6.0	5.0	5.48
X	3.7	8.11	3.25
U	2.0	15.0	1.78
K	1.3	23.08	1.13
Q	0.7	42.86	0.64
KuGARS			
Ku (Spw2)	2.14	13.99	1.51

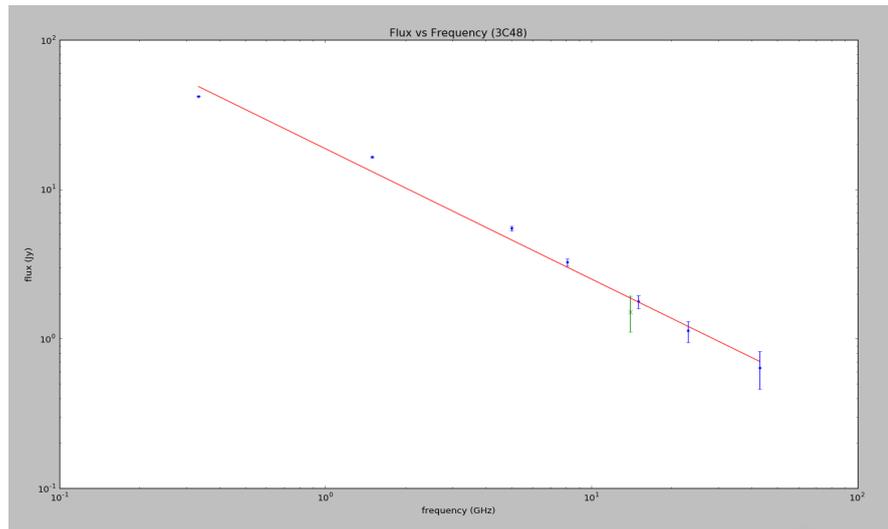


FIGURE 4.3: Flux of 3C48 at different frequencies on the VLA calibrator manual (blue dot) and the Ku-band (green cross).

GHz, the peak flux was 1510mJy/beam while the rms noise was 13.9mJy. These gave a signal to noise ratio of 108.8σ .

4.1.1.2 Phase Calibrator

The source used as phase calibrator was J1922+1530 (Fig. 4.4). With a wavelength of 0.0214 metres, and a longest baseline of 10284.7 metres, the resolution and image size used in the imaging of the phase calibrator were 0.08 arcseconds and 480 square pixels, respectively (Fig. 4.5).

In Fig. 4.5 we show an image of the phase calibrator in KuGARS spectral window 2 (13.99 GHz). We find that the image is again, as expected, a point source. Here, in the same way as in § 4.1.1.1, we determine the fidelity of the image by way of comparison to the VLA calibrator manual and sensitivity calculator.

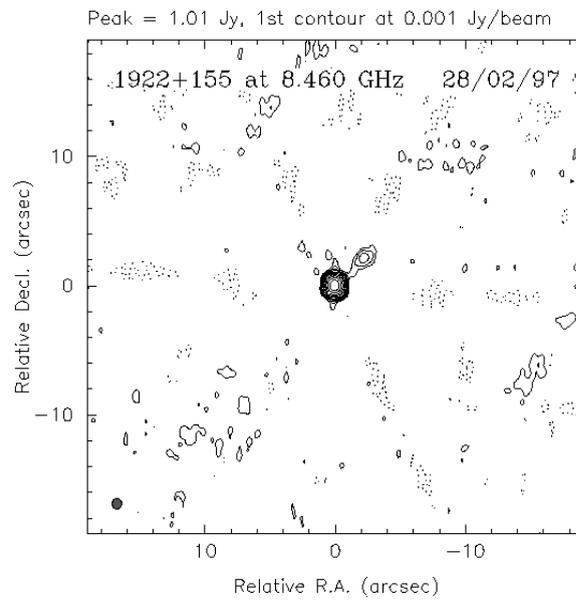


FIGURE 4.4: VLA Calibrator Manual image of J1922+1530 at 8.46 GHz in 1997

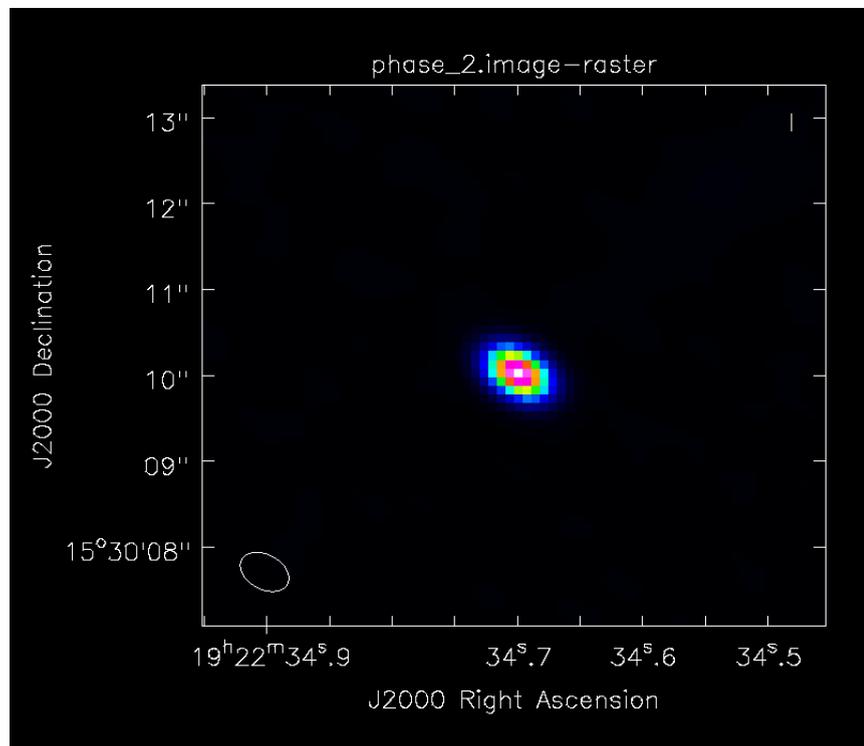


FIGURE 4.5: KuGARS block 5pm phase calibrator, J1922+1530 at 13.99 GHz.

TABLE 4.2: Fluxes of J1922+1530 across multiple frequencies from the VLA calibrator manual and KUGARS.

VLA Calibrator Manual			
Band	Wavelength (cm)	Frequency (GHz)	Flux Density(Jy)
L	20	1.5	0.5
C	6.0	5.0	0.8
X	3.7	8.11	1.0
U	2.0	15.0	0.7
Q	0.7	42.86	0.28
KuGARS			
Ku (Spw2)	2.14	13.99	0.29

Unlike with 3C48, we do not see a particularly obvious trend in the values in table 4.2. This is because the sources observed for phase calibrators (quasars) are variable in time. We, however, use graphical methods to further analyse the trend. In Fig. 4.6 we show that the Ku-band flux varies within 10% error of those from the VLA calibrator manual.

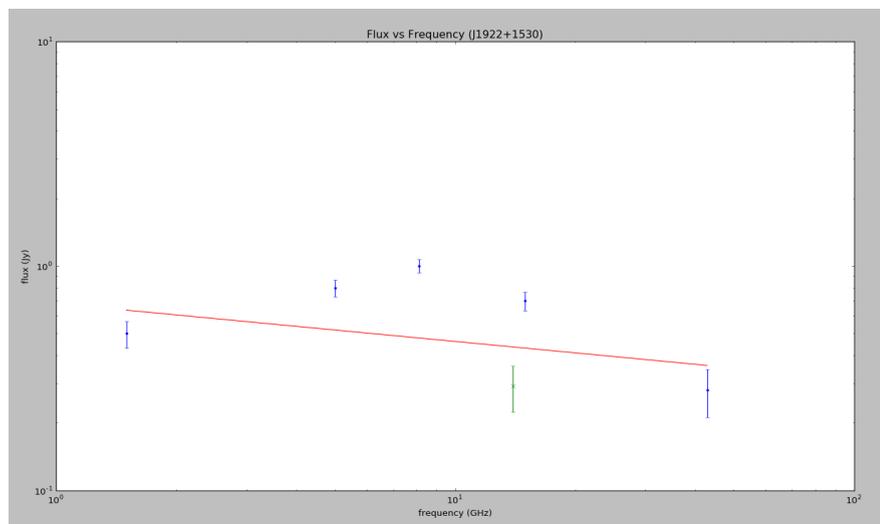


FIGURE 4.6: Flux of J1922+1530 at different frequencies on the VLA calibrator manual (blue dots) and the Ku-band (green cross).

Just as with 3C48, the peak flux and rms noise were determined by running IMSTAT on the final map. At 13.99 GHz, the peak flux was 290mJy/beam while the rms noise was 2.25mJy. These gave a signal to noise ratio of 129.5σ .

4.2 Target-Noise Evaluation and Statistics

4.2.1 Spectral Window 2 (13.99 GHz)

In addition to the properties in §3.5.1.2, the VIEWER GUI was used to determine the rms noise in the final image. This was done by drawing a region box over an area with no visible emission. The value found here was 0.19mJy. The same calculation was made using IMSTAT which takes into account the entire map and includes bright emission. This returned a value of 0.207mJy for the rms noise whereas the VLA sensitivity calculator's prediction for this was 0.120mJy. Included in IMSTAT was the flux value of the brightest source, 4.71mJy. Using the IMSTAT values, we get a value of 22.8σ for signal to noise.

4.3 Source Catalogue

In § 3.5.1.2 we describe the cleaning process that resulted in the images in Fig. 3.32. The sources were then outlined in table 3.2. We used AEGEAN (Hancock et al., 2012), an automated source finding algorithm, to identify and catalogue the sources over the same region in our image of spw 2.

Fig. 4.7 was created in the interactive imaging tool, DS9. The sources listed in table 4.3 were imported as a catalogue list in .vot file format, whereas the image, which was exported from casa as a FITS image, was imported into DS9 in the same format. The sources in the catalogue list were then overlaid on the FITS image as circles. AEGEAN, which primarily focuses on compact sources, uses an island-source identification method wherein an extended source is first identified as an island and then further divided into components or peaks. In table 4.3 the outlined format for the source identifier is (isl,src) showing the island number followed by the component. The rows of the table have been sectioned to distinguish between the respective islands: it shows a single row for compact/point sources and multiple rows for extended sources. Ergo, the total integrated flux of the extended source in question would be the sum of the components in the island.

For the purpose of the discussion, we have shown the 57 components on the 35 islands closest to W49A - that is, the sources in the map described and imaged in § 3.5.1.2.

TABLE 4.3: Source positions and fluxes in the region surrounding W49A.

```

=====
#Aegean version 2.0b158-(2017-05-31)
#on dataset: 15A-043_sb5pm_spw2.fits
#(isl,src) RA DEC Peak err S_int err
# Jy/beam Jy/beam Jy Jy
=====
(0001,00) 19:09:51.36 +09:03:15.70 0.000739 0.000351 0.001747 0.000866
(0001,01) 19:09:51.39 +09:03:15.73 0.000736 0.000351 0.000603 0.000355
(0001,02) 19:09:51.43 +09:03:16.19 0.000733 0.000351 0.000753 0.000419
(0001,03) 19:09:51.45 +09:03:16.02 0.000798 0.000351 0.007885 0.003501
-----
(0002,00) 19:09:42.20 +09:03:18.55 0.000788 0.000111 0.006729 0.001309
-----
(0003,00) 19:10:32.49 +09:03:21.99 0.001256 0.000142 0.010399 0.002059
-----
(0004,00) 19:09:54.29 +09:03:39.14 0.001070 0.000124 0.008754 0.001258
-----
(0005,00) 19:10:30.65 +09:03:39.12 0.001359 0.000195 0.006554 0.001527
(0005,01) 19:10:30.59 +09:03:39.62 0.000906 0.000195 0.001905 0.000587
(0005,02) 19:10:30.55 +09:03:40.13 0.001180 0.000195 0.004252 0.000990
-----
(0006,00) 19:09:39.33 +09:03:56.12 0.000639 0.000082 0.006492 0.001224
-----
(0007,00) 19:09:39.25 +09:04:16.98 0.000690 0.000110 0.001494 0.000275
-----
(0008,00) 19:10:00.94 +09:04:32.65 0.001248 0.000200 0.004863 0.000960
-----
(0009,00) 19:09:58.86 +09:04:50.31 0.001126 0.000136 0.011739 0.001981
-----
(0010,00) 19:10:21.79 +09:05:01.37 0.002520 0.000427 0.010959 0.001900
(0010,01) 19:10:21.79 +09:05:03.18 0.002563 0.000427 0.009345 0.001667
(0010,02) 19:10:21.78 +09:05:02.14 0.002132 0.000427 0.005409 0.001144
(0010,03) 19:10:21.75 +09:05:04.16 0.001802 0.000427 0.013148 0.003185
(0010,04) 19:10:21.72 +09:05:02.74 0.001429 0.000427 0.002091 0.000654
-----
(0011,00) 19:10:11.04 +09:05:20.17 0.002002 0.000241 0.004233 0.000571
-----
(0012,00) 19:10:42.77 +09:05:28.75 0.000724 0.000104 0.005567 0.001089
-----
(0013,00) 19:10:02.86 +09:05:40.53 0.001186 0.000190 0.003140 0.000595
-----
(0014,00) 19:10:16.36 +09:06:06.89 0.004055 0.000198 0.011937 0.000626
-----
(0015,00) 19:10:17.41 +09:06:09.45 0.001379 0.000226 0.003456 0.000633
-----
(0016,00) 19:10:13.22 +09:06:11.16 0.003404 0.000219 0.009792 0.000686
=====

```

TABLE 4.3: Continued

(0017,00)	19:10:13.37	+09:06:12.67	0.003084	0.000505	0.002975	0.000511
(0017,01)	19:10:13.43	+09:06:13.04	0.002827	0.000505	0.016160	0.002906
(0017,02)	19:10:13.57	+09:06:11.63	0.001818	0.000505	0.003298	0.000924
(0017,03)	19:10:13.52	+09:06:11.71	0.002322	0.000505	0.007079	0.001550
(0017,04)	19:10:13.56	+09:06:13.68	0.002378	0.000505	0.006341	0.001356
(0017,05)	19:10:13.60	+09:06:10.98	0.001866	0.000505	0.012597	0.003412
(0017,06)	19:10:13.68	+09:06:12.51	0.001446	0.000505	0.002150	0.000756
(0017,07)	19:10:13.67	+09:06:11.86	0.001815	0.000505	0.004100	0.001149
(0017,08)	19:10:13.55	+09:06:12.70	0.001768	0.000505	0.013679	0.003915
(0017,09)	19:10:13.64	+09:06:12.69	0.001345	0.000505	0.005949	0.002237
(0018,00)	19:10:12.89	+09:06:11.74	0.004397	0.000214	0.009159	0.000476
(0019,00)	19:10:13.39	+09:06:11.81	0.001829	0.000270	0.004584	0.000764
(0020,00)	19:10:13.15	+09:06:12.73	0.003728	0.000254	0.005452	0.000396
(0021,00)	19:10:13.16	+09:06:18.61	0.002922	0.000231	0.007127	0.000613
(0022,00)	19:10:19.13	+09:06:29.25	0.001339	0.000236	0.003466	0.000729
(0023,00)	19:10:02.42	+09:06:59.76	0.001238	0.000134	0.013035	0.002483
(0024,00)	19:10:27.64	+09:07:08.25	0.001053	0.000182	0.002605	0.000537
(0025,00)	19:10:22.80	+09:07:28.35	0.001263	0.000174	0.005943	0.000973
(0026,00)	19:10:25.62	+09:07:51.14	0.000885	0.001147	0.003306	0.004367
(0026,01)	19:10:25.70	+09:07:51.59	0.000973	0.001147	0.004951	0.005951
(0027,00)	19:10:00.78	+09:07:59.87	0.000910	0.000117	0.006706	0.001814
(0028,00)	19:10:26.80	+09:08:00.42	0.001315	0.000207	0.004141	0.000763
(0029,00)	19:10:07.65	+09:08:05.92	0.001393	0.000182	0.004463	0.000812
(0030,00)	19:09:42.79	+09:08:07.09	0.000719	0.000082	0.011131	0.002930
(0031,00)	19:10:28.66	+09:08:15.38	0.001199	0.000150	0.005486	0.000863
(0031,01)	19:10:28.54	+09:08:14.49	0.001073	0.000150	0.003523	0.000639
(0032,00)	19:10:05.98	+09:08:17.70	0.001156	0.000159	0.005386	0.001007
(0033,00)	19:10:29.65	+09:08:23.38	0.001155	0.000183	0.005233	0.001192
(0033,01)	19:10:29.74	+09:08:24.05	0.001003	0.000183	0.002916	0.000873
(0034,00)	19:10:31.47	+09:08:38.04	0.001119	0.000115	0.008910	0.001092
(0035,00)	19:10:32.46	+09:08:45.82	0.000903	0.000512	0.002584	0.001527
(0035,01)	19:10:32.40	+09:08:45.62	0.001105	0.000512	0.008788	0.004159

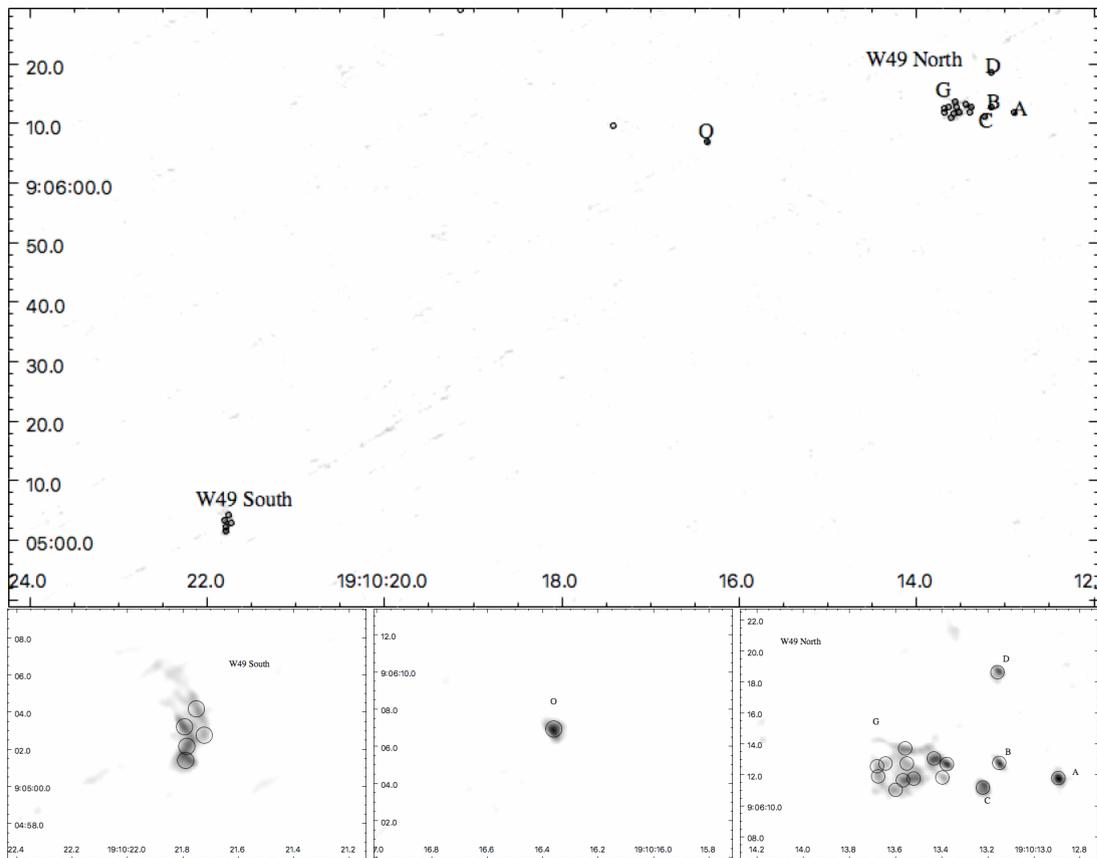


FIGURE 4.7: Source overlay plot of the imaged mosaic in spectral window 2, plotted in inverted grey scale with each source represented by a circle. We have followed the naming convention of Dreher et al. (1984). The sources overlaid on this plot are presented in table 4.3.

Chapter 5

Discussion

In this chapter, we compare the results from KuGARS to the literature. We put into perspective the resolution achieved in the KuGARS observations by overlaying KuGARS contours on GLIMPSE three-colour maps, CORNISH sources, Herschel at $70\mu\text{m}$, and the 3.6 cm map from De Pree et al. (1997). We explore the nature of the detected radio emission based on the sensitivity of each band in the GLIMPSE three-colour maps, and confirm our emission in CORNISH, Herschel, and De Pree et al. (1997). Finally, we overlay objects from the SIMBAD database to see what has previously been found in these regions.

The KuGARS contours used in the comparisons through this chapter were set using the 3, 6, 9, and 18 σ detection limits. In Fig. 5.1 we show these contours on our KuGARS image with the colour set to grey scale. We have followed the Dreher et al. (1984) naming convention and we refer to the sources under discussion by this convention here and throughout (cf. De Pree et al. 1997).

5.1 GLIMPSE Three Colour Maps

5.1.1 Description of GLIMPSE

The Galactic Legacy Infrared Mid-Plane Survey Extraordinaire (GLIMPSE; Benjamin et al. 2003 ; Churchwell et al. 2009), is a multiwavelength survey of two-thirds of the inner Galactic disk with a spatial resolution of ~ 2 arcseconds. The GLIMPSE survey was carried out using the InfraRed Array Camera (IRAC) at 3.6, 4.5, 5.8 and $8.0\mu\text{m}$ on the Spitzer space telescope.

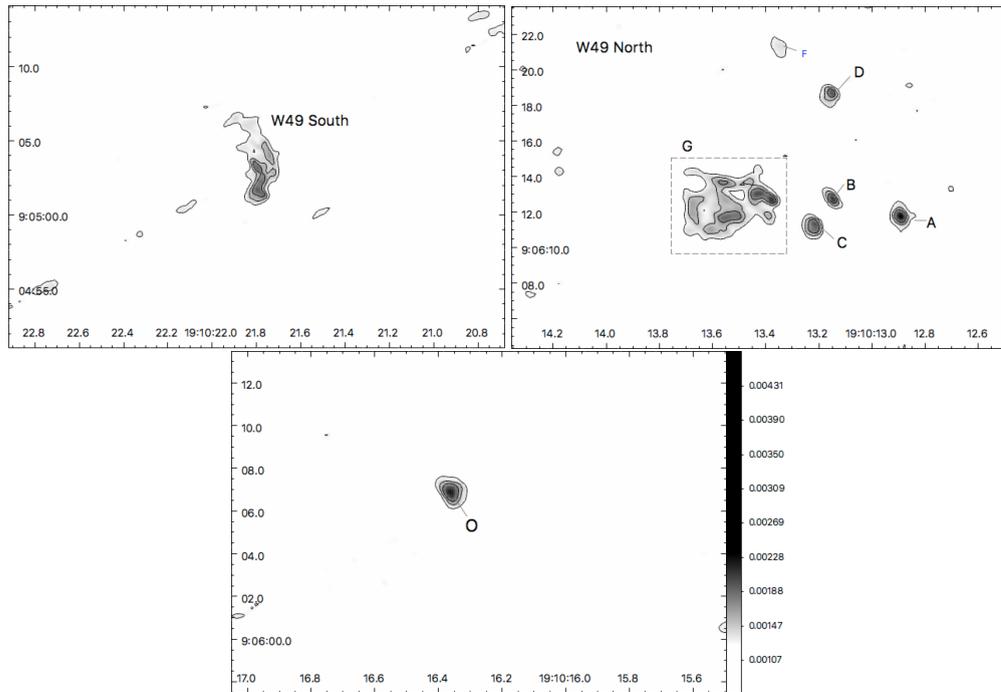


FIGURE 5.1: Contours overlaid on an inverted grey scale image of the KuGARS map of W49A. The contours have been set at the 3, 6, 9, and 18 σ detection limits. We have followed the naming convention used in Dreher et al. (1984).

Each of these IRAC bands highlights different physical properties from the others. Therefore, studying them allows us to determine the origin of the radio emission. The 3.6- μm band is dominated by field stars, the 4.5- μm band emerges from shocked molecular hydrogen and CO bandheads, consequently indicating the possible presence of an outflow. On the other hand, the strong polycyclic aromatic hydrocarbon (PAH) features that are excited in the photon dominated regions (also referred to as photo-dissociation regions) (PDRs) that lie in a thin shell of neutral gas located between an ionization front and cold molecular gas are prevalent in both the 5.8- and 8.0- μm (Hindson et al., 2012).

5.1.2 KuGARS Sources Overlaid On GLIMPSE

In order to properly explore the nature of the detected radio emission, we take advantage of each of the IRAC band's sensitivity to different physical properties. We have created three-colour images using the 4.5, 5.8, and 8.0 μm IRAC bands in blue, green and red, respectively (cf. Hindson et al. 2012; Cohen et al. 2007; Urquhart et al. 2009). In Fig. 5.2 we present a three-colour image of the region surrounding W49A. The green rectangle shows the area of sky imaged in our first KuGARS observations presented in Fig. 3.32. Evidently this is a dynamic and complex region with a lot of bright emission.

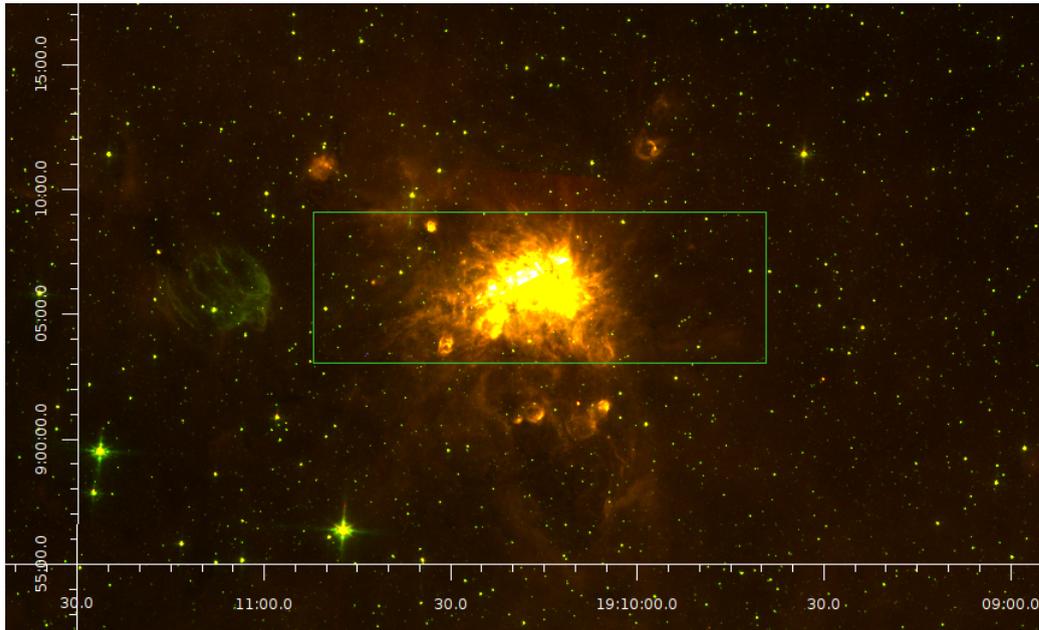


FIGURE 5.2: Three-colour GLIMPSE image at 99.5% colour scale using the IRAC 4.5, 5.8, and $8.0\mu\text{m}$ bands in blue, green and red, respectively. The green rectangle is an overlay of the region presented in Fig. 3.32

The red emission ($8.0\mu\text{m}$) indicates heated dust and also emission from large excited molecules known as polycyclic aromatic hydrocarbons; this shows the interface between molecular gas and an ionising radiation field. In Fig. 5.2, the red emission is seen in the envelope of emission around W49A. The green emission ($5.8\mu\text{m}$) is an intermediate band, showing some of the hottest parts of the field as well as some of the stellar emission. In the same manner as with the red emission, we see an example of the green emission in Fig. 5.2 positioned to the left of W49A. The blue emission ($4.5\mu\text{m}$) comprises of high energy and mostly shows stellar sources. Again, this can be seen in Fig. 5.2 which has several compact blue sources throughout the image.

The KuGARS survey has a spatial resolution of ~ 0.45 arcseconds while GLIMPSE has a spatial resolution of ~ 2 arcseconds. To bring the KuGARS resolution into perspective we overlay KuGARS contours on GLIMPSE three-colour images (Fig. 5.3). We set the colour scale to min-max to avoid overexposure. This colour setting allows us to see the brightest regions. The KuGARS contours used here were set at the 3 , 6 , 9 and 18σ detection levels. From visual inspection of these images, we find that the detected emission can be separated into two categories: Extended radio emission associated with dense $8.0\mu\text{m}$ GLIMPSE emission (red), and compact radio emission associated with no obvious concentration of GLIMPSE emission.

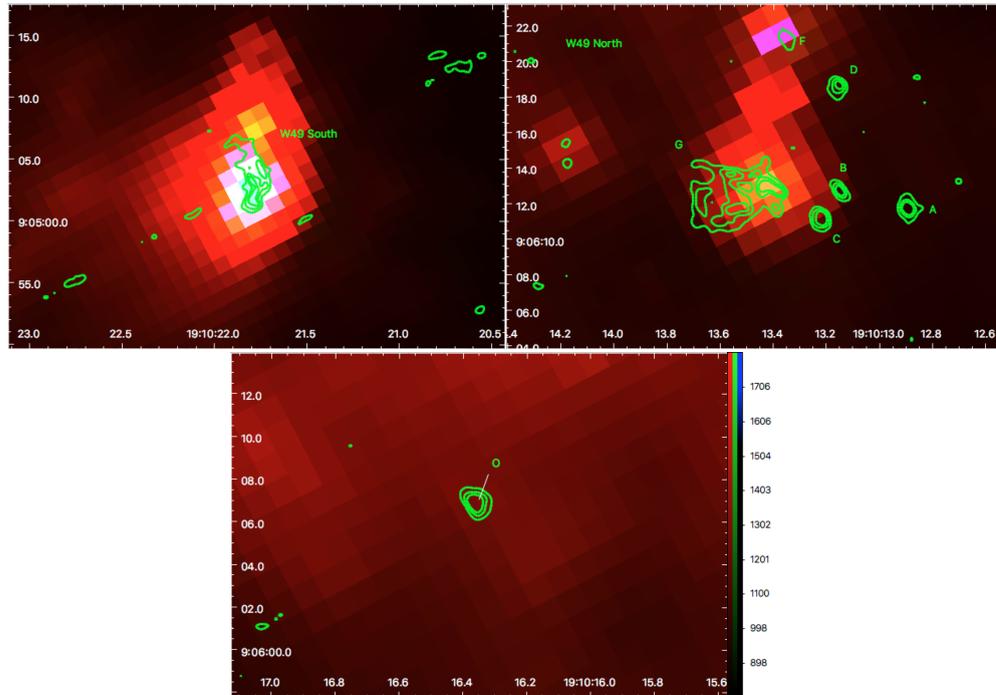


FIGURE 5.3: Contours of the brightest KuGARS detections in the region surrounding W49A overlaid on a three-colour GLIMPSE map with the min-max colour scale using the IRAC 4.5, 5.8, and $8.0\mu\text{m}$ bands in blue, green and red, respectively.

In W49 South (Fig. 5.3: top left) our radio continuum emission is associated with bright $8.0\mu\text{m}$ emission (red). This is indicative of a dusty region. In W49 North (Fig. 5.3: top right) we have extended radio emission (source G) associated with another dusty component. Here, we also see isolated radio emission sources (A, B, D) which do not have distinct mid-IR emission associated with them. Finally, in Fig. 5.3: bottom (source O), we see another radio emission feature in bright diffuse emission but has no associated compact source in the mid-IR, much like the isolated radio emission in W49 North. Note, however, that in brighter contrasts these are embedded in a denser mid-IR environment with no distinct mid-IR association.

We find that our radio emission is largely embedded within dusty clouds of mid-IR emission. Whereas this indeed indicates star formation, our analysis will require a comparison with similar surveys, particularly in the radio, to better understand the nature of the emission we have detected. Therefore, to confirm our radio emission, we compare our sources to the CORNISH survey as well as the 3.6 cm image from De Pree et al. (1997) and to better understand the nature of these peculiar features, we compare to Herschel at $70\mu\text{m}$ and look at the SIMBAD database.

5.2 Comparison of KuGARS with the De Pree et al. (1997) 3.6 cm observations.

De Pree et al. (1997) carried out multifrequency, multiconfiguration VLA observations of the massive Galactic star-forming region W49A. The radio continuum observations were carried out at 0.7, 1.3 and 3.6 cm. Here, we compare KuGARS to the highest sensitivity images (3.6 cm). These observations were carried out in the B, C and D configurations of the VLA and phase self-calibrated. At this sensitivity, they detected 45 distinct continuum sources exhibiting a wide variety of sizes and morphological types.

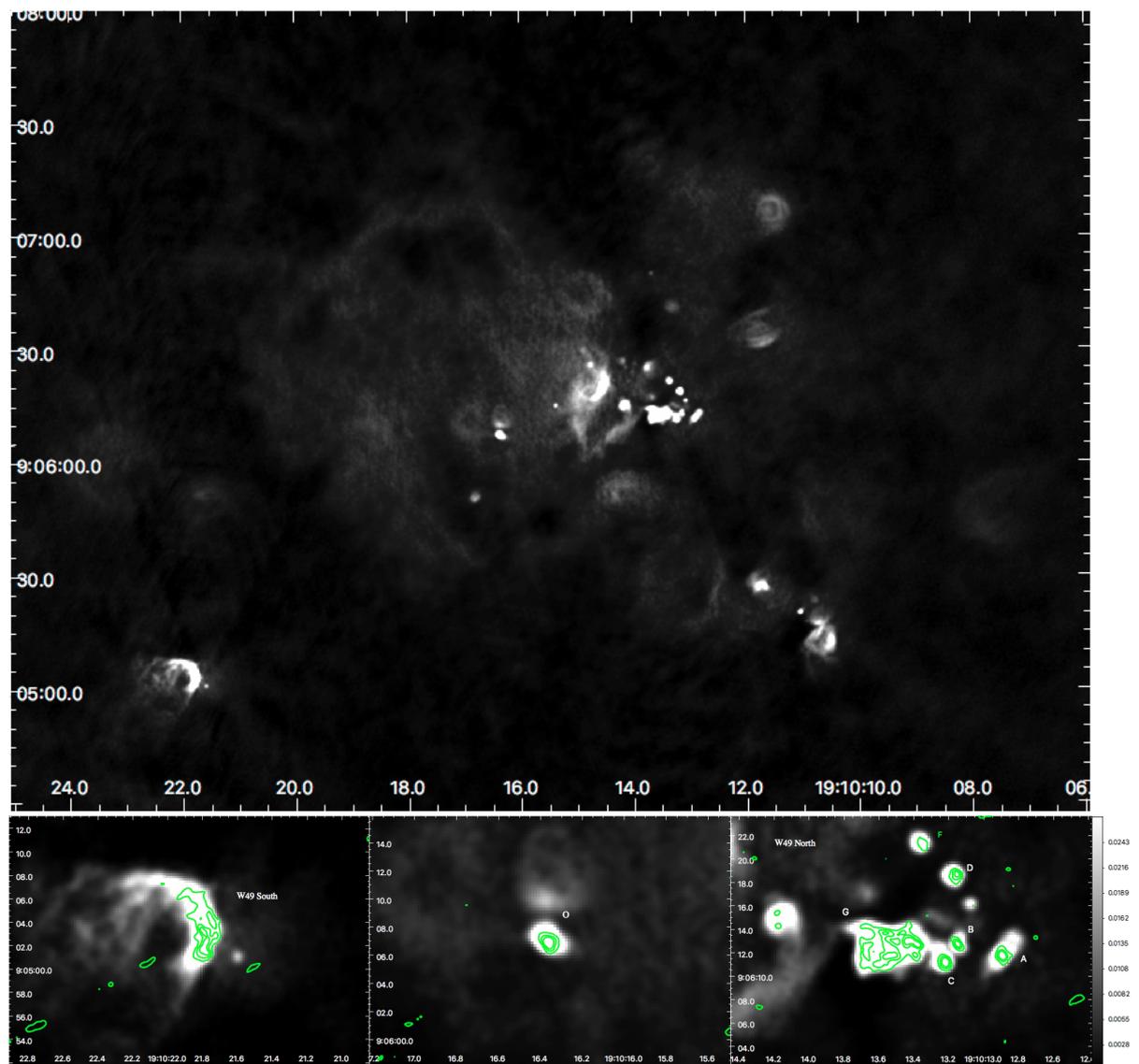


FIGURE 5.4: Contours from KuGARS sources in the region around W49A overlaid on the 3.6 cm De Pree et al. 1997 image. The contours are set at the 3 , 6 , 9 and 18σ detection limits. The top shows the image in full view. The lower images are close-ups of the sources with the highest emission overlaid with KuGARS contours.

5.2.1 Overlapping Sources

The 2.1 cm KuGARS and 3.6 cm De Pree et al. (1997) images have several overlapping sources. We find that the 2.1 cm image is missing a lot of extended emission flux but has a higher resolution than the 3.6 cm image, and thus shows much finer detail on compact sources. Furthermore, it is important to note that these two images have different UV coverages - whereas the 3.6 cm observations were carried out in the B, C and D configuration, the 2.1 cm observations were carried out in the B configuration, consequently contributing to the missing flux. In Fig. 5.4, we have overlaid the 3.6 cm image with contours from the 2.1 cm image. The contours are set at the 3, 6, 9, and 18 σ detection limits. We find that the brightest of the overlapping sources are: W49 North (A, B, C, D, and G), O, and W49 South. De Pree et al. (1997) classified the continuum source detections in the 3.6 cm image according to the criteria of Wood and Churchwell (1989) as either: (1) spherical or unresolved, (2) cometary, (3) shell-like, (4) irregular. Therefore, the overlapping source detections were classified as: A (type 4), B (type 1), C (type 4), D (type 1), O (type 1), W49 South (type 2), and G (extended/unclassified).

5.3 Comparison of KuGARS with CORNISH

The Coordinated Radio and Infrared Survey for High-Mass Star Formation (CORNISH: Hoare et al. (2012); Purcell et al. (2013)), is the highest resolution (1.5 arcseconds) radio continuum survey of the Galactic plane at 5 GHz. The survey, which focused on the northern GLIMPSE region ($10^\circ < l < 65^\circ$), has detected 3062 sources above a 7σ level; the catalogue consists of HII regions, planetary nebulae, and radio galaxies. Here, we compare our source detections to CORNISH by overlaying contours from our images and sources from our catalogue on a ~ 3 arcminute image around W49A from the CORNISH image server¹.

CORNISH is a good basis for comparison with KuGARS. At 5 GHz, the survey lies within the most effective frequency range (1.4-5 GHz) to find objects which are predominantly non-thermal or optically thin at these frequencies (Thompson and Goedhart, 2016). We see an overlap between several of the KuGARS and CORNISH sources in Fig. 5.5. Furthermore, the high resolution of KuGARS is evident. We see some of the unresolved CORNISH sources being resolved in KuGARS. A similar comparison can be made in the GLIMPSE images in §5.1.2, De Pree et al. (1997) in §5.4, and the Herschel images in §5.5. In all three instances, the KuGARS

¹<http://cornish.leeds.ac.uk>

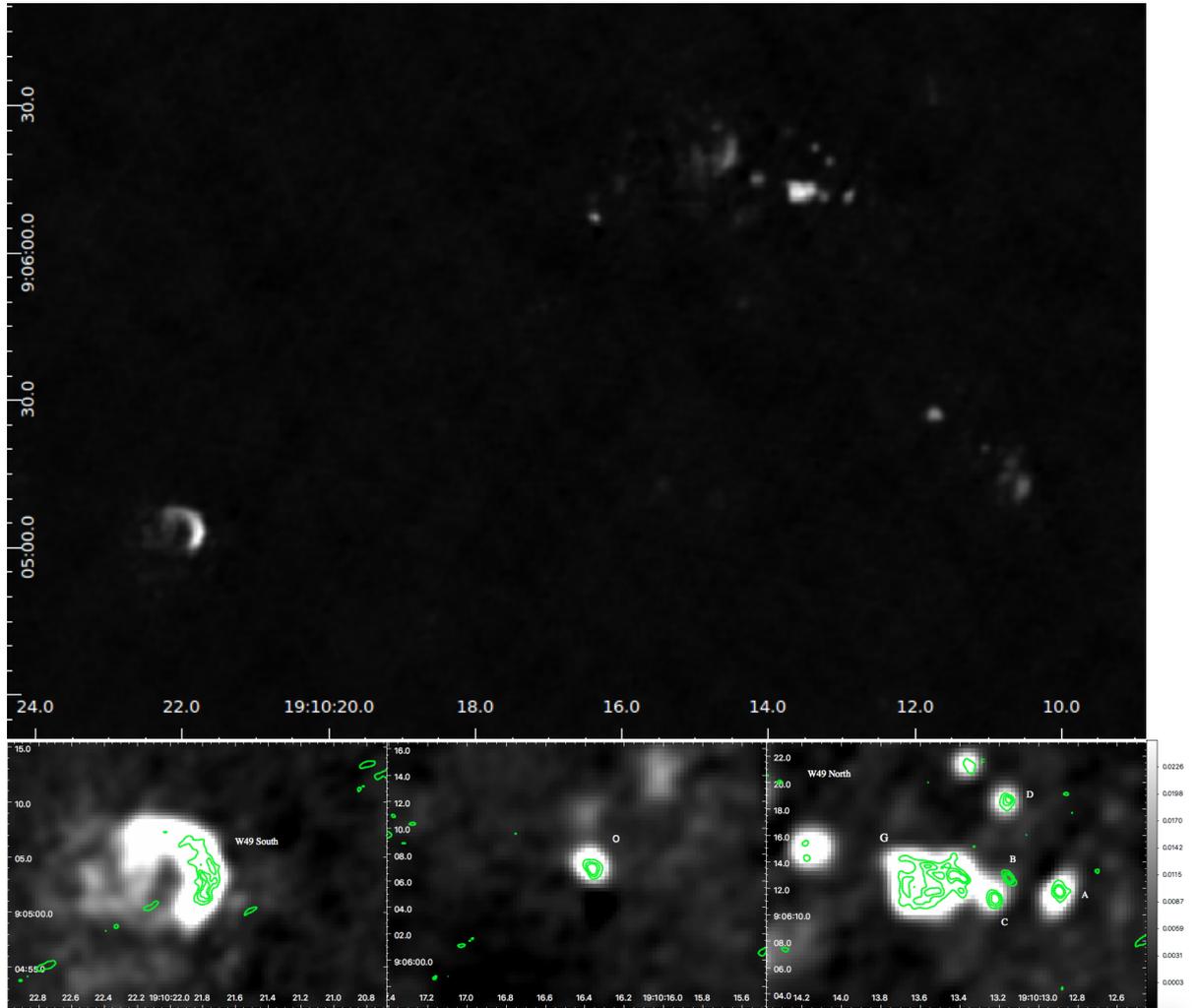


FIGURE 5.5: Contours from KuGARS sources in the region around W49A overlaid on a CORNISH image. The contours are set at the 3, 6, 9 and 18σ detection limits. The top image shows a full view of the CORNISH sources around W49A. The lower images are close-ups of the sources with the highest emission overlaid with KuGARS contours.

contours are deeply embedded in the emission of the image being analysed. It is interesting to note that source B in Fig. 5.5 is not detected by CORNISH. De Pree et al. (1997) detected this source and classified it as a type 1 HII region. In their analysis at the three frequencies (43 GHz, 23 GHz, 8.3 GHz) they found that this source has a strongly rising radio continuum spectrum ($\alpha \sim 1.1$) and is associated with broad radio recombination lines at all three frequencies ($\Delta V_{H92\alpha} = 56.4 \pm 5.4 \text{ km s}^{-1}$, $\Delta V_{H66\alpha} = 60.6 \pm 8.1 \text{ km s}^{-1}$, $\Delta V_{H52\alpha} = 47.5 \pm 2.8 \text{ km s}^{-1}$). This source is an indication that there may be several others like it in KuGARS not observed in CORNISH and other low frequency surveys.

5.4 KuGARS Detections Compared with the Literature

We use the SIMBAD astronomical database to survey each region from § 5.1.2 for previous detections in the literature. We find the following sources: HII regions, stars, molecular clouds, IR objects, radio objects, millimetre radio objects, YSOs, and masers (Fig. 5.6). Based on the SIMBAD sources found, we can draw an inference about the stellar activity in the respective regions.

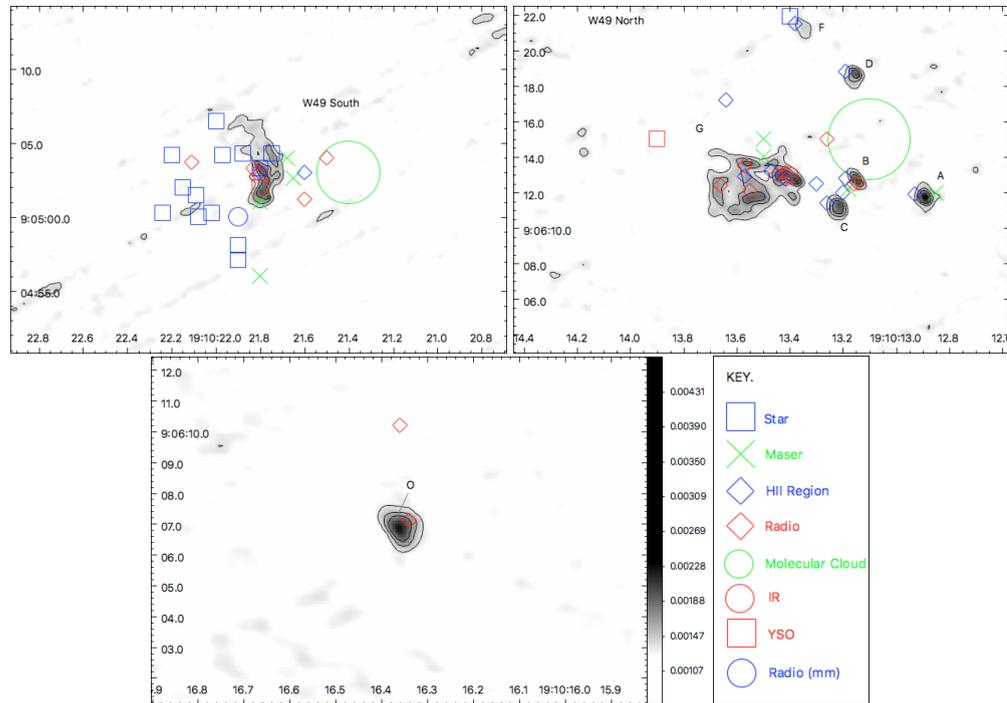


FIGURE 5.6: The sources in the region around W49A overlaid with KuGARS contours and SIMBAD sources on three colour GLIMPSE images set to the min-max colour scale using the IRAC 4.5, 5.8, and $8.0\mu\text{m}$ bands in blue, green and red, respectively.

In W49 South (Fig. 5.6: top left), we see a YSO (red square) centred on the brightest peak of the KuGARS contour. We see that this peak is embedded in a cocoon of dust in the corresponding GLIMPSE image in Fig. 5.3, therefore, implying early stages of star formation. Furthermore, the peak has two masers (green cross) on its border as well as an HII region farther to the right. The presence of these star formation tracers within a cocoon of dust leads us to believe that this is a star forming region. W49 North (Fig. 5.6: top right) has three HII regions within the contours of the large extended source, G. This too is surrounded by a cocoon of dust associated with $8\mu\text{m}$ emission in its GLIMPSE counterpart. We further find that the extended emission source, G, has two masers on the north borders of its outermost contour. This, evidently, is a star forming region. In as much as the aforementioned sources show direct evidence of star

formation, there are some sources that do not. The three radio continuum sources that appeared to have no mid-IR emission in §5.1.2 have SIMBAD HII regions, all with a slight offset. The source farthest from the centre has, in addition to the HII region, a maser detection. In the isolated region O (Fig. 5.6: bottom), the KuGARS source only matches an isolated SIMBAD radio source. However, we know from De Pree et al. (1997) that this has radio recombination lines and therefore must be an HII region.

Having looked at the sources previously found in these regions, W49 South (Fig. 5.6: top left) and W49 North (Fig. 5.6: top right) appear to be star forming regions, we, however, only find a definitive inference on the nature of source O (Fig. 5.6: bottom) in De Pree et al. (1997) which classifies it as spherical HII region. We, therefore, look at a HERSCHEL $70\mu\text{m}$ image of W49A.

5.5 Comparison of KuGARS With Herschel $70\mu\text{m}$ Data

The $70\mu\text{m}$ Herschel data is indicative of heated dust from star formation (Molinari et al., 2010). We, therefore, use it to further determine whether we have star forming regions in the KuGARS data.

Each of the regions in Fig. 5.7 is associated with $70\mu\text{m}$ emission as can be seen in the top image. Therefore, in the closeups the colour scale has been changed to highlight the brightest pixel in the respective regions. As can be seen, W49 South (Fig. 5.7: left) has its brightest pixels at the peak of the radio continuum contour. The region around source O (Fig. 5.7: middle) has the brightest pixels in the region centred on the radio continuum contours. This is indicative of star formation, therefore since source O is an embedded $70\mu\text{m}$ source with a CORNISH 5 GHz counterpart, it is likely to be a young HII region that is not in the SIMBAD catalogue. In W49 North (Fig. 5.7: right) again we see bright pixels in the same position as the bright emission in Fig. 5.3: right. With this, we find that the $70\mu\text{m}$ data infers our sources are star forming regions.

We have begun a multiwavelength analysis for the KuGARS sources to determine their nature. Although this analysis is at an early stage, it nevertheless reveals the rich star forming nature of W49A. In GLIMPSE and Herschel, we have associated our detections with heated dust emission, which is indicative of star formation. In SIMBAD, we have associated our detections with multiple star formation tracers, therefore, confirming the star forming nature of our sources. In CORNISH and De Pree et al. (1997), we have found that the objects we have observed are HII regions of various morphologies. Furthermore, we have found that one of these HII regions

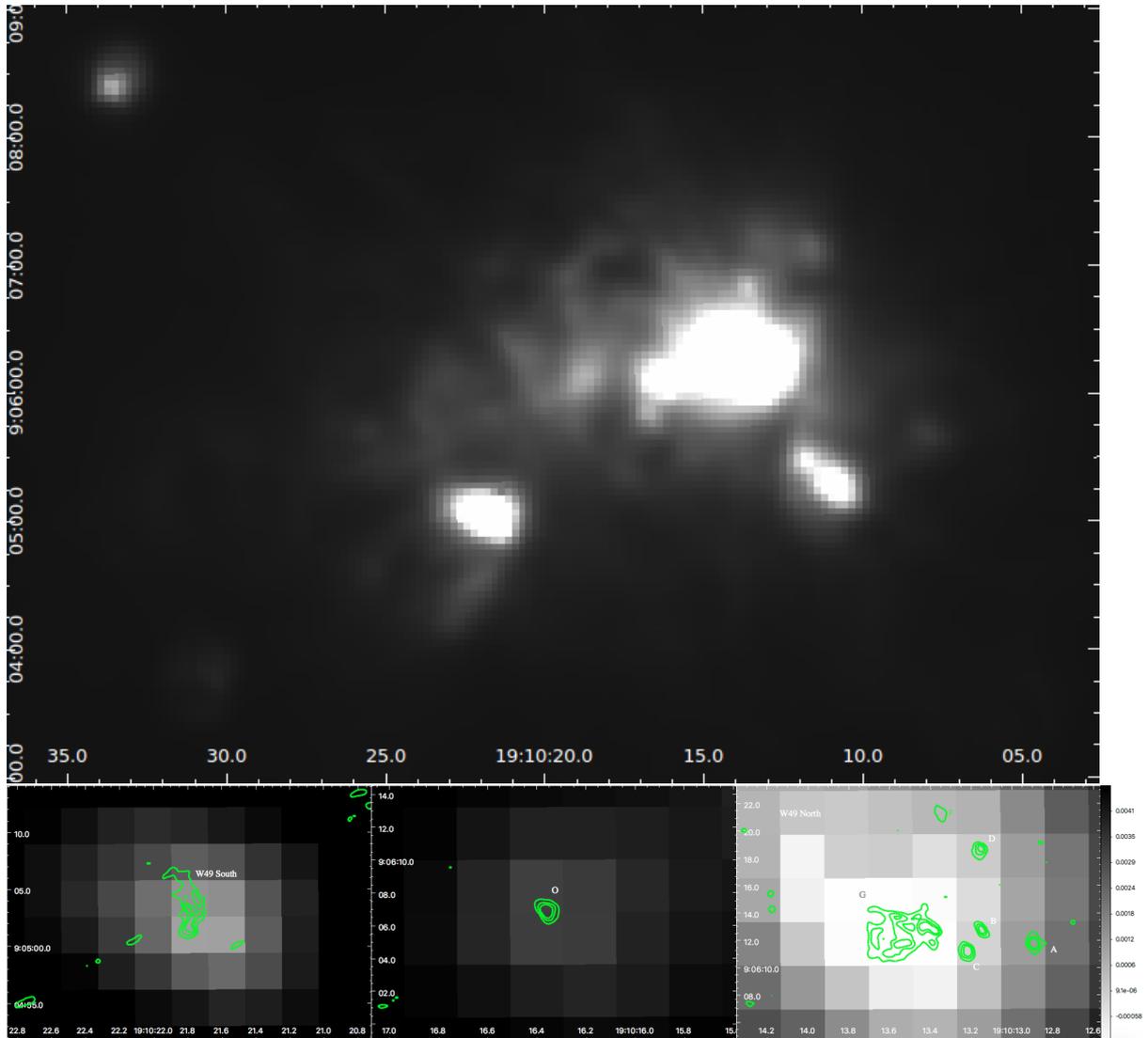


FIGURE 5.7: W49A in Herschel at $70\mu\text{m}$. The top image shows the complete field of view under discussion. The bottom images show closeups of the three regions of interest overlaid with KuGARS contours. Note: the colour scales in the closeup images have been adjusted to highlight the brightest pixels in each image. Of the three, the brightest pixels were found in the region containing W49A (right), the least bright were in the isolated source (middle).

(source B) has a strongly rising radio continuum spectrum ($\alpha \sim 1.1$) and has broad radio recombination lines with line widths ($> 45\text{km s}^{-1}$). This source is most likely a HC HII region. Since it has not been detected in CORNISH (i.e. a low frequency survey), there may be several others like it in KuGARS.

In addition to this analysis, we have attempted to compare fluxes derived from KuGARS with CORNISH and De Pree et al. (1997). However, we found that the KuGARS fluxes are systematically lower than CORNISH and De Pree et al. (1997) as can be seen in Figs. 5.4 and 5.5. The De Pree et al. (1997) and CORNISH images are much deeper and more sensitive to extended

emission than KuGARS. This can be seen by the presence of extended CORNISH and De Pree et al. (1997) emission outside the 3σ KuGARS contour. The KuGARS data are thus likely to underestimate the true flux of each object. Future improvements in cleaning and self-calibration of the data are required to adequately compare with CORNISH and De Pree et al. (1997) data.

Chapter 6

Conclusion

The primary goals in this thesis were to develop data reduction procedures for KuGARS data, detect and identify HII regions and put together a source catalogue. This study is essentially being carried out in order to improve our understanding of massive star formation.

In Chapters 2 and 3, we presented interferometry and the development of the data reduction procedures for KuGARS. Data reduction in CASA is not an arbitrary process, therefore, even for new techniques the data reduction will follow the same general steps and principle. The steps in our data reduction were, however, tailored specifically towards KuGARS. Because of the large volume of data, we focused on a pilot of the region around W49A and a subset of the complete frequency range. In chapter 3 we show detailed steps on how the calibration and imaging of the data at 13.99 GHz was carried out. In each step, data plots have been presented to aid the description of the calibration. The calibration was then verified in the quality assurance section of Chapter 4, where the frequently observed calibrators were imaged and compared to the literature. The flux of the flux calibrator (3C48) was 1.51Jy at 13.99 GHz with a signal to noise ratio of 108.8. The flux of the phase calibrator (J1922+1530) was 0.29Jy at 13.99 GHz with a signal to noise ratio of 129.5. These flux values were within 10% error of those recorded in the VLA calibrator manual at various frequencies.

In Chapter 3 we present a list of the 18 brightest sources around W49A at 13.99 GHz. In Chapter 4 we use the source finder AEGEAN to present the sources and brightest peaks found in the field around W49A at 13.99 GHz. We have found 35 islands with a total of 57 distinct components. These sources were clustered or found in three particular regions: the central

”ring” (W49 North), an isolated spherical source, and W49 South. Further analysis was carried out to determine the nature of some of the objects in Chapter 5

Chapter 5 was focused on analysing and determining the nature of the objects found in the KuGARS observations around W49A at 13.99 GHz. Here, we compared these results to the literature. We put into perspective the resolution achieved in the KuGARS observations, explored the nature of the detected radio emission, and cross matched the KuGARS sources with the literature to see what has previously been found. The nature of the radio emission when compared to GLIMPSE was separated into two categories: Extended radio emission embedded within a dusty cocoon of $8\mu\text{m}$ GLIMPSE emission, and isolated radio emission associated with no obvious concentration of GLIMPSE emission. In our comparison with Herschel at $70\mu\text{m}$ we found our sources deeply embedded in $70\mu\text{m}$ emission which is indicative of heated dust from star formation. Thus, both GLIMPSE and Herschel associate our sources with star formation. A survey of the sources previously found in the literature using the SIMBAD astronomical database found the following sources: HII regions, stars, molecular clouds, IR objects, radio objects, millimetre radio objects, YSOs, and masers. As our aim is to study massive star formation, we were particularly interested in objects tracing massive star formation that overlapped with our sources. We found that W49 North and W49 South had previous detections of HII regions, masers, and YSOs, thus confirming their star forming nature, but found Source O associated with only a radio source. This gave no particular indication of the nature of the source. However, we found that De Pree et al. (1997) classified this source as an UC HII region. Our comparisons with the radio observations, CORNISH and De Pree et al. (1997), revealed that our major detections are HII regions. We have also identified a region that is clearly detected at high frequencies (KuGARS and De Pree et al. 1997) but is not seen at 5 GHz (CORNISH). This is exactly the kind of object we were hoping to find. Evidently, this multiwavelength analysis, although at an early stage, shows evidence of star formation in our sources and thus reveals the rich star forming nature W49A and demonstrates the ability of KuGARS to detect steep positive spectrum objects.

Chapter 7

Future Work

7.1 Spectral Lines

Spectral lines are a phenomenon that presents itself as either narrow emission or absorption features in the spectrum of a gaseous source. The first spectral lines to be detected are what are now known as the “Fraunhofer lines”. These were first detected in astronomical objects in the spectra of the Sun and Stars, and since then the emission of spectral lines from cosmic objects is an essential tool in astronomy. The importance of this lies in that each line has a unique frequency. Thus, the object emitting the radiation, be it an atom, ion or molecule can be identified (Gordon and Sorochenko, 2009). Some of the more common examples of spectral lines include the $\lambda = 21\text{cm}$ hyperfine line of interstellar hydrogen (HI) and recombination lines of ionized hydrogen (HII) as well as recombination lines of heavier atoms such Helium, Lithium or Carbon. Note that, “spectral lines are intrinsically quantum phenomena. They form an idealised concept and much like perfectly straight lines in geometry, they do not exist in the real world. The fundamental characteristics of radio spectral lines from interstellar atoms and molecules can be derived from fairly simple applications of quantum mechanics and thermodynamics. Ergo, spectral lines make powerful diagnostics of physical and chemical conditions in astronomical objects”¹.

Evidently, there are as many spectral lines in the electromagnetic spectrum as there are cosmic atoms and molecules. However, the ones of interest to us in this case are a special class resulting from transitions between highly excited atomic levels. The idea is that these lines appear after the recombination of ions and electrons to form atoms, leaving the electrons with high principal

¹<http://www.cv.nrao.edu/course/astr534/Recombination.html>

quantum numbers. As these newly bound electrons drop levels, they lose energy in each drop by radiating it away in the form of a spectral line. When these lines appear in the radio regime, they are called “radio recombination lines” (Gordon and Sorochenko, 2009).

In KuGARS, we aim at investigating Radio Recombination lines in HII regions, specifically HC HII regions. This will involve the analysis of RRLs detected in the survey and will as a consequence involve determining the physical properties (electron temperature and density) of UC and HC H II regions detected in KuGARS - particularly deriving physical properties for thermal bremsstrahlung.

Aims:

- Detect and analyse radio recombination lines in the transitions from $H81\alpha$ to $H74\alpha$
- Derive physical properties for thermal bremsstrahlung

Other spectral line studies will include detecting and analysing,

- 12.2 GHz CH_3OH masers
- 13.4 GHz OH masers and
- the 14.4 GHz H_2CO line in absorption.

7.2 Data Reduction and Future Analyses

We have made significant advances in the development of a data reduction pipeline for KuGARS, as is evident in the contents of this thesis. In our analysis we have shown that our sources are all associated with star formation. In particular, our comparisons with CORNISH and De Pree et al. (1997) reveal that our major detections are HII regions. Furthermore, we see a source detected in KuGARS and De Pree et al. (1997) but not in CORNISH (5 GHz). With a strongly rising radio continuum spectrum ($\alpha \sim 1.1$) and broad radio recombination lines ($> 45\text{kms}^{-1}$), this source is most likely a HC HII region. We, therefore, have demonstrated KuGARS ability to observe star formation and detect these rare objects. Nonetheless, we need to improve the procedures and develop further steps to include multiple spectral windows and explore self-calibration. As can be seen in Figs. 5.4 and 5.5, we are underestimating the fluxes, and therefore need to improve the cleaning.

We have advanced to the point that we are able to identify radio objects, and we are able to identify potential HC HII regions since we can see objects that CORNISH cannot see. However, we cannot determine the spectral indices of the observed objects and therefore cannot fully determine whether they are in fact HC HII regions or other types of objects. Nonetheless, with our improved cleaning methods, we will be able to adequately compare our flux measurements with CORNISH and De Pree et al. (1997). With our developments on the pipeline to include multiple spectral windows, we will be able to determine in-band spectral indices which has the advantage that there is no time gap as is the case with the respective observations in this analysis: ~ 10 years between De Pree and CORNISH and ~ 8 years between CORNISH and KuGARS. Thus, we will be able to confidently determine the types of objects observed.

We will, therefore, need to experiment with different strategies for cleaning including self-calibration in order to make sure we are recovering all the flux. With this we will be able to compare fluxes with CORNISH and De Pree et al. (1997). Furthermore, we will develop the procedures further to include more spectral windows. The data was split into five 0.7 GHz bands in order to allow us to determine in-band spectral indices. Therefore, our future goals and developments include:

- Improving our general data reduction steps
- Developing self-calibration procedures
- Developing data reduction procedures for multiple spectral windows
- Reducing the remaining data
- Developing procedures for spectral index analysis
- Investigating modifications to standard procedures for On the Fly Mapping
- Exploring the full extent of steep positive spectrum objects such HC HII regions in KuGARS.

Appendix A

LISTOBS Output

TABLE A.1: listobs of the entire block

```

=====
MeasurementSet Name: /local/home/mubela/Data_Reduction/JVLAdata/new_blocks/78sb/spw2/new_run2/15A-043.78.spw2.ms      MS Version 2
=====
Observer: Dr. Mark T. Thompson      Project: uid://evla/pdb/30098882
Observation: EVLA
Data records: 2622321      Total elapsed time = 8641 seconds
Observed from 29-Mar-2015/16:19:31.0 to 29-Mar-2015/18:43:32.0 (UTC)

ObservationID = 0      ArrayID = 0
Date      Timerange (UTC)      Scan  FldId FieldName      nRows  SpwIds  Average Interval(s)  ScanIntent
29-Mar-2015/16:19:31.0 - 16:20:30.0      2      0 0137+331=3C48      20709  [0] [1] [SYSTEM_CONFIGURATION#UNSPECIFIED]
      16:24:51.0 - 16:25:19.0      5      0 0137+331=3C48      9828  [0] [1] [SYSTEM_CONFIGURATION#UNSPECIFIED]
      16:25:20.0 - 16:30:19.0      6      0 0137+331=3C48      104949 [0] [1] [CALIBRATE_FLUX#UNSPECIFIED]
      16:35:19.0 - 16:35:48.0      9      1 J2253+1608      10179 [0] [1] [SYSTEM_CONFIGURATION#UNSPECIFIED]
      16:35:49.0 - 16:40:47.0      10     1 J2253+1608      104598 [0] [1] [CALIBRATE_BANDPASS#UNSPECIFIED]
      16:46:32.0 - 16:47:01.0      13     2 J1922+1530      10179 [0] [1] [SYSTEM_CONFIGURATION#UNSPECIFIED]
      16:47:02.0 - 16:48:11.0      14     2 J1922+1530      24219 [0] [1] [CALIBRATE_AMPLI#UNSPECIFIED,CALIBRATE_PHASE#UNSPECIFIED]
      16:48:12.0 - 16:48:51.0      15     3 kugars_sp41_st      13689 [0] [1] [SYSTEM_CONFIGURATION#UNSPECIFIED]
      16:48:51.0 - 16:49:01.0      17     4 OTFDUMMY      3510  [0] [1] [SYSTEM_CONFIGURATION#UNSPECIFIED]
      16:49:01.0 - 16:49:11.0      18     5 1905349+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:49:11.0 - 16:49:21.0      19     6 1905453+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:49:21.0 - 16:49:31.0      20     7 1905557+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:49:31.0 - 16:49:41.0      21     8 1906061+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:49:41.0 - 16:49:51.0      22     9 1906165+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:49:51.0 - 16:50:01.0      23    10 1906269+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:50:01.0 - 16:50:11.0      24    11 1906374+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:50:11.0 - 16:50:21.0      25    12 1906478+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:50:21.0 - 16:50:31.0      26    13 1906582+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:50:31.0 - 16:50:41.0      27    14 1907086+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:50:41.0 - 16:50:51.0      28    15 1907190+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:50:51.0 - 16:51:01.0      29    16 1907294+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:51:01.0 - 16:51:11.0      30    17 1907398+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
      16:51:11.0 - 16:51:21.0      31    18 1907503+090430      3510  [0] [1] [OBSERVE_TARGET#UNSPECIFIED]
=====

```

TABLE A.1: Continued

16:51:21.0 - 16:51:31.0	32	19 1908007+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:51:31.0 - 16:51:41.0	33	20 1908111+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:51:41.0 - 16:51:51.0	34	21 1908215+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:51:51.0 - 16:52:01.0	35	22 1908319+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:52:01.0 - 16:52:11.0	36	23 1908423+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:52:11.0 - 16:52:21.0	37	24 1908527+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:52:21.0 - 16:52:31.0	38	25 1909032+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:52:31.0 - 16:52:41.0	39	26 1909136+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:52:41.0 - 16:52:51.0	40	27 1909240+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:52:51.0 - 16:53:01.0	41	28 1909344+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:53:01.0 - 16:53:11.0	42	29 1909448+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:53:11.0 - 16:53:21.0	43	30 1909552+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:53:21.0 - 16:53:31.0	44	31 1910057+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:53:31.0 - 16:53:41.0	45	32 1910161+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:53:41.0 - 16:53:51.0	46	33 1910265+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:53:51.0 - 16:54:01.0	47	34 1910369+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:54:01.0 - 16:54:11.0	48	35 1910473+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:54:11.0 - 16:54:21.0	49	36 1910577+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:54:21.0 - 16:54:31.0	50	37 1911081+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:54:31.0 - 16:54:41.0	51	38 1911186+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:54:41.0 - 16:54:51.0	52	39 1911290+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:54:51.0 - 16:55:01.0	53	40 1911394+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:55:01.0 - 16:55:11.0	54	41 1911498+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:55:11.0 - 16:55:21.0	55	42 1912002+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:55:21.0 - 16:55:31.0	56	43 1912106+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:55:31.0 - 16:55:41.0	57	44 1912211+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:55:41.0 - 16:55:51.0	58	45 1912315+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:55:51.0 - 16:56:01.0	59	46 1912419+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:56:01.0 - 16:56:11.0	60	47 1912523+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:56:11.0 - 16:56:21.0	61	48 1913027+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:56:21.0 - 16:56:31.0	62	49 1913131+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:56:31.0 - 16:56:41.0	63	50 1913235+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]

TABLE A.1: Continued

16:56:41.0 - 16:56:51.0	64	51 1913340+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:56:51.0 - 16:57:01.0	65	52 1913444+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:57:01.0 - 16:57:11.0	66	53 1913548+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:57:11.0 - 16:57:21.0	67	54 1914052+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:57:21.0 - 16:57:31.0	68	55 1914156+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:57:31.0 - 16:57:41.0	69	56 1914260+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:57:41.0 - 16:57:51.0	70	57 1914364+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:57:51.0 - 16:58:01.0	71	58 1914469+090430	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
16:58:02.0 - 16:59:11.0	72	2 J1922+1530	24219	[0]	[1]	[CALIBRATE_AMPLI#UNSPECIFIED,CALIBRATE_PHASE#UNSPECIFIED]
16:59:12.0 - 16:59:51.0	73	59 kugars_sp42_st	13689	[0]	[1]	[SYSTEM_CONFIGURATION#UNSPECIFIED]
16:59:51.0 - 17:00:01.0	75	60 OTFDUMMY	3510	[0]	[1]	[SYSTEM_CONFIGURATION#UNSPECIFIED]
17:00:01.0 - 17:00:11.0	76	61 1905386+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:00:11.0 - 17:00:21.0	77	62 1905490+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:00:21.0 - 17:00:31.0	78	63 1905594+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:00:31.0 - 17:00:41.0	79	64 1906099+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:00:41.0 - 17:00:51.0	80	65 1906203+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:00:51.0 - 17:01:01.0	81	66 1906307+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:01:01.0 - 17:01:11.0	82	67 1906411+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:01:11.0 - 17:01:21.0	83	68 1906515+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:01:21.0 - 17:01:31.0	84	69 1907019+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:01:31.0 - 17:01:41.0	85	70 1907124+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:01:41.0 - 17:01:51.0	86	71 1907228+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:01:51.0 - 17:02:01.0	87	72 1907332+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:02:01.0 - 17:02:11.0	88	73 1907436+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:02:11.0 - 17:02:21.0	89	74 1907540+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:02:21.0 - 17:02:31.0	90	75 1908044+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:02:31.0 - 17:02:41.0	91	76 1908148+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:02:41.0 - 17:02:51.0	92	77 1908253+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:02:51.0 - 17:03:01.0	93	78 1908357+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:03:01.0 - 17:03:11.0	94	79 1908461+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:03:11.0 - 17:03:21.0	95	80 1908565+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]

(nRows = Total number of rows per scan)

TABLE A.1: Continued

17:03:21.0 - 17:03:31.0	96	81	1909069+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:03:31.0 - 17:03:41.0	97	82	1909173+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:03:41.0 - 17:03:51.0	98	83	1909278+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:03:51.0 - 17:04:01.0	99	84	1909382+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:04:01.0 - 17:04:11.0	100	85	1909486+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:04:11.0 - 17:04:21.0	101	86	1909590+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:04:21.0 - 17:04:31.0	102	87	1910094+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:04:31.0 - 17:04:41.0	103	88	1910198+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:04:41.0 - 17:04:51.0	104	89	1910302+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:04:51.0 - 17:05:01.0	105	90	1910407+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:05:01.0 - 17:05:11.0	106	91	1910511+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:05:11.0 - 17:05:21.0	107	92	1911015+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:05:21.0 - 17:05:31.0	108	93	1911119+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:05:31.0 - 17:05:41.0	109	94	1911223+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:05:41.0 - 17:05:51.0	110	95	1911327+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:05:51.0 - 17:06:01.0	111	96	1911431+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:06:01.0 - 17:06:11.0	112	97	1911536+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:06:11.0 - 17:06:21.0	113	98	1912040+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:06:21.0 - 17:06:31.0	114	99	1912144+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:06:31.0 - 17:06:41.0	115	100	1912248+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:06:41.0 - 17:06:51.0	116	101	1912352+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:06:51.0 - 17:07:01.0	117	102	1912456+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:07:01.0 - 17:07:11.0	118	103	1912561+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:07:11.0 - 17:07:21.0	119	104	1913065+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:07:21.0 - 17:07:31.0	120	105	1913169+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:07:31.0 - 17:07:41.0	121	106	1913273+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]

(nRows = Total number of rows per scan)

TABLE A.1: Continued

17:07:41.0 - 17:07:51.0	122	107	1913377+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:07:51.0 - 17:08:01.0	123	108	1913481+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:08:01.0 - 17:08:11.0	124	109	1913585+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:08:11.0 - 17:08:21.0	125	110	1914090+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:08:21.0 - 17:08:31.0	126	111	1914194+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:08:31.0 - 17:08:41.0	127	112	1914298+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:08:41.0 - 17:08:51.0	128	113	1914402+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:08:51.0 - 17:09:01.0	129	114	1914506+090616	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:09:02.0 - 17:10:11.0	130	2	J1922+1530	24219	[0]	[1]	[CALIBRATE_AMPLI#UNSPECIFIED,CALIBRATE_PHASE#UNSPECIFIED]
17:10:12.0 - 17:10:51.0	131	115	kugars_sp43_st	13689	[0]	[1]	[SYSTEM_CONFIGURATION#UNSPECIFIED]
17:10:51.0 - 17:11:01.0	133	116	OTFDUMMY	3510	[0]	[1]	[SYSTEM_CONFIGURATION#UNSPECIFIED]
17:11:01.0 - 17:11:11.0	134	117	1905424+090802	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:11:11.0 - 17:11:21.0	135	118	1905528+090802	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:11:21.0 - 17:11:31.0	136	119	1906032+090802	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:11:31.0 - 17:11:41.0	137	120	1906136+090802	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
17:11:41.0 - 17:11:51.0	138	121	1906240+090802	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
.							
.							
[TRUNCATED]							
.							
.							
18:41:32.0 - 18:41:42.0	593	558	1914390+092024	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
18:41:42.0 - 18:41:52.0	594	559	1914494+092024	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
18:41:52.0 - 18:42:02.0	595	560	1914599+092024	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
18:42:02.0 - 18:42:12.0	596	561	1915103+092024	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
18:42:12.0 - 18:42:22.0	597	562	1915207+092024	3510	[0]	[1]	[OBSERVE_TARGET#UNSPECIFIED]
18:42:23.0 - 18:43:32.0	598	2	J1922+1530	24219	[0]	[1]	[CALIBRATE_AMPLI#UNSPECIFIED,CALIBRATE_PHASE#UNSPECIFIED]

(nRows = Total number of rows per scan)

TABLE A.1: Continued

Fields: 563

ID	Code Name	RA	Decl	Epoch	SrcId	nRows
0	NONE 0137+331=3C48	01:37:41.299431	+33.09.35.13299	J2000	0	135486
1	NONE J2253+1608	22:53:57.747932	+16.08.53.56089	J2000	1	114777
2	NONE J1922+1530	19:22:34.699188	+15.30.10.03213	J2000	2	303264
3	NONE kugars_sp41_st	19:05:34.911380	+09.04.30.30108	J2000	3	13689
4	NONE OTFDUMMY	19:05:24.496285	+09.04.30.30108	J2000	4	3510
5	NONE 1905349+090430	19:05:34.911380	+09.04.30.30108	J2000	5	3510

[TRUNCATED]

561	NONE 1915103+092024	19:15:10.321886	+09.20.24.30108	J2000	561	3510
562	NONE 1915207+092024	19:15:20.736980	+09.20.24.30108	J2000	562	3510

Spectral Windows: (1 unique spectral windows and 1 unique polarization setups)

SpwID	Name	#Chans	Frame	Ch0(MHz)	ChanWid(kHz)	TotBW(kHz)	CtrFreq(MHz)	BBC	Num	Corrs
0	EVLA_KU#A1C1#2	64	TOPO	13926.765	2000.000	128000.0	13989.7651	10	RR	RL LR LL

Sources: 563

ID	Name	SpwId	RestFreq(MHz)	SysVel(km/s)
0	0137+331=3C48	0	14950	0
1	J2253+1608	0	14950	0
2	J1922+1530	0	14950	0
3	kugars_sp41_st	0	14950	0
4	OTFDUMMY	0	14950	0
5	1905349+090430	0	14950	0

[TRUNCATED]

561	1915103+092024	0	14950	0
562	1915207+092024	0	14950	0

TABLE A.1: Continued

Antennas: 27:

ID	Name	Station	Diam.	Long.	Lat.	Offset from array center (m)			ITRF Geocentric coordinates (m)		
						East	North	Elevation	x	y	z
0	ea01	N12	25.0 m	-107.37.09.0	+33.54.30.0	-107.1732	870.2603	-7.3307	-1601110.040800	-5041488.081600	3555597.436600
1	ea02	N16	25.0 m	-107.37.10.9	+33.54.48.0	-155.8505	1426.6391	-9.3851	-1601061.955600	-5041175.881400	3556058.032500
2	ea03	W28	25.0 m	-107.39.20.2	+33.52.46.6	-3473.3006	-2324.6512	23.0349	-1604865.655200	-5042190.032000	3552962.363500
3	ea04	E04	25.0 m	-107.37.00.8	+33.53.59.7	102.7988	-63.7797	-2.6157	-1601068.805000	-5042051.934600	3554824.840100
4	ea05	E20	25.0 m	-107.35.43.6	+33.53.29.9	2082.1626	-987.0179	9.4300	-1599340.812800	-5043150.959000	3554065.220400
5	ea06	N20	25.0 m	-107.37.13.2	+33.55.09.5	-214.2279	2092.8875	-10.7846	-1601004.701400	-5040802.793900	3556610.139200
6	ea07	N32	25.0 m	-107.37.22.0	+33.56.33.6	-441.7289	4689.9702	-16.9402	-1600781.046100	-5039347.430600	3558761.525300
7	ea08	E28	25.0 m	-107.34.29.3	+33.53.04.9	3732.7709	-1757.3245	21.4270	-1597899.901200	-5044068.679300	3553432.455400
8	ea09	W08	25.0 m	-107.37.21.6	+33.53.53.0	-432.1156	-272.1458	-1.4994	-1601614.091200	-5042001.656900	3554652.514300
9	ea10	E36	25.0 m	-107.33.20.2	+33.52.34.3	5761.3884	-2704.6607	33.0163	-1596127.715100	-5045193.742400	3552652.431600
10	ea11	W12	25.0 m	-107.37.37.4	+33.53.44.2	-835.3874	-544.2181	0.5592	-1602044.909700	-5042025.788200	3554427.830700
11	ea12	N36	25.0 m	-107.37.25.6	+33.57.07.6	-533.8103	5740.7449	-18.9619	-1600690.596500	-5038758.724500	3559632.064800
12	ea13	W04	25.0 m	-107.37.10.8	+33.53.59.1	-152.8649	-83.7944	-2.4632	-1601315.895500	-5041985.311370	3554808.312700
13	ea14	E08	25.0 m	-107.36.48.9	+33.53.55.1	407.8280	-206.0296	-3.2233	-1600801.931400	-5042219.381700	3554706.431200
14	ea15	E24	25.0 m	-107.35.13.4	+33.53.18.1	2858.1769	-1349.1349	13.7075	-1598663.092300	-5043581.378100	3553767.008200
15	ea16	W24	25.0 m	-107.38.49.0	+33.53.04.0	-2673.3313	-1784.6199	10.4898	-1604008.735300	-5042135.843900	3553403.698000
16	ea18	W36	25.0 m	-107.40.32.6	+33.52.06.0	-5328.6226	-3577.7610	7.0250	-1606841.958500	-5042279.678700	3551913.023700
17	ea19	W16	25.0 m	-107.37.57.4	+33.53.33.0	-1348.7132	-890.6192	1.2955	-1602592.854200	-5042054.990800	3554140.704000
18	ea20	N04	25.0 m	-107.37.06.5	+33.54.06.1	-42.6102	132.8507	-3.5389	-1601173.967700	-5041902.666400	3554987.529200
19	ea21	E32	25.0 m	-107.34.01.5	+33.52.50.3	4701.6616	-2209.7365	25.2061	-1597053.119800	-5044604.698200	3553058.977000
20	ea22	N24	25.0 m	-107.37.16.1	+33.55.37.7	-290.3773	2961.8599	-12.2422	-1600930.088900	-5040316.392900	3557330.385800
21	ea23	E16	25.0 m	-107.36.09.8	+33.53.40.0	1410.0331	-673.4694	-0.7949	-1599926.107500	-5042772.964400	3554319.789100
22	ea24	W32	25.0 m	-107.39.54.8	+33.52.27.2	-4359.4446	-2923.1293	11.7722	-1605808.640900	-5042230.090000	3552459.205500
23	ea25	W20	25.0 m	-107.38.21.4	+33.53.19.5	-1963.5551	-1305.5777	4.0943	-1603249.673100	-5042091.416000	3553797.806000
24	ea26	E12	25.0 m	-107.36.31.7	+33.53.48.5	848.7020	-411.6090	-2.7582	-1600416.513100	-5042462.423500	3554536.052600
25	ea27	N08	25.0 m	-107.37.07.5	+33.54.15.8	-68.9087	433.1982	-5.0729	-1601147.938500	-5041733.821100	3555235.956600
26	ea28	N28	25.0 m	-107.37.18.7	+33.56.02.5	-357.6272	3729.2179	-15.4017	-1600863.699800	-5039885.312600	3557965.308800

Bibliography

- Benjamin, R.A., Churchwell, E., Babler, B.L., et al., 2003. GLIMPSE. I. An SIRTf Legacy Project to Map the Inner Galaxy. *PASP*, 115:953.
- Bonnell, I.A., Bate, M.R., and Zinnecker, H., 1998. On the formation of massive stars. *MNRAS*, 298:93.
- Braatz, J., 2014. Bandpass Calibration: Hands-on. *NRAO*.
- Churchwell, E., Babler, B.L., Meade, M.R., et al., 2009. The Spitzer/GLIMPSE Surveys: A New View of the Milky Way. *PASP*, 121:213.
- Cohen, M., et al., et al., et al., 2007. Absolute diffuse calibration of IRAC through mid-infrared and radio study of HII regions. *MNRAS*, 374:979.
- Condon, J.J., 1998. Radio Surveys. *IAU*, 179:19.
- De Pree, C., Mehringer, D.M., and Goss, W., 1997. Multifrequency, High-Resolution Radio Recombination Line Observations of the Massive Star-forming Region W49A. *ApJ*, 482:307.
- De Pree, C.G., Peters, T., Mac Low, M.M., et al., 2014. Flickering of 1.3 cm Sources in Sgr B2: Toward a Solution to the Ultracompact H II Region Lifetime Problem. *ApJ*, 781:L36.
- Dreher, J.W., Johnston, K.J., Welch, W.J., et al., 1984. Ultracompact structure in the H II region W49N. *ApJ*, 283:632.
- Gaume, R.A., Goss, W.M., Dickel, H.R., et al., 1995. The Pre-Main-Sequence Evolution of Intermediate-Mass Stars. *ApJ*, 438:2:776.
- Ginsburg, A., Darling, J., Battersby, C., et al., 2011. Galactic H₂CO Densitometry. I. Pilot Survey of Ultracompact H II Regions and Methodology. *ApJ*, 736:149.
- Gordon, M.A. and Sorochenko, R.L., 2009. Radio Recombination Lines. *ASSL*, 282.

- Green, J.A., McClure-Griffiths, N.M., Caswell, J.L., et al., 2012. MAGMO: coherent magnetic fields in the star-forming regions of the Carina-Sagittarius spiral arm tangent. *MNRAS*, 425:2530.
- Hancock, P.J., Murphy, T., Gaensler, B.M., et al., 2012. Compact continuum source finding for next generation radio surveys. *MNRAS*, 422:1812.
- Hindson, L., Thompson, M.A., Urquhart, J.S., et al., 2012. The G305 star-forming complex: a wide-area radio survey of ultracompact H II regions. *MNRAS*, 421:3418.
- Hoare, M.G., Kurtz, S.E., Lizano, S., et al., 2007. Ultracompact HII Regions and the Early Lives of Massive Stars. *Protostars and Planets*, V:181.
- Hoare, M.G., Purcell, C.R., Churchwell, E.B., et al., 2012. The Coordinated Radio and Infrared Survey for High-Mass Star Formation (The CORNISH Survey). I. Survey Design. *PASP*, 124:939.
- Högbom, J., 1974. Aperture Synthesis With A Non-Regular Distribution of Interferometer Baselines. *A&AS*, 15:417.
- Keto, E., 2003. The Formation of Massive Stars by Accretion through Trapped Hypercompact H II Regions. *ApJ*, 599:1196.
- Krumholz, M.R., 2016. Notes On Star Formation. *The Open Astrophysics Bookshelf*.
- Kuiper, R., Klahr, H., Beuther, H., et al., 2012. A solution to the radiation pressure problem in the formation of massive stars. *arXiv*, page 1211.7064.
- Kurtz, S., 2000. Ultracompact H II Regions: New Challenges. *Astrophysical Plasmas: Codes, Models, and Observations*, 9:169.
- Kurtz, S., 2002. Ultracompact H II Regions. *ASP*, 267:81.
- Kurtz, S., 2005a. Hypercompact HII Regions. *IAU*, 227:111.
- Kurtz, S., 2005b. The Young Massive Star Environment. *IAU*, 231:47.
- Kurtz, S., Cesaroni, R., Churchwell, E., et al., 2000. Hot Molecular Cores and the Earliest Phases of High-Mass Star Formation. *Protostars and Planets*, IV:299.
- Kurtz, S. and Franco, J., 2002. Ultracompact H II Regions. *RevMexAA(SC)*, 12:16.

- Lumsden, S.I., Hoare, M.G., Urquhart, J.S., et al., 2013. The Red MSX Source Survey: The Massive Young Stellar Population of Our Galaxy. *ApJS*, 208:17pp.
- Mangum, J., Emerson, D., and Greisen, E., 2000. The On-The-Fly Imaging Technique. *ASP*, 217:179.
- Mercer, E.P., Clemens, D.P., et al., et al., 2005. New Star Clusters Discovered in the GLIMPSE Survey. *ApJ*, 635:560.
- Messineo, M., Davies, B., Ivanov, V.D., et al., 2009. Near-Infrared Spectra of Galactic Stellar Clusters Detected on Spitzer/GLIMPSE Images. *ApJ*, 147:471.
- Mezger, P.G. and Henderson, A.P., 1967. Galactic H II Regions. I. Observations of Their Continuum Radiation at the Frequency 5 GHz. *ApJ*, 147:471.
- Molinari, S., et al., et al., et al., 2010. Clouds, filaments, and protostars: The Herschel Hi-GAL Milky Way. *A&A*, 518:L100.
- Motte, F., Bontemps, S., and Louvet, F., 2017. High-Mass Star and Massive Cluster Formation in the Milky Way. *eprint arXiv:1706.00118*.
- Palla, F. and Stahler, S.W., 1993. The Pre-Main-Sequence Evolution of Intermediate-Mass Stars. *ApJ*, 418:414.
- Pandian, J.D. and Goldsmith, Paul F. and Deshpande, A.A., 2007. The Arecibo Methanol Maser Galactic Plane Survey. I. Data. *ApJ*, 656:255.
- Purcell, C.R., Hoare, M. G. and Cotton, W.D., Lumsden, S.L., et al., 2013. The Coordinated Radio and Infrared Survey for High-mass Star Formation. II. Source Catalog. *ApJS*, 205:1 (22pp).
- Reid, M.J., Menten, K.M., Brunthaler, A., et al., 2014. Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way. *ApJ*, 783:130.
- Sanchez-Monge, A., et al., and et al., 2013. . *ApJ*, 721.
- Sewilo, M., Churchwell, E., Kurtz, S., et al., 2004. Broad Radio Recombination Lines from Hypercompact H II Regions. *ApJ*, 605:285.
- Sjouwerman, L.O. and Mills, E.A.C., 2013. Galactic kU-band Thermal Survey (GUTS). *arXiv*, 1312.6710.

- Stahler, S.W., Palla, F., and Ho, P.T.P., 2000. The Formation of Massive Stars. *Protostars and Planets*, IV:327.
- Stahler, S.W., Shu, F.H., and Taam, R.E., 1980a. The evolution of protostars. I - Global formulation and results. *ApJ*, 241:637.
- Stahler, S.W., Shu, F.H., and Taam, R.E., 1980b. The evolution of protostars. II - The hydrostatic core. *ApJ*, 242:226.
- Thompson, A.R., Moran, J.M., and Swenson, George. W., J., 2017. Interferometry and Synthesis in Radio Astronomy, 3rd Edition. *A&A*.
- Thompson, M.A. and Goedhart, S., 2016. MeerGAL: the MeerKAT Galactic Plane Survey. *PoS*.
- Urquhart, J.S., Hoare, M.G., Lumsden, S.L., et al., 2009. The RMS survey. H₂O masers towards a sample of southern hemisphere massive YSO candidates and ultra compact HII regions. *A&A*, 507:795.
- Urquhart, J.S., Thompson, M.A., Moore, T.J.T., et al., 2013. ATLASGAL - properties of compact H II regions and their natal clumps. *MNRAS*, 435:400.
- Wood, D.O.S. and Churchwell, E., 1989. The Morphologies and Physical Properties of Ultra-compact H II Regions. *ApJS*, 69:831.